BeBeC-2018-D12

SPARSE HOLOGRAPHY FROM ITERATED BAYESIAN FOCUSING

Jérôme Antoni[1], Thibaut Lemagueresse[2] and Quentin Leclère[1]
[1]Univ Lyon, INSA-Lyon, Laboratoire Vibrations Acoustique, F-69621 Villeurbanne, France
[2]MicrodB, F-69134, Ecully, France

Abstract

In a previous work, a unified vision of some acoustic imaging methods has been given within a Bayesian framework. One advantage of the so-called "Bayesian focusing" approach is that it introduces an aperture function that acts as a lens and thus significantly improves the reconstruction results in terms of spatial resolution, but also of quantification over a much larger frequency interval than allowed by conventional methods. This is particularly remarkable when the aperture function is allowed to become very narrow in the case of sparse sources. The aim of the present paper it to demonstrate that the aperture function – which was previously manually tuned by the used – can be automatically estimated, together with the source distribution, in the same inverse problem. The intuition is to appraise an unknown aperture function from the energy distribution of a first estimate of the sources. This provides an updating rule of the aperture function, where the posterior in the current iteration is then used to refine the prior in the next iteration. The resulting al-gorithm is an iterated version of the Bayesian focusing approach, which can be formalized as an Expectation-Maximization algorithm with proved convergence. Some byproducts of the iterated Bayesian focusing approach – which are more painfully reached by other approaches – are 1) to provide a technique for the automatic setting of the regularization parameter, 2) to easily handle multiple snapshots, 3) to possibly apply on the cross-spectral matrix of the measurements, and 4) to easily allow the grouping of frequencies for the wideband analysis of sources that are stationary in space.

1 INTRODUCTION

The reconstruction of sound sources from acoustic array measurements is known as a difficult inverse problem. Several methods have been proposed over the years with sustained efforts to

reach improved performance. An unavoidable limitation arises from the fact that reconstructing a continuous acoustical field from a limited number of discrete and remote measurements, as returned by an array of microphones, has no unique solution in general. This issue is particularly critical in near-field configurations, where the spectrum of the propagation operator is characterized by a fast exponential decay and is therefore hardