



**HAL**  
open science

# Empirical Bayesian regularization of the inverse acoustic problem

Antonio A. Pereira, Jérôme Antoni, Q. Leclere

► **To cite this version:**

Antonio A. Pereira, Jérôme Antoni, Q. Leclere. Empirical Bayesian regularization of the inverse acoustic problem. *Applied Acoustics*, 2015, pp.11-29. 10.1016/j.apacoust.2015.03.008 . hal-01176879

**HAL Id: hal-01176879**

**<https://hal.science/hal-01176879v1>**

Submitted on 16 Jul 2015

**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.

# Empirical Bayesian regularization of the inverse acoustic problem

A. Pereira<sup>a,\*</sup>, J. Antoni<sup>a</sup>, Q. Leclère<sup>a</sup>

<sup>a</sup>Laboratoire Vibrations Acoustique, INSA Lyon, 25 bis avenue Jean Capelle F-69621 Villeurbanne Cedex, FRANCE

---

## Abstract

This paper answers the challenge as how to automatically select a good regularization parameter when solving inverse problems in acoustics. A Bayesian solution is proposed that consists either in directly finding the most probable value of the regularization parameter or, indirectly, in estimating it as the ratio of the most probable values of the noise and source expected energies. It enjoys several nice properties such as ease of implementation and low computational complexity (the proposed algorithm boils down to searching for the global minimum of a 1D cost function). Among other advantages of the Bayesian approach, it makes possible to appraise the sensitivity of the reconstructed acoustical quantities of interest with respect to regularization, a performance that would be otherwise arduous to achieve.

### Keywords:

Inverse problems, regularization, Tikhonov regularization, Bayesian probabilities, source identification, acoustical holography.

---

## 1. Introduction

The inverse acoustic problem aims at reconstructing an acoustic quantity of interest (e.g. parietal pressure, particle velocity, acoustical intensity) from a limited number of remote measurements – as typically returned by an array of microphones or probes. As well-known, this is an ill-posed problem in the sense of Hadamard – i.e. it may have no solution at all or the solution may not be unique and it may be extremely sensitive to slight changes in the initial conditions [1] – for its exact solution would require measuring the complete field over a surface enclosing the source(s) of interest. As a consequence, solutions of an ill-posed inverse acoustic problems are typically found unstable with respect to very small perturbations in the measurements. Since ill-posedness fundamentally results from unavoidable loss of information during the measurement process, the usual cure is to *regularize* – i.e. to modify – the inverse operator so as to control the magnitude or energy of the expected solution within plausible limits [2–11]. In practice, the resort to regularization is just as essential as it is difficult and, in many aspects, it appertains as much to an art as to exact science.

The prevailing approach in acoustics and vibration is surely the popular Tikhonov regularization (control of the energy of the solution) [1, 12–17]. A critical aspect of Tikhonov regularization – actually shared by most regular-

---

\*Corresponding author. E-mail address: antonio.pereira@insa-lyon.fr.

ization techniques – is how to automatically determine the amount of regularization to be introduced in the system, which translates into the determination of a “regularization parameter”. Several strategies have been developed in this perspective, however, at the present time, there is still no absolutely universal method that is robust enough and always produces good results. Amongst the parameter choice methods used in the field of acoustics and vibration, the Generalized Cross Validation (GCV) [18] and the L-curve [19] seem to prevail largely, although other methods have been investigated such as the Normalized Cumulative Periodogram (NCP) [20, 21] and the Morozov discrepancy principle. The latter depends on a good estimate of the measurement noise level, that may not be available in practice. NCP is a relatively recent method whose idea is to track the aspect of the residual error as the regularization parameter changes and select the parameter for which the residual can be considered as close as possible to white noise. Several papers in the literature provide comparisons of different parameter choice methods, either applied in acoustics [10, 22–26] or in vibrations [27]. A general conclusion is that the behavior of each method is very problem-dependent and no consensus on which one is the best has been reached. Indeed, the inverse acoustic problem is sometimes so much ill-posed that choosing a proper regularization strategy can make a real difference. Recent publications in acoustics propose different regularization techniques, such as iterative methods (beneficial when dealing with large-scale problems) [28], Tikhonov regularization in its general form (i.e. by the use of discrete smoothing norms) [2], and a sparse regularization techniques [29–31], to cite only a few. Most of them still depend on either a regularization parameter that must be optimally tuned or on a stopping rule for the iterative methods. In a more general context (i.e. outside the field of acoustics), reference [32] provides an extensive comparison of several parameter choice methods by means of a large simulation study.

This paper introduces a Bayesian approach to regularization that is conceptually rather different from former methodologies that have been prevailing in inverse acoustics. The key idea is to conceive regularization as the introduction of prior information to compensate for the loss of information resulting from the measurement process; this is achieved in the form of a probability density function that characterizes all physically plausible values of the solution before the inverse problem is solved. The solution of the inverse problem – including the reconstruction of the acoustical quantity of interest plus the optimal tuning of the regularization parameter – is then found as the most probable values that comply with both the measurements and the prior information. In the special case of a Gaussian prior – which is investigated only in this paper – the proposed Bayesian regularization scheme boils down to the same structure as Tikhonov regularization, yet with the definite advantage of providing rigorous criteria for automatically tuning the regularization parameter. It is shown in this paper that the proposed Bayesian regularization enjoys several advantages as compared to other criteria traditionally used in inverse acoustics such as GCV and the L-curve:

- for a large range of acoustical configurations (simulation and experiments), it generally returns a regularized solution which is (in the least-square sense) closer to the optimal one,
- for a large range of physical/acoustical parameters (level of noise, frequency range, degree of ill-posedness) it is generally more robust to (strong) additive noise,

- it lends itself to easy implementation for it boils down to searching for the *global minimum* of a 1-D cost function,
- it is fully automatic and does not involve any tuning parameter (it actually *returns* the noise level and expected source energy as byproducts).

These advantages surely deserve a thorough introduction of Bayesian regularization to the acoustical community, even though the mathematical apparatus required may seem far from the acoustical discipline. As a consequence, the first part of the paper (in particular section 2) is presented as a tutorial before some novel theoretical and experimental results are introduced in other sections. In spite of several precursory Refs. [33–44], the Bayesian regularization does not seem to have attracted much attention in acoustics (possibly because many of the former early works came with complex iterative algorithms). It is part of the present paper aim to partly fill in this gap. The paper also highlights several important properties which, according to the authors' knowledge, have never been recognized before. Part of the present material was first published in Ref. [45], which aimed at finding an optimal basis for source reconstruction and demonstrating the benefit of taking prior information into account within a Bayesian framework. Herein, the focus is on Bayesian regularization only and its generalization to any reconstruction basis, be it optimal or not. The paper contains several original results listed hereafter.

- A theoretical proof is given about the existence of a global minimum of the Bayesian regularization criterion; this property is of prime practical importance since it confers good robustness to regularization (e.g. as compared to GCV and the L-curve for which a global minimum does not exist in general); in addition, it makes possible an automated practice of regularization.
- The posterior probability density function of the regularization parameter is given in the case of complex-valued data (i.e. Fourier transformed data); this is found useful to assess the errors due to regularization in all acoustical quantities of interest and, as far as the authors know, a sensitivity analysis to regularization is demonstrated here for the first time.
- Physical interpretation of Bayesian regularization is given in terms of energy (first principle of thermodynamics) which, hopefully, will participate to bridge the gap between an abstract probabilistic theory and the intuition gained from physics.
- Extensive experimental results are given, both on numerical and on actual data, that clearly demonstrate the supremacy of Bayesian regularization over the GCV and L-curve methods.

The paper is organized as follows. The second section first introduces general facts about the direct acoustic problem and then addresses the inverse problem within the Bayesian probabilistic framework. One objective of this section is to introduce the notations and probabilistic premises necessary for the remaining of the paper, and preferably so in a self-contained treatment of the acoustic inverse problem. As mentioned above, it should be read as a tutorial on the

Bayesian approach to inverse (acoustic) problems. The third section addresses the issue of Bayesian regularization, where several new results are established after scrupulously following the Bayesian program. Theoretical developments are accompanied by discussions relating to the properties and practical aspects of the proposed algorithms. The fourth section is an attempt to demonstrate a physical meaning of the proposed Bayesian regularization criteria in terms of thermodynamics. The fifth section addresses the important question as how sensitive the inverse problem is to regularization, to which the Bayesian framework is shown to provide a rather unique answer. Finally, the sixth section is devoted to comparing Bayesian regularization with the state-of-the-art methods, thus demonstrating its asserted superiority.

## 2. Bayesian approach to the inverse acoustic problem: a tutorial

The object of this section is to cast the inverse acoustic problem within the Bayesian probabilistic framework. This is not only necessary to introduce the fundamental ideas and notations to be used in derivation of the Bayesian regularization criterion in section 3 (the central result of the paper), but it also offers upstream justification to the classical cost function used in inverse acoustics and its *ad hoc* Tikhonov regularization. Since most of the results presented in this section can be recovered by compiling the Bayesian literature on linear models, it should be read as a tutorial. A general reference on the Bayesian approach to inverse problems is [46]. Moreover, the treatment of the inverse problem could be seen as dual of Bayesian linear regression [47–50] after exchanging the role of the explanatory variables and of the regression coefficients<sup>1</sup>.

### 2.1. General statement of the inverse acoustic problem

Broadly speaking, the inverse acoustic problem of interest herein amounts to reconstructing a source distribution or “source field” (e.g. parietal pressure or normal velocity) given a finite number of measurements, as typically returned by an array of microphones (or possibly velocity or pressure-velocity probes). More formally, let  $q(\mathbf{r})$ ,  $\mathbf{r} \in \Gamma$ , be the source field of interest and  $\Gamma$  its spatial domain, and let  $p(\mathbf{r}_i)$  be the measurement at location  $\mathbf{r}_i$ ,  $i = 1, \dots, M$ , where  $M$  is the number of probes. The direct problem states how these quantities are related together from physical principles. Assuming a frequency domain representation, this work makes use of the integral equation representation,

$$p(\mathbf{r}_i) = \int_{\Gamma} G(\mathbf{r}_i|\mathbf{r})q(\mathbf{r})d\Gamma(\mathbf{r}) + v_i, \quad i = 1, \dots, M, \quad (1)$$

where  $G(\mathbf{r}_i|\mathbf{r})$  is the (known) Green’s function of the medium between positions  $\mathbf{r}_i$  and  $\mathbf{r}$ , and  $v_i$  accounts for measurement noise and possible modeling errors. The objective of the inverse problem is to recover an estimate of  $q(\mathbf{r})$  given the noisy measurements  $\{p(\mathbf{r}_i); i = 1, \dots, M\}$  from the set of equations (1).

---

<sup>1</sup>The objectives of linear regression are, strictly speaking, oriented towards solving the direct and not the inverse problem. However, the inverse problem can be tackled after exchanging the roles of the explanatory variables and of the regression coefficients. Although this might seem artificial at first glance because the regression coefficients are, from a physical point of view, not variables but deterministic “transfer functions”, it causes no problem in the Bayesian framework where all parameters are regarded as random.

Whether explicit or not, the usual approach is first to parameterize the unknown source. In most instances, this amounts to expanding the source field onto a spatial basis,

$$q(\mathbf{r}) \approx \sum_{k=1}^K c_k \phi_k(\mathbf{r}), \quad K \geq M, \quad (2)$$

where  $\phi_k(\mathbf{r})$ ,  $\mathbf{r} \in \Gamma$  are known basis functions which, without loss of generality, will be considered normalized such that  $\int_{\Gamma} |\phi_k(\mathbf{r})|^2 d\Gamma(\mathbf{r}) = 1$ . **Typical basis functions used in acoustics are for instance, the Fourier basis (plane waves) and the spherical harmonics (spherical waves).** Plugging Eq. (2) into Eq. (1), the direct model now becomes

$$\mathbf{p} = \mathbf{H}\mathbf{c} + \mathbf{n}, \quad (3)$$

where vectors  $\mathbf{p} \in \mathbb{C}^M$ ,  $\mathbf{c} \in \mathbb{C}^K$ , and  $\mathbf{n} \in \mathbb{C}^M$  have  $p(\mathbf{r}_i)$ ,  $c_i$ , and  $v_i$  for their  $i$ -th element, respectively, and  $[\mathbf{H}]_{ij} = \int_{\Gamma} G(\mathbf{r}_i|\mathbf{r})\phi_j(\mathbf{r})d\Gamma(\mathbf{r})$ ,  $i = 1, \dots, M$ ,  $j = 1, \dots, K$ . The inverse problem is thus turned discrete and basically consists in identifying  $K$  unknown coefficients  $c_k$  (note that the problem of jointly identifying the “best” basis functions  $\phi_k(\mathbf{r})$  was addressed in Ref. [45]). Note that the apparent simplicity of the matrix equation (3) is only structural and does not alleviate its intrinsic complexity in terms of *i*) availability of the Green’s functions  $G(\mathbf{r}_i|\mathbf{r})$  leading to the transfer functions  $[\mathbf{H}]_{ij}$  and *ii*) the ill-conditioning of the resulting transfer matrix  $\mathbf{H}$ . Indeed, the strong ill-conditioning of the acoustic inverse problem is probably why its difficulty singles out as compared to other domains.

So far, the questions still remain as which cost function should be minimized to solve for  $\mathbf{c}$  in Eq. (3) and how the inverse problem should be regularized when matrix  $\mathbf{H}$  is ill-conditioned – see e.g. Refs. [24, 51]. Definite answers are provided by the probabilistic approach introduced in the next subsection.

## 2.2. A probabilistic solution to inverse acoustics

The essence of the Bayesian approach resumed in this section is to endow all quantities of interest with probability density functions (pdf’s). On the one hand, this provides a careful description of measurement noise, giving rise to the so-called “likelihood function”; on the other hand it makes possible the specification of physically admissible fluctuations in the reconstructed source field through the definition of a “prior” pdf, a precaution which intrinsically gives rise to a regularization mechanism. The combination of the likelihood function and the prior then returns the “posterior” pdf, that is the “inverse probability” of the vector of coefficients  $\mathbf{c}$  given measurements  $\mathbf{p}$ . This completely solves the inverse problem, in particular by providing “point” as well as “interval” estimates to the after-sought source field. Each of these steps is detailed hereafter.

### 2.2.1. Encoding experimental errors with the likelihood function

The most natural variable to endow with a pdf is the measurement noise  $v_i$  entering in the right hand side of Eq. (1). There are many reasons why a Gaussian pdf can be reasonably assumed here. One is simply dictated by experimental observations, where Gaussianity results from the superposition of a large number of errors as formally proved by the Central Limit theorem [52]. Another reason stems from the passage to the frequency domain results by

application of a (typically long) Fourier transform which, again due to the Central Limit theorem, implies convergence to Gaussianity. Still another reason is because a Gaussian law is the “worst case” (most dispersed) one could imagine when modeling random fluctuations with given mean value and covariance matrix according the Maximum Entropy principle [53].

In this paper  $\mathbf{n}$  is assumed with zero mean, i.e.  $\mathbb{E}\{\mathbf{n}\} = \mathbf{0}$ , where  $\mathbb{E}$  stand for the expected value; this is because any residual bias should be captured by the direct model as is inherent to the definition of “measurement” noise. The noise covariance is defined as  $\mathbb{E}\{\mathbf{nn}^H\} = \beta^2 \mathbf{\Omega}_n$  where  $^H$  stands for the conjugate transpose operator,  $\beta^2$  is the (usually *unknown*) expected noise energy, and  $\mathbf{\Omega}_n$  is a matrix of *a priori* correlation coefficients reflecting the nature of the noise field such that  $\text{Trace}\{\mathbf{\Omega}_n\} = \dim(\mathbf{n}) = M$  (in many cases one may simply take  $\mathbf{\Omega}_n = \mathbf{I}$  equal to the identity matrix). The *a priori* choice of  $\mathbf{\Omega}_n = \mathbf{I}$  should be viewed here as the most neutral, in the sense that no information is introduced *a priori*. In case the experimenter may have some *a priori* knowledge on the nature of the noise field, different noise models (such as described in Ref. [54, chap. 3]) may be employed. Therefore, keeping in mind that  $\mathbf{n}$  is a complex-valued vector, its pdf (denoted by brackets  $[\mathbf{n}|\beta^2]$  for notational simplicity, where the conditioning on  $\beta^2$  is explicitly reminded for reasons to become clear later in section 3) reads

$$[\mathbf{n}|\beta^2] = \mathcal{N}_C(\mathbf{0}, \beta^2 \mathbf{\Omega}_n) = \frac{1}{\pi^M \beta^{2M} |\mathbf{\Omega}_n|} \exp(-\beta^{-2} \|\mathbf{n}\|_{\mathbf{\Omega}_n}^2), \quad (4)$$

where  $\mathcal{N}_C$  stands for “complex Gaussian”,  $|\mathbf{\Omega}_n|$  is the determinant of matrix  $\mathbf{\Omega}_n$  and, by notational convention,  $\|\mathbf{n}\|_{\mathbf{\Omega}_n}^2 = \mathbf{n}^H \mathbf{\Omega}_n^{-1} \mathbf{n}$ . The reader not familiar with the pdf of complex-valued data are referred to Refs. [55, 56] for a good support and to Ref. [55] for an explanation on how the expression in Eq. (4) is obtained. In turn, this implies that  $\mathbf{p}$  in Eq. (1) has the complex Gaussian pdf

$$[\mathbf{p}|\mathbf{c}, \beta^2] = \mathcal{N}_C(\mathbf{H}\mathbf{c}, \beta^2 \mathbf{\Omega}_n). \quad (5)$$

The above pdf viewed as a function of variables  $c_k$ 's is referred to as the “likelihood function” [53]. It describes the “direct” probability of the values of  $\mathbf{p}$  allowed by model (1) given the random fluctuations of the measurement noise.

### 2.2.2. Construction of the prior probability density function

The assignment of a pdf to coefficients  $c_k$ 's is conceptually less obvious, especially because these are usually considered as deterministic parameters outside the Bayesian approach. However, one should keep in mind that these are unknowns with no predefined values so that their estimates could just as well be considered as possible outcomes of a random process. Concerning the coefficients  $c_k$ 's, two different situations must be distinguished: their possible outcomes *before* measurements are taken, the set of which reflects vague information possessed *a priori* by the experimenter, and possible outcomes *after* the measurements are taken, which should evidence much less variability than the former ones. The former are described by the *prior pdf*  $[\mathbf{c}]$  (investigated in this paragraph) and the latter by the *posterior pdf*  $[\mathbf{c}|\mathbf{p}]$  (investigated in the next paragraph).

In the present work, the prior pdf reflects a distribution of physically admissible values for the source field [57, 58]. This is rather a weak constraint (the Bayesian framework offering the possibility of considering many other types of

constraints), but general enough for our purpose. Let us therefore assume that the complex-valued source field is *a priori* of zero mean (in the ensemble average) and characterized by the covariance function,

$$C_q(\mathbf{r}, \mathbf{r}') \triangleq \mathbb{E}\{q(\mathbf{r})q(\mathbf{r}')^*\} = \alpha^2 \sum_{k,l=1}^K [\mathbf{\Omega}_c]_{kl} \phi_k(\mathbf{r})\phi_l^*(\mathbf{r}'). \quad (6)$$

The last expression (wherein \* stands for the complex conjugate operator) results from inserting Eq. (2) and posing  $\alpha^2 \mathbf{\Omega}_c = \mathbb{E}\{\mathbf{c}\mathbf{c}^H\}$  with  $\alpha^2$  a (typically unknown) scaling factor on the source energy and  $\mathbf{\Omega}_c$  a known matrix of correlation coefficients (whose exact setting depends on the problem) normalized such that  $\text{Trace}\{\mathbf{\Omega}_c\} = \dim(\mathbf{c}) = K$ . For instance a typical choice (implicit to deterministic approaches) is to assume that values at two different positions on the source field are uncorrelated (i.e. a spatially white source field), leading to  $\mathbf{\Omega}_c = \mathbf{I}_K$ , the identity matrix of dimension  $K$ , which implies that

$$\int_{\Gamma} \mathbb{E}\{|q(\mathbf{r})|^2\} d\Gamma(\mathbf{r}) = K\alpha^2. \quad (7)$$

A possibly more relevant choice is to enforce *a priori* the spatial whiteness of the source field – so as to enhance spatial resolution; the interested reader may refer to Appendix A for a detailed explanation on how this is achieved.

From now on, many choices are possible for a prior pdf  $[\mathbf{c}]$  that complies with the above covariance function. The simplest choice is again to select a complex Gaussian:

$$[\mathbf{c}|\alpha^2] = \mathcal{N}_C(\mathbf{0}, \alpha^2 \mathbf{\Omega}_c) = \frac{1}{\pi^K \alpha^{2K} |\mathbf{\Omega}_c|} \exp\left(-\alpha^{-2} \|\mathbf{c}\|_{\mathbf{\Omega}_c}^2\right) \quad (8)$$

(where conditioning on  $\alpha^2$  is explicitly reminded). The Gaussian choice in Eq. (4) is natural enough not to be arguable, but this is less the case for the Gaussian prior in Eq. (8). This issue is well-known in the Bayesian literature where it has been proved that the benefit of the *presence* of a vague prior pdf (whatever its exact shape) usually largely counterbalances the bias introduced by a specific prior shape as compared to another one [53, 59]. Basically, the prior information brought by such a pdf is to consider that the higher the source norm, the weaker is its probability of existence. The advantage of the above choice is to lead to tractable calculation – as shown in the next subsection – yet the same approach could be followed with other prior shapes provided the resulting integrals were numerically evaluated or analytically approached [59].

### 2.2.3. Posterior probability density function and MAP estimate

The solution to the inverse problem is finally returned by the posterior pdf  $[\mathbf{c}|\mathbf{p}; \alpha^2]$  which assigns probabilities to possible values of coefficients  $c_k$ 's once the measurements are taken. Bayes rule indicates how this pdf is obtained from updating the prior pdf  $[\mathbf{c}|\alpha^2]$  with the likelihood function  $[\mathbf{p}|\mathbf{c}, \beta^2]$ . Namely,

$$[\mathbf{c}|\mathbf{p}, \alpha^2, \beta^2] = \frac{[\mathbf{p}|\mathbf{c}, \beta^2][\mathbf{c}|\alpha^2]}{[\mathbf{p}|\alpha^2, \beta^2]}, \quad (9)$$

where

$$[\mathbf{p}|\alpha^2, \beta^2] = \int [\mathbf{p}|\mathbf{c}, \beta^2][\mathbf{c}|\alpha^2] d\mathbf{c} \quad (10)$$



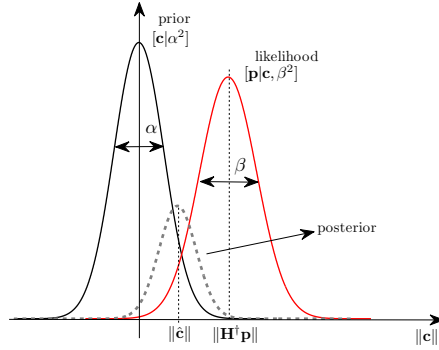


Figure 1: Schematic 1D representation of the Bayesian framework. The figure shows the different terms explicitly depending on  $\mathbf{c}$  in Bayes' rule (see Eq. (9)). In the case of a uniform prior, the solution depends uniquely on the likelihood and, for the case treated in this paper, it returns the solution (maximum likelihood) in terms of the pseudo-inverse of  $\mathbf{H}$  (noted  $\mathbf{H}^\dagger$ ). The consideration of a Gaussian prior has “biased” our search giving rise to the posterior, whose maximum (MAP estimate) returns the regularized solution expressed by  $\hat{\mathbf{c}}$ .

is the “evidence” which will play a major role in deriving the Bayesian regularization criterion later on in section 3. The particularity of the Bayes rule is to express an “inverse” probability in terms of “direct” probabilities. Once the full pdf  $[\mathbf{c}|\mathbf{p}, \alpha^2, \beta^2]$  is known, a relevant point estimate of the unknown vector  $\mathbf{c}$  is returned by the most probable value after observing the data, i.e. the maximum of the posterior pdf in Eq. (9) – the so-called “maximum a posteriori estimate” (MAP). The expression for the posterior pdf is obtained by substituting for the complex Gaussians found in Eqs. (5) and (8) into Eq. (9), leading to a quadratic cost function whose minimum returns the estimate of  $\mathbf{c}$ . This has already been done in Ref. [45] for the general case of an unknown basis function  $\phi_k(\mathbf{r})$ , and thus the mathematical details are omitted here. The standard result is then obtained:

$$\hat{\mathbf{c}} = \boldsymbol{\Omega}_c \mathbf{H}^H (\mathbf{H} \boldsymbol{\Omega}_c \mathbf{H}^H + \eta^2 \boldsymbol{\Omega}_n)^{-1} \mathbf{p} \quad (11)$$

where

$$\eta^2 = \frac{\beta^2}{\alpha^2}. \quad (12)$$

Equation (11) is an important step towards the goal of this paper. Indeed, the consideration of a Gaussian prior through the Bayesian framework has “mechanically” produced a regularized solution of the Tikhonov type – see e.g. Refs. [1, 51]. A simple illustration of the Bayesian framework followed hitherto is described in Fig. 1, representing the different terms in Bayes' rule. The significance of the regularization parameter  $\eta^2$  given in Eq. (12) is clearly that of a noise-to-signal ratio (NSR) (ratio of expected noise energy  $\beta^2$  to expected source energy  $\alpha^2$ ). A similar interpretation is found in early references on classical NAH [10], outside the Bayesian framework, and in statistical approaches such as SONAH [60].

It is reminded that no assumption have been made concerning either the acoustical propagation or the nature of sources. Although the approach requires the knowledge of the Green's function of the medium, it is not restricted to a particular propagation type (e.g. free-field, half-space, closed space).

### 2.3. Filtration of eigenvalues

A helpful interpretation of the regularized solution (11) results from the singular-value decomposition

$$\mathbf{\Omega}_n^{-1/2} \mathbf{H} \mathbf{\Omega}_c^{1/2} = \sum_{k=1}^M s_k \mathbf{u}_k \mathbf{v}_k^H \quad (13)$$

where  $s_1 \geq s_2 \geq \dots \geq s_M \geq 0$  and  $\mathbf{U} = [\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_M] \in \mathbb{C}^{M \times M}$  and  $\mathbf{V} = [\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_M] \in \mathbb{C}^{K \times M}$  form two unitary matrices (such that  $\mathbf{U}^H \mathbf{U} = \mathbf{V}^H \mathbf{V} = \mathbf{I}_M$ ). Substituting into Eq. (11), one has the simple result

$$\hat{\mathbf{c}} = \mathbf{\Omega}_c^{1/2} \mathbf{V} \left[ \frac{s_k}{s_k^2 + \eta^2} \right] \mathbf{U}^H \mathbf{\Omega}_n^{-1/2} \mathbf{p} \quad (14)$$

where  $[s_k]$  symbolizes a diagonal matrix with generic diagonal element  $s_k$ . As well-known, this is the structure of a generalized pseudo-inverse of  $\mathbf{H}$  where the smallest singular values  $s_k$  are progressively filtered out from the inversion, i.e.  $s_k/(s_k^2 + \eta^2) \approx 1/s_k$  if  $s_k \gg \eta^2$  and  $s_k/(s_k^2 + \eta^2) \approx 0$  if  $s_k \ll \eta^2$  so as to prevent instability due to over-amplification of measurement noise or simply numerical overflow when  $\mathbf{H}$  is ill-conditioned. In other words,  $\eta^2$  is clearly to be construed as a ‘‘cutoff’’ frequency indicating where  $s_k^2$  passes below the NSR, a terminology especially relevant when the unitary basis  $\mathbf{V}$  happens to contain the complex exponentials of a Discrete Fourier transform as encountered in classical Near-Field Acoustical Holography (NAH) [10]. Yet, common to all regularization strategies is the difficulty in tuning optimally the value of that parameter, a practice that is not less essential than critical. The goal of the present paper is to propose a solution that enjoys many advantages as compared to the current state-of-the-art in acoustics.

### 3. Regularization within the empirical Bayesian framework

The regularization structure formulated in the previous section is known to work well provided the regularization parameter  $\eta^2$  is correctly set. Obviously, rare are the situations where that value is known in advance and numerous strategies have been proposed as how to set it, either from empirical, physical, or statistical rules. Within the Bayesian framework, the purist approach would be to see it a ‘‘nuisance parameter’’ and to integrate it out, that is to recover a regularized version of the source field from the marginalized posterior  $[\mathbf{c}|\mathbf{p}] = \int [\mathbf{c}|\mathbf{p}, \alpha^2, \beta^2][\alpha^2, \beta^2] d\alpha^2 d\beta^2$ . The major difficulty that arises with this strategy is that it generally leads to integrals that are analytically intractable. Thus, not only is an equivalent to the closed-form expression (14) lost, but computationally very demanding methods (such as Markov Chains Monte Carlo) have to be resorted to. This is still unacceptable today in many inverse acoustic problems where industrial feasibility is of concern (remember that the inverse problem usually has to be solved for all frequency bins of a discrete Fourier transform). An alternative solution known as ‘‘empirical Bayes’’ in the literature is to replace the unknown hyperparameters by estimates obtained from the data. Apart from being legitimate in its own right [61], empirical Bayes can also be seen as a very good approximation to the full Bayesian approach discussed above [53, chap. 10]. Within the full Bayesian framework, one example is reference [62], which deals with hyperparameter estimation for image restoration using a MCMC algorithm. In the context of regularization, the empirical Bayesian approach is discussed to obtain estimates of hyperparameters such as  $\alpha^2$  and  $\beta^2$  in [33] in a quite general setting, in

[34, 35, 37] as the “evidence” method and in [39, 40] as the “marginalized MAP (of type 3)” method. In particular, reference [37] handles the case of a Gaussian likelihood and a Gaussian assumption on the prior pdf, as studied here. Some early references tackling a similar issue are [41] in the context of spline smoothing and [42]. References that explicitly give an estimate of the regularization parameter  $\eta^2$  are [43] and [44] for a Bayesian interpretation of the latter. Yet, reference [63] also provides a Bayesian interpretation of regularization and a maximum likelihood estimation of the regularization parameter, which is compared to the GCV method in the context of image deblurring.

### 3.1. The Bayesian program

In the empirical Bayesian approach, the interpretation of the regularization parameter as a NSR suggests estimating it directly from the data. As discussed above, different strategies are conceivable. A first one is to estimate  $\eta^2$  indirectly as the ratio of the most probable values of  $\alpha^2$  and  $\beta^2$  given the measurements, i.e.

$$\hat{\eta}_{\text{Joint}}^2 = \frac{\hat{\beta}^2}{\hat{\alpha}^2} \text{ where } (\hat{\alpha}^2, \hat{\beta}^2) = \text{Argmax}[\alpha^2, \beta^2 | \mathbf{p}]. \quad (15)$$

The above strategy has already been briefly described in Ref. [45]. A second strategy, which is presented here for the first time, is to find directly the marginal posterior pdf of  $\eta^2$  and then to select that value with maximum probability of occurrence, i.e.

$$\hat{\eta}_{\text{MAP}}^2 = \text{Argmax}[\eta^2 | \mathbf{p}] \text{ where } [\eta^2 | \mathbf{p}] = \int [\alpha^2, \beta^2 = \alpha^2 \eta^2 | \mathbf{p}] \left| \frac{\partial \beta^2}{\partial \eta^2} \right| d\alpha^2 \quad (16)$$

with  $|\partial \beta^2 / \partial \eta^2| = \alpha^2$  standing for the Jacobian of the change of variables  $(\alpha^2, \beta^2) \mapsto (\alpha^2, \eta^2)$ . This solution is more in the “Bayesian spirit” since it integrates out the intermediary variable  $\alpha^2$  when it is not of direct interest.

The two strategies are addressed hereafter in a unified treatment, for they require the same “ingredients”: first the evaluation of the likelihood function  $[\mathbf{p} | \alpha^2, \beta^2]$ , second the definition of the prior  $[\alpha^2, \beta^2]$ , and third the evaluation of the posterior pdf  $[\alpha^2, \beta^2 | \mathbf{p}]$ . The likelihood function  $[\mathbf{p} | \alpha^2, \beta^2]$  is obtained from marginalizing the likelihood function  $[\mathbf{p} | \mathbf{c}, \beta^2]$  of section 2.2.1 over  $\mathbf{c}$ , which is nothing else than the evidence defined by Eq. (10). This is obtained by substituting Eqs. (5) and (8) into Eq. (10). For the detailed mathematical treatments, the reader is referred to Appendix B or Refs. [47–50].

The Bayesian framework offers the possibility of including any prior information the experimenter may have on either  $\alpha^2$  or  $\beta^2$  before the measurements are taken, practice that is known as a hierarchical Bayes approach [53]. Although the Bayesian literature is very vast on how to define prior pdf’s, first principles are considered here to distinguish two plausible scenarii.

1. Scenario 1: In the simplest case one may simply set  $[\alpha^2, \beta^2] \propto 1$  ( $\propto$  stands for the “proportional” sign) in Eq. (19) meaning that all outcomes are assumed *a priori* equiprobable, choice that may be viewed as the “worst case”. Then the empirical Bayesian approach is basically equivalent to maximum likelihood.
2. Scenario 2: In a more luckily configuration, the experimenter may have some vague knowledge about the expected noise energy  $\beta^2$ , for instance related to the known instrumentation dynamics, preliminary measurements

(e.g. with the acoustic sources switched off), or known sensor sensitivity. Note that such types of errors do not cover modeling errors which are unavoidable in model (1) and may even predominate [58]. A simple choice is therefore to endow the noise energy with an inverse Gamma pdf which enforces positive values:

$$[\alpha^2, \beta^2] \propto [\beta^2] = \mathcal{G}^{-1}(a, b) \propto \beta^{-2a} \exp(-b\beta^{-2}), \quad a \geq 0, b \geq 0, \quad (17)$$

where parameters  $a$  and  $b$  are tuned to set the central position and width of the pdf. The inverse Gamma choice is motivated by the fact that it leads to analytically tractable expressions when combined with a Gaussian likelihood, to which it is a conjugate pdf [48]. Particular choices are  $(a = 1, b = 0)$  when all outcomes of  $\beta^2$  are assumed *a priori* equiprobable on a logarithmic scale [59] and  $(a = 0, b > 0)$  for the Maximum Entropy solution when solely the mean value  $\mathbb{E}\{\beta^{-2}\} = b^{-1}$  is imposed. In general, the mean of the inverse Gamma is  $\mathbb{E}\{\beta^2\} = b/(a - 2)$ ,  $a > 2$ , its mode is  $b/a$ , and its variance  $\text{Var}\{\beta^2\} = b^2/(a - 2)^2(a - 3)$ ,  $a > 3$ . Therefore, if  $\mathbb{E}\{\beta^2\}$  and the coefficient of variation  $\gamma = \text{Var}\{\beta^2\}^{-1/2}/\mathbb{E}\{\beta^2\}$  are both known *a priori*, the inverse Gamma is to be parameterized as

$$a = 3 + \gamma^{-2} \quad (18a)$$

$$b = (1 + \gamma^{-2})\mathbb{E}\{\beta^2\}. \quad (18b)$$

By way of an example, a phase error mismatch on microphone  $i$  with standard deviation  $\sigma$  will produce multiplicative noise,  $p(\mathbf{r}_i) = p_0(\mathbf{r}_i) \exp(j\phi)$ , where  $p_0(\mathbf{r}_i)$  stands for the noise-free pressure value; assuming  $\sigma$  small enough,  $p(\mathbf{r}_i) \simeq p_0(\mathbf{r}_i) + v_i$  with  $v_i = jp_0(\mathbf{r}_i)\phi$ . Thus, one has  $\mathbb{E}\{\beta^2\} = |p_0(\mathbf{r}_i)|^2\sigma^2 \simeq |p(\mathbf{r}_i)|^2\sigma^2$ ; finally, given a user-defined value of  $\gamma$ ,  $a = 3 + \gamma^{-2}$ , and  $b = (1 + \gamma^{-2})|p(\mathbf{r}_i)|^2\sigma^2$ .

Other scenarii with prior knowledge on the expected source energy  $\alpha^2$  are less plausible and will not be considered in this paper, although the Bayesian framework perfectly allows to take them into account. The next results are derived for the case of scenario 1, which is the case employed in the numerical simulations and experimental set-up in the following sections. Results for the case of scenario 2 may be readily obtained by substituting for the inverse Gamma pdf in Eq. (17) into Eq. (19) and proceeding similarly as in the case of scenario 1.

### 3.2. First strategy: Joint probability density function of noise and source energies

Given all the above preliminaries, one is now in a position to evaluate the joint pdf of  $(\alpha^2, \beta^2)$  given the measurements. Using again the Bayes' rule it is written as:

$$[\alpha^2, \beta^2 | \mathbf{p}] \propto [\mathbf{p} | \alpha^2, \beta^2][\alpha^2, \beta^2], \quad (19)$$

where the setting  $[\alpha^2, \beta^2] \propto 1$  is considered from now on. The maximization of the above equation, that is, the MAP estimate, leads to the following cost function to be minimized with respect to  $\alpha^2$  and  $\beta^2$

$$\mathbf{J}_{\text{Joint}}(\alpha^2, \beta^2) = \sum_{k=1}^M \ln(\alpha^2 s_k^2 + \beta^2) + \sum_{k=1}^M \frac{|y_k|^2}{\alpha^2 s_k^2 + \beta^2}, \quad (20)$$

as conveniently expressed in terms of the eigen-elements (13) and where  $y_k$  is the  $k$ -th element of vector

$$\mathbf{y} = \mathbf{U}^H \boldsymbol{\Omega}_n^{-1/2} \mathbf{p}, \quad (21)$$

that is the projection of the measurements on the array subspace. One peculiarity is that the sufficient statistics arising in Eq. (20) is no longer the vector of complex pressures,  $\mathbf{p}$ , but actually the covariance matrix of the latter,  $\mathbf{p}\mathbf{p}^H$  – i.e. it is the only measurement required to compute  $|y_k|^2$  in Eq. (20). This leaves the interesting option of averaging together several measurements in order to get more accurate estimates whenever the acoustical field is statistically stationary. Specifically, let us assume that multiple snapshots  $\{\mathbf{p}_j; j = 1, \dots, N\}$  are available, for instance as a result of segmenting the time series into short-time blocks before applying the Fourier transform. The only modification is in the likelihood function (5) where the joint pdf  $[\{\mathbf{p}_j; j = 1, \dots, N\}|\mathbf{c}; \beta^2]$  should be considered instead. Assuming independent snapshots, the latter is the same as the product of individual pdf's, i.e.  $\prod_{j=1}^N [\mathbf{p}_j|\mathbf{c}; \beta^2] = \prod_{j=1}^N \mathcal{N}_C(\mathbf{H}\mathbf{c}_j, \beta^2 \boldsymbol{\Omega}_n)$ . This results in all the same equations as obtained hitherto, but with  $|y_k|^2$  replaced by its average on  $N$  snapshots,

$$\langle |y_k|^2 \rangle = \mathbf{u}_k^H \boldsymbol{\Omega}_n^{-1/2} \underbrace{\left( \frac{1}{N} \sum_{j=1}^N \mathbf{p}_j \mathbf{p}_j^H \right)}_{\mathbf{S}_{pp}} \boldsymbol{\Omega}_n^{-1/2} \mathbf{u}_k, \quad (22)$$

where  $\mathbf{S}_{pp}$  stands for the (empiric) correlation matrix of measurements and  $\mathbf{u}_k$  for the  $k$ -th column of matrix  $\mathbf{U}$ . This general configuration will be assumed from now on, a special case being  $N = 1$  when no averaging is requested.

The minimization of cost function  $\mathbf{J}_{\text{Joint}}(\alpha^2, \beta^2)$  is an easy exercise. Let us first introduce the change of variables  $(\alpha^2, \beta^2) \mapsto (\alpha^2, \eta^2)$ . Then, setting the derivative of

$$\mathbf{J}_{\text{Joint}}(\alpha^2, \eta^2) = \sum_{k=1}^M \ln(s_k^2 + \eta^2) + \frac{1}{\alpha^2} \left( \sum_{k=1}^M \frac{\langle |y_k|^2 \rangle}{s_k^2 + \eta^2} \right) + M \ln \alpha^2 \quad (23)$$

with respect to  $\alpha^2$  to zero, one immediately gets the MAP estimate

$$\hat{\alpha}^2 = \frac{1}{M} \left( \sum_{k=1}^M \frac{\langle |y_k|^2 \rangle}{s_k^2 + \eta^2} \right). \quad (24)$$

Unlike the source energy, there is no closed-form solution for the regularization parameter  $\eta^2$ . Substituting for  $\hat{\alpha}^2$  in  $\mathbf{J}_{\text{Joint}}(\alpha^2, \eta^2)$ , one then has

$$\hat{\eta}_{\text{Joint}}^2 = \text{Argmin } \mathbf{J}_{\text{Joint}}(\eta^2) \quad (25)$$

with

$$\begin{aligned} \mathbf{J}_{\text{Joint}}(\eta^2) &\triangleq \mathbf{J}_{\text{Joint}}(\hat{\alpha}^2, \eta^2) - M \\ &= \sum_{k=1}^M \ln(s_k^2 + \eta^2) + M \ln \hat{\alpha}^2. \end{aligned} \quad (26)$$

This is a 1-D minimization problem which may be easily carried out starting with a rough grid search method followed by a refined gradient descent or dichotomy method. In turn, the MAP estimate of the noise energy, if needed, is  $\hat{\beta}^2 = \hat{\alpha}^2 \times \hat{\eta}_{\text{Joint}}^2$ .

### 3.3. Second strategy: Marginal probability density function of the regularization parameter

As discussed in section 3.1, another laudable strategy to estimate the regularization parameter from the data is to directly find its posterior pdf,  $[\eta^2|\mathbf{p}]$ , without requiring the intermediate estimations of the noise and source energies,  $\hat{\alpha}^2$  and  $\hat{\beta}^2$ . The exact expression of  $[\eta^2|\mathbf{p}]$  is carried out by marginalization of the joint pdf  $[\alpha^2, \beta^2|\mathbf{p}]$  in Eq. (19). The mathematical steps describing how this is carried out is presented in Appendix C for the interested reader. The following MAP estimate of the regularization parameter is obtained:

$$\hat{\eta}_{\text{MAP}}^2 = \text{Argmin } \mathbf{J}_{\text{MAP}}(\eta^2) \quad (27)$$

with

$$\begin{aligned} \mathbf{J}_{\text{MAP}}(\eta^2) &\triangleq \sum_{k=1}^M \ln(s_k^2 + \eta^2) + \left(M - \frac{2}{N}\right) \ln \hat{\alpha}^2 \\ &= \mathbf{J}_{\text{Joint}}(\eta^2) - \frac{2}{N} \ln(\hat{\alpha}^2(\eta^2)), \end{aligned} \quad (28)$$

where it is emphasized in the second line that  $\hat{\alpha}^2$  is a function of  $\eta^2$ , as given in Eq. (24), and  $N$  is the number of snapshots as defined prior to Eq. (22). The estimate of the regularization parameter  $\hat{\eta}_{\text{MAP}}^2$  is returned as the minimum of cost function (28).

### 3.4. Discussion

Equations (28) and (26) are important results of the (empirical) Bayesian approach undertaken in this paper. Thus, before one proceeds further, several remarks are in order at this juncture.

#### 3.4.1. Comparison of strategies

It is anticipated that the two estimates  $\hat{\eta}_{\text{Joint}}^2$  and  $\hat{\eta}_{\text{MAP}}^2$  found in sections 3.2 and 3.3 must be closely related – since  $\mathbf{J}_{\text{MAP}} = \mathbf{J}_{\text{Joint}} - (2/N) \ln \hat{\alpha}^2$  according to Eq. (28) – although they have no reason to be *identical* for the MAP estimate of a ratio is generally not equal to the ratio of the MAP estimates. Indeed, since  $-(2/N) \ln \hat{\alpha}^2$  is a monotonically increasing function of  $\eta^2$ , it comes that

$$\hat{\eta}_{\text{MAP}}^2 \leq \hat{\eta}_{\text{Joint}}^2, \quad (29)$$

meaning that  $\hat{\eta}_{\text{MAP}}^2$  in general yields a “less regularized” solution than  $\hat{\eta}_{\text{Joint}}^2$ . In addition, the term  $-(2/N) \ln \hat{\alpha}^2$  acts as a penalty that forbids solutions (i.e. source fields) with zero energy  $\alpha^2$ . This is confirmed by the asymptotic behavior of  $\mathbf{J}_{\text{Joint}}$  and  $\mathbf{J}_{\text{MAP}}$  as  $\eta^2 \rightarrow \infty$ , given as follows:

$$\mathbf{J}_{\text{Joint}}(\eta^2) \sim CM, \quad \text{as } \eta^2 \rightarrow \infty \quad (30a)$$

$$\mathbf{J}_{\text{MAP}}(\eta^2) \sim \frac{2}{N} \ln \eta^2, \quad \text{as } \eta^2 \rightarrow \infty, \quad (30b)$$

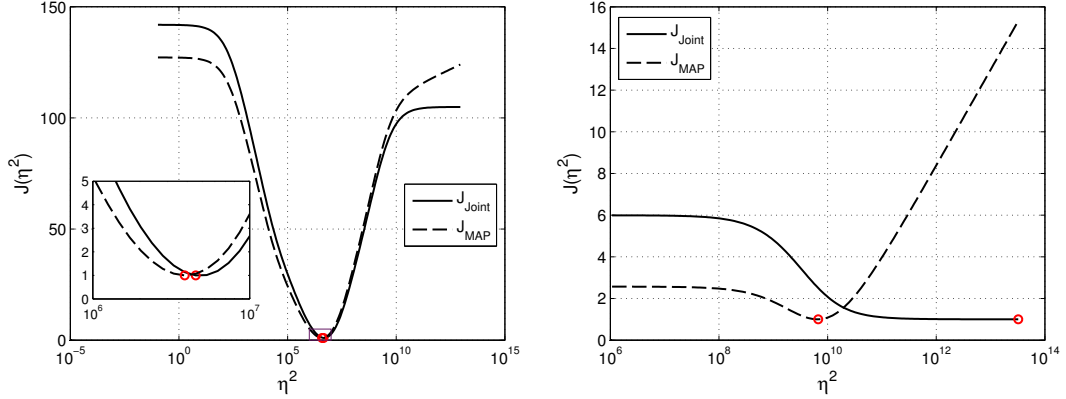


Figure 2: In most practical situations  $\hat{\eta}_{\text{Joint}}^2$  and  $\hat{\eta}_{\text{MAP}}^2$  are found nearly equal, at least when the numbers of sensors or snapshots is large enough (see left panel). One exception is when the SNR is so poor that  $\mathbf{J}_{\text{Joint}}$  tends to see only noise, thus rejecting its minimum to infinity and returning an unbounded value of the regularization parameter. Such a situation is forbidden by  $\mathbf{J}_{\text{MAP}}$  which, by construction, advantageously forces a finite value of  $\eta^2$  (see right panel).

with  $C$  a constant. Notice that  $\mathbf{J}_{\text{Joint}}$  tends to a constant while  $\mathbf{J}_{\text{MAP}}$  still grows to infinity. This property might be advantageous in some low SNR configurations where  $\mathbf{J}_{\text{Joint}}$  tends to “see” only noise and thus returns an infinite value for  $\hat{\eta}_{\text{Joint}}^2$ .

However, in most experimental instances it has been verified by the authors that  $\hat{\eta}_{\text{Joint}}^2$  and  $\hat{\eta}_{\text{MAP}}^2$  are found nearly equal, especially when the product  $MN$  of the numbers of microphones and snapshots is large. This is formally justified by the fact that

$$\underbrace{\mathbf{J}_{\text{MAP}}(\eta^2)}_{\sim \mathcal{O}(M)} = \underbrace{\mathbf{J}_{\text{Joint}}(\eta^2)}_{\sim \mathcal{O}(M)} - \underbrace{\frac{2}{N} \ln \hat{\alpha}^2}_{\sim \mathcal{O}(N^{-1})} \xrightarrow{MN \rightarrow \infty} \mathbf{J}_{\text{Joint}}(\eta^2) \quad (31)$$

where  $\mathcal{O}$  means “on the order of”. Figure 2 illustrates the above remarks by means of an example selected from the simulations to be presented in Sec. 6.

### 3.4.2. A posteriori estimate of the signal-to-noise ratio

Following the discussion in Sec. 2.3, it is interesting to remark that an estimate of the signal-to-noise ratio (SNR) *a posteriori* may be obtained from  $\hat{\eta}_{\text{Joint}}^2$  or  $\hat{\eta}_{\text{MAP}}^2$ . Taking into account the propagation from the source field to the microphone array it follows that

$$\text{SNR} = \frac{1}{\hat{\eta}^2 M} \sum_{k=1}^M s_k^2, \quad (32)$$

where  $M$  is the number of microphones and  $s_k$  the singular values as defined in Eq. (13). The above estimate may be compared, for instance, to a measured or expected signal-to-noise ratio over the array of microphones.

### 3.4.3. Accounting for the inverse Gamma prior on the noise energy

For the sake of completeness, we briefly provide hereafter the modified cost functions to account for the inverse Gamma prior on the noise energy  $\beta^2$ . Omitting the intermediate steps, Eq. (20) is now given as

$$\mathbf{J}_{\text{Joint}}(\alpha^2, \beta^2) = \sum_{k=1}^M \ln(\alpha^2 s_k^2 + \beta^2) + \sum_{k=1}^M \frac{\langle |y_k|^2 \rangle}{\alpha^2 s_k^2 + \beta^2} + \frac{a}{N} \ln \beta^2 + \frac{b}{N} \beta^{-2}, \quad (33)$$

with  $a$  and  $b$  the parameters of the inverse Gamma pdf in Eq. (17). Introducing the change of variables  $(\alpha^2, \beta^2) \mapsto (\alpha^2, \eta^2)$  and minimizing with respect to  $\alpha^2$  gives

$$\hat{\alpha}^2 = \frac{1}{M + \frac{a}{N}} \left( \sum_{k=1}^M \frac{\langle |y_k|^2 \rangle}{s_k^2 + \eta^2} + \frac{b}{N\eta^2} \right). \quad (34)$$

The cost function to be minimized with respect to  $\eta^2$  then follows

$$\mathbf{J}_{\text{Joint}}(\eta^2) = \sum_{k=1}^M \ln(s_k^2 + \eta^2) + \left( M + \frac{a}{N} \right) \ln \hat{\alpha}^2 + \frac{a}{N} \ln \eta^2, \quad (35)$$

from which the modified  $\mathbf{J}_{\text{MAP}}(\eta^2)$  directly follows from the relation in Eq. (28).

### 3.5. Properties

The Bayesian estimators of the regularization parameters introduced in the former section happen to enjoy several interesting properties which are listed in this subsection.

#### 1. Uniqueness of solution.

**Proposition 1.** *The cost functions  $\mathbf{J}_{\text{Joint}}(\eta^2)$  and  $\mathbf{J}_{\text{MAP}}(\eta^2)$  in Eqs. (26) and (28) have at most one minimum with probability one when the product  $MN$  of the numbers of microphones and snapshots becomes large.*

PROOF. : see Appendix D

This proposition is of fundamental importance, since it establishes the robustness of the proposed regularization strategy: as opposed to most alternative strategies, it guaranties that the estimation of  $\eta^2$  will not be trapped into local minima (as might happen with GCV or the L-curve for instance). An explanation of this nice property is to view  $\exp\{-\mathbf{J}_{\text{MAP}}(\eta^2)\}$  as proportional to a unimodal pdf. Note that Property 1 is also valid for (small) finite values of  $MN$  with overwhelming probability – see Appendix D – a fact that has always been observed by the authors in a large variety of different configurations. However, Proposition 1 might allow for  $\mathbf{J}_{\text{Joint}}(\eta^2)$  to possibly have no minimum at all, a situation likely to reflect inconsistency in model (1) or its associated likelihood function (5) (e.g. presence of strong modeling errors). This property is presented for the first time and thus consists one of the novel results in this paper.

#### 2. Ease of implementation.

The proposed regularization strategy leads to simple implementation. First, the explicit evaluation of the inverse problem is to be evaluated only once through the calculation of the eigen-elements  $\mathbf{U}$ ,



$\mathbf{V}$ , and  $\{s_k; k = 1, \dots, M\}$ , required in Eq. (14). The only extra work is to minimize  $\mathbf{J}_{\text{Joint}}(\eta^2)$  or  $\mathbf{J}_{\text{MAP}}(\eta^2)$  to which Proposition 1 allows the use of efficient numerical algorithms. Second, the regularization strategy applies just as well to measurement setups that record either the complex data  $\mathbf{p}$  or solely their correlation matrix,  $\mathbf{S}_{pp}$ , since the latter is the sufficient statistics entering into the evaluation of  $\eta^2$  as discussed in section 3.2.

