



HAL
open science

Mapping uncertainties involved in sound source reconstruction with a cross-spectral-matrix-based Gibbs sampler

Jérôme Antoni, Charles Vanwynsberghe, Thibaut Le Magueresse, Simon Bouley, Laurent Gilquin

► **To cite this version:**

Jérôme Antoni, Charles Vanwynsberghe, Thibaut Le Magueresse, Simon Bouley, Laurent Gilquin. Mapping uncertainties involved in sound source reconstruction with a cross-spectral-matrix-based Gibbs sampler. *Journal of the Acoustical Society of America*, 2019, 146 (6), pp.4947-4961. 10.1121/1.5138930 . hal-03211968

HAL Id: hal-03211968

<https://hal.science/hal-03211968>

Submitted on 3 Sep 2021

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

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

Mapping uncertainties involved in sound source reconstruction with a cross-spectral-matrix-based Gibbs sampler

Jérôme Antoni, Charles Vanwysberghe, Thibaut Le Magueresse, Simon Bouley, and Laurent Gilquin

Citation: *The Journal of the Acoustical Society of America* **146**, 4947 (2019); doi: 10.1121/1.5138930

View online: <https://doi.org/10.1121/1.5138930>

View Table of Contents: <https://asa.scitation.org/toc/jas/146/6>

Published by the [Acoustical Society of America](#)

A promotional banner for a special issue of JASA. The background is a dark blue gradient with a blurred image of a 3D printer nozzle printing a red, lattice-structured part. On the left, the JASA logo is displayed in white, with the full name 'THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA' underneath. To the right of the logo, the text 'Special Issue: Additive Manufacturing and Acoustics' is written in yellow and white. At the bottom left, a yellow button contains the text 'Submit Today!' in black.

JASA
THE JOURNAL OF THE
ACOUSTICAL SOCIETY OF AMERICA

Special Issue:
Additive Manufacturing and Acoustics

Submit Today!

Mapping uncertainties involved in sound source reconstruction with a cross-spectral-matrix-based Gibbs sampler

Jérôme Antoni,^{1,a)} Charles Vanwynsberghe,^{1,b)} Thibaut Le Magueresse,² Simon Bouley,^{1,c)} and Laurent Gilquin¹

¹Université de Lyon, Institut National des Sciences Appliquées de Lyon, Laboratoire Vibrations Acoustique, F-69621 Villeurbanne, France

²MicrodB, F-69134 Ecully, France

(Received 6 March 2019; revised 21 August 2019; accepted 26 August 2019; published online 31 December 2019)

The reconstruction of sound sources by using inverse methods is known to be prone to estimation errors due to measurement noise, model mismatch, and poor conditioning of the inverse problem. This paper introduces a solution to map the estimation errors together with the reconstructed sound sources. From a Bayesian perspective, it initializes a Gibbs sampler with the Bayesian focusing method. The proposed Gibbs sampler is shown to converge within a few iterations, which makes it realistic for practical purposes. It also turns out to be very flexible in various scenarios. One peculiarity is the capability to directly operate on the cross-spectral matrix. Another one is to easily accommodate sparse priors. Eventually, it can also account for uncertainties in the microphone positions, which reinforces the regularization of the inverse problem.

© 2019 Acoustical Society of America. <https://doi.org/10.1121/1.5138930>

[ZHM]

Pages: 4947–4961

I. INTRODUCTION

The recovery of acoustic sources, be it for localization or identification purposes, is of prime importance in various fields of acoustics. This topic has been active in research for several decades, with new algorithms being regularly published in an effort to constantly push forward the limits of the reconstruction. A persistent issue of many state of the art methods is that they must be content with point estimates, without indication of their inherent uncertainty. On the one hand, this prevents the user from assessing the reliability of the results. This situation is critical as the reconstruction of source distributions is an inherently ill-posed inverse problem without guarantee of the existence, uniqueness, and stability of a solution.^{1,2} On the other hand, knowing the uncertainty in the results is also a prerequisite to allowing comparisons between competitive methods. Solutions to accompany the reconstruction of sources together with their inherent uncertainties are therefore needed.

The present paper aims at assessing uncertainties due to estimation errors produced from the presence and interaction of measurement noise, model mismatch, and poor conditioning of the inverse problem. To the best of the authors' knowledge, the analysis of estimation errors induced in inverse methods dedicated to the identification of sound sources has been rarely addressed. This is notwithstanding the vast literature that exists in other related fields, such as in geophysics inversion,^{3–5} structural dynamics,^{6,7} and underwater acoustics, where the concern is more related to the environmental uncertainties that affect the propagation—and therefore the subsequent inversion.^{8–12} In contrast, this work

is concerned with the reconstruction of distributions of sound sources—i.e., pressure, velocity, intensity—which marks both conceptual and technical differences with other approaches found in related fields (e.g., Refs. 13–15).

Contrary to the situation with direct measurement methods—such as sound intensity—whose estimation errors are well covered in the literature (see, e.g., Refs. 16–18), the propagation of errors (i.e., the effect of the errors on the uncertainty of the inversion) in inverse methods such as NAH, ESM, HELS¹⁹ remains mathematically difficult. Reference 20 investigated propagation of uncertainties in NAH by using first-order perturbation and arrived at expressions of the variance and bias of the source estimations under different configurations. However, this type of approach—also known as the delta method in statistics²¹—is known to underestimate the actual level of uncertainties since it ignores higher-order terms. In addition, it rapidly becomes intractable when nonlinear transform are involved, unless rough approximations are used—for instance, the effect of regularization or the statistical tests introduced later in Sec. IV would be tedious to set up with this approach.

An interesting class of inverse methods that naturally provides the posterior probability density functions (PDFs) of the estimates—i.e., the PDFs conditioned on the data—is based on Bayesian inference.^{2,22,23} This turns especially useful since all estimators of position—such as the mean value or the maximum *a posteriori* estimate (MAP)—as well as of statistical dispersion—such as standard deviations or confidence intervals—can theoretically be deduced from posterior PDFs. In the present context, this was shown, for instance, in the Bayesian focusing method introduced in Ref. 24 and further exploited in Ref. 25 to set up confidence intervals on the regularization parameter.

In theory, Bayesian inference makes it possible to propagate uncertainties in all quantities of interest in the inverse

^{a)}Electronic mail: jerome.antoni@insa-lyon.fr

^{b)}Also at: Lab-STICC UMR 6285, CNRS, ENSTA Bretagne, 2 rue François Verny, 29200 Brest, France.

^{c)}Also at: MicrodB, F-69134 Ecully, France.

problems, yet closed-form expressions are not always available or do not generally coincide with known PDFs. In such instances, it becomes necessary to resort to numerical techniques, such as the Markov Chains Monte Carlo (MCMC) algorithm, to directly sample in the posterior PDF.

The aim of this paper is to explore such a solution based on Bayesian inference using the Gibbs sampler,^{26,27} one popular MCMC algorithm. As far as the authors know, this has never been addressed before in the present context, although it appears in related fields concerned with uncertainties.^{28–32} As far as the reconstruction of source distributions is of concern, Ref. 33 actually introduced a Gibbs sampler, yet more as an inferential mechanism than as a means to assess statistical dispersion in the results. A similar remark applies to Ref. 34 which makes use of a particle filter, a special form of a sequential Gibbs sampler.

The paper first shows in Sec. III that a Gibbs sampler is easily set up in the Bayesian setting introduced in Ref. 24. In particular, an original version is provided that directly takes as an input the cross-spectral matrix (CSM). This considerably reduces the amount of data to be recorded when measuring time-stationary acoustical fields. It is also necessary with commercial data acquisition systems that can only record the CSM.

This paper further proposes a Gibbs sampler with rapid convergence provided that it is initialized with the results of the Bayesian focusing method²⁴ and its hyperparameters are tuned with the empirical Bayesian method of Ref. 25. This is necessary to make it a realistic solution in practice, since MCMC algorithms are otherwise computationally intensive methods.

Providing uncertainties on the spatial distributions of the sources also opens new questions as for their representation. The paper proposes some solutions in Sec. IV. The first one is to set up a threshold on the source coefficients computed from the statistical dispersion of the Markov chains and to display only the significant results. The second one is to map the coefficient of determination together with the source distribution, so as to return a degree of confidence between 0 and 100% as a function of space.

Finally, the present work also demonstrates some interesting byproducts of the Gibbs sampler, that illustrates its high flexibility. First, it is easily amenable to handle non-homogeneous errors on the sensors, a configuration referred to as heteroscedastic. This turns out to be useful, for example, to detect the presence of defective sensors. Second, it is able to recover sparse sources by exploiting the heteroscedasticity of the source coefficients. This is equivalent to adopting a non-Gaussian prior in the Bayesian inference, in the same vein as in Ref. 35, for instance, yet with a different algorithmic solution. It may also be related to a relevance vector machine (RVM).³⁶ Third, the Gibbs sampler easily includes parametric uncertainties in the propagator, which pertains to model mismatch errors.^{20,37} Although this is not the main focus of the paper, the concept is rapidly illustrated with uncertainties in the positions or in the selection of the sensors. The Gibbs sampler is then shown to provide increased regularization in the inverse problem, even in the case of mistuned hyperparameters.

The main novelties introduced in the paper are briefly as follows.

- The first one is to propose a methodology for assessing uncertainties in the reconstruction of sound sources and, in particular, to map them together with the reconstructed sources.
- As a means to achieve this objective, the second novelty is a Gibbs sampler that directly operates on the CSM, a characteristic that is crucial to making the approach feasible in some industrial scenarios.
- The CSM-based Gibbs sampler can also easily handle sparse priors and regularize the inversion in the case of sensor position uncertainties.

The following notations are used throughout the document. Vectors are denoted by bold lower-case letters (e.g., \mathbf{p}); unless otherwise stated, the elements of a vector are assumed arranged columnwise. Matrices are denoted by bold upper-case letters (e.g., \mathbf{A}). Symbol H stands for the transpose conjugate of a matrix or a vector. \mathbf{I}_N stands for the identity matrix with dimension N and is simply noted \mathbf{I} when there is no possible confusion. \mathbf{e}_k is the k th column of \mathbf{I} . $\mathbf{I}_{M \times N}$ ($M \geq N$) stands for the identity matrix \mathbf{I}_M truncated to its N first columns, i.e., $\mathbf{I}_{M \times N} = \{\mathbf{e}_1, \dots, \mathbf{e}_N\}$ with $\dim(\mathbf{e}_k) = M$. The notation $\text{dg}\{a_k\}$ stands for the diagonal matrix composed of diagonal elements a_k . $\text{tr}\{\mathbf{A}\}$ is the trace of matrix \mathbf{A} . Operator $\text{vec}\{\mathbf{A}\}$ vectorizes matrix \mathbf{A} by stacking its columns one underneath the other. Curly braces around an indexed element means that the whole set of elements, for all values of the index, is considered: for instance, if $i \in \{1, \dots, I\}$, then $\{a_i\}$ is short for $\{a_i, i = 1, \dots, I\}$. When needed, random variables will be denoted by upper-case letters and their realizations by lower-case letters (e.g., the random variable X takes a value x). The expected value of the random variable X is denoted by the operator $\mathbb{E}\{X\}$. Square brackets are used to denote a PDF (e.g., $[x]$ denotes the PDF of the random variable X evaluated at x). The notation $[x|y]$ stands for the conditional PDF of X given $Y = y$ evaluated at $X = x$. The complex-valued Gaussian PDF of a random vector \mathbf{X} with mean value $\boldsymbol{\mu}$, covariance matrix $\boldsymbol{\Omega}$, and evaluated at \mathbf{x} is noted $\mathcal{N}_{\mathbb{C}}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Omega})$ or simply $\mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \boldsymbol{\Omega})$. The inverse Gamma PDF of a random variable X with shape parameter a , scale parameter b , and evaluated at x is noted $\mathcal{IG}(x; a, b)$ or simply $\mathcal{IG}(a, b)$.

The paper is organized as follows. Section II first sets up the direct problem and formulates its inversion as a Bayesian inference. Next, Sec. III introduces the Gibbs sampler and its extension based on the CSM. Section IV then introduces techniques for post-processing the results of the Gibbs sampler and for displaying uncertainties. Finally, Sec. V illustrates the proposed methodology on numerical and lab experiments.

II. PROBABILISTIC MODELING OF THE INVERSE PROBLEM

This section introduces a solution to the inverse problem, together with its variables, and proceeds to its layout by means of a hierarchical Bayesian model.

A. Parameterization of the direct problem

The general scenario considered in this paper involves an array of M microphones that delivers pointwise measurements of the acoustic pressures at positions \mathbf{r}_m , $m = 1, \dots, M$. The pressure signals are seen as realizations of stationary and ergodic stochastic processes in time. All acoustical variables are expressed in the frequency domain, at a frequency f , after application of the Fourier transform to a series of snapshots $i = 1, \dots, I$. The snapshots are designed to be independent and, in virtue of the signal stationarity, they are also identically distributed. The Fourier coefficient of the acoustic pressure assigned to snapshot i and microphone m reads $p_{m,i}$ (for the sake of simplicity, the explicit dependence on f is dropped out from the notation). The collection of Fourier coefficients, for all microphones at snapshot i , is arranged in vector $\mathbf{p}_i \in \mathbb{C}^M$. The cross-spectral matrix (CSM) is then defined as

$$\mathbf{S}_{pp} = \frac{1}{I} \sum_{i=1}^I \mathbf{p}_i \mathbf{p}_i^H. \quad (1)$$

The sensed acoustic field is assumed to be radiated by a set of sources, seen as stochastic processes statistically stationary in time, whose distribution is described by the vector of coefficients $\mathbf{q}_i \in \mathbb{C}^K$. The radiated pressure at the microphones then reads

$$\mathbf{p}_i = \mathbf{H} \mathbf{q}_i + \mathbf{n}_i \quad (2)$$

where \mathbf{H} stands for the transfer matrix and \mathbf{n}_i for additive errors accounting for all uncontrolled sources of errors (e.g., measurement noise, instrumentation noise, modeling errors). The noise process \mathbf{n}_i is assumed stationary. Equation (2) is directly given in its matrix form as it comes out after discretizing the integral formulation of the direct problem. The interpretation of the source coefficients q_{ki} 's will thus depend on how the discretization is performed. For instance, in the ESM method, the quantity q_{ki} will be the source strength of the k th monopole, and in NAH it will be the amplitude of the k th plane wave. For simplicity, index k will refer to a spatial position in the rest of the paper.

The aim of the present paper is to infer the source coefficients \mathbf{q}_i from the measured CSM \mathbf{S}_{pp} . In this view, a probabilistic approach based on a hierarchical Bayesian model is set up.

B. Bayesian hierarchical inference

The inverse problem includes the unknowns \mathbf{q}_i . They are first assigned prior PDFs that reflect the range of possible values the experimenter expect them to take before incorporating the experimental data. Next, once experimental data are available, the posterior PDF of the unknown variables are deduced from experimental data using Bayes' theorem, from which point estimates and measures of statistical dispersion (e.g., variance, intervals) can be eventually inferred.

Since the Bayesian method used herein has been introduced previously,^{24,25,35} it is only briefly summarized.

1. Likelihood

The first ingredient in the Bayesian formulation is the likelihood function, defined from the PDF of the data given all other parameters including the unknown source coefficients. According to model (2), the latter is of the same type as the PDF of the additive error \mathbf{n}_i . There are at least two reasons that justify a Gaussian distribution here. First, the noise errors result from the summation of a multitude of sources and are therefore subjected to the central limit theorem. Second, \mathbf{n}_i being a Fourier coefficient, it converges quickly to a complex Gaussian even if the original PDF of the signal in the time domain is not Gaussian.³⁸ By definition of errors, their mean is zero—otherwise it would be an explanatory variable in the model. In the present work, since no dependence structure is *a priori* known for the errors in space, the covariance matrix is taken diagonal with elements β_m^2 , $m = 1, \dots, M$. Finally, since the snapshots are independent and identically distributed, the conditional PDF of the noise is $[\{\mathbf{n}_i\} | \{\beta_m^2\}] = \prod_{i=1}^I \mathcal{N}_{\mathbb{C}}(\mathbf{n}_i; 0, \text{dg}\{\beta_m^2\})$. Consequently, the likelihood function reads

$$[\{\mathbf{p}_i\} | \{\mathbf{q}_i\}, \{\beta_m^2\}] = \prod_{i=1}^I \mathcal{N}_{\mathbb{C}}(\mathbf{p}_i; \mathbf{H} \mathbf{q}_i, \text{dg}\{\beta_m^2\}). \quad (3)$$

2. Hyperpriors in the likelihood

Since in general the variances β_m^2 's at microphones $m = 1, \dots, M$ are also unknown hyperparameters, they are themselves assigned a prior PDF in the form of an inverse gamma law, $[\beta_m^2] = \mathcal{IG}(\beta_m^2; a_\beta, b_\beta)$, with given shape and scale parameters a_β and b_β . The choice of an inverse gamma law, apart from being flexible enough, is motivated by simplicity. On the one hand, it preserves a positive support ($\beta_m^2 \geq 0$) and, on the other hand, it is conjugated with the Gaussian^{26,27} (conjugacy will substantially simplify the inference in the posterior PDF in Sec. III A). Other choices are obviously possible at this stage, yet they might not result in significantly different inferential results as long as the prior is set “vague” enough (i.e., with a large variance). The general case with heterogeneity of variances in space will be referred to as *heteroscedastic*. The particular case when all variances are assumed equal, $\beta_m^2 = \beta^2, \forall m$, will be referred to as *homoscedastic*.

3. Source priors

The next and essential ingredient in the Bayesian formulation is the prior PDF on the source coefficients \mathbf{q}_i . There are several possibilities at this juncture depending on the problem at hand. A relevant choice should reflect the user's knowledge or expectation on the sources before the data are observed. A common prior PDF is the complex Gaussian, which is the less informative PDF according to the maximum entropy principle³⁹ when nothing else than the expected mean value and the covariance of the source coefficients is *a priori* known. In the present work, the mean is taken equal to zero since the phase of the Fourier coefficients is not known in advance. In the absence of further knowledge about spatial dependence, the covariance matrix is

taken diagonal with elements α_k^2 , $k = 1, \dots, K$. Therefore, and since snapshots are independent and identically distributed,

$$[\{\mathbf{q}_i\}|\{\alpha_k^2\}] = \prod_{i=1}^I \mathcal{N}_C(\mathbf{q}_i; 0, \text{dg}\{\alpha_k^2\}). \quad (4)$$

a. Homoscedasticity and heteroscedasticity of the sources in space. The particular case where all variances are assumed equal, $\alpha_k^2 = \alpha^2$, $\forall k$, will be referred to as *homoscedasticity* of the sources. It tends to diffuse the power of the sound sources densely and uniformly all over the source domain. In inverse problems, where one expects to recover sources that are sparsely distributed in space with high concentrations of energy in narrow regions of space, one should allow the individual variances α_k^2 to take different values. This case is referred to as *heteroscedasticity* of the sources. To see this, let us consider the prior PDF averaged in space with respect to variances α_k^2 ,

$$[q_i] = \sum_{k=1}^K [q_i|\alpha_k^2] \cdot \frac{1}{K}. \quad (5)$$

This produces a new space-invariant prior PDF, $[q_i]$, in the form a scaled mixture of Gaussians (SMoG) with uniform weights $1/K$. An SMoG prior enforces the spatial sparsity of the reconstructed sources by compelling most of them to switch off and allowing only a few to remain active.

b. Hyperpriors. Eventually, for the same reasons as for the errors variances, the prior PDF of the variances $[\alpha_k^2]$ is taken in the form of an inverse Gamma law, $[\alpha_k^2] = \mathcal{IG}(\alpha_k^2; a_x, b_x)$, with given shape and scale parameters a_x and b_x . The values of the latter will directly control the degree of spatial sparsity induced in Eq. (5). Incidentally, the limit of the sum (5) when the grid $\{k = 1, \dots, K\}$ tends to a continuum is

$$[q_i] = \int \mathcal{N}_C(q_i; 0, v) \mathcal{IG}(v; a_x, b_x) dv, \quad (6)$$

which is recognized as a multivariate complex Student- t , a prior PDF that enforces highly sparse sources (see, e.g., Ref. 35).

4. Model uncertainties

Finally, the transfer matrix \mathbf{H} can also be assigned unknown parameters $\boldsymbol{\theta}$ with a given PDF $[\boldsymbol{\theta}]$. In the present paper, $\boldsymbol{\theta}$ will be considered as a vector of nuisance parameters that model some uncontrolled uncertainties in the transfer functions, yet without interest in their exact value. This will turn out useful in Sec. III C to reflect uncertainty in the position of the sensors or to perform a random selection of a subset of sensors. More generally, $\boldsymbol{\theta}$ could also model random uncertainties in the propagation medium (e.g., Refs. 8, and 11), yet this is outside the scope of the present work.

Since all the aforementioned probabilities are interlaced, they define a Bayesian hierarchical model where the

parameters of a PDF are themselves seen as random variables which are assigned PDFs in a higher level. This is illustrated in Fig. 1(a).

C. MAP estimation (Bayesian focusing)

The principle of the Bayesian focusing method of Ref. 24 is now briefly resumed since it is used in this work to initialize a Gibb's sampler [Fig. 1(b)]. It consists of finding the MAP estimate based on the likelihood introduced in Eq. (3) with homoscedastic errors, a complex Gaussian prior with homoscedastic sources, and without the uncertainty $\boldsymbol{\theta}$. The inverse problem then finds the closed-form solution

$$\begin{aligned} \hat{\mathbf{q}}_i &= \text{Arg max}_{\mathbf{q}_i} [\mathbf{q}_i|\hat{\alpha}^2, \hat{\beta}^2, \{\mathbf{p}_i\}] \\ &= \text{Arg max}_{\mathbf{q}_i} [\mathbf{p}_i|\mathbf{q}_i, \hat{\beta}^2] [\mathbf{q}_i|\hat{\alpha}^2] \\ &= \left(\mathbf{H}^H \mathbf{H} + (\hat{\beta}^2/\hat{\alpha}^2) \mathbf{I}_K \right)^{-1} \mathbf{H}^H \mathbf{p}_i, \end{aligned} \quad (7)$$

where, according to a technique known as empirical Bayes, the unknown hyperparameters α^2 and β^2 are replaced by their MAP estimates on the marginalized posterior,

$$\begin{aligned} (\hat{\alpha}^2, \hat{\beta}^2) &= \text{Arg max}_{(\alpha^2, \beta^2)} [\alpha^2, \beta^2|\{\mathbf{p}_i\}] \\ &= \text{Arg max}_{(\alpha^2, \beta^2)} [\alpha^2, \beta^2] \\ &\quad \times \int [\{\mathbf{p}_i\}|\{\mathbf{q}_i\}, \beta^2] [\{\mathbf{q}_i\}|\alpha^2] d\{\mathbf{q}_i\}. \end{aligned} \quad (8)$$

Reference 25 gives details to solve Eq. (8) using one-dimensional optimization. In addition, Ref. 35 proposes an iterative version of the Bayesian focusing method to find the MAP estimate in the case of a SMoG prior PDF that enforces spatial sparsity.

III. BAYESIAN INFERENCE WITH THE GIBBS SAMPLER

Based on the Bayesian hierarchical formulation introduced in Sec. II, the object is to infer the full posterior PDF

$$\begin{aligned} &[\{\mathbf{q}_i\}, \{\beta_m^2\}, \{\alpha_k^2\}|\{\mathbf{p}_i\}] \\ &= \int [\{\mathbf{q}_i\}, \{\beta_m^2\}, \{\alpha_k^2\}, \boldsymbol{\theta}|\{\mathbf{p}_i\}] [\boldsymbol{\theta}] d\boldsymbol{\theta} \end{aligned} \quad (9)$$

of all unknown variables, with marginalization over the nuisance variable $\boldsymbol{\theta}$, from which point estimates (e.g., mean values) as well as statistical dispersion (e.g., variances) of the source coefficients \mathbf{q}_i will ultimately be deduced. There are generally no closed-form solutions for the latter, yet one can resort to Monte Carlo numerical techniques to directly sample in the posterior PDF (9). The Gibbs sampler is a popular MCMC algorithm to do so, which is particularly well suited to the hierarchical method of Fig. 1.

A. Standard Gibbs sampler

The principle of the Gibbs sampler is to create a Markov chain that has the posterior PDF (9) as its stationary

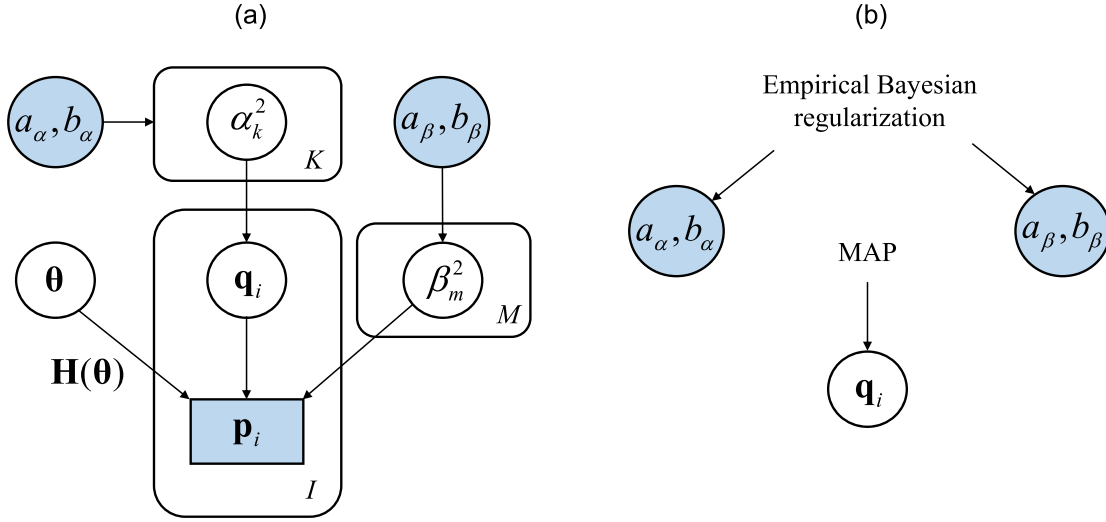


FIG. 1. (Color online) (a) Bayesian hierarchical model of the inverse problem using plate notation (Ref. 40). Circles indicate random variables and hyperparameters. The square in the lowest level identifies the data. Filled-in areas indicate fixed values. The letter in the lower right corner of a plate indicates the number of repetitions of the enclosed variable. (b) Initialization of the Gibbs sampler with the MAP solution of Bayesian focusing (Ref. 24) and the hyperparameters estimated from empirical Bayesian regularization (Ref. 25).

distribution. This is accomplished by iterating successive draws from the conditional distributions $[\mathbf{q}_i|\text{rest}]$, $[\beta_m^2|\text{rest}]$, $[\alpha_k^2|\text{rest}]$, where “|rest” means conditioned to all other variables (as listed in Fig. 1). The hierarchy between the random variables actually leads to some simplifications in the expression of the PDFs, due to the dependencies between the nodes in the graphical model of Fig. 1. Specifically, let $x^{[n]}$ denote the value that random variable X takes at the n th iteration of the Gibbs sampler and let symbol \leftarrow mean “sample from”; then, by representing direct dependencies only, the sketch of the algorithm given in the case of heteroscedastic errors and sources is as follows:

- $\boldsymbol{\theta}^{[n]} \leftarrow [\boldsymbol{\theta}]$,
- $\mathbf{q}_i^{[n]} \leftarrow [\mathbf{q}_i | \{\beta_m^2\}^{[n-1]}, \{\alpha_k^2\}^{[n-1]}, \boldsymbol{\theta}^{[n]}, \{\mathbf{p}_i\}]$,
- $\beta_m^{2,[n]} \leftarrow [\beta_m^2 | \{\mathbf{q}_i\}^{[n]}, \boldsymbol{\theta}^{[n]}, \{\mathbf{p}_i\}]$,
- $\alpha_k^{2,[n]} \leftarrow [\alpha_k^2 | \{\mathbf{q}_i\}^{[n]}]$.

Note that variables updated in the previous steps are systematically used in the next step (index n instead of $n - 1$). The proposed sequence of scans is arbitrary and can be modified, even during the iterations.⁴¹ Using the prior PDFs introduced in Sec. II B, the above conditional PDFs can now be related to known distributions which are easy to sample (superscripts n and $n - 1$ are momentarily dropped to simplify the notation).

Namely, one has for the source coefficients $[\mathbf{q}_i|\text{rest}]$ proportional to a Gaussian $\mathcal{N}_{\mathbb{C}}(\mathbf{q}_i; \boldsymbol{\mu}_{q_i}, \boldsymbol{\Omega}_{q_i})$ with

$$\begin{cases} \boldsymbol{\mu}_{q_i} = \mathbf{B}\mathbf{p}_i \\ \mathbf{B} = \boldsymbol{\Omega}_q \mathbf{H}(\boldsymbol{\theta})^H \text{dg}\{\beta_m^{-2}\} \\ \boldsymbol{\Omega}_q = \left(\mathbf{H}(\boldsymbol{\theta})^H \text{dg}\{\beta_m^{-2}\} \mathbf{H}(\boldsymbol{\theta}) + \text{dg}\{\alpha_k^{-2}\} \right)^{-1} \end{cases} \quad (10)$$

(see Appendix A).

In addition, under the heteroscedastic configuration, $[\beta_m^2|\text{rest}]$ is proportional to an inverse Gamma $\mathcal{IG}(\beta_m^2; a_\beta + I, b_\beta + b_{\beta_m}^{\text{post}})$ with $b_{\beta_m}^{\text{post}} = \sum_{i=1}^I |\mathbf{e}_m^H(\mathbf{p}_i - \mathbf{H}\mathbf{q}_i)|^2$. As for the

homoscedastic configuration, $[\beta^2|\text{rest}]$ is found proportional to $\mathcal{IG}(\beta^2; a_\beta + MI, b_\beta + \sum_{m=1}^M b_{\beta_m}^{\text{post}})$.

Similarly, for heteroscedastic sources, $[\alpha_k^2|\text{rest}]$ is found proportional to $\mathcal{IG}(\alpha_k^2; a_\alpha + I, b_\alpha b_{\alpha_k}^{\text{post}})$ with $b_{\alpha_k}^{\text{post}} = \sum_{i=1}^I |\mathbf{e}_k^H \mathbf{q}_i|^2$ and, for homoscedastic sources, $[\alpha^2|\text{rest}] \propto \mathcal{IG}(\alpha^2; a_\alpha + KI, b_\alpha + \sum_{k=1}^K b_{\alpha_k}^{\text{post}})$.

Algorithm 1 based on these equations is given hereafter.

B. CSM-based Gibbs sampler

This section introduces a modification of the above Gibbs algorithm that aims at significantly lowering its computational cost. A limitation of algorithm 1 (Table I) is to rapidly become over costly in scenarios with long time records, since it involves as many Markov chains as they are

TABLE I. (Algorithm 1) Standard Gibbs sampler.

1: Initialize $\beta_m^{2,[0]}$, $\alpha_k^{2,[0]}$, and $\mathbf{q}_i^{[0]}$
2: $n \leftarrow 1$
3: repeat
4: $\boldsymbol{\theta}^{[n]} \leftarrow [\boldsymbol{\theta}]$
5: Compute $\boldsymbol{\mu}_{q_i}^{[n]}$ and $\boldsymbol{\Omega}_q^{[n]}$
6: $\mathbf{q}_i^{[n]} \leftarrow \mathcal{N}_{\mathbb{C}}(\mathbf{q}_i; \boldsymbol{\mu}_{q_i}^{[n]}, \boldsymbol{\Omega}_q^{[n]})$
7: Compute $b_{\alpha_k}^{\text{post},[n]}$
8: if heteroscedastic errors then
9: $\beta_m^{2,[n]} \leftarrow \mathcal{IG}(\beta_m^2; a_\beta + I, b_\beta + b_{\beta_m}^{\text{post},[n]})$
10: else if homoscedastic errors then
11: $\beta^{2,[n]} \leftarrow \mathcal{IG}(\beta^2; a_\beta + MI, b_\beta + \sum_{m=1}^M b_{\beta_m}^{\text{post},[n]})$
12: end if
13: Compute $b_{\alpha_k}^{\text{post},[n]}$
14: if heteroscedastic sources then
15: $\alpha_k^{2,[n]} \leftarrow \mathcal{IG}(\alpha_k^2; a_\alpha + KI, b_\alpha + \sum_{k=1}^K b_{\alpha_k}^{\text{post},[n]})$
16: else if homoscedastic sources then
17: $\alpha^{2,[n]} \leftarrow \mathcal{IG}(\alpha^2; a_\alpha + I, b_\alpha + b_{\alpha_k}^{\text{post},[n]})$
18: end if
19: $n \leftarrow n + 1$
20: until convergence

snapshots I . In addition, it involves the evaluation of all the Fourier coefficients from the measured time frames, $\{\mathbf{p}_i\}_{i=1}^I$ in Eq. (10), thus requiring large storage capacity. As a matter of fact, several data acquisition systems do not even record the set of Fourier coefficients, but only the CSM of Eq. (1). Therefore, there is a need for a version of the Gibbs sampler that takes the CSM as its input. This is also motivated by the fact that, under the stationary Gaussian model of the measurements, the CSM is a “sufficient” statistic that contains all the information required to solve the inverse problem.

The idea of the CSM-based Gibbs sampler is based on recognizing that the source coefficients \mathbf{q}_i only intervene quadratically in the Gibbs sampler, through the quantities $\mathbf{S}_{qq} = (1/I)\sum_{i=1}^I \mathbf{q}_i \mathbf{q}_i^H$ and $\mathbf{S}_{pq} = (1/I)\sum_{i=1}^I \mathbf{p}_i \mathbf{q}_i^H$. To see this, let us rephrase Eq. (10) as

$$\mathbf{q}_i = \boldsymbol{\mu}_{q_i} + \boldsymbol{\Omega}_q^{1/2} \boldsymbol{\xi}_i \text{ where } \boldsymbol{\xi}_i \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_K). \quad (11)$$

Substituting the above in the expression of \mathbf{S}_{qq} , one has

$$\begin{aligned} \mathbf{S}_{qq} &= \underbrace{\frac{1}{I} \sum_{i=1}^I \boldsymbol{\mu}_{q_i} \boldsymbol{\mu}_{q_i}^H}_{\mathcal{O}(1)} + \underbrace{\boldsymbol{\Omega}_q^{1/2} \left(\frac{1}{I} \sum_{i=1}^I \boldsymbol{\xi}_i \boldsymbol{\xi}_i^H \right) \boldsymbol{\Omega}_q^{H/2}}_{\mathcal{O}(1)} \\ &\quad + \underbrace{\frac{1}{I} \sum_{i=1}^I \left(\boldsymbol{\mu}_{q_i} \boldsymbol{\xi}_i^H \boldsymbol{\Omega}_q^{H/2} + \boldsymbol{\Omega}_q^{1/2} \boldsymbol{\xi}_i \boldsymbol{\mu}_{q_i}^H \right)}_{\mathcal{O}(1/I)} \end{aligned} \quad (12)$$

where $\mathcal{O}(1/I)$ [$\mathcal{O}(1)$] means that a quantity grows, in the root means square sense, proportionally to $1/I$ (1) with the number of snapshots I . Substituting $\boldsymbol{\mu}_{q_i}$ by its expression (10), the first term in the right-hand side of Eq. (11) is $\mathbf{B} \mathbf{S}_{pp} \mathbf{B}^H$, the second one converges to $\boldsymbol{\Omega}_q^{1/2} \mathbf{I}_K \boldsymbol{\Omega}_q^{H/2} = \boldsymbol{\Omega}_q$ and the last one converges to zero since it involves the average $\sum_{i=1}^I \mathbf{p}_i \boldsymbol{\xi}_i^H / I$ of independent random variables \mathbf{p}_i and $\boldsymbol{\xi}_i$. Therefore, the following approximation holds:

$$\mathbf{S}_{qq} \simeq \mathbf{B} \mathbf{S}_{pp} \mathbf{B}^H + \frac{1}{I} \boldsymbol{\Omega}_q^{1/2} \mathbf{W} \boldsymbol{\Omega}_q^{H/2}, \quad (13)$$

where \mathbf{W} is a random matrix following a complex Wishart PDF $\mathcal{W}_{\mathbb{C}}(I, \mathbf{I}_R)$.⁴² In other words, \mathbf{S}_{qq} , which initially followed a noncentral complex Wishart PDF, has been replaced by a central complex Wishart PDF plus an offset. Similarly,

$$\mathbf{S}_{pq} \simeq \mathbf{S}_{pp} \mathbf{B}^H \quad (14)$$

and

$$\frac{1}{I} \sum_{i=1}^I (\mathbf{p}_i - \mathbf{H} \mathbf{q}_i)(\mathbf{p}_i - \mathbf{H} \mathbf{q}_i)^H \simeq \mathbf{P} \mathbf{S}_{pp} \mathbf{P}^H + \mathbf{W}_{pp} / I \quad (15)$$

with $\mathbf{P} = \mathbf{I}_M - \mathbf{H} \mathbf{B}$ and $\mathbf{W}_{pp} = (\mathbf{H} \boldsymbol{\Omega}_q^{1/2}) \mathbf{W} (\mathbf{H} \boldsymbol{\Omega}_q^{1/2})^H$ (see Appendix B). Putting all these results together, the CSM-based Gibbs sampler consists in iterating the following successive draws listed in algorithm 2 (Table II) where $b_{\beta_m}^{\text{post}} = \mathbf{e}_m^H (I \cdot \mathbf{P} \mathbf{S}_{pp} \mathbf{P}^H + \mathbf{W}_{pp}) \mathbf{e}_m$ and $b_{z_k}^{\text{post}} = I \cdot \mathbf{e}_k^H \mathbf{S}_{qq} \mathbf{e}_k$.

It is noteworthy that the reduction in complexity is of a factor I as compared to algorithm 1, which is substantial

TABLE II. (Algorithm 2) CSM-based Gibbs sampler.

```

1: Initialize  $\beta_m^{2,[0]}$ ,  $\alpha_k^{2,[0]}$  and  $\mathbf{S}_{qq}^{[0]}$ 
2:  $n \leftarrow 1$ 
3: repeat
4:    $\boldsymbol{\theta}^{[n]} \leftarrow [\boldsymbol{\theta}]$ 
5:    $\mathbf{W}^{[n]} \leftarrow \mathcal{W}_{\mathbb{C}}(0, \mathbf{I}_R)$ 
6:   Compute  $\mathbf{B}^{[n]}$  and  $\boldsymbol{\Omega}_q^{[n]}$ 
7:    $\mathbf{S}_{qq}^{[n]} = \mathbf{B}^{[n]} \mathbf{S}_{pp} (\mathbf{B}^{[n]})^H + (\boldsymbol{\Omega}_q^{[n]})^{1/2} (\mathbf{W}^{[n]} / I) (\boldsymbol{\Omega}_q^{[n]})^H / 2$ 
8:   Compute  $b_{\beta_m}^{\text{post},[n]}$ 
9:   if heteroscedastic errors then
10:     $\beta_m^{2,[n]} \leftarrow \mathcal{IG}(\beta_m^2; a_{\beta} + I, b_{\beta} + b_{\beta_m}^{\text{post},[n]})$ 
11:   else if homoscedastic errors then
12:     $\beta^{2,[n]} \leftarrow \mathcal{IG}(\beta^2; a_{\beta} + MI, b_{\beta} + \sum_{m=1}^M b_{\beta_m}^{\text{post}})$ 
13:   end if
14:   Compute  $b_{z_k}^{\text{post},[n]}$ 
15:   if heteroscedastic sources then
16:     $\alpha_k^{2,[n]} \leftarrow \mathcal{IG}(\alpha_k^2; a_{\alpha} + I, b_{\alpha} + b_{z_k}^{\text{post},[n]})$ 
17:   else if homoscedastic sources then
18:     $\alpha^{2,[n]} \leftarrow \mathcal{IG}(\alpha^2; a_{\alpha} + KI, b_{\alpha} + \sum_{k=1}^K b_{z_k}^{\text{post}})$ 
19:   end if
20:    $n \leftarrow n + 1$ 
21: until convergence

```

when I is large. Besides, the generation of random Wishart matrices comes with an efficient algorithm based on the Bartlett decomposition.⁴³

C. Sensor marginalization and sensor dropout

As explained in Sec. II B, the inclusion of unknown parameters $\boldsymbol{\theta}$ makes it possible to marginalize the posterior PDF over nuisance parameters [see Eq. (9)]. Among the many perspectives that this approach offers, the present paper focuses on uncertainties in the sensor positions. The motivation is practical: while the inverse problem is known to be extremely sensitive to small perturbations in the data (without regularization it would try to fit the data exactly by allowing large departures in the source coefficients), it has to be solved with sensor positions that are always provided with limited precision. Hence, it is expected that taking this uncertainty into account will implicitly regularize the inversion. The idea is reminiscent of the theory developed in Ref. 44 and is introduced here as a conjecture that will be experimentally verified in Sec. V. It is compliant with the discrepancy principle, which states that the amount of regularization in the inverse problem should be tuned so as to match the uncertainty in the measurements.

Two strategies are proposed. The first one is to encode each sensor position \mathbf{r}_m with a random error $\boldsymbol{\theta}_m$ around its expected (or “nominal”) value \mathbf{r}_m^0 , i.e., $\mathbf{r}_m = \mathbf{r}_m^0 + \boldsymbol{\theta}_m$. For convenience, $\boldsymbol{\theta}_m$ might be distributed like a zero mean Gaussian. The random variables $\{\boldsymbol{\theta}_m\}_{m=1}^M$ can additionally be made dependent in order to model uncertainties in the position of the array, such as uncontrolled rotations and translations (see, e.g., Ref. 45). The Green functions in matrix $\mathbf{H}(\{\boldsymbol{\theta}_m\})$ are then constantly updated with random draws of $\boldsymbol{\theta}_m$ in the iterations of the Gibbs sampler. This strategy will be referred to as “sensor marginalization.”

The second strategy follows a similar goal, yet from a different perspective reminiscent to cross-validation.¹ It

aims at constraining the solution of the inverse problem to be almost invariant with respect to the distribution of the sensors. After all, if one sensor is removed from the array, the reconstructed source distribution is ideally expected to remain the same. This scenario is easily simulated in the Gibbs sampler by randomly removing a subset \mathcal{M} of M_0 sensors ($M_0 \ll M$) from the array at each iteration. The $\{\theta_m\}_{m=1}^M$ are then random indicator variables used to sample M_0 rows of matrix \mathbf{H} , to be removed at iteration n , with replacement at iteration $n+1$. In the case of heteroscedastic errors, the values of $\beta_m^{2,[n]}$, $m \in \mathcal{M}$ corresponding to the removed sensors are either not updated (i.e., $\beta_m^{2,[n]} = \beta_m^{2,[n-1]}$) or replaced by the average of the other variances [i.e., $\beta_m^{2,[n]} = \sum_{k \notin \mathcal{M}} \beta_k^{2,[n]} / (M - M_0)$]. This strategy will be referred to as “sensor dropout.”

Simulations verify that the two proposed strategies return very similar results. The first one requires the setting of the variance σ_θ^2 , whereas the second one requires the setting of the number M_0 of sensors to drop out. However, sensor dropout is computationally less demanding since it does not require the re-evaluation of the Green functions in matrix \mathbf{H} at each iteration of the Gibbs sampler and it involves the manipulation of matrices of dimension $M - M_0$ instead of M .

D. Initialization

Being based on a random exploration of the posterior PDF (9), it is important to properly initialize the Gibbs sampler in the regions of highest probabilities so as to accelerate its convergence to the stationary PDF (9).

Since μ_{q_i} in Eq. (10) corresponds to the MAP estimate given in Eq. (7), it is recommended to initialize the Markov chains with the latter. In turn, the values of β^2 and α^2 (homoscedastic assumption at initialization) are well initialized with the estimates returned by the empirical Bayesian regularization criterion (8). This is illustrated in Fig. 1(b). By way of an example, let $\hat{\alpha}^2$ be the empirical Bayes estimate of α^2 . Then a_x and b_x can be set so that the prior PDF $[a_x^2 | b_x^2]$ has its mode centered on $\hat{\alpha}^2$ and a standard deviation equal to a given multiple (or fraction) of $\hat{\alpha}^2$. Note that having accurate values of β^2 and α^2 at initialization is crucial in order to properly regularize the inversion involved in the calculation of $\Omega_q^{[1]}$ in algorithm 1 or 2 and therefore prevents numerical divergence of the algorithm from the onset.

Intensive simulations have shown that the Gibbs sampler so initialized can then converge after only a few iterations (see, for instance, Fig. 3 in Sec. V). This reduces the burn-in, i.e., the number of samples burnt for reaching convergence.

IV. POST-PROCESSING OF THE RESULTS

A. Propagation of errors

One of the definite advantages of MCMC algorithms is to evaluate uncertainties in any estimate. For instance, the statistical dispersion of any function $f(\mathbf{q}_i)$ of the source coefficients is assessed by evaluating $f(\mathbf{q}_i^{[n]})$ on the samples $\{\mathbf{q}_i^{[n]}\}$ of the Markov chain. Algorithm 1 immediately returns

this when drawing samples from $[\mathbf{q}_i | \text{rest}]$ and algorithm 2 indirectly returns it by means of Eq. (11).

Again, if quadratic quantities only are of interest, it is more efficient to draw samples $\mathbf{S}_{qq}^{[n]}$ rather than $\mathbf{q}_i^{[n]}$. This involves sampling a noncentral complex Wishart matrix with noncentrality parameter $I \cdot \mathbf{B} \mathbf{S}_{pp} \mathbf{B}^H$. Although possible,⁴⁶ it may be more convenient to approximate \mathbf{S}_{qq} as in Eq. (13) and to sample the central complex Wishart PDF $\mathcal{W}_C(0, \mathbf{I}_R)$ by taking advantage of the Bartlett decomposition.⁴³ For example, several quadratic quantities of interest, such as acoustic intensity and acoustic power, can be formulated as a scalar value $\mathcal{Q} = \mathbf{c}^H \mathbf{S}_{qq} \mathbf{d}$ for some vectors \mathbf{c} and \mathbf{d} . Then, one has the following samples:

$$\mathcal{Q}^{[n]} = \underbrace{\mathbf{c}^H \mathbf{B}^{[n]} \mathbf{S}_{pp} (\mathbf{B}^{[n]})^H \mathbf{d}}_{\mathcal{Q}_0} + \underbrace{\mathbf{c}^H (\Omega_q^{[n-1]})^{1/2} (\mathbf{W}^{[n]} / I) (\Omega_q^{[n-1]})^{H/2} \mathbf{d}}_{\mathcal{Q}_r}, \quad (16)$$

where $\mathbf{W}^{[n]} \sim \mathcal{W}_C(0, \mathbf{I}_R)$, $\mathcal{Q}_0^{[n]}$ returns the central value of the chain $\{\mathcal{Q}^{[n]}\}$ (notwithstanding variability in the Markov chain $\{\mathbf{B}^{[n]}\}$) and $\mathcal{Q}_r^{[n]}$ reflects the stochastic errors of the estimation. Algorithm 2 directly outputs the above two quantities. If point estimates are of concern, it might be convenient to replace the chain $\{\mathcal{Q}^{[n]}\}$ by its average over iterations, $\langle \mathcal{Q}^{[n]} \rangle = N^{-1} \sum_{n=1}^N \mathcal{Q}^{[n]}$, where care must be taken to start the summation after the burn-in samples corresponding to the transient part of the chain before it has reached convergence.

B. Imaging the errors

Imaging the uncertainties, e.g., estimation errors, of the reconstructed source distribution is a challenging task since it involves a representation of high-dimension quantities.

A first and natural solution is to display only quantities which are statistically significant and otherwise to clip insignificant coefficients. The proper approach to do so would require setting up a statistical test to check whether the mean value of the source coefficient q_{ki} at “position” k is statistically nil (null hypothesis) or not, based on the estimated source power

$$\begin{aligned} \mathcal{P}_k &= \frac{1}{I} \sum_{i=1}^I \langle |q_{ki}^{[n]}|^2 \rangle \\ &= \mathbf{e}_k^H \langle \mathbf{S}_{qq}^{[n]} \rangle \mathbf{e}_k = \underbrace{\mathbf{e}_k^H \mathbf{B}^{[n]} \mathbf{S}_{pp} (\mathbf{B}^{[n]})^H \mathbf{e}_k}_{\mathcal{P}_{0,k}} \\ &\quad + \underbrace{\mathbf{e}_k^H (\Omega_q^{[n-1]})^{1/2} (\mathbf{W}^{[n]} / I) (\Omega_q^{[n-1]})^{H/2} \mathbf{e}_k}_{\mathcal{P}_{r,k}} \end{aligned} \quad (17)$$

obtained by setting $\mathbf{c} = \mathbf{d} = \mathbf{e}_k$ in Eq. (16). Here, a simplified approach is proposed by introducing an empirical threshold on \mathcal{P}_k defined as the posterior quantile on the reconstructed source distribution under the null hypothesis. It is proved in Appendix C that an approximation to the $(1-p)$ posterior quantile on $\mathcal{P}_{0,k}$ is

$$\mathcal{P}_{0,k,1-p} = \mathcal{P}_{r,k} \cdot \left(\frac{\chi_{2I,1-p}^2}{2I} - 1 \right), \quad (18)$$

with $\chi_{2I,1-p}^2$ the quantile of the chi-squared with $2I$ degrees of freedom. All source coefficients below the above bound are simply clipped or, equivalently, not displayed. As illustrated in Sec. V, this solution might be good enough for practical purposes.

The second solution is to display a map of the coefficient of determination of the source coefficients \mathbf{q}_i on the source domain, together with the map of the reconstructed source distribution. The coefficient of determination is defined as

$$R_k = \frac{\mathcal{P}_{0,k}}{\mathcal{P}_{0,k} + \mathcal{P}_{r,k}} \in [0, 1] \quad (19)$$

and is interpreted as the fraction of power explained by the reconstructed source at position k . The value of R_k indicates the degree of confidence in the result on a scale from 0 to 1. It results from Eq. (18) that an approximation to the $(1-p)$ posterior quantile under the null hypothesis is

$$R_{k,1-p} = 1 - \frac{2I}{\chi_{2I,1-p}^2}. \quad (20)$$

C. Test of homoscedasticity of the errors in space

The Gibbs sampler has been shown to be flexible enough to easily handle homoscedastic or heteroscedastic errors. In practice, it may be useful to check which of the two configurations is most realistic given the measured data. A statistical test is again easily set up based on the samples of the Markov chains. It follows the lines of Lindley's method⁴⁷ applied to the Bayesian analysis of variance (see, e.g., Ref. 48).

The principle consists in checking whether the squared distance between the M chains is significantly different from the variance within the chains. Care should be taken to properly pre-process the Markov chain samples in order to meet the assumptions of the test. First, the logarithmic transformation $\{X_{n,m}\} = \{\ln(\beta_m^{2,[n]})\}$ is applied to Gaussianize the data. Next, the effective sample size $N_{\text{eff}} = N/(1 + 2\sum_{n>0}\rho_n)$ (with ρ_n the normalized autocovariance of $\{X_{n,m}\}$) is calculated in order to account for sequential correlation in the chain.²⁶ Then, it is proved in Appendix D that the hypothesis of homoscedasticity of errors is rejected if the F-ratio

$$F = \frac{M(N_{\text{eff}} - 1)N_{\text{eff}}}{M - 1} \frac{\sum_{m=1}^M |\bar{X}_m - \bar{\bar{X}}|^2}{\sum_{m=1}^M \sum_{n=1}^{N_{\text{eff}}} |X_{n,m} - \bar{X}_m|^2}, \quad (21)$$

with $\bar{X}_m = (1/N)\sum_{n=1}^N X_{n,m}$ and $\bar{\bar{X}} = (1/M)\sum_{m=1}^M \bar{X}_m$ is found greater than $F_{M-1, M(N_{\text{eff}}-1), 1-p}$, the $(1-p)$ quantile of the Fisher distribution with $M-1$ and $M(N_{\text{eff}}-1)$ degrees of freedom.

V. EXPERIMENTAL RESULTS

This section demonstrates the potential of the proposed CSM-based Gibbs sampler through a few experiments. First, a numerical experiment is considered to illustrate the technical aspects. Next, the feasibility of the approach is validated on real measurements. The latter experiment also considers errors due to uncertainty in the sensor positions.

A. Numerical experiment

In order to ease the interpretation of the results, a simple one-dimensional setup is considered. The source distribution consists of two pistons located at $x_1 = -0.5$ m and $x_2 = 0.5$ m, both with diameter of 0.5 m and driven by two independent white Gaussian noises with RMS values of 1 and 0.2 m/s, respectively. The analysis is carried out in a frequency band centered at $f = 1000$ Hz. The sound field is captured with an uniform linear array (ULA) composed of $M = 20$ microphones, with spacing 0.1 m and range distance $z = 0.5$ m to the source plane. The experimental setup is schemed in Fig. 5(a). The SNR is set to 20 dB on all microphones and the number of snapshots used in the estimation of the CSM is $I = 200$. In modelling the inverse problem, the source distribution is expanded on the canonical basis, so that each source coefficient q_{ik} represents an elementary source strength at "position" k , with a total of $K = 61$ coefficients. With the present experimental setup, this results in a severely ill-posed problem. As observed from Picard's plot¹ in Fig. 2, it appears that only 11 out of 20 of the projected data $|y_m|^2 = \mathbf{u}_m^H \mathbf{S}_{pp} \mathbf{u}_m$ are above the eigenvalues s_m^2 obtained from the singular value decomposition $\mathbf{H} = \sum_{m=1}^M s_m^2 \mathbf{u}_m \mathbf{v}_m^H$ of the transfer matrix \mathbf{H} ; the largest gap between $|y_m|^2$ and s_m^2 is about 6 orders of magnitude. Clearly the inverse problem would not be solvable without regularization. The source and the errors variances are assumed homoscedastic and heteroscedastic, respectively. The Gibbs sampler is initialized from the outputs of Bayesian focusing, as explained in Sec. III D. It is run for at least 100 iterations. The Markov chains of the Gibbs sampler are displayed in Fig. 3. The chains quickly converge to their stationary distribution after only a few iterations thanks to the good initialization provided by Bayesian focusing. It is also seen that the Markov chains of the variances $\beta_m^{2,[n]}$ are well clustered, which indicates that homoscedasticity should be accepted. This observation is further confirmed by displaying the histograms of

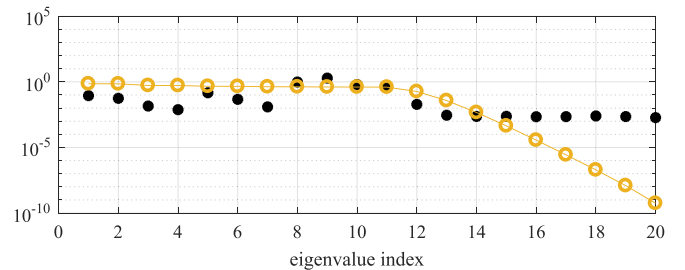


FIG. 2. (Color online) Picard plot comparing the projected data $|y_m|^2 = \mathbf{u}_m^H \mathbf{S}_{pp} \mathbf{u}_m$ (black bullets) with the eigenvalues s_m^2 (orange circles), where $\mathbf{H} = \sum_{m=1}^M s_m^2 \mathbf{u}_m \mathbf{v}_m^H$ is the singular value decomposition of the transfer matrix \mathbf{H} .

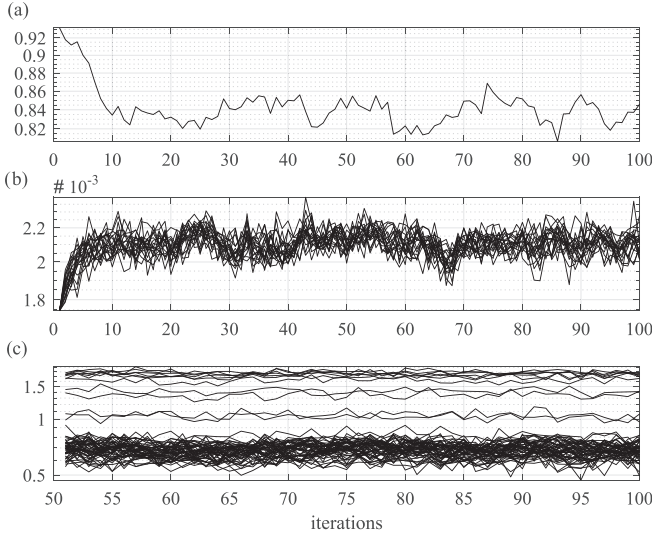


FIG. 3. Markov chains of (a) the source variance $\alpha^{2,[n]}$, (b) the errors variance on the microphones $\beta_m^{2,[n]}$, $m = 1, \dots, 20$, and (c) the spectra of the source coefficients $S_{q_i, k}^{[n]}$, $k = 1, \dots, 61$.

the errors variances in Fig. 4(a). The F-ratio of Eq. (21) is 1.1 (with $N_{eff} = 130$), to be compared to a 0.95 posterior quantile of 1.6 under the null hypothesis.

The reconstructed source distribution is displayed by means of quantity $\mathcal{P}_{0,k}$ defined in Eq. (18) in Fig. 5(a), together with the MAP returned by Bayesian focusing and the 0.95 posterior quantile given by Eq. (18). As expected, $\mathcal{P}_{0,k}$ almost coincides with the MAP estimate. Yet, the Gibbs sampler indicates that most of the estimated source coefficients stay below the 0.95 posterior quantile when $I=200$ snapshots are used, which is similarly confirmed by the coefficient of determination R_k in Fig. 5(b). The conclusion is therefore that the reconstruction of the second piston is not statistically significant and should be interpreted with

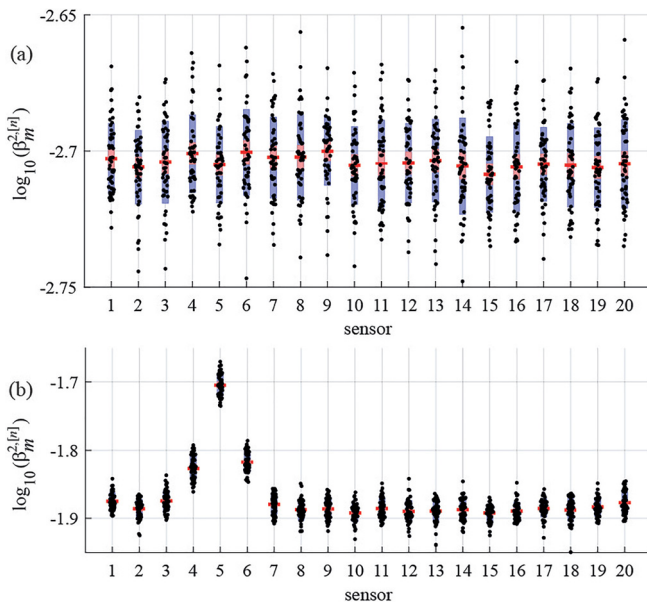


FIG. 4. (Color online) Histograms of the errors variances $\log_{10}(\beta_m^{2,[n]})$ for microphones $m = 1, \dots, 20$ with 1-standard-error-of-the-mean interval in pink (light gray), 1-standard-deviation interval in blue (dark gray), and average value in red (dash). (a) Homoscedastic and (b) heteroscedastic scenarios.

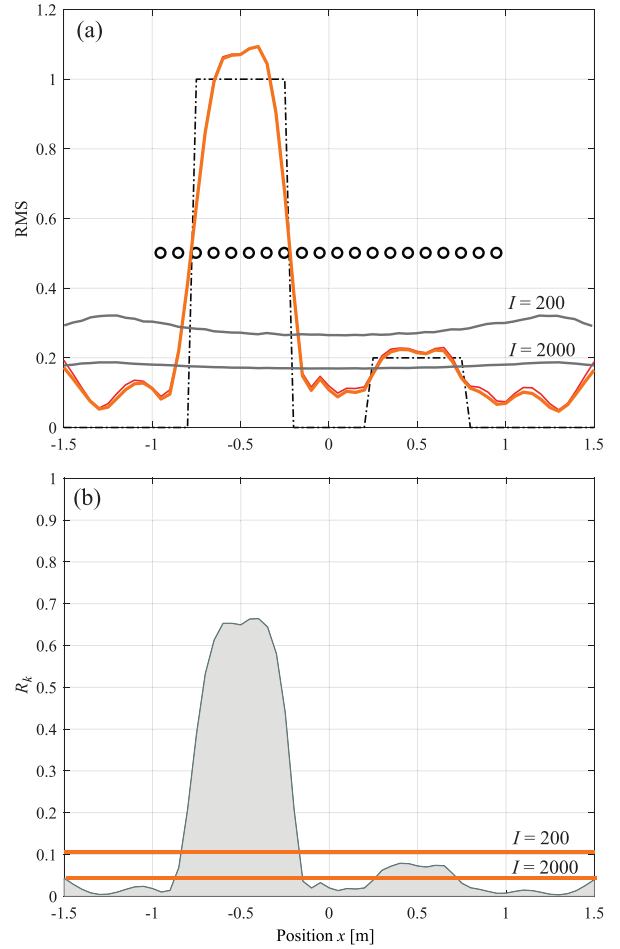


FIG. 5. (a) (Color online) Reconstructed source distribution $\mathcal{P}_{0,k}$ (RMS values). The dashed black line shows the theoretical distribution made of two pistons driven by incoherent signals and the black circles show the array of microphones. The thick orange curve and the thin red curve are the MCMC (homoscedastic sources) and MAP estimates, respectively. The gray curves are the thresholds given by the 0.95 posterior quantile under the null hypothesis with $I=200$ and 2000 snapshots. (b) Coefficient of determination R_k corresponding to (a) with its threshold given by the 0.95 posterior quantile under the null hypothesis.

caution. One remedy is to increase the number of snapshots, e.g., to $I=2000$, which then lowers the threshold below the level of the source coefficients pertaining to the second piston.

For the sake of comparison, the same experiment is also run with a standard deviation of the noise five times as large on the fifth microphone. Figure 4(b) shows that the Gibbs sampler easily detects the presence of non-homogeneous errors. The F-ratio of Eq. (21) is now 171.4, to be compared to the same 0.95 posterior quantile of 1.6 as before, under the null hypothesis. In this case, the best solution is simply to remove the measurement coming from the defective microphone.

Eventually, comparisons are made on the numerical data between the results of the CMS-based and the standard Gibbs samplers. Let $\mathcal{P}_{0,k}^{CMS}$, $\mathcal{P}_{0,k}^{std}$, and $\mathcal{P}_{0,k}^{true}$ stand for the source distribution estimated from the CMS-based Gibbs sampler, the standard Gibbs sampler, and the true distribution, respectively. The relative mean difference, $\mathcal{D}_I = \sum_{k=1}^K |\mathcal{P}_{0,k}^{CMS} - \mathcal{P}_{0,k}^{std}| / \sum_{k=1}^K \mathcal{P}_{0,k}^{true}$, is displayed in Fig. 6 for

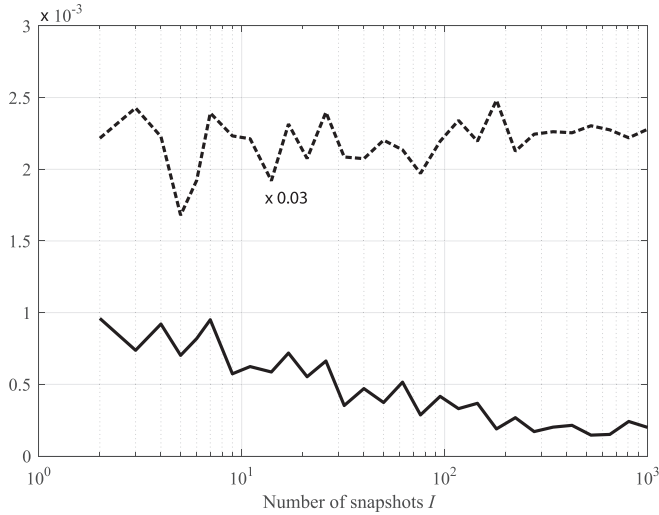


FIG. 6. Relative mean difference \mathcal{D}_I (black line) between the estimates of $\mathcal{P}_{0,k}$ from the CMS-based and the standard Gibbs samplers as a function of the number of snapshots I used in the algorithms, together with the standard deviation of the posterior of $\mathcal{P}_{0,k}$ (dotted line).

increasing numbers of snapshots I . It is seen that \mathcal{D}_I asymptotically decreases to zero when I grows, as expected. The mean difference is remarkably small even for low values of I , about 1 : 1000 the true distribution. It is much smaller than the bias (not shown in the figure) and is two magnitudes lower than the standard deviation of the posterior of $\mathcal{P}_{0,k}$ (dotted line). The latter is in turn almost identical for the two algorithms, with a difference of the same order as \mathcal{D}_I .

A last numerical experiment is carried out with the diameter of the two pistons becoming infinitesimal so as to simulate two (incoherent) point sources. The inverse problem is solved with heteroscedasticity of the source coefficients in order to enforce the recovery of spatially sparse sources. The working frequency is $f=1500$ Hz and the distance range $z=0.05$ m so as to place the inversion in a difficult configuration. Figure 7 illustrates the reconstructed source distribution. Bayesian focusing without sparsity enforcement is clearly unable to properly retrieve the two point sources, whereas the MCMC estimate $\mathcal{P}_{0,k}$ shows a remarkable result in this respect.

B. Lab experiment

The proposed CMS-based Gibbs sampler is now illustrated using experimental measurements. The experimental setup shown in Fig. 8 is comprised of three small acoustical sources produced by the expansion of air at the outlet of three tubes connected to a common compression chamber. The latter is excited with white noise up to 5 kHz. A uniform rectangular array (URA) with $M=6 \times 5$ microphones is placed at a range distance $z=0.2$ m parallel to the source plane. The spacing between the microphones is 0.1 m. Then aim is to reconstruct the distribution of acoustical intensity, $\mathcal{I}_{0,k}$ at “position” k in the source plane, as returned by Eq. (16) by properly setting \mathbf{c} and \mathbf{d} and averaging the $N=50$ last samples of the Markov chain. The CSM is computed with $I=100$ snapshots with a frequency resolution of 1 Hz. As in the numerical experiment, the results of Bayesian

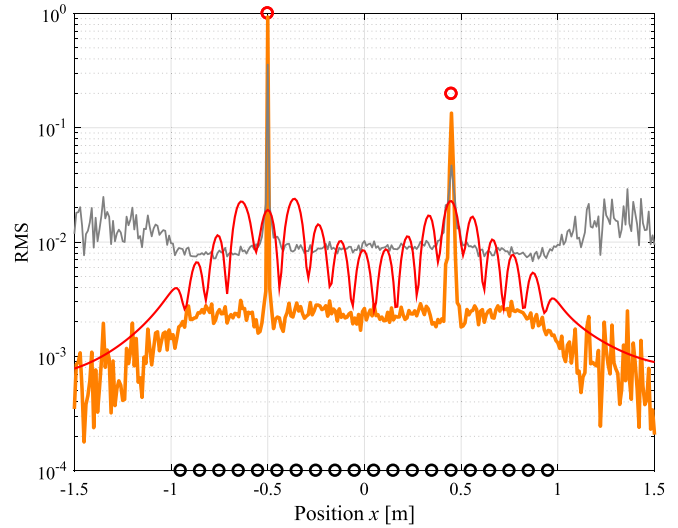


FIG. 7. (Color online) Reconstructed source distributions $\mathcal{P}_{0,k}$ (RMS values) in the case of two incoherent point sources marked by the red circles. The black circles show the array of microphones. The thick orange curve and the thin red curve are the MCMC (heteroscedastic sources) and MAP (without sparse prior) estimates, respectively. The gray curve is the threshold given by the 0.95 posterior quantile under the null hypothesis.

focusing together with Bayesian empirical regularization initialize the Gibbs sampler. The number of iterations is kept limited to $N=100$ with a burn-in of 50 samples in order to not inflate the overall computational burden.

For the sake of completeness, models with homoscedastic and heteroscedastic sources are compared. For both of them, the statistical test of Sec. IV C indicated that the hypothesis of homoscedastic errors could be accepted based on an 0.95 posterior quantile. The results are first displayed at the working frequency $f=1850$ Hz, where the PSDs of the pressure signals show a peak of high energy. Only the source coefficients above the 0.95 posterior quantile under the null hypothesis given by Eq. (18) are displayed. The corresponding threshold on the coefficient of determination is equal to 1.39. The reconstructed acoustical intensity $\mathcal{I}_{0,k}$ for the two

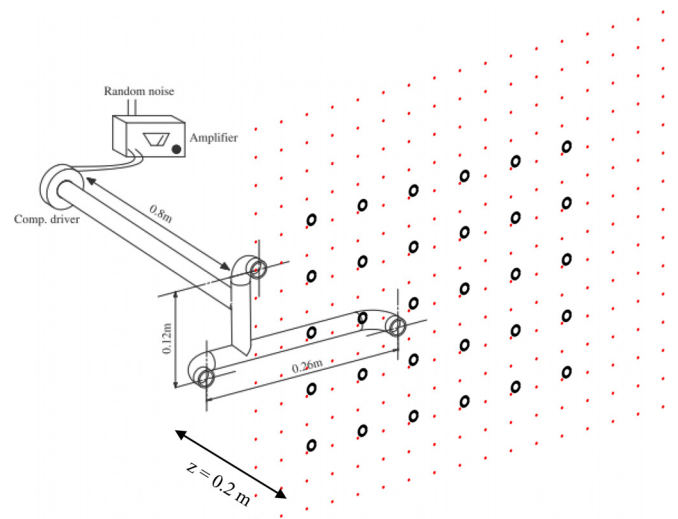


FIG. 8. (Color online) Experimental setup showing the generation of the acoustical sources and their sensing with an URA (black circles stand the microphone positions and red dots for the discretization of the source domain).

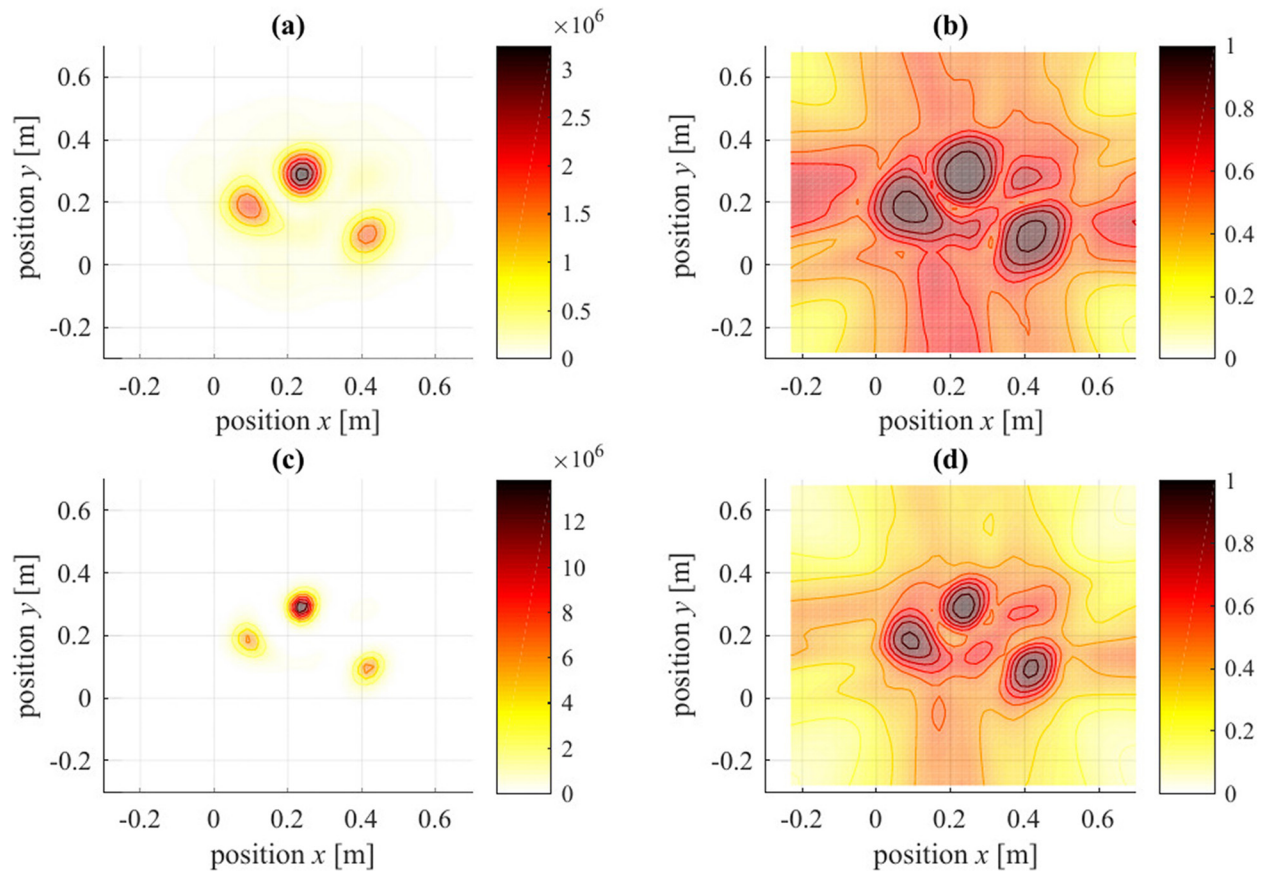


FIG. 9. (Color online) Reconstructed acoustical intensity $\mathcal{I}_{0,k}$ (left column, in W/m^2) in the source plane and coefficient of determination R_k (right column) at $f=1800$ Hz for (a), (b) homoscedastic and (c), (d) heteroscedastic sources.

models is displayed in Fig. 9, together the coefficient of determination R_k . It is seen that the three sources are well reconstructed, yet since the heteroscedastic model enforces spatial sparsity, it ends up in better spatial resolution, with estimated distributions close to point sources. The coefficient of determination also indicates good confidence in the results, even if the heteroscedastic model is accompanied by a slightly lower coefficient of determination outside the source support sets [see Fig. 9(d)].

Next, the reconstruction of the acoustical intensity is attempted at frequency $f=1270$ Hz corresponding to a dip of energy in the PSDs of the pressure signals. The result is displayed in Fig. 10(a) together with the coefficient of determination in Fig. 10(b). The latter clearly indicates low confidence: the maximum value of R_k reaches 0.145 and stays below the 0.95 posterior quantile equal to 1.38 under the null hypothesis. Therefore, it is safe to simply ignore the estimation at this frequency.

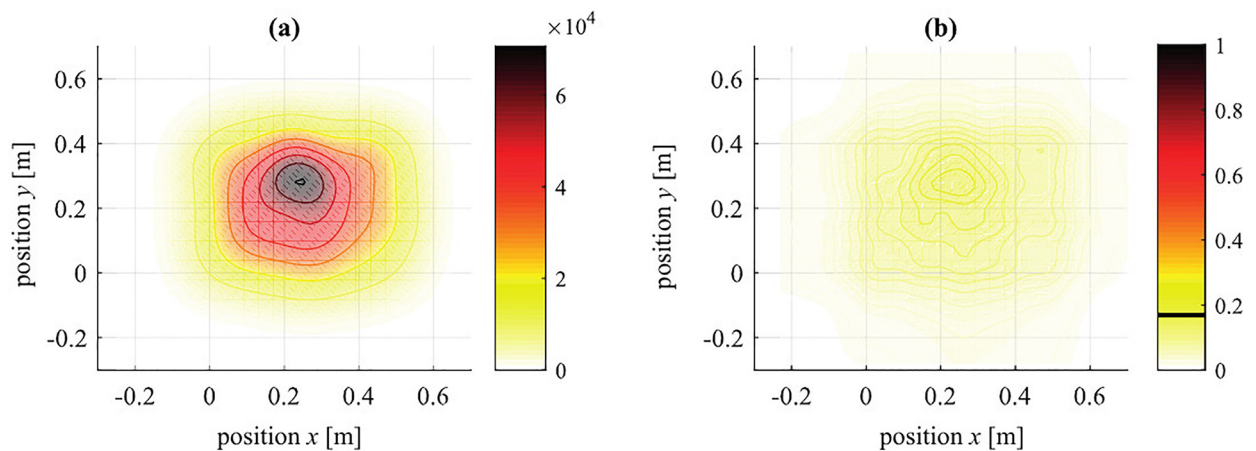


FIG. 10. (Color online) (a) Reconstructed acoustical intensity $\mathcal{I}_{0,k}$ (W/m^2) in the source plane and (b) coefficient of determination R_k at $f=1270$ Hz (homoscedastic sources).

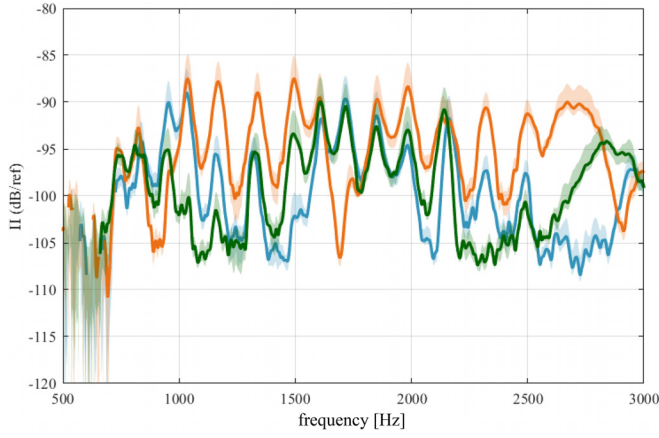


FIG. 11. (Color online) Acoustical power spectra $\pi_s(f)$ of the three sources $s = 1, 2, 3$ obtained from the Markov chains with their 95% credible intervals.

Finally, the algorithm is run for 2500 frequency lines in the excited frequency range from 500 to 3000 Hz. The acoustical power of each source is then estimated by integrating the acoustical intensity $\mathcal{I}_{0,k}$ in a circular region around its center, for each frequency f . The so obtained acoustical power spectra $\pi_s(f)$, $s = 1, 2, 3$ are displayed in Fig. 11 together with their credible intervals computed from 0.5 and 0.95 quantiles of the noncentral chi-squared

$$\frac{\mathcal{I}_{r,k}}{2I} \cdot \chi_{2l}^2(\delta) \quad (22)$$

with noncentrality parameter $\delta = I \cdot \mathcal{I}_{0,k} / \mathcal{I}_{r,k}$. These intervals show regions where the estimates are less reliable, in particular below 750 Hz where the excitation is low. At the same time, they are much narrower than the variations in the spectra, which can be reliably interpreted as resonances of the tube connected to the compression chamber. It is emphasized at this stage that obtaining intervals on the acoustical power spectra (in dB scale) is easily obtained from the Markov chains, as explained in Sec. IV A. In contrast, obtaining them from other means such as analytical expressions would be much more tedious, or even impossible in general without approximation when the result of interest is transformed through a nonlinear operation.

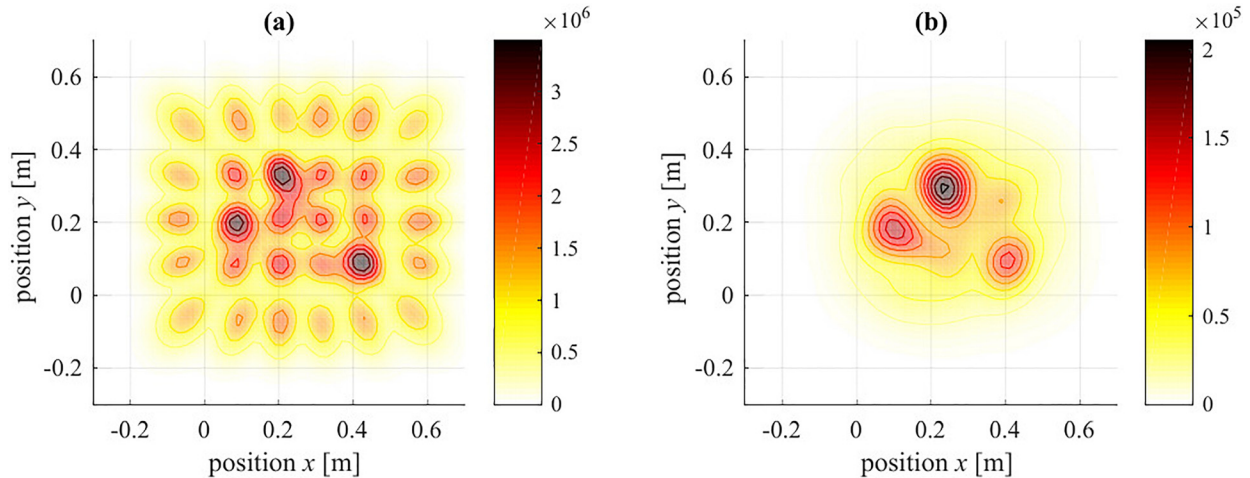


FIG. 13. (Color online) (a) Reconstructed acoustical intensity $\mathcal{I}_{0,k}$ (W/m^2) at $f = 1850$ Hz with hyperparameters α^2 and β^2 kept fixed to mistuned values in the Markov chains. (b) Same as in (a) with microphone dropout.

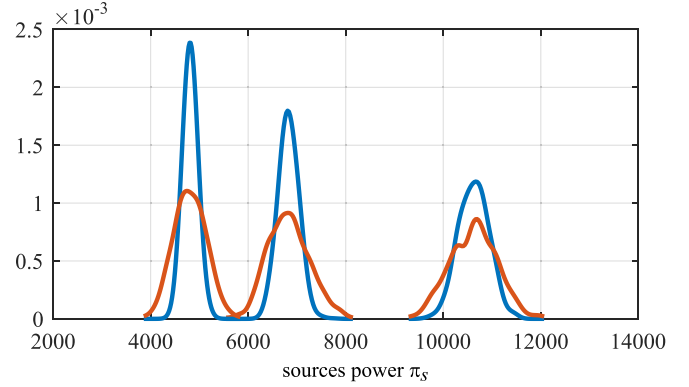


FIG. 12. (Color online) Histogram of the acoustical power spectra $\pi_s(f)$ of the three sources $s = 1, 2, 3$ at $f = 1850$ Hz. The red (light gray) histograms are obtained from the three Markov chains $\{\mathbf{S}_{qq}^{[n]}\}$, $\{\alpha^{2,[n]}\}$, $\{\beta^{2,[n]}\}$ of the Gibbs sampler and the blue (dark gray) one from the first chain only with frozen values of α^2 and β^2 . (Histograms computed with Hamming kernel smoothing with bandwidth equal to 100 units.)

Another result that can hardly be obtained by other means is the total statistical dispersion due to all sources of uncertainties in the inverse problem, including those relating to the estimation of hyperparameters such as α_k^2 and β_m^2 . This is demonstrated here by comparing the histograms of the acoustical power spectra $\pi_s(f)$, $s = 1, 2, 3$ at $f = 1850$ Hz obtained from the Markov chains $\{\mathbf{S}_{qq}^{[n]}\}$, $\{\alpha^{2,[n]}\}$, $\{\beta^{2,[n]}\}$ with the histograms obtained from the chain $\{\mathbf{S}_{qq}^{[n]}\}$, wherein the values of α^2 and β^2 have been replaced by their average over all iterations. As seen in Fig. 12, the first strategy only can account for all sources of errors, thus leading to wider histograms than the second strategy. The difference is especially noticeable at frequencies where regularization is strong, thus with high values of β^2 and of its variance.

Being severely ill-posed, the present inverse problem eventually illustrates well the benefit of the marginalization or dropout strategies of Sec. III C. With this end in view, the Gibbs sampler is run with hyperparameters α^2 and β^2 purposely fixed to mistuned values so as to quickly produce an unstable estimate, as depicted in Fig. 13(a). Yet, this is easily fixed by dropping microphones randomly (here 2 out of 20),

as seen in Fig. 13(b). A nearly similar result is obtained by assigning a small random error around the sensor positions. However, the price to pay for this robust regularization is necessarily a loss in the spatial resolution.

VI. CONCLUSION

This paper has investigated a solution, based on running Markov chains of a cross-spectral-matrix-based Gibbs sampler, to assess estimation errors involved in the reconstruction sound sources. The Gibbs sampler can converge within a few iterations when initialized with Bayesian focusing, which makes it a realistic tool in spite of its high computational cost. This suggests using it as a post-processing tool, subsequent to Bayesian focusing.

In addition, this paper has proposed means to display the estimation errors. One solution that seems particularly relevant is to map the coefficient of determination, as a function of space, together with the reconstructed source distribution.

The Gibbs sampler designed in this paper has shown great flexibility to accommodate various scenarios. In particular, it can easily handle non-homogeneous errors on the sensors, enforce the sparsity of the reconstructed sources, or regularize the inverse problem against uncertainties in the sensor positions. This opens new prospects to be further investigated in future research.

ACKNOWLEDGMENTS

This work was performed within the framework of the Labex CeLyA of Université de Lyon, operated by the French National Research Agency (Grant No. ANR-10-LABX-0060/ANR-11-IDEX-0007), and supported by the national project FUI LUG2.

APPENDIX A: STANDARD GIBBS SAMPLER

1. Expression of $[\mathbf{q}_i|\text{rest}]$

Using Bayes rule, one has

$$\begin{aligned} [\mathbf{q}_i|\text{rest}] &= [\mathbf{q}_i|\{\beta_m^2\}, \{\alpha_k^2\}, \boldsymbol{\theta}, \{\mathbf{p}_i\}] \\ &\propto [\mathbf{p}_i|\{\beta_m^2\}, \boldsymbol{\theta}, \{\mathbf{q}_i\}][\mathbf{q}_i|\{\alpha_k^2\}] \\ &\propto \mathcal{N}_{\mathbb{C}}(\mathbf{p}_i; \mathbf{H}(\boldsymbol{\theta})\mathbf{q}_i, \text{dg}\{\beta_m^2\})\mathcal{N}_{\mathbb{C}}(\mathbf{q}_i; \mathbf{0}; \text{dg}\{\alpha_k^2\}). \end{aligned} \quad (\text{A1})$$

The above product of Gaussians is itself proportional to the Gaussian $\mathcal{N}_{\mathbb{C}}(\mathbf{q}_i; \mu_{\mathbf{q}}, \boldsymbol{\Omega}_q)$ defined in Eq. (10) (see, e.g., Ref. 49).

2. Expression of $[\beta_m^2|\text{rest}]$ in the heteroscedastic case

$$\begin{aligned} [\beta_m^2|\text{rest}] &= [\beta_m^2|\{\mathbf{q}_i\}, \boldsymbol{\theta}, \{\mathbf{p}_i\}] \\ &\propto [\{\mathbf{p}_i\}|\beta_m^2, \boldsymbol{\theta}, \{\mathbf{q}_i\}][\beta_m^2] \\ &= \prod_{i=1}^I \mathcal{N}_{\mathbb{C}}(\mathbf{p}_i; \mathbf{H}(\boldsymbol{\theta})\mathbf{q}_i, \text{dg}\{\beta_m^2\})\mathcal{IG}(\beta_m^2; a_\beta, b_\beta). \end{aligned} \quad (\text{A2})$$

Thanks to the conjugacy of the inverse Gamma with the Gaussian, the above product is itself proportional to the

inverse Gamma $\mathcal{IG}(\beta_m^2; a_\beta + I, b_\beta + b_{\beta_m}^{\text{post}})$ with $b_{\beta_m}^{\text{post}}$ as defined in Sec. III A.

3. Expression of $[\beta^2|\text{rest}]$ in the homoscedastic case

$$\begin{aligned} [\beta^2|\text{rest}] &= [\beta^2|\{\mathbf{q}_i\}, \boldsymbol{\theta}, \{\mathbf{p}_i\}] \\ &\propto \prod_{i=1}^I \mathcal{N}_{\mathbb{C}}(\mathbf{p}_i; \mathbf{H}(\boldsymbol{\theta})\mathbf{q}_i, \beta^{-2}\mathbf{I}_M) \\ &\quad \times \prod_{m=1}^M \mathcal{IG}(\beta^2; a_\beta, b_\beta) \\ &\propto \mathcal{IG}\left(\beta^2; a_\beta + MI, b_\beta + \sum_{m=1}^M b_{\beta_m}^{\text{post}}\right). \end{aligned} \quad (\text{A3})$$

4. Expression of $[\alpha_k^2|\text{rest}]$ in the heteroscedastic case

$$\begin{aligned} [\alpha_k^2|\text{rest}] &= [\alpha_k^2|\{\mathbf{q}_i\}, \boldsymbol{\theta}, \{\mathbf{p}_i\}] \\ &\propto [\{\mathbf{q}_i\}|\{\alpha_k^2\}][\alpha_k^2] \\ &= \prod_{i=1}^I \mathcal{N}_{\mathbb{C}}(\mathbf{q}_i; \mathbf{0}; \text{dg}\{\alpha_k^2\})\mathcal{IG}(\alpha_k^2; a_\alpha, b_\alpha) \\ &\propto \mathcal{IG}\left(\alpha_k^2; a_\alpha + I, b_\alpha + b_{\alpha_k}^{\text{post}}\right) \end{aligned} \quad (\text{A4})$$

with $b_{\alpha_k}^{\text{post}}$ as defined in Sec. III A and where the last line again results from the conjugacy of the inverse Gamma prior with the Gaussian.

5. Expression of $[\alpha^2|\text{rest}]$ in the homoscedastic case

$$\begin{aligned} [\alpha^2|\text{rest}] &= [\alpha^2|\{\mathbf{q}_i\}, \boldsymbol{\theta}, \{\mathbf{p}_i\}] \\ &\propto \prod_{i=1}^I \mathcal{N}_{\mathbb{C}}(\mathbf{q}_i; \mathbf{0}; \alpha^2\mathbf{I}_K) \prod_{k=1}^K \mathcal{IG}(\alpha^2; a_\alpha, b_\alpha) \\ &\propto \mathcal{IG}\left(\alpha^2; a_\alpha + KI, b_\alpha + \sum_{k=1}^K b_{\alpha_k}^{\text{post}}\right). \end{aligned} \quad (\text{A5})$$

APPENDIX B: CSM-BASED GIBBS SAMPLER

Proof of Eq. (14),

$$\begin{aligned} \mathbf{S}_{pq} &= \frac{1}{I} \sum_{i=1}^I \mathbf{p}_i \mathbf{q}_i^H = \frac{1}{I} \sum_{i=1}^I \mathbf{p}_i (\mu_{\mathbf{q}_i} + \boldsymbol{\Omega}_q^{1/2} \boldsymbol{\xi}_i)^H \\ &= \frac{1}{I} \sum_{i=1}^I \mathbf{p}_i \mathbf{p}_i^H \mathbf{B}^H + \frac{1}{I} \sum_{i=1}^I \mathbf{p}_i \boldsymbol{\xi}_i^H \boldsymbol{\Omega}_q^{H/2}, \end{aligned} \quad (\text{B1})$$

where the first term on the right-hand side is recognized as $\mathbf{S}_{pp} \mathbf{B}^H$ and the second one is seen to converge to zero since the random variables \mathbf{p}_i and $\boldsymbol{\xi}_i$ are independent.

Proof of Eq. (15),

$$\begin{aligned} &\frac{1}{I} \sum_{i=1}^I (\mathbf{p}_i - \mathbf{H}\mathbf{q}_i)(\mathbf{p}_i - \mathbf{H}\mathbf{q}_i)^H \\ &= \frac{1}{I} \sum_{i=1}^I (\mathbf{p}_i \mathbf{p}_i^H + \mathbf{H}\mathbf{q}_i \mathbf{q}_i^H \mathbf{H}^H - \mathbf{p}_i \mathbf{q}_i^H \mathbf{H}^H - \mathbf{H}\mathbf{q}_i \mathbf{p}_i^H) \\ &= \mathbf{S}_{pp} + \mathbf{H}\mathbf{S}_{qq} \mathbf{H}^H - \mathbf{S}_{pq} \mathbf{H}^H - \mathbf{H}\mathbf{S}_{pq}^H. \end{aligned} \quad (\text{B2})$$

The rest of the proof is to substitute expressions (13) and (14) for \mathbf{S}_{qq} and \mathbf{S}_{pq} .

APPENDIX C: THRESHOLD ON SOURCE DISTRIBUTION

Let us momentarily consider the case where \mathbf{B} and $\mathbf{\Omega}_q$ are known. Then, the posterior PDF of $q_{ki} = \mathbf{e}_k^H \mathbf{q}_i$ given \mathbf{B} and $\mathbf{\Omega}_q$ is the complex Gaussian with mean $\mathbf{e}_k^H \mathbf{B} \mathbf{p}_i$ and covariance $\mathbf{e}_k^H \mathbf{\Omega}_q^{1/2} \mathbf{W} \mathbf{\Omega}_q^{H/2} \mathbf{e}_k$. In turn, the quadratic average of the source coefficients has the posterior PDF

$$\left[\frac{1}{I} \sum_{i=1}^I |q_{ki}|^2 \mid \mathbf{S}_{pp}, \mathbf{B}, \mathbf{\Omega}_q \right] = \frac{1}{2I} \chi_{2I}^2(\delta_k) \mathbf{e}_k^H \mathbf{\Omega}_q \mathbf{e}_k \quad (\text{C1})$$

with $\chi_{2I}^2(\delta_k)$ the noncentral chi-squared distribution with $2I$ degrees of freedom and noncentrality parameter $\delta_k = \mathbf{e}_k^H \mathbf{B} \mathbf{S}_{pp} \mathbf{B}^H \mathbf{e}_k$. Under the null hypothesis, $\delta_k = 0$ and the above PDF then involves a standard chi-squared distribution. In this case, the $(1-p)$ posterior quantile on $\sum_{i=1}^I |q_{ki}|^2 / I$ is $\chi_{2I, 1-p}^2 \mathbf{e}_k^H \mathbf{\Omega}_q \mathbf{e}_k / (2I)$. This defines a threshold under which one expects a fraction $(1-p)$ of the values to fall. The last step is to use these results as approximations when \mathbf{B} and $\mathbf{\Omega}_q$ are unknown. The proper way to deal with this situation would be to marginalize the posterior with respect to these variables. Instead, a technique known as empirical Bayes is used, which consists in replacing them by their posterior means. Then $\sum_{i=1}^I |q_{ki}|^2 / I$ is replaced by \mathcal{P}_k as given in Eq. (17) and the threshold by $(\chi_{2I, 1-p}^2 / (2I)) \mathbf{e}_k^H \langle (\mathbf{\Omega}_q^{1/2})^{[n]} \mathbf{W}^{[n]} (\mathbf{\Omega}_q^{H/2})^{[n]} \rangle \mathbf{e}_k$.

APPENDIX D: TEST OF HOMOSCEDASTICITY OF ERRORS

Since the posterior PDF of $X_{n,m}$ is Gaussian with mean independent of index m under the null hypothesis, the average $\bar{X}_m = (1/N) \sum_{n=1}^N X_{n,m}$ is also *a posteriori* Gaussian with a mean μ and a variance $\sigma^2 / N_{\text{eff}}$. Similarly, the overall average $\bar{X} = (1/M) \sum_{m=1}^M \bar{X}_m$ is *a posteriori* Gaussian with mean μ and a variance $\sigma^2 / (MN_{\text{eff}})$. Thus, the sum of squares $\sum_{m=1}^M |\bar{X}_m - \bar{X}|^2$, which reflects the squared distance between the chains, is distributed like $(\sigma^2 / N_{\text{eff}}) \chi_{M-1}^2$. In addition, the sum of the squares $\sum_{m=1}^M \sum_{n=1}^N |X_{n,m} - \bar{X}_m|^2$, which reflects the variance within the chains, is distributed like $\sigma^2 \cdot \chi_{M(N_{\text{eff}}-1)}^2$. Therefore, the F-ratio given by Eq. (21) follows a Fisher distribution $F_{M-1, M(N_{\text{eff}}-1)}$. The null hypothesis is thus rejected if $F > F_{M-1, M(N_{\text{eff}}-1), 1-p}$.

¹P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1998).

²A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2004).

³J. Dettmer, S. E. Dosso, and J. C. Osler, "Bayesian evidence computation for model selection in non-linear geoacoustic inference problems," *J. Acoust. Soc. Am.* **128**(6), 3406–3415 (2010).

⁴W. Sanders, "Estimation of error bounds in geoacoustic inversions," *J. Acoust. Soc. Am.* **113**(4), 2191–2191 (2003).

⁵S. E. Dosso, H. Dong, and K. Duffaut, "Model selection for profile structure in Bayesian geoacoustic inversion," *J. Acoust. Soc. Am.* **141**(5), 3470–3470 (2017).

⁶M. Blau, "Indirect measurement of multiple excitation force spectra by FRF matrix inversion: Influence of errors in statistical estimates of FRFs and response spectra," *Acustica* **85**(4), 464–479 (1999).

- ⁷E. Zhang, J. Antoni, and P. Feissel, "Bayesian force reconstruction with an uncertain model," *J. Sound Vib.* **331**(4), 798–814 (2012).
- ⁸K. R. James and D. R. Dowling, "A probability density function method for acoustic field uncertainty analysis," *J. Acoust. Soc. Am.* **118**(5), 2802–2810 (2005).
- ⁹D. Battle, "Applications of Markov chain Monte Carlo methods in ocean acoustics," *J. Acoust. Soc. Am.* **119**(5), 3343–3343 (2006).
- ¹⁰S. E. Dosso, "Bayesian inversion methods in ocean geoacoustics," *J. Acoust. Soc. Am.* **121**(5), 3170–3171 (2007).
- ¹¹S. E. Dosso and M. J. Wilmut, "Bayesian source tracking in an uncertain ocean environment," *J. Acoust. Soc. Am.* **125**(4), 2730–2730 (2009).
- ¹²S. E. Dosso, "Bayesian localization, tracking, and environmental inference," *J. Acoust. Soc. Am.* **139**(4), 2194–2195 (2016).
- ¹³Z. Michalopoulou, "A Gibbs sampling approach to matched-field source localization and deconvolution," *J. Acoust. Soc. Am.* **108**(5), 2646–2646 (2000).
- ¹⁴P. L. Nielsen and S. E. Dosso, "Bayesian inversion of propagation and reverberation data," *J. Acoust. Soc. Am.* **120**(5), 3356–3356 (2006).
- ¹⁵Z.-H. Michalopoulou, "Multiple source localization using a maximum *a posteriori* Gibbs sampling approach," *J. Acoust. Soc. Am.* **120**(5), 2627–2634 (2006).
- ¹⁶M. Q. Wu and M. J. Crocker, "A statistical analysis of the estimation error of sound power measured by sound intensity techniques," *J. Acoust. Soc. Am.* **80**(S1), S119–S119 (1986).
- ¹⁷T. Loyau and J. Pascal, "Statistical errors in the estimation of the magnitude and direction of the complex acoustic intensity vector," *J. Acoust. Soc. Am.* **97**(5), 2942–2962 (1995).
- ¹⁸J. Ghan, B. Cazzolato, and S. Snyder, "Statistical errors in the estimation of time-averaged acoustic energy density using the two-microphone method," *J. Acoust. Soc. Am.* **115**(3), 1179–1184 (2004).
- ¹⁹M. R. Bai, J.-G. Ih, and J. Benesty, *Acoustic Array Systems: Theory, Implementation, and Application*, 1st ed. (Wiley-IEEE Press, New York, 2013).
- ²⁰G. P. Carroll, "The effect of sensor placement errors on cylindrical near-field acoustic holography," *J. Acoust. Soc. Am.* **105**(4), 2269–2276 (1999).
- ²¹A. W. van der Vaart, *Asymptotic Statistics*, in *Cambridge Series in Statistical and Probabilistic Mathematics* (Cambridge University Press, Cambridge, 1998), pp. 25–34.
- ²²J. A. Scales and L. Tenorio, "Prior information and uncertainty in inverse problems," *Geophysics* **66**, 389 (2001).
- ²³A. M. Stuart, "Inverse problems: A Bayesian perspective," *Acta Numerica* **19**, 451–559 (2010).
- ²⁴J. Antoni, "A Bayesian approach to sound source reconstruction: Optimal basis, regularization, and focusing," *J. Acoust. Soc. Am.* **131**(4), 2873–2890 (2012).
- ²⁵A. Pereira, Q. Leclère, and J. Antoni, "A theoretical and experimental comparison of the equivalent source method and a Bayesian approach to noise source identification," in *Proceedings of Berlin Beamforming Conference*, Berlin (2012), pp. 1–12.
- ²⁶A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, 2nd ed. (Chapman and Hall/CRC, New York, 2004).
- ²⁷S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, *Handbook of Markov Chain Monte Carlo* (CRC, Boca Raton, FL, 2011).
- ²⁸S. E. Dosso, "Quantifying uncertainty in geoacoustic inversion. I. A fast Gibbs sampler approach," *J. Acoust. Soc. Am.* **111**(1), 129–142 (2002).
- ²⁹S. E. Dosso and P. L. Nielsen, "Quantifying uncertainty in geoacoustic inversion. II. Application to broadband, shallow-water data," *J. Acoust. Soc. Am.* **111**(1), 143–159 (2002).
- ³⁰C. Huang, P. Gerstoft, and W. S. Hodgkiss, "Influence of data uncertainty on matched-field geoacoustic inversion," *J. Acoust. Soc. Am.* **115**(5), 2409–2409 (2004).
- ³¹R. T. Muehleisen, "Use of the Monte Carlo method for uncertainty analysis of acoustic models and measurements," *J. Acoust. Soc. Am.* **118**(3), 1890–1890 (2005).
- ³²S. E. Dosso and M. J. Wilmut, "Uncertainty estimation in simultaneous Bayesian tracking and environmental inversion," *J. Acoust. Soc. Am.* **124**(1), 82–97 (2008).
- ³³E. Zhang, J. Antoni, B. Dong, and H. Snoussi, "Bayesian space-frequency separation of wide-band sound sources by a hierarchical approach," *J. Acoust. Soc. Am.* **132**(5), 3240–3250 (2012).
- ³⁴M. R. Bai, A. Agarwal, C.-C. Chen, and Y.-C. Wang, "Bayesian approach of nearfield acoustic reconstruction with particle filters," *J. Acoust. Soc. Am.* **133**(6), 4032–4043 (2013).

- ³⁵J. Antoni, T. L. Magueresse, Q. Leclère, and P. Simard, "Sparse acoustical holography from iterated Bayesian focusing," *J. Sound Vib.* **446**, 289–325 (2019).
- ³⁶M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.* **1**, 211–244 (2001).
- ³⁷M. S. Brandstein, J. E. Adcock, and H. F. Silverman, "Microphone–array localization error estimation with application to sensor placement," *J. Acoust. Soc. Am.* **99**(6), 3807–3816 (1996).
- ³⁸D. R. Brillinger, *Time Series: Data Analysis and Theory* (Holt, Rinehart & Winston, New York, 1975).
- ³⁹H. Gzyl, *The Method of Maximum Entropy*, Vol. 29 of *Series on Advances in Mathematics for Applied Sciences* (World Scientific, Singapore, 1995).
- ⁴⁰S. Höhna, T. A. Heath, B. Boussau, M. J. Landis, F. Ronquist, and J. P. Huelsenbeck, "Probabilistic graphical model representation in phylogenetics," *Syst. Biol.* **63**(5), 753–771 (2014).
- ⁴¹R. A. Levine and G. Casella, "Optimizing random scan Gibbs samplers," *J. Multivar. Anal.* **97**(10), 2071–2100 (2006).
- ⁴²N. R. Goodman, "The distribution of the determinant of a complex Wishart distributed matrix," *Ann. Math. Statist.* **34**(1), 178–180 (1963).
- ⁴³M. S. Bartlett, "XX.—On the theory of statistical regression," *Proc. R. Soc. Edinburgh* **53**, 260–283 (1934).
- ⁴⁴X. Wang, B. Quost, J.-D. Chazot, and J. Antoni, "Estimation of multiple sound sources with data and model uncertainties using the EM and evidential EM algorithms," *Mech. Syst. Sign. Process.* **66-67**, 159–177 (2016).
- ⁴⁵L. Gilquin, S. Bouley, J. Antoni, T. L. Magueresse, and C. Marteau, "Sensitivity analysis of two inverse methods: Conventional beamforming and Bayesian focusing," *J. Sound Vib.* **455**, 188–202 (2019).
- ⁴⁶D. E. Johnson and V. Hegemann, "Procedures to generate random matrices with noncentral distributions," *Commun. Stat.* **3**(7), 691–699 (1974).
- ⁴⁷D. V. Lindley, *Introduction to Probability and Statistics from a Bayesian Viewpoint* (Cambridge University Press, Cambridge, 1965).
- ⁴⁸P. M. Lee, *Bayesian Statistics—An Introduction* (Arnold, London, 2004).
- ⁴⁹S. Roweis, "Gaussian identities," <https://cs.nyu.edu/~roweis/notes/gaussid.pdf> (Last viewed 08/07/2019).