3. *Physical insight.* In contrast to other approaches, the proposed algorithm returns both an estimate of noise and source energies,  $\alpha^2$  and  $\beta^2$ , in addition to their ratio. This may provide useful physical insight into the final solution of the inverse acoustic problem, for instance to test for the presence of modeling errors by comparing the value of  $\hat{\beta}^2$  to the maximum of the expected measurement noise level, or if some figures of merit such as the Akaike Information Criterion or Bayesian Information Criterion are to be computed to test the validity of model (1) [64].
4. *Evaluation of sensitivity to regularization.* Finally, the probabilistic apparatus attached to the determination of the regularization parameter allows the propagation of estimation errors to any acoustical quantity of interest. In other words, it provides a definite answer to the crucial question: “how sensitive is the reconstructed source field with respect to the setting of the regularization parameter”. As far as the authors know, this question has rarely been answered despite of its considerable importance and the criticism of regularization in practice. Due to its relevance, this point will be further investigated in section 5.

#### 4. Physical interpretations in terms of energy

In contrast to most regularization strategies used in inverse acoustics, the proposed Bayesian solution is offspring of a formal probabilistic approach. Very interestingly, one interpretation stems from the conservation of energy (1st thermodynamic principle).

##### 4.1. Energy balance

One simple way to understand the tenants of the proposed regularization strategy is to consider Eq. (21) that yields the projection,  $\{y_k; k = 1, \dots, M\}$ , of the data onto the array subspace. Taking the expected value of energy with respect to the prior source coefficients and measurement noise, one gets

$$\begin{aligned}
\mathbb{E}\{\langle |y_k|^2 \rangle\} &= \mathbf{u}_k^H \boldsymbol{\Omega}_n^{-1/2} \mathbb{E}\{\mathbf{S}_{pp}\} \boldsymbol{\Omega}_n^{-1/2} \mathbf{u}_k \\
&= \mathbf{u}_k^H \boldsymbol{\Omega}_n^{-1/2} \left( \alpha^2 \mathbf{H} \boldsymbol{\Omega}_c \mathbf{H}^H + \beta^2 \boldsymbol{\Omega}_n \right) \boldsymbol{\Omega}_n^{-1/2} \mathbf{u}_k \\
&= \alpha^2 s_k^2 + \beta^2, \quad k = 1, \dots, M,
\end{aligned} \tag{36}$$

where the last line was arrived at after using Eq. (13). This equation establishes the energy balance between the expected energy of measurements before they are actually collected and the source and noise energies,  $\alpha^2$  and  $\beta^2$ . In practice,  $\mathbb{E}\{\langle |y_k|^2 \rangle\}$  is not available, however the set of equations (36) could still be solved approximately after substituting it by the actual measurements  $\langle |y_k|^2 \rangle$ . Thus least squares could be a solution to get the “slope”  $\alpha^2$  and the

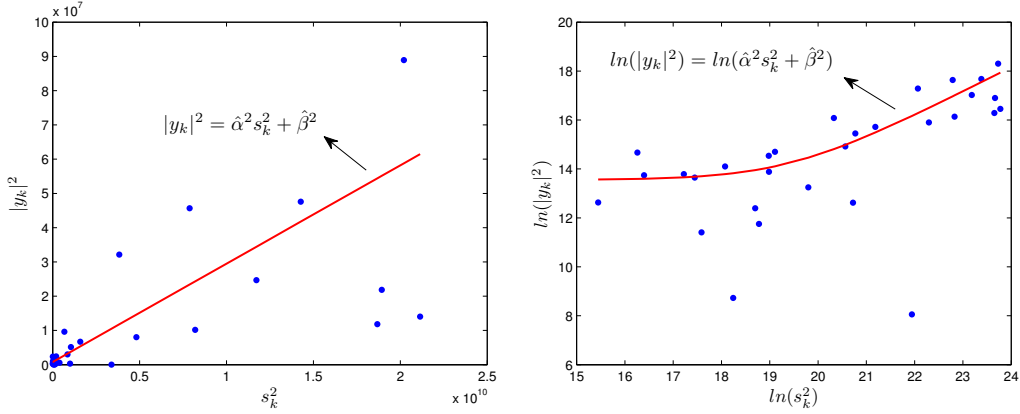


Figure 3: The estimation of the source and noise energies amounts to finding the slope and intercept, respectively, of the regression line passing through the cluster of measurement points  $\{(s_k^2; \langle |y_k|^2 \rangle); k = 1, \dots, M\}$  complying with the conservation of energy law  $\mathbb{E}\{\langle |y_k|^2 \rangle\} = \alpha^2 s_k^2 + \beta^2$ . To improve visualization the right panel shows the plot of the natural logarithm of both the singular values  $s_k^2$  and the coefficients  $|y_k|^2$ .

“intercept”  $\beta^2$ , but with no guarantee of positiveness of the latter (quadratic) values. A much better strategy would be to estimate these parameters from the Maximum Likelihood principle after recognizing that  $y_k$  has a complex Gaussian pdf  $\mathcal{N}_C(0, \alpha^2 s_k^2 + \beta^2)$ . When the prior is uniform, this actually happens to boil down to the MAP estimates of (25). This is illustrated in Fig. 3 by way of an example extracted from simulations to be presented in Sec. 6.

## 5. How sensitive is the acoustic inverse problem to regularization?

As mentioned several times in the paper, the art of regularization is as risky as essential for successfully solving the acoustic inverse problem. Therefore, it is crucial to assess the effect small modifications in the regularization parameter will have on the reconstructed source field. In short, the question to be addressed is how sensitive the reconstruction is to the actual setting of the regularization parameter? As far as the authors known, this issue has never been investigated in the present field of interest despite its practical importance. One refraining difficulty is that  $\eta^2$  intervenes in a very non-linear way into the estimation of  $q(\mathbf{r})$ .

Because the Bayesian frameworks treats  $\eta^2$  as a random variable endowed with a pdf, it offers a rather unique answer to the present issue.

### 5.1. Exact solution from Markov Chain Monte Carlo (MCMC) sampling

The posterior pdf  $[\eta^2 | \mathbf{p}]$  found in Proposition 6 fully answers the goal of quantifying the uncertainty stemming from regularization: by allowing to sample values of  $\eta^2$ , it makes possible to propagate the variability due to regularization to any acoustical quantity of interest functionally depending on the regularization parameter, such as the source field, its quadratic flux, the acoustical intensity, or the acoustical power. Since such a pdf does not seem to pertain to a standard family, one has to resort to MCMC methods, such as the popular Metropolis-Hasting algorithm, to sample it [53]. This is surely a minor price to pay for exactness.

### 5.2. Gaussian approximation

Should one not require the full posterior pdf  $[\eta^2|\mathbf{p}]$  but simply characterize it by its variance, then the following simple Gaussian approximation may be useful:

**Proposition 2.**

$$[\eta^2|\mathbf{p}] \approx \mathcal{N}(\hat{\eta}^2; \sigma^2) \quad \text{with} \quad \sigma^2 = \left( \frac{d^2 \mathbf{J}_{MAP}(\eta^2)}{(d\eta^2)^2} \right)_{\eta^2=\hat{\eta}^2}^{-1}, \quad (37)$$

$\hat{\eta}^2$ , and  $\mathbf{J}_{MAP}(\eta^2)$  given in Eqs. (27) and (28), respectively.

PROOF. : see Appendix E

The second derivative of  $\mathbf{J}_{MAP}(\eta^2)$  arising in Eq. (37) may be evaluated analytically or, more simply, numerically. In the case where the product  $MN$  is large (so that the convergence in Eq. (31) holds true),  $\sigma^2$  is given in closed-form by Eq. (D.2) of Appendix D. Such an approximation has the benefit of ease of implementation, for it requires little more than a random generator of Gaussian variables to propagate errors stemming from regularization to any reconstructed acoustical quantity.

### 5.3. Cramer-Rao lower bound

Interestingly enough, the inverse of the expected value of  $\sigma^{-2}$  in Proposition 2 returns to the Cramer-Rao lower bound [59] of the regularization parameter  $\eta^2$ , that is the lowest possible variance that can be attained by any unbiased estimate  $\hat{\eta}^2$ . Specifically, when the product  $MN$  of the numbers of sensors and snapshots is large,

**Proposition 3.**

$$\text{Var}\{\eta^2|\mathbf{p}\} \geq \frac{1}{\mathbb{E}\left(\frac{d^2 \mathbf{J}_{MAP}(\eta^2)}{(d\eta^2)^2}\right)_{\eta^2=\hat{\eta}^2}} \xrightarrow{MN \rightarrow \infty} \frac{1}{(1 - B_2) \sum_{k=1}^M \left(\frac{1}{s_k^2 + \eta^2}\right)^2} \quad (38)$$

where

$$B_2 = \frac{\left(\sum_{k=1}^M \frac{1}{s_k^2 + \eta^2}\right)^2}{M \sum_{k=1}^M \frac{1}{(s_k^2 + \eta^2)^2}}. \quad (39)$$

PROOF. : see Appendix F

Factor  $B_2$  entering into Eq. (38) is to be interpreted as a measure of bandwidth of the sequence  $(s_k^2 + \eta^2)^{-1}$ ; indeed,

$$\frac{1}{M} \leq B_2 \leq 1, \quad (40)$$

where the lower bound is reached when all but one of the  $(s_k^2 + \eta^2)$ 's are nil and the upper bound when all  $(s_k^2 + \eta^2)$ 's are identical. This latter situation is especially critical for  $B_2 = 1$  implies an infinite variance in Eq. (38): practically, it corresponds to the experimental configuration where the *same* equation  $\langle |y_k|^2 \rangle \approx \alpha^2 s_k^2 + \beta^2$ ,  $k = 1, \dots, M$  is observed  $M$  times, thus making it impossible to estimate the two unknowns  $\alpha^2$  and  $\beta^2$  and their ratio  $\eta^2$ . This is fully compliant with the interpretation of section 4.1. At the same time, this particular configuration also corresponds to a well-posed inverse problem where virtually no regularization should be required, thus lessening the effect of a large variance around  $\eta^2 = 0$  on the reconstructed acoustical quantity of interest.

## 6. Numerical comparison to the state-of-the-art methods in acoustics

In order to examine the performance of the Bayesian regularization criterion, numerical experiments of acoustic inverse problems are presented in this section. Two reference papers on this field are selected as benchmarks [22, 24], illustrating a wide range of source reconstruction configurations. In both papers, the performance of GCV and the L-curve was compared in the framework of numerical acoustic simulations. It was shown that the behavior of both methods varies significantly with the simulation parameters (i.e. level of measurement noise, frequency or distance between microphone array and source surface) and no prevailing method could be indicated. Our aim here is to check the effectiveness of the proposed Bayesian regularization criterion applied to the cases investigated in the aforementioned references. For that purpose, the geometry of the acoustic problem treated in each reference is reproduced and is briefly recalled in the next subsections.

The direct problem employed in the simulations has been described in Sec. (2) and a free-field propagation is assumed, with Green's function given by:

$$G(\mathbf{r}_i|\mathbf{r}) = \frac{e^{-jk\|\mathbf{r}_i-\mathbf{r}\|}}{4\pi\|\mathbf{r}_i-\mathbf{r}\|}, \quad (41)$$

where  $k = \omega/c$  is the acoustic wavenumber,  $\omega$  is the angular frequency and  $c$  is the speed of sound. The solution of the inverse problem for the source coefficients ( $\hat{\mathbf{c}}$ ) is given by Eq. (14) and the different parameter selection methods will be used to adjust the regularization parameter  $\eta^2$  therein. The implemented cost functions for the GCV and L-curve methods are exactly the same as those described in references [22, 24], and a detailed analysis of each can be found in refs. [18, 65].

An indicator based on the knowledge of the exact solution of the inverse problem ( $\mathbf{c}$ ) is obtained by computation of the mean squared error (MSE) [22] between  $\mathbf{c}$  and solutions for all potential regularization parameters. A cost function can be written as:

$$\mathbf{J}_{MSE} = \|\hat{\mathbf{c}}(\eta^2) - \mathbf{c}\|. \quad (42)$$

The optimal regularization parameter is thus returned by the minimum of the cost function:

$$\eta_{MSE}^2 = \text{Argmin } \mathbf{J}_{MSE}(\eta^2). \quad (43)$$

This indicator illustrates the ‘‘best we can do’’ scenario and the effectiveness of each method is evaluated as the relative error to the optimal solution, as follows:

$$\varepsilon_{BA} = \frac{\|\hat{\mathbf{c}}_{BA} - \hat{\mathbf{c}}_{MSE}\|}{\|\hat{\mathbf{c}}_{MSE}\|}, \quad \varepsilon_{LC} = \frac{\|\hat{\mathbf{c}}_{LC} - \hat{\mathbf{c}}_{MSE}\|}{\|\hat{\mathbf{c}}_{MSE}\|}, \quad \varepsilon_{GCV} = \frac{\|\hat{\mathbf{c}}_{GCV} - \hat{\mathbf{c}}_{MSE}\|}{\|\hat{\mathbf{c}}_{MSE}\|}, \quad (44)$$

where, from now on, we use the notation BA for the presented Bayesian criterion and LC for the L-curve. The noise term entering on the right hand side of Eq. (3), used to simulate the acoustic pressure, consists of multiplicative and additive perturbations, with signal-to-noise ratio (SNR) ranging from 40 dB to 6 dB. The noise term corresponding to the  $i$ -th microphone is given by:

$$v_i = 10^{-\text{SNR}/20} \left( \gamma e^{j\theta} p_0(\mathbf{r}_i) + \delta e^{j\phi} \sqrt{\frac{\|\mathbf{p}_0\|^2}{M}} \right), \quad (45)$$

where  $\gamma$  and  $\delta$  are zero mean Gaussian random variables with  $\text{Var}(\gamma) = \text{Var}(\delta) = 1$ ,  $\theta$  and  $\phi$  are random variables uniformly distributed between 0 and  $2\pi$ ,  $\mathbf{p}_0$  is the vector of noise-free pressure and  $p_0(\mathbf{r}_i)$  its  $i$ -th component. The employed frequency band ranges from 100 Hz to 2500 Hz and the simulations are carried over 500 random trials of measurement noise for each frequency and SNR. A MATLAB<sup>®</sup> implementation of the codes used for the simulations is available online [66].

### 6.1. Case 1: Square system ( $M = K$ )

The geometry of this problem is reproduced from Ref. [22] and is sketched on the left panel of Fig. 4. A planar array of  $9 \times 9$  microphones is placed at  $z_h$  from a vibrating surface modeled as a distribution of  $9 \times 9$  monopole sources. The inter-source spacing is set to be identical to the inter-microphone spacing ( $r_s = r_m$ ). Simply one point source is placed at the center of the source surface with volume velocity equal to unity. The condition number of the transfer matrix for this scenario is shown on the right panel of Fig. 4 for three different distances to the source plane. We observe that as the distance is increased, the conditioning of the matrix is increased (especially at low frequencies) and that the problem is mainly ill-conditioned at low frequencies. The results of the simulations are given by the

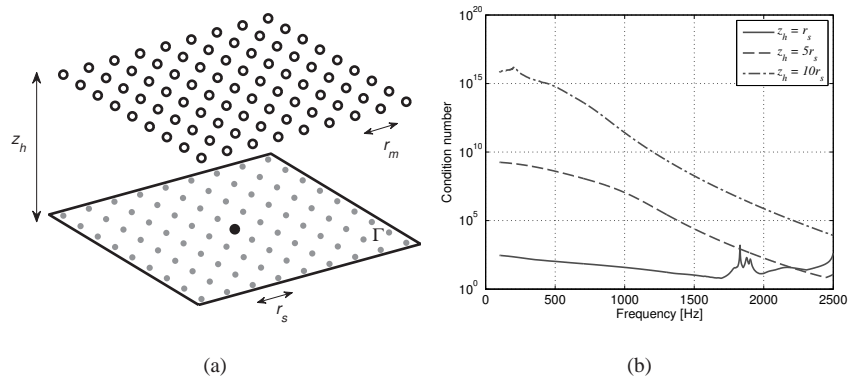


Figure 4: (a) Geometry of the problem for the first scenario, showing the discretized source surface, the microphone array and the simulated point source placed at the center of the surface. (b) Condition number for three different distances from the source plane ( $z_h$ ).

average of the indicator (Eq. (44)) over all random trials. Figure 5 shows the results for three distances from the array to the source surface respectively equal to  $z_h = r_s$  (first row),  $z_h = 5r_s$  (second row) and  $z_h = 10r_s$  (third row), with  $r_s = r_m = 12$  cm. Note that the results for the non-regularized case (NR) are added on the last column. We observe that GCV gives satisfactory results when the array is placed close to the source, however, when the microphone array is moved farther away (increasing on the condition number) it provides very poor results for a wide frequency range and all levels of noise, showing that GCV is very sensitive to the conditioning of the problem. In fact, although GCV returns reasonable regularization for some noise ensembles at those range, it occasionally fails to do so and this failure leads to completely anomalous solutions, because the regularization parameter  $\eta_{GCV}^2$  is too small. This is confirmed by the histogram plot shown in Fig. 6, which corresponds to the 500 trials for a SNR of 30 dB, frequency of 1000 Hz and distant  $z_h = 5r_s$  from the source surface. It shows the ratio between the regularization parameter estimated by each

method and the optimal one. The closer the ratio is to 1, the better is the estimate. We note that the GCV's histogram presents a tail towards very small values, which means a severe underestimation of the regularization parameter. This behavior was observed for all range of SNR and frequencies which GCV gives high relative errors  $\varepsilon_{GCV}$ .

The results of the L-curve do not present a common trend for the three studied cases, however it is clearly not effective for high levels of measurement noise (low SNR) and at very low frequencies. On the other hand, the results returned by the Bayesian regularization criterion are satisfactory for the whole set of tested acoustical configurations, with errors rarely exceeding 20% of the optimal MSE solution. We remark that the results returned by the cost functions  $\mathbf{J}_{\text{Joint}}$  and  $\mathbf{J}_{\text{MAP}}$  were very similar, therefore just the case of  $\mathbf{J}_{\text{Joint}}$  is presented here.

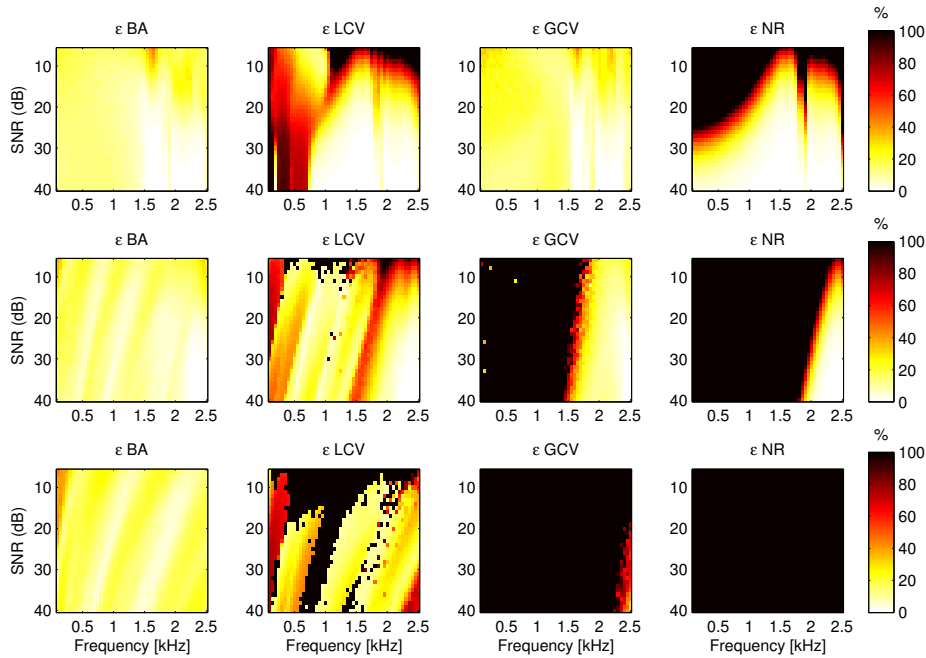


Figure 5: Average value of the relative error to the optimal (MSE) solution over 500 realizations of measurement noise. Each row corresponds to a given distance from the array to the source surface. First row:  $z_h = r_s$ , second row:  $z_h = 5r_s$  and third row:  $z_h = 10r_s$ .

## 6.2. Case 2: Under-determined system ( $M \ll K$ )

The geometry of the second case is reproduced from Ref. [24] and is depicted in Fig. 7. A planar array of  $6 \times 5$  microphones is placed at a distance  $z_h$  from a source surface which is modeled by a grid of  $28 \times 23$  monopole-like sources. The inter-microphone and inter-source spacings are respectively set to 10 cm and 2 cm in both  $x$  and  $y$  directions. This configuration models an under-determined scenario, with the number of measurement positions much lesser than the number of unknown source coefficients. In this case, the simulation is done by randomly placing 3 monopoles on the source distribution and assigning random complex strength to each of them. The simulated acoustic pressure is then perturbed with the noise model given by Eq. (45). The same indicator (relative error to the MSE solution) is used and the results are presented in Fig. 8. We note that the GCV presents similar behavior to the

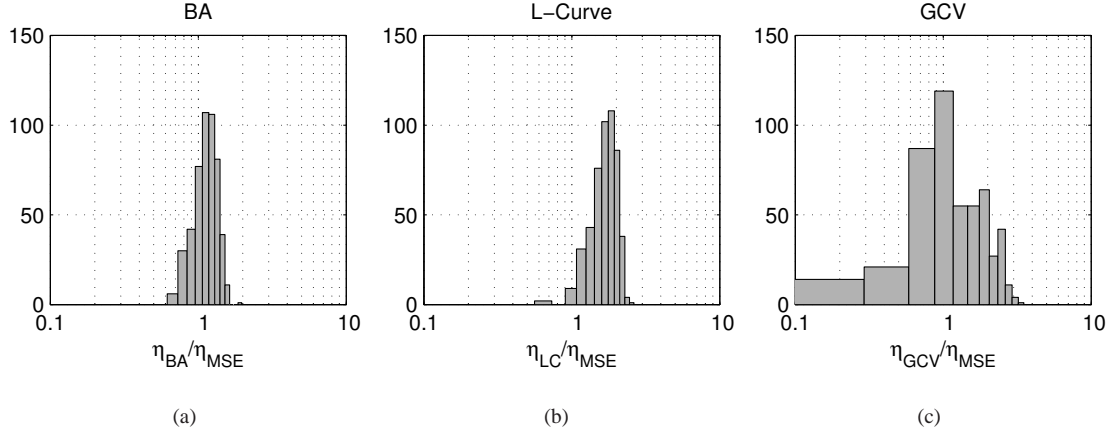


Figure 6: Histogram plot of the ratio between estimated and optimal regularization parameters for Bayesian regularization: (a) BA, (b) L-curve and (c) GCV. They correspond to 500 realizations of measurement noise with a SNR of 30 dB, frequency 1000 Hz and distance source-array  $z_h = 5r_{ss}$ .

first scenario, producing satisfactory results when the array is placed relatively close to the source (except at low frequencies and low SNR) and poor results when it is moved farther away (higher condition number). The L-curve seems to treat this case better than the previous one except at very low frequencies and a frequency band depending on the array-source distance. Again, the Bayesian regularization criterion is able to produce satisfactory results for all source-array distances and over the full range of SNR levels and frequencies. We can notice, however, that the relative error  $\varepsilon_{BA}$  at high frequencies and high SNR are slightly greater than the errors for the non-regularized (NR) and L-curve (LC) cases. It is apparent that no regularization is the best option at those ranges, indeed, the L-curve had no corner for those cases and the employed algorithm applies no regularization. On the other hand, the Bayesian cost function still exhibits a minimum and the solution is slightly oversmoothed. Although, as it will be shown in the next section, introducing a small amount of regularization when the problem is not very ill-conditioned (high frequencies on the simulations) has a small impact on the reconstructed acoustic quantities.

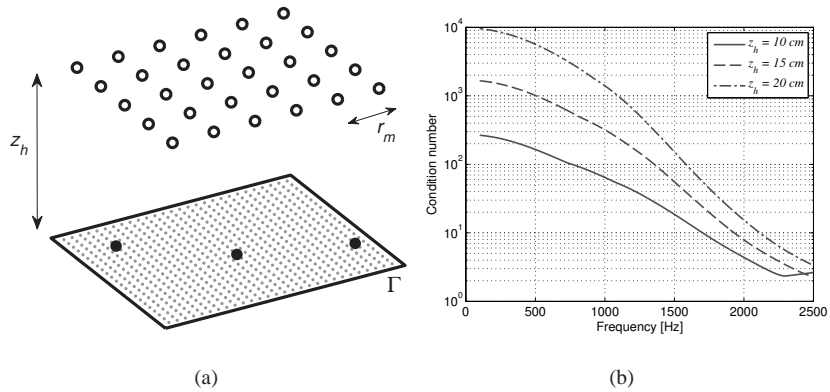


Figure 7: (a) Geometry of the problem for the second scenario. It shows the microphone array, the discretized source surface and 3 point sources randomly placed at the source surface.; (b) Condition number of the transfer matrix for distance  $z_h$  equals to 10, 15 and 20 cm.

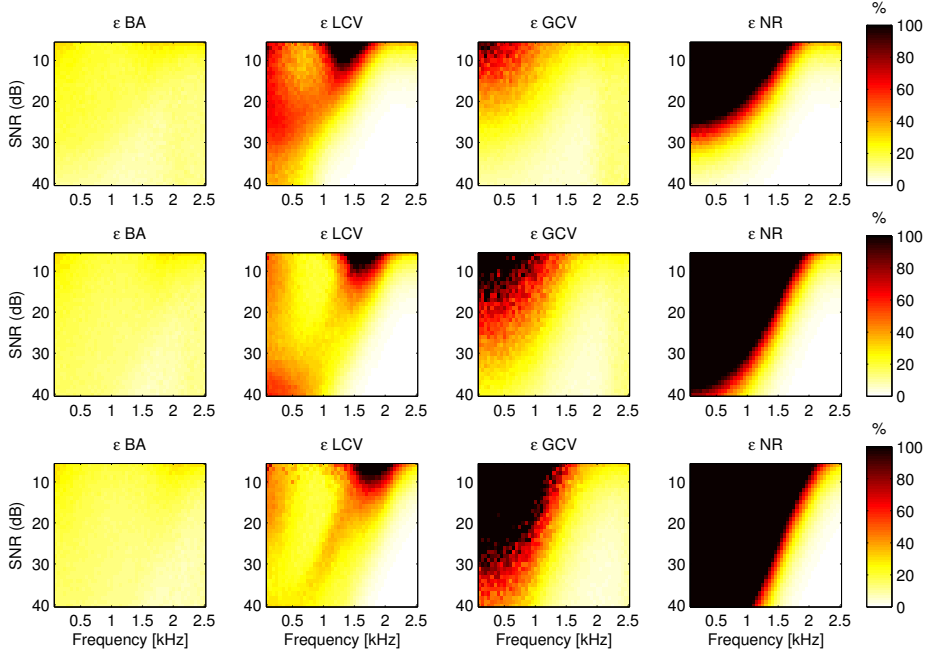


Figure 8: Average value of the relative error to the optimal (MSE) solution over 500 realizations of measurement noise. Each row corresponds to a given distance from the array to the source surface. First row:  $z_h = 10$  cm, second row:  $z_h = 15$  cm and third row:  $z_h = 20$  cm.

## 7. Estimation of confidence intervals

As previously stated in Sec. 5, the Bayesian framework being a probabilistic approach, it allows one to compute parameters of the pdf's assigned to the problem unknowns. One particularly attractive is the posterior pdf of the regularization parameter  $[\eta^2|\mathbf{p}]$ , whose variability can be propagated to any acoustic quantity of interest in order to provide confidence intervals, for instance. Two possible ways of performing this task were discussed in Sec. 5, one by a simple Gaussian approximation of the posterior  $[\eta^2|\mathbf{p}]$  and a second by its exact evaluation using MCMC methods. Both procedures are applied here by means of an example selected from the scenario described in Sec. 6.1. It corresponds to the configuration with the array placed at a distance  $z_h = r_s$  from the source surface and with a SNR of 30 dB. A point estimate of the regularization parameter delivered by the Bayesian criterion is shown in Fig. 9(a) along with its variability (filled gray area) estimated by a MCMC procedure. Figure 9(a) also shows, in dashed black lines, the confidence interval obtained by approximating the posterior pdf  $[\eta^2|\mathbf{p}]$  by a Gaussian distribution. The latter is obtained by numerical evaluation of the second derivative in Eq. (37), which is related to the variance ( $\sigma^2$ ) of the distribution. We can note that the intervals given by the Gaussian approximation agree well with those computed by MCMC up to approximately 800 Hz. Above those frequencies the approximation of the posterior  $[\eta^2|\mathbf{p}]$  by a Gaussian distribution does not hold anymore, illustrating the limits of this approximation. The discontinuities above 1500 Hz are related to oscillations on the condition number of the transfer matrix used in this example (see right panel of Fig. 4 for  $z_h = r_s$ ), since the regularization criterion alternates between a regularized and non regularized solution.



The estimate of source coefficients is obtained by Eq. (14), with the *a priori* assumption that  $\mathbf{\Omega}_n = \mathbf{I}$  and  $\mathbf{\Omega}_c = \mathbf{I}$ , that is, no additional information concerning the spatial correlation on either the measured field or the source field is introduced, practice that is implicit to deterministic methods. We emphasize the fact that these are *a priori* assumptions, and thus do not impede one to obtain a reconstructed source field having a degree of spatial correlation between sources. The basis functions stemming from a singular value decomposition of the propagation operator (see ref. [45]) in Eq. (41) is then used to recover the source field as given by Eq. (2). An integration over the source surface at each frequency gives the source spectrum (see Fig. 9(b)) for the Bayesian regularization criterion (red) and the optimal one (black). The variability of the regularization parameter is then propagated to the reconstructed source field as shown by the filled gray region in Fig. 9(b). It is interesting to note that the reconstructed volume velocity is much more sensitive to the setting of the regularization parameter at low frequencies. Indeed, a small variation on the regularization parameter at low frequencies leads to higher uncertainties on the reconstructed source spectrum. On the other hand, a large variation on  $\eta^2$  at higher frequencies had little effect on the reconstructed quantity. This analysis can be further extended to the acoustic power integrated over the source surface, as shown in Fig. 9(c). We can see that the estimated acoustic power using the Bayesian regularization criterion is fairly close to the optimal (MSE) solution. Moreover, we note narrow confidence intervals on the acoustic power ( $\pm 0.5$  dB), meaning that small variations on the regularization parameter generates small uncertainties on this quantity. We remark that this result is not to be interpreted as if the regularization was not important to the reconstruction, it actually shows how sensitive is the reconstruction of these acoustic quantities to small variations around the point estimate of the regularization parameter.

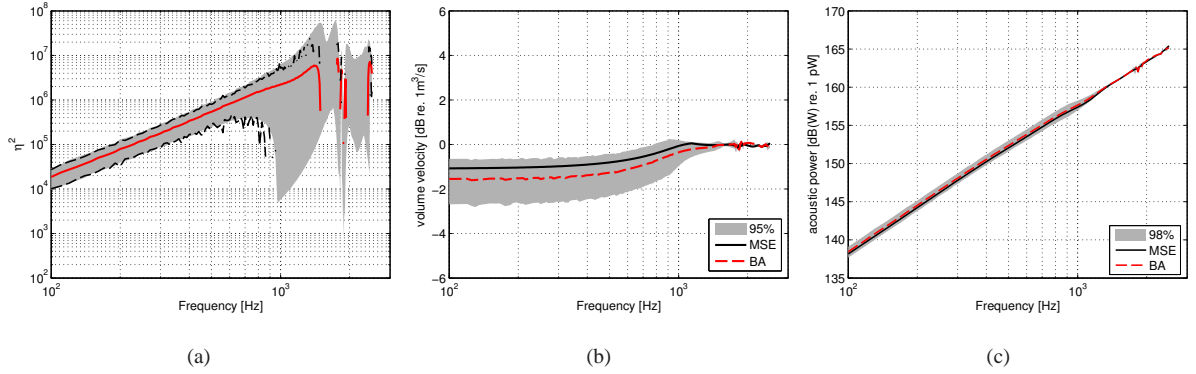


Figure 9: (a) Point estimate of the regularization parameter  $\eta^2$  (—) with confidence intervals estimated by a MCMC procedure (shaded region) and by a Gaussian approximation (—). (b) The reconstructed source spectrum using the Bayesian regularization criterion (—) with a 95% confidence interval (shaded region) and the optimal MSE solution (—). (c) Global acoustic power as integrated over the source surface for the Bayesian regularization criterion (—) with a 98% confidence interval (shaded region) and the optimal (MSE) one (—).

## 8. Experimental results

This section illustrates an application of the Bayesian regularization scheme in an experimental set-up. Before describing the experimental set-up we briefly summarize the *a priori* assumptions in the model. Similarly to the previous simulation examples, we assume that there is no *a priori* spatial correlation over the measured pressure field, that is,  $\mathbf{\Omega}_n$  equals the identity matrix. Other structures for  $\mathbf{\Omega}_n$  could be modeled or even measured, for instance, by performing a measurement with the physical sources switched-off. This idea has been tested in Ref. [45], although in terms of qualitative results the gain was shown to be marginal. The choice of a prior structure to  $\mathbf{\Omega}_c$  (i.e. spatial correlation coefficients of sources) is less obvious in practical cases in which little information concerning the sources is available before making the experiments. Based on the observation that physical sources are often spatially compact, one could for instance, introduce an *a priori* that points on the source field close to each other are correlated, such as to enforce *a priori* the continuity of the source field. We finally stress that the setting of  $\mathbf{\Omega}_n$  or  $\mathbf{\Omega}_c$  is strongly dependent on the practical application. For instance, in aeroacoustics one could use an exponential decay structure for the spatial correlation of the background noise, in case the latter is originated by a turbulent flow. A complete study on the influence of different settings for  $\mathbf{\Omega}_n$  and  $\mathbf{\Omega}_c$  are outside the scope of this paper and is left as a perspective of future work. A reference on the modeling of covariance matrices is [67]. The results presented hereafter are obtained with the simple setting  $\mathbf{\Omega}_c = \mathbf{I}$ , i.e., no *a priori* spatial correlation is introduced.

The source of interest is a driver unit connected to a piping system with three outlets (diameter 2.2 cm), as shown in Fig. 10. This source is a good approximation of three correlated monopoles in the frequency range of interest (200-2000 Hz). A planar array of 6×5 microphones (inter-microphone spacing of 10 cm) is placed at 20 cm from the plane comprising the three outlets. A fictitious source surface (on the plane of outlets) with dimensions 80×70 cm<sup>2</sup> is discretized with a constant spacing of 4 cm. The virtual sources are modeled as point sources and the Green's function given in Eq. (41) is used to compute the propagation to the array of microphones. Equation (14) is again used to solve for the source coefficients with the regularization parameter  $\eta^2$  set by the different approaches. The experiment was carried out in a semi-anechoic room of dimensions 5×3.40×2 m<sup>3</sup> and the source was driven with white random noise. The particle velocity and the sound pressure were measured at 5 cm from each opening using a Microflown *p-u* intensity probe, allowing an estimation of their acoustic power. The radiated acoustic power is estimated by integrating the measured acoustic intensity over a spherical surface of radius 5 cm and centered at each opening.

Figure 11 shows the reconstructed source field (volume velocity) at 945 Hz using the regularization parameter returned by the Bayesian criterion (BA), the L-curve and GCV. It can be seen that GCV fails to compute a reasonable solution, on the other hand, the reconstruction obtained by the L-curve and the Bayesian criterion is fairly good. The cost function at this frequency for each method is shown in Fig. 12. We notice that the L-curve and the Bayesian criterion yield similar regularization parameter, yet in a more convincing way for the latter (note the presence of local maxima on the L-curve's curvature). We can also observe that the minimum of the GCV function (not visible with

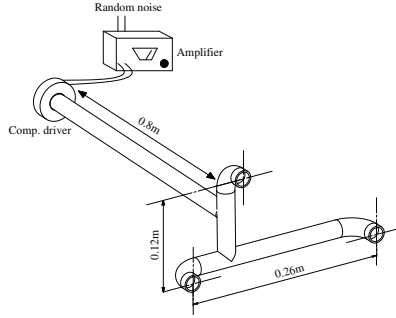


Figure 10: Acoustic source of interest.

the employed scale) is located towards smaller values of  $\eta^2$ , which led to the undersmoothed solution shown in Fig. 11 (right hand side). The reconstructed source field at 650 Hz is shown in Fig. 13. In this case, the GCV and the Bayesian criterion yield similar results, on the contrary, the L-curve criterion led to a rather over-smoothed solution. We remark that at this discrete frequency only two of the three sources radiate significant energy, as shall be confirmed by the measurements to be presented next (see Fig. 16). The cost functions corresponding to this case are represented in Fig. 14. We note that GCV and the presented Bayesian criterion returned similar regularization parameters. On the other hand, the L-curve's curvature exhibits two local maxima very close to each other in level. A better alternative in this particular case, would rather be to select the local maximum slightly on the left of the global maximum.

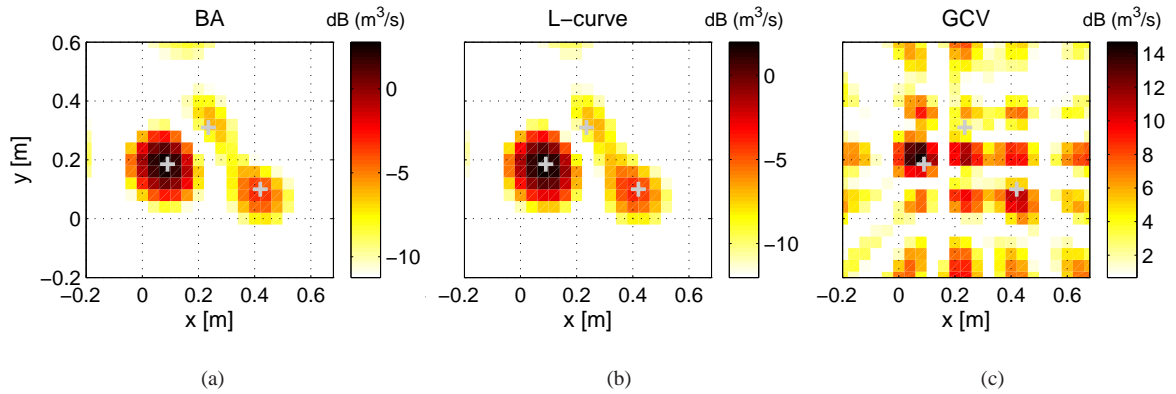


Figure 11: Reconstructed source field (volume velocity) at 945 Hz corresponding to each regularization criterion. (a) Bayesian regularization (BA); (b) L-curve and (c) GCV. A dynamic range of 15 dB is used for all.

The acoustic power radiated by the reconstructed source field was computed in a similar manner as in Ref. [68]. It essentially consists of summing the power radiated by each virtual point source in isolation and a term that is the power generated by each source in working against the induced pressure generated by neighbor sources (see, for example, Ref. [69]). Figure 15 compares the reconstructed global acoustic power (as integrated over the source surface) to the reference  $p$ - $u$  measurements. As can be seen, the GCV results diverge at some frequencies, which are related to a severe underestimation of the regularization parameter. For the L-curve, although the results are globally satisfactory,

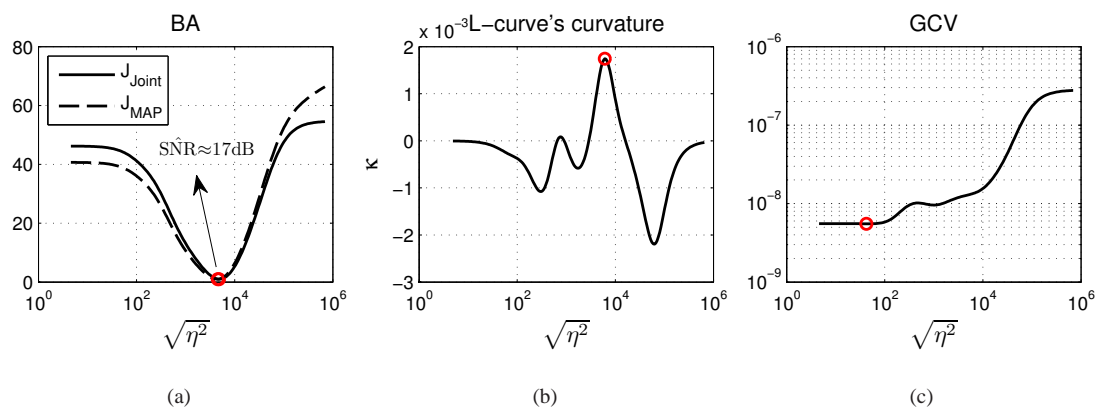


Figure 12: Cost functions associated to each regularization criterion at the frequency of 945 Hz. (a) the Bayesian regularization criteria. (b) curvature of the L-curve (whose corner is the point with maximum curvature). (c) the GCV cost function.

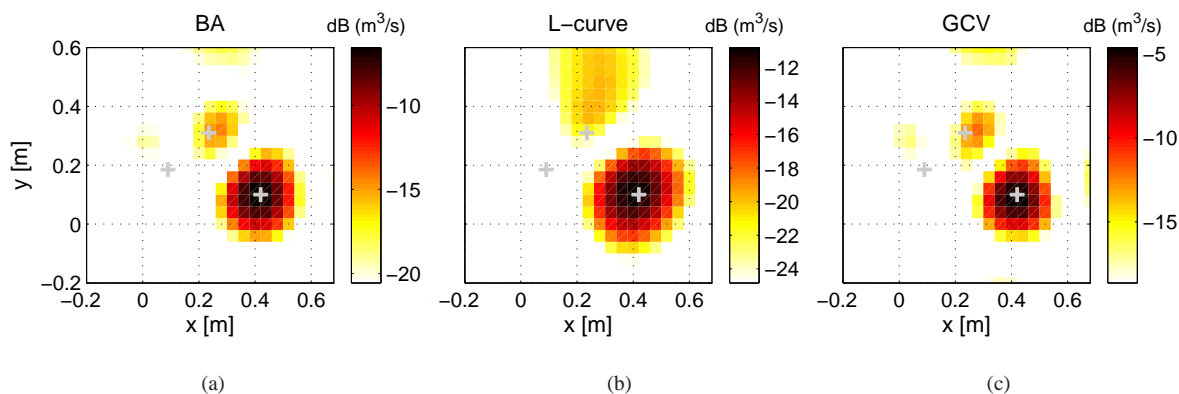


Figure 13: Reconstructed source field (volume velocity) at 650 Hz corresponding to each regularization criterion. (a) Bayesian regularization (BA); (b) L-curve and (c) GCV. A dynamic range of 15 dB is used for all.

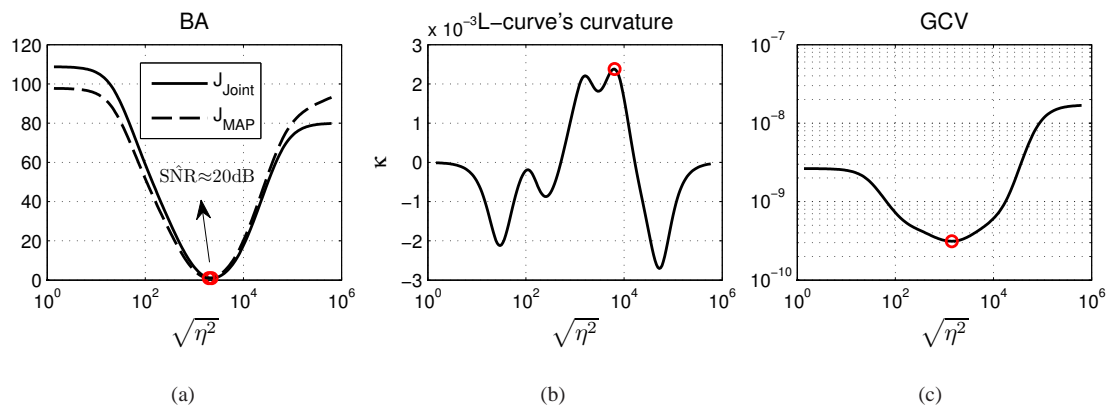


Figure 14: Cost functions associated to each regularization criterion at the frequency of 650 Hz. (a) the Bayesian regularization criteria. (b) curvature of the L-curve (whose corner is the point with maximum curvature). (c) the GCV cost function.

we notice few discontinuities for instance around 650 and 970 Hz. Further investigation has shown that they are due to alternations among competing local maxima of curvature evolving with frequency. In turn, the estimate given by the Bayesian criterion is found more stable in the whole frequency range, illustrating its robustness. The acoustic power radiated by each opening has also been predicted and compared to the measurements (see Fig. 16). The predictions were obtained by carefully choosing an integration area around each identified source, as shown on the top left panel of Fig. 16. One can notice that the three sources are well located around their real positions and their contributions in terms of acoustic power are fairly well predicted.

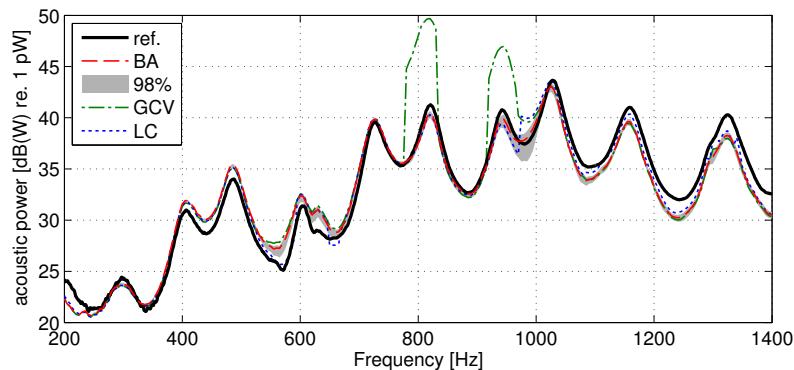


Figure 15: Estimate of the total acoustic power for different regularization strategies. The reference (ref.) was obtained from  $p$ - $u$  measurements. The figure also shows a 98% Bayesian confidence interval around the point estimate returned by the proposed Bayesian criterion (BA).

## 9. Conclusion

A vast literature covers the issue of regularization on which critically depends the solution of inverse acoustic problems. Taking a rather different route than traditional approaches encountered in acoustics, the paper demonstrates the advantage of using a solution based on (empirical) Bayesian regularization. As evidenced by numerical simulations and experimental results, the proposed regularization approach presents an improved performance as compared to the state-of-the-art in acoustics. The gap seems significant possibly because the inverse acoustical problem is so much ill-posed. Technically speaking, one important result of the paper is to deliver the posterior probability density function of the regularization parameter as well as the joint posterior probability density function of the source and noise energies. The optimal regularization parameter is then directly estimated as the maximum (most probable value) of the former or, indirectly, as the ratio of the most probable values of noise and source energies (SNR ratio). In practice, the two estimates are found very similar and virtually indistinguishable when the numbers of sensors or snapshots is high. Although necessary to prove its main results, the probabilistic apparatus deployed in the paper is ultimately transparent to the end user. Indeed, the automatic selection of the optimal regularization parameter amounts to searching for the global minimum of a 1D cost function, which lends itself to easy implementation. Numerous experimental results have demonstrated the superiority of the proposed Bayesian regularization to GCV

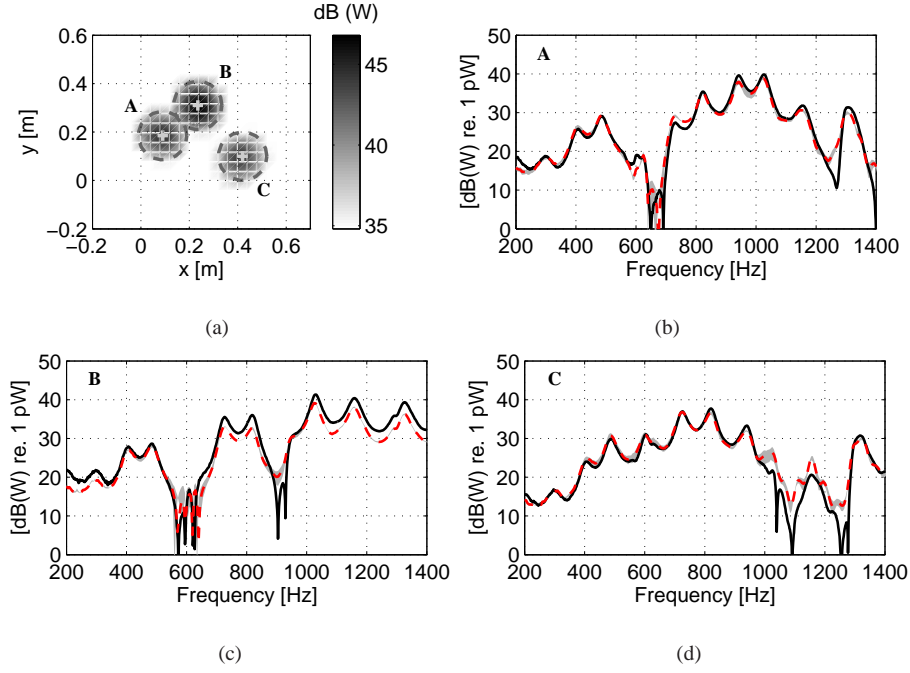


Figure 16: (a) reconstructed acoustic power map integrated over the frequency band 200-1400 Hz. The + symbols are the real source positions and the dashed circles are the integration area used to compute the partial PSD of each source. The other panels show the PSD of each source identified by the proposed method (— —) and obtained from  $p$ - $u$  measurements (—). (b) PSD of source A. (c) PSD of source B. (d) PSD of source C.

and the L-curve, whose use have been prevailing in acoustics. Another, rather unique, advantage of the approach is to allow sampling the posterior probability density function of the regularization parameter, from which variability due to regularization can be propagated to any reconstructed acoustical quantity. This opens many perspectives for future research. In conclusion, it seems that Bayesian regularization enjoys so many nice properties that there are few, if any, reasons why not to use it.

## Acknowledgements

This work was conducted in the framework of the LabEx CeLyA (“Centre Lyonnais d’Acoustique”, ANR-10-LABX-60)

## Appendix A.

**Proposition 4.** Let  $\mathbf{\Omega}_\phi$  be a  $K \times K$  matrix whose generic element is given by  $[\mathbf{\Omega}_\phi]_{kl} = \int_{\Gamma} \phi_k(\mathbf{r}) \phi_l^*(\mathbf{r}) d\Gamma(\mathbf{r})$ . Then

$$\mathbf{\Omega}_c = \frac{\mathbf{\Omega}_\phi^{-1}}{\text{Trace}\{\mathbf{\Omega}_\phi^{-1}\}} \quad (\text{A.1})$$

minimizes the scale length of the covariance function (6) – i.e. it makes it the closest possible to a delta function in the mean-square sense.

PROOF. : Let us find a Hermitian matrix  $\mathbf{A}$  in  $C_q(\mathbf{r}, \mathbf{r}') = \sum_{k=1}^K A_{kl} \phi_k(\mathbf{r}) \phi_l^*(\mathbf{r}')$  such that it minimizes

$$\begin{aligned} \|\delta(\mathbf{r} - \mathbf{r}') - C_q(\mathbf{r}, \mathbf{r}')\|^2 &\triangleq \int_{\Gamma} \int_{\Gamma} |\delta(\mathbf{r} - \mathbf{r}') - C_q(\mathbf{r}, \mathbf{r}')|^2 d\Gamma(\mathbf{r}) d\Gamma(\mathbf{r}') \\ &= 1 - 2 \int_{\Gamma} C_q(\mathbf{r}, \mathbf{r}) d\Gamma(\mathbf{r}) + \int_{\Gamma} \int_{\Gamma} |C_q(\mathbf{r}, \mathbf{r}')|^2 d\Gamma(\mathbf{r}) d\Gamma(\mathbf{r}') \\ &= 1 - 2 \sum_{k,l=1}^K A_{kl} \int_{\Gamma} \phi_k(\mathbf{r}) \phi_l^*(\mathbf{r}) d\Gamma(\mathbf{r}) + \sum_{k,l=1}^K \sum_{i,j=1}^K A_{kl} A_{ij}^* \int_{\Gamma} \int_{\Gamma} \phi_k(\mathbf{r}) \phi_l^*(\mathbf{r}') \phi_i^*(\mathbf{r}) \phi_j(\mathbf{r}') d\Gamma(\mathbf{r}) d\Gamma(\mathbf{r}'). \end{aligned} \quad (\text{A.2})$$

Setting the derivative with respect to  $A_{kl}$  to zero then returns

$$\int_{\Gamma} \phi_k(\mathbf{r}) \phi_l^*(\mathbf{r}) d\Gamma(\mathbf{r}) = \sum_{i,j=1}^K A_{ij}^* \int_{\Gamma} \int_{\Gamma} \phi_k(\mathbf{r}) \phi_l^*(\mathbf{r}') \phi_i^*(\mathbf{r}) \phi_j(\mathbf{r}') d\Gamma(\mathbf{r}) d\Gamma(\mathbf{r}'). \quad (\text{A.3})$$

Putting  $\int_{\Gamma} \phi_k(\mathbf{r}) \phi_l^*(\mathbf{r}) d\Gamma(\mathbf{r}) = [\mathbf{\Omega}_{\phi}]_{kl}$ , one then arrives at the system of equations  $\mathbf{\Omega}_{\phi} = \mathbf{\Omega}_{\phi} \mathbf{A} \mathbf{\Omega}_{\phi}$ . In all problems of interest matrix  $\mathbf{\Omega}_{\phi}$  will be invertible (if not, this means that order  $K$  in Eq. (2) can be reduced to a lower value until invertibility is met), so that finally  $\mathbf{A} = \mathbf{\Omega}_{\phi}^{-1}$ .

The setting recommended in Proposition (4) is actually equivalent to using the *orthogonal* spatial basis  $\{\psi_k(\mathbf{r}) = \sum_{l=1}^K \phi_l(\mathbf{r}) [\mathbf{\Omega}_c^{-1/2}]_{lk}; k = 1, \dots, K\}$  with the identity covariance matrix. The fact that it produces the finest spatial resolution – an important result in practice – turns out obvious within the Bayesian formalism, whilst it might not be so from a deterministic approach. It is interesting to note that the result in Proposition (4) is independent of the choice of the decomposition basis.

## Appendix B.

### Proposition 5.

$$[\mathbf{p}|\alpha^2, \beta^2] = \mathcal{N}_C(\mathbf{0}, \alpha^2 \mathbf{H} \mathbf{\Omega}_c \mathbf{H}^H + \beta^2 \mathbf{\Omega}_n) \quad (\text{B.1a})$$

$$= \frac{\exp\left(-\alpha^{-2} \sum_{k=1}^M \frac{|y_k|^2}{s_k^2 + \eta^2}\right)}{\pi^M \alpha^{2M} |\mathbf{\Omega}_n| \prod_{k=1}^M (s_k^2 + \eta^2)} \quad (\text{B.1b})$$

where  $y_k$  is the  $k$ -th element of vector

$$\mathbf{y} = \mathbf{U}^H \mathbf{\Omega}_n^{-1/2} \mathbf{p}, \quad (\text{B.2})$$

the projection of the measurements on the array subspace.

PROOF. : This proof concerns a standard result in the Bayesian treatment of linear regression and can be found in classical references such as [47–50]. It is reproduced here for the sake of completeness. The goal is to compute integral

$$[\mathbf{p}|\alpha^2, \beta^2] = \int [\mathbf{p}|\mathbf{c}, \beta^2] [\mathbf{c}|\alpha^2] d\mathbf{c} = \int \mathcal{N}_C(\mathbf{H}\mathbf{c}, \beta^2 \mathbf{\Omega}_n) \mathcal{N}_C(\mathbf{0}, \alpha^2 \mathbf{\Omega}_c) d\mathbf{c}. \quad (\text{B.3})$$

A direct calculation based on expanding the product of the Gaussians and “completing the square” is fastidious, although not difficult. A more direct proof proceeds as follows. Let us first suppose that  $(K - M)$  extra measurements are taken (while keeping the number of coefficients  $c_k$ 's constant), thus producing the extended vectors  $\mathbf{p}_a \in \mathbb{C}^K$ ,  $\mathbf{n}_a \in \mathbb{C}^K$ , and matrices  $\mathbf{\Omega}_{n,a} \in \mathbb{C}^{K \times K}$  and  $\mathbf{H}_a \in \mathbb{C}^{K \times K}$ , where  $\mathbf{H}_a$  is supposed of full-rank. Solving for this new problem requires computing the integral

$$\int [\mathbf{p}_a | \mathbf{c}, \beta^2] [\mathbf{c} | \alpha^2] d\mathbf{c} = \int [\mathbf{n}_a = \mathbf{p}_a - \mathbf{H}_a \mathbf{c} | \beta^2] [\mathbf{H}_a \mathbf{c} | \alpha^2] d(\mathbf{H}_a \mathbf{c}) = \int [\mathbf{p}_a | \mathbf{H}_a \mathbf{c}, \beta^2] [\mathbf{H}_a \mathbf{c} | \alpha^2] d(\mathbf{H}_a \mathbf{c}) \quad (\text{B.4})$$

wherein  $[\mathbf{n}_a = \mathbf{p}_a - \mathbf{H}_a \mathbf{c} | \beta^2]$  stands for the pdf of  $\mathbf{n}_a$  evaluated at  $\mathbf{p}_a - \mathbf{H}_a \mathbf{c}$  and the change of variable  $\mathbf{c} \mapsto \mathbf{H}_a \mathbf{c}$  was used. From first principles of probability calculus, integral (B.4) is recognized as the pdf of the sum of two random variables, i.e.  $\mathbf{p}_a = \mathbf{H}_a \mathbf{c} + \mathbf{n}_a$ , with pdf's  $[\mathbf{H}_a \mathbf{c} | \alpha^2] = \mathcal{N}_C(\mathbf{0}, \alpha^2 \mathbf{H}_a \mathbf{\Omega}_c \mathbf{H}_a^H)$  and  $[\mathbf{n}_a | \beta^2] = \mathcal{N}_C(\mathbf{0}, \beta^2 \mathbf{\Omega}_{n,a})$ , respectively. The sum of two independent Gaussian random variables is again a Gaussian [48], with mean equal to the sum of their individual means and similarly for its covariance. Therefore,

$$[\mathbf{p}_a | \alpha^2, \beta^2] = \mathcal{N}_C(\mathbf{0}, \alpha^2 \mathbf{H}_a \mathbf{\Omega}_c \mathbf{H}_a^H + \beta^2 \mathbf{\Omega}_{n,a}). \quad (\text{B.5})$$

The last point is to marginalize the above pdf over the  $(K - M)$  extra measurements. As well-known for Gaussian distributions, this is simply  $\mathcal{N}_C(\mathbf{0}, \alpha^2 \mathbf{H} \mathbf{\Omega}_c \mathbf{H}^H + \beta^2 \mathbf{\Omega}_n)$  [48]. Next, expression (B.1b) is easily worked out using the eigen-elements in Eq. (13).

## Appendix C.

### Proposition 6.

$$[\eta^2 | \mathbf{p}] \propto \left( \left( \sum_{k=1}^M \frac{\langle |y_k|^2 \rangle}{s_k^2 + \eta^2} \right)^{MN-2} \prod_{k=1}^M (s_k^2 + \eta^2)^N \right)^{-1}, \quad MN > 1. \quad (\text{C.1})$$

PROOF. : The integral to be evaluated is

$$\begin{aligned} [\eta^2 | \mathbf{p}] &= \int [\eta^2, \alpha^2 | \mathbf{p}] d\alpha^2 = \int [\alpha^2, \beta^2 | \mathbf{p}] \underbrace{\left| \frac{d\beta^2}{d\eta^2} \right|}_{\alpha^2} d\alpha^2 \\ &\propto \int [\mathbf{p} | \alpha^2, \beta^2] [\alpha^2, \beta^2] \alpha^2 d\alpha^2 \\ &\propto \int \mathcal{N}_C(\mathbf{0}, \alpha^2 (\mathbf{H} \mathbf{\Omega}_c \mathbf{H}^H + \eta^2 \mathbf{\Omega}_n)) |_{\beta^2 = \alpha^2 \eta^2} \alpha^2 d\alpha^2 \end{aligned} \quad (\text{C.2})$$

where it was assumed that  $[\alpha^2, \beta^2] \propto 1$  as discussed in section 3.1. Now, using Eqs. (B.1b) (with  $N$  snapshots) and (B.1b),

$$[\eta^2 | \mathbf{p}] \propto \frac{1}{\prod_{k=1}^M (s_k^2 + \eta^2)^N} \int_0^\infty \frac{\exp(-MN \hat{\alpha}^2 / \alpha^2)}{\alpha^{2(MN-1)}} d\alpha^2 \quad (\text{C.3})$$

with  $\hat{\alpha}^2$  as given in Eq. (24). The integral in the above equation is  $\Gamma(MN - 2)(MN \hat{\alpha}^2)^{-(MN-2)}$ ,  $MN > 1$ , thus proving Eq. (C.1).



## Appendix D.

The proof essentially consists in showing that all stationary points have positive curvatures with overwhelming probability, i.e. that the second derivative of  $\mathbf{J}_{\text{Joint}}(\eta^2)$  is positive with a very high probability whenever the first derivative vanishes (on the opposite, the existence of several local minima would necessarily involve intermediate maxima with significant probability). First, stationary points are easily found to satisfy the following equation

$$\frac{d\mathbf{J}_{\text{Joint}}(\eta^2)}{d\eta^2} = \sum_{k=1}^M \frac{1}{s_k^2 + \eta^2} - \sum_{k=1}^M \frac{\langle |y_k|^2 \rangle}{\hat{\alpha}^2 (s_k^2 + \eta^2)^2} = 0 \quad (\text{D.1})$$

with  $\hat{\alpha}^2$  as given in Eq. (24). Next, the second derivative at those stationary points is found equal to

$$\frac{d^2\mathbf{J}_{\text{Joint}}(\eta^2)}{(d\eta^2)^2} = 2 \sum_{k=1}^M \frac{\langle |y_k|^2 \rangle}{\hat{\alpha}^2 (s_k^2 + \eta^2)^6} - \sum_{k=1}^M \frac{1}{(s_k^2 + \eta^2)^4} - \frac{1}{M} \left( \sum_{k=1}^M \frac{1}{(s_k^2 + \eta^2)^2} \right)^2. \quad (\text{D.2})$$

It is unfortunately impossible to prove the above expression is always positive, whatever the values of the  $\langle |y_k|^2 \rangle$ 's,  $s_k^2$ 's, and  $\eta^2$ . However, one may prove the probability it takes negative values is overwhelmingly small. To see this, let us first assume that  $\alpha^2 \simeq \hat{\alpha}^2$ , which is reasonable enough when  $MN \gg 1$ . Next, let us introduce

$$S_q \triangleq \sum_{k=1}^M \frac{1}{(s_k^2 + \eta^2)^q} \text{ and } B_q \triangleq \frac{1}{M} \frac{S_q^2}{S_{2q}}. \quad (\text{D.3})$$

Thus, the probability of a negative second derivative reads

$$\mathbb{P} \left( \sum_{k=1}^M \frac{\langle |y_k|^2 \rangle}{\alpha^2 (s_k^2 + \eta^2)^6} < \frac{S_4}{2} (1 + B_2) \right), \quad (\text{D.4})$$

or, equivalently,

$$\mathbb{P} \left( e^{-\lambda \sum_{k=1}^M \frac{\langle |y_k|^2 \rangle}{\alpha^2 (s_k^2 + \eta^2)^6}} \geq e^{-\frac{\lambda}{2} S_4 (1 + B_2)} \right) \leq \frac{\mathbb{E} \left\{ e^{-\lambda \sum_{k=1}^M \frac{\langle |y_k|^2 \rangle}{\alpha^2 (s_k^2 + \eta^2)^6}} \right\}}{e^{-\frac{\lambda}{2} S_4 (1 + B_2)}} \quad (\text{D.5})$$

for any  $\lambda \geq 0$ , wherein the upper bound follows from Markov's inequality [70]. Now, it results from Proposition 2 that  $2N\langle |y_k|^2 \rangle / (\alpha^2 (s_k^2 + \eta^2))$  are independent and identically distributed Chi-2 random variables with  $2N$  degrees of freedom,  $\chi_{2N}^2$ . Thus,

$$\begin{aligned} \mathbb{E} \left\{ e^{-\lambda \sum_{k=1}^M \frac{\langle |y_k|^2 \rangle}{\alpha^2 (s_k^2 + \eta^2)^6}} \right\} &= \prod_{k=1}^M \mathbb{E} \left\{ e^{-\frac{\lambda}{2N} \frac{\chi_{2N}^2}{(s_k^2 + \eta^2)^2}} \right\} = \prod_{k=1}^M \left( 1 + \frac{\lambda}{N(s_k^2 + \eta^2)^2} \right)^{-N} \\ &= e^{-N \sum_{k=1}^M \ln \left( 1 + \frac{\lambda}{N(s_k^2 + \eta^2)^2} \right)} \leq e^{-\lambda \sum_{k=1}^M (s_k^2 + \eta^2)^{-2} \left( 1 - \frac{\lambda}{2N(s_k^2 + \eta^2)^2} \right)}, \end{aligned} \quad (\text{D.6})$$

where the last term in the first line stems from the definition of the moment generating function of a Chi-2 random variable and the last inequality in the second line from the fact that  $\ln(1 + x) \geq x - 1/2x^2$  [70]. Therefore, the upper bound in Eq. D.5 becomes

$$e^{-\frac{\lambda}{2} S_4 (1 - B_2 - \frac{\lambda}{N} \frac{S_8}{S_4})}. \quad (\text{D.7})$$

Since  $\lambda$  is arbitrary, one may set it up so as to minimize the above upper bound. Setting the derivative of Eq. (D.7) with respect to  $\lambda$  to zero, one finds,  $\lambda = 1/2(1 - B_2)MN B_4/S_4$ . Plugging in Eq. (D.7) one finally arrives at

$$\mathbb{P}\left(\frac{d\mathbf{J}_{\text{Joint}}(\eta^2)}{d\eta^2} < 0\right) \leq e^{-\frac{1}{8}(1-B_2)^2 MN B_4} \xrightarrow{MN \rightarrow \infty} 0. \quad (\text{D.8})$$

This probability is seen to decrease exponentially fast to zero with the product  $MN$  of the numbers of microphones and snapshots. It also depends on  $B_2$  and  $B_4$  which, according to Eqs. (D.3), are to be interpreted as measures of bandwidth of the sequence  $(s_k^2 + \eta^2)^{-1}$ ; in particular, a constant-valued sequence is especially critical because  $B_2 = 1$  jeopardizes the convergence to zero of the upper bound in inequality (D.4) – which is reminiscent to having an infinite variance in Proposition 3. However, it can be readily verified that expression (D.1) is positive under this particular circumstance.

Following the same lines, a similar result is proved for estimate  $\hat{\eta}_{\text{MAP}}^2$ ; namely,

$$\mathbb{P}\left(\frac{d\mathbf{J}_{\text{Joint}}(\eta^2)}{d\eta^2} < 0\right) \leq e^{-\frac{1}{8} \frac{(1-B_2-3\epsilon+2\epsilon^2)^2 MN B_4}{(1-\epsilon)^2}} \xrightarrow{MN \rightarrow \infty} 0. \quad (\text{D.9})$$

with  $\epsilon = 2/(MN)$ .

## Appendix E.

Let us expand  $\mathbf{J}_{\text{MAP}}(\eta^2)$  in Eq. (28) into a second order Taylor series about  $\hat{\eta}_{\text{MAP}}^2$ :

$$\mathbf{J}_{\text{MAP}}(\eta^2) = \mathbf{J}_{\text{MAP}}(\hat{\eta}_{\text{MAP}}^2) + \frac{|\eta^2 - \hat{\eta}_{\text{MAP}}^2|^2}{2} \left( \frac{d^2 \mathbf{J}_{\text{MAP}}(\hat{\eta}_{\text{MAP}}^2)}{(d\eta^2)^2} \right)_{\eta^2 = \hat{\eta}_{\text{MAP}}^2} + \mathcal{O}(|\eta^2 - \hat{\eta}_{\text{MAP}}^2|^3) \quad (\text{E.1})$$

where  $d\mathbf{J}_{\text{MAP}}(\eta^2)/d\eta^2|_{\eta^2 = \hat{\eta}_{\text{MAP}}^2} = 0$  (by definition of  $\hat{\eta}_{\text{MAP}}^2$ ) has been used and  $\mathcal{O}$  means ‘‘on the order of’’. Therefore,

$$[\eta^2 | \mathbf{p}] \propto \exp(-\mathbf{J}_{\text{MAP}}(\eta^2)) \approx \exp\left(-\frac{|\eta^2 - \hat{\eta}_{\text{MAP}}^2|^2}{2} \frac{d^2 \mathbf{J}_{\text{MAP}}(\hat{\eta}_{\text{MAP}}^2)}{(d\eta^2)^2}\right), \quad (\text{E.2})$$

which proves Proposition 2.

## Appendix F.

The goal is to find the expected value of expression (D.1). Assuming as before that  $\alpha^2 \simeq \hat{\alpha}^2$  when  $MN \gg 1$  and keeping in mind that  $\langle |y_k|^2 \rangle \sim \chi_{2N}^2/(2N)$ ,

$$\mathbb{E}\{\langle |y_k|^2 \rangle\} = \alpha^2 (s_k^2 + \eta^2). \quad (\text{F.1})$$

Therefore

$$\mathbb{E}\left\{\frac{d^2 \mathbf{J}_{\text{Joint}}(\eta^2)}{(d\eta^2)^2}\right\} = 2S_4 - S_4 - \frac{1}{M} S_2^2 = S_4(1 - B_2), \quad (\text{F.2})$$

with  $S_2$ ,  $S_4$ , and  $B_2$  as defined in Eq. (D.3).

## References

- [1] A. N. Tikhonov, V. I. A. Arsenin, *Solutions of ill-posed problems*, Scripta series in mathematics, Winston, 1977.
- [2] P.-A. Gauthier, C. Camier, Y. Pasco, A. Berry, E. Chambatte, R. Lapointe, M.-A. Delalay, Beamforming regularization matrix and inverse problems applied to sound field measurement and extrapolation using microphone array, *Journal of Sound and Vibration* 330 (24) (2011) 5852 – 5877.
- [3] Y. H. Hong, H.-K. Kim, H. S. Lee, Reconstruction of dynamic displacement and velocity from measured accelerations using the variational statement of an inverse problem, *Journal of Sound and Vibration* 329 (23) (2010) 4980 – 5003.
- [4] B.-K. Kim, J.-G. Ih, Design of an optimal wave-vector filter for enhancing the resolution of reconstructed source field by near-field acoustical holography (nah), *The Journal of the Acoustical Society of America* 107 (6) (2000) 3289–3297.
- [5] Y. Kim, P. A. Nelson, Spatial resolution limits for the reconstruction of acoustic source strength by inverse methods, *Journal of Sound and Vibration* 265 (3) (2003) 583 – 608.
- [6] K. Saijyou, C. Okawara, Regularization method for measurement of structural intensity using nearfield acoustical holography, *The Journal of the Acoustical Society of America* 117 (4) (2005) 2039–2045.
- [7] K. Saijyou, Regularization method for the application of k-space data extrapolation to near-field acoustical holography, *The Journal of the Acoustical Society of America* 116 (1) (2004) 396–404.
- [8] A. N. Thite, 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) (2003) 433 – 451.
- [9] N. Valdivia, E. G. Williams, Krylov subspace iterative methods for boundary element method based near-field acoustic holography, *The Journal of the Acoustical Society of America* 117 (2) (2005) 711–724.
- [10] E. G. Williams, Regularization methods for near-field acoustical holography, *The Journal of the Acoustical Society of America* 110 (4) (2001) 1976–1988.
- [11] P. A. G. Zavala, W. D. Roeck, K. Janssens, J. R. F. Arruda, P. Sas, W. Desmet, Generalized inverse beamforming with optimized regularization strategy, *Mechanical Systems and Signal Processing* 25 (3) (2011) 928 – 939.
- [12] H. G. Choi, A. N. Thite, D. J. Thompson, A threshold for the use of tikhonov regularization in inverse force determination, *Applied Acoustics* 67 (7) (2006) 700 – 719.
- [13] E. Jacquelin, A. Bennani, P. Hamelin, Force reconstruction: analysis and regularization of a deconvolution problem, *Journal of Sound and Vibration* 265 (1) (2003) 81 – 107.
- [14] M. Khalil, S. Adhikari, A. Sarkar, Linear system identification using proper orthogonal decomposition, *Mechanical Systems and Signal Processing* 21 (8) (2007) 3123 – 3145.
- [15] J. Luo, X. D. Li, An inverse aeroacoustic problem on rotor wake/stator interaction, *Journal of Sound and Vibration* 254 (2) (2002) 219 – 229.
- [16] P. A. Nelson, S. H. Yoon, Estimation of acoustic source strength by inverse methods: Part i, conditioning of the inverse problem, *Journal of Sound and Vibration* 233 (4) (2000) 639 – 664.
- [17] S. H. Yoon, P. A. Nelson, Estimation of acoustic source strength by inverse methods: Part ii, experimental investigation of methods for choosing regularization parameters, *Journal of Sound and Vibration* 233 (4) (2000) 665 – 701.
- [18] 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.
- [19] P. C. Hansen, D. P. O’Leary, The use of the l-curve in the regularization of discrete ill-posed problems, *SIAM Journal on Scientific Computing* 14 (1993) 1487–1503.
- [20] P. C. Hansen, M. E. Kilmer, R. H. Kjeldsen, Exploiting residual information in the parameter choice for discrete ill-posed problems, *BIT Numerical Mathematics* 46 (2006) 41–59.
- [21] B. W. Rust, D. P. O’Leary, Residual periodograms for choosing regularization parameters for ill-posed problems, *Inverse Problems* 24 (3) (2008) 034005.

- [22] Y. Kim, P. A. Nelson, Optimal regularisation for acoustic source reconstruction by inverse methods, *Journal of Sound and Vibration* 275 (3-5) (2004) 463 – 487.
- [23] J. Gomes, P. C. Hansen, A study on regularization parameter choice in near-field acoustical holography, in: *Proceedings of Acoustics'08*, Paris, France, 2008.
- [24] Q. Leclère, Acoustic imaging using under-determined inverse approaches: Frequency limitations and optimal regularization, *Journal of Sound and Vibration* 321 (3-5) (2009) 605 – 619.
- [25] A. Schuhmacher, J. Hald, K. B. Rasmussen, P. C. Hansen, Sound source reconstruction using inverse boundary element calculations, *The Journal of the Acoustical Society of America* 113 (1) (2003) 114–127.
- [26] M. B. S. Magalhaes, R. A. Tenenbaum, Sound sources reconstruction techniques: A review of their evolution and new trends, *Acta Acustica united with Acustica* 90 (2) (2004) 199–220.
- [27] H. G. Choi, A. N. Thite, D. J. Thompson, Comparison of methods for parameter selection in tikhonov regularization with application to inverse force determination, *Journal of Sound and Vibration* 304 (3-5) (2007) 894 – 917.
- [28] N. P. Valdivia, E. G. Williams, P. C. Herdic, B. Houston, Surface decomposition method for near-field acoustic holography, *The Journal of the Acoustical Society of America* 132 (1) (2012) 186–196.
- [29] G. Chardon, L. Daudet, A. Peillot, F. Ollivier, N. Bertin, R. Gribonval, Near-field acoustic holography using sparse regularization and compressive sampling principles, *The Journal of the Acoustical Society of America* 132 (3) (2012) 1521–1534.
- [30] N. Chu, A. Mohammad-Djafari, J. Picheral, Robust bayesian super-resolution approach via sparsity enforcing a priori for near-field aeroacoustic source imaging, *Journal of Sound and Vibration* 332 (18) (2013) 4369 – 4389.
- [31] N. Chu, J. Picheral, A. Mohammad-Djafari, N. Gac, A robust super-resolution approach with sparsity constraint in acoustic imaging, *Applied Acoustics* 76 (2014) 197 – 208.
- [32] F. Bauer, M. A. Lukas, Comparing parameter choice methods for regularization of ill-posed problems, *Mathematics and Computers in Simulation* 81 (9) (2011) 1795 – 1841.
- [33] A. M. Thompson, J. Kay, On some bayesian choices of regularization parameter in image restoration, *Inverse Problems* 9 (6) (1993) 749.
- [34] D. J. C. MacKay, Bayesian interpolation, *Neural Computation* 4 (1991) 415–447.
- [35] D. J. C. MacKay, Comparison of approximate methods for handling hyperparameters, *Neural Computation* 11 (5) (1999) 1035–1068.
- [36] D. J. C. MacKay, Hyperparameters: optimize, or integrate out, in: *In Maximum Entropy and Bayesian Methods*, Santa Barbara, Kluwer, 1996, pp. 43–60.
- [37] R. Molina, A. K. Katsaggelos, J. Mateos, Bayesian and regularization methods for hyperparameter estimation in image restoration, *Image Processing, IEEE Transactions on* 8 (2) (1999) 231 –246.
- [38] A. Mohammad-Djafari, On the estimation of hyperparameters in bayesian approach of solving inverse problems, in: *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 5, 1993, pp. 495 –498.
- [39] A. Mohammad-Djafari, Joint estimation of parameters and hyperparameters in a bayesian approach of solving inverse problems, in: *Proceedings of International Conference on Image Processing*, Vol. 1, 1996, pp. 473–476 vol.2.
- [40] A. Mohammad-Djafari, A full bayesian approach for inverse problems, in: *In Maximum Entropy and Bayesian Methods*, Kluwer Academic Publishers, 1996, pp. 135–143.
- [41] G. Wahba, A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem, *The Annals of Statistics* 13 (4) (1985) 1378–1402.
- [42] S. F. Gull, Bayesian inductive inference and maximum entropy, in: G. Erickson, C. Smith (Eds.), *Maximum-Entropy and Bayesian Methods in Science and Engineering*, Vol. 31-32 of *Fundamental Theories of Physics*, Springer Netherlands, 1988, pp. 53–74.
- [43] N. P. Galatsanos, A. K. Katsaggelos, Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation, *Image Processing, IEEE Transactions on* 1 (3) (1992) 322–336.
- [44] G. Archer, D. M. Titterton, On some bayesian/regularization methods for image restoration, *Image Processing, IEEE Transactions on* 4 (7) (1995) 989–995.

- [45] J. Antoni, A bayesian approach to sound source reconstruction: Optimal basis, regularization, and focusing, *The Journal of the Acoustical Society of America* 131 (4) (2012) 2873–2890.
- [46] J. Idier, *Bayesian Approach to Inverse Problems*, 1st Edition, John Wiley & Sons Inc, 2001.
- [47] D. V. Lindley, A. F. M. Smith, Bayes estimates for the linear model, *Journal of the Royal Statistical Society. Series B (Methodological)* (1972) 1–41.
- [48] S. J. Press, *Applied Multivariate Analysis: Using Bayesian And Frequentist Methods Of Inference*, Dover Books on Mathematics Series, Dover Publications, 2005, chap. 8.6.
- [49] L. D. Broemeling, *Bayesian analysis of linear models*, Vol. 60, CRC Press, 1984.
- [50] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Rubin, *Bayesian Data Analysis*, 2nd Edition, Chapman & Hall/CRC Texts in Statistical Science, Taylor & Francis, 2003, chap. 14.
- [51] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1998.
- [52] H. Cramer, *Mathematical methods of statistics*, Princeton University Press, 1946.
- [53] C. Robert, *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*, Springer Texts in Statistics, Springer, 2001.
- [54] P. C. Hansen, *Discrete Inverse Problems: Insight and Algorithms*, Fundamentals of algorithms, Society for Industrial and Applied Mathematics (SIAM), 2010.
- [55] D. Fuhrmann, *Complex random variables and stochastic processes*, *Digital Signal Processing Handbook*, CRC Press, 1999.
- [56] P. J. Schreier, L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*, Cambridge University Press, 2010.
- [57] J.-D. Chazot, E. Zhang, J. Antoni, Acoustical and mechanical characterization of poroelastic materials using a bayesian approach, *The Journal of the Acoustical Society of America* 131 (6) (2012) 4584–4595.
- [58] E. Zhang, J. Antoni, P. Feissel, Bayesian force reconstruction with an uncertain model, *Journal of Sound and Vibration* 331 (4) (2012) 798 – 814.
- [59] P. M. Lee, *Bayesian Statistics: An Introduction*, 3rd Edition, Wiley, 2004.
- [60] J. Hald, Basic theory and properties of statistically optimized near-field acoustical holography, *The Journal of the Acoustical Society of America* 125 (4) (2009) 2105–2120.
- [61] B. P. Carlin, T. A. Louis, *Bayes and Empirical Bayes Methods for Data Analysis*, 1st Edition, Chapman & Hall, 1996.
- [62] F. Orioux, J.-F. Giovannelli, T. Rodet, A. Abergel, Estimating hyperparameters and instrument parameters in regularized inversion illustration for Herschel/SPIRE map making, *A&A* 549 (2013) A83.
- [63] N. Fortier, G. Demoment, Y. Goussard, GCV and ML methods of determining parameters in image restoration by regularization: Fast computation in the spatial domain and experimental comparison, *Journal of Visual Communication and Image Representation* 4 (2) (1993) 157 – 170.
- [64] J. E. Cavanaugh, Unifying the derivations for the akaike and corrected akaike information criteria, *Statistics & Probability Letters* 33 (2) (1997) 201 – 208.
- [65] P. C. Hansen, Analysis of discrete ill-posed problems by means of the L-curve, *SIAM Rev.* 34 (4) (1992) 561–580.
- [66] <http://lva.insa-lyon.fr/data-jsv-2013-pereira>.
- [67] J. Barnard, R. McCulloch, X.-L. Meng, Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage, *Statistica Sinica* 10 (4) (2000) 1281–1312.
- [68] J.-L. L. Carrou, Q. Leclère, F. Gautier, Some characteristics of the concert harp’s acoustic radiation, *The Journal of the Acoustical Society of America* 127 (5) (2010) 3203–3211.
- [69] F. J. Fahy, *Foundations of engineering acoustics*, Academic Press, 2000, chap. 6, pp. 129-134.
- [70] S. Boucheron, G. Lugosi, P. Massart, *Concentration inequalities: a nonasymptotic theory of independence*, Oxford University Press, 2013.

## List of Figures

- Figure 1: Schematic 1D representation of the Bayesian framework. The figure shows the different terms explicitly depending on  $\mathbf{c}$  in Bayes' rule (see Eq. (9)). In the case of a uniform prior, the solution depends uniquely on the likelihood and, for the case treated in this paper, it returns the solution (maximum likelihood) in terms of the pseudo-inverse of  $\mathbf{H}$  (noted  $\mathbf{H}^\dagger$ ). The consideration of a Gaussian prior has "biased" our search giving rise to the posterior, whose maximum (MAP estimate) returns the regularized solution expressed by  $\hat{\mathbf{c}}$ .
- Figure 2: In most practical situations  $\hat{\eta}_{\text{Joint}}^2$  and  $\hat{\eta}_{\text{MAP}}^2$  are found nearly equal, at least when the numbers of sensors or snapshots is large enough (see left panel). One exception is when the SNR is so poor that  $\mathbf{J}_{\text{Joint}}$  tends to see only noise, thus rejecting its minimum to infinity and returning an unbounded value of the regularization parameter. Such a situation is forbidden by  $\mathbf{J}_{\text{MAP}}$  which, by construction, advantageously forces a finite value of  $\eta^2$  (see right panel).
- Figure 3: The estimation of the source and noise energies amounts to finding the slope and intercept, respectively, of the regression line passing through the cluster of measurement points  $\{(s_k^2; \langle |y_k|^2 \rangle); k = 1, \dots, M\}$  complying with the conservation of energy law  $\mathbb{E}\{\langle |y_k|^2 \rangle\} = \alpha^2 s_k^2 + \beta^2$ . To improve visualization the right panel shows the plot of the natural logarithm of both the singular values  $s_k^2$  and the coefficients  $|y_k|^2$ .
- Figure 4: (a) Geometry of the problem for the first scenario, showing the discretized source surface, the microphone array and the simulated point source placed at the center of the surface. (b) Condition number for three different distances from the source plane ( $z_h$ ).
- Figure 5: Average value of the relative error to the optimal (MSE) solution over 500 realizations of measurement noise. Each row corresponds to a given distance from the array to the source surface. First row:  $z_h = r_s$ , second row:  $z_h = 5r_s$  and third row:  $z_h = 10r_s$ .
- Figure 6: Histogram plot of the ratio between estimated and optimal regularization parameters for Bayesian regularization: (a) BA, (b) L-curve and (c) GCV. They correspond to 500 realizations of measurement noise with a SNR of 30 dB, frequency 1000 Hz and distance source-array  $z_h = 5r_{ss}$ .
- Figure 7: (a) Geometry of the problem for the second scenario. It shows the microphone array, the discretized source surface and 3 point sources randomly placed at the source surface.; (b) Condition number of the transfer matrix for distance  $z_h$  equals to 10, 15 and 20 cm.
- Figure 8: Average value of the relative error to the optimal (MSE) solution over 500 realizations of measurement noise. Each row corresponds to a given distance from the array to the source surface. First row:  $z_h = 10$  cm, second row:  $z_h = 15$  cm and third row:  $z_h = 20$  cm.
- Figure 9: (a) Point estimate of the regularization parameter  $\eta^2$  (—) with confidence intervals estimated by a MCMC procedure (shaded region) and by a Gaussian approximation (---). (b) The reconstructed source spectrum using the Bayesian regularization criterion (---) with a 95% confidence interval (shaded region) and the

optimal MSE solution (—). (c) Global acoustic power as integrated over the source surface for the Bayesian regularization criterion (—) with a 98% confidence interval (shaded region) and the optimal (MSE) one (—).

Figure 10: Acoustic source of interest.

Figure 11: Reconstructed source field (volume velocity) at 945 Hz corresponding to each regularization criterion. (a) Bayesian regularization (BA); (b) L-curve and (c) GCV. A dynamic range of 15 dB is used for all.

Figure 12: Cost functions associated to each regularization criterion at the frequency of 945 Hz. (a) the Bayesian regularization criteria. (b) curvature of the L-curve (whose corner is the point with maximum curvature). (c) the GCV cost function.

Figure 13: Reconstructed source field (volume velocity) at 650 Hz corresponding to each regularization criterion. (a) Bayesian regularization (BA); (b) L-curve and (c) GCV. A dynamic range of 15 dB is used for all.

Figure 14: Cost functions associated to each regularization criterion at the frequency of 650 Hz. (a) the Bayesian regularization criteria. (b) curvature of the L-curve (whose corner is the point with maximum curvature). (c) the GCV cost function.

Figure 15: Estimate of the total acoustic power for different regularization strategies. The reference (ref.) was obtained from  $p$ - $u$  measurements. The figure also shows a 98% Bayesian confidence interval around the point estimate returned by the proposed Bayesian criterion (BA).

Figure 16: (a) reconstructed acoustic power map integrated over the frequency band 200-1400 Hz. The + symbols are the real source positions and the dashed circles are the integration area used to compute the partial PSD of each source. The other panels show the PSD of each source identified by the proposed method (—) and obtained from  $p$ - $u$  measurements (—). (b) PSD of source A. (c) PSD of source B. (d) PSD of source C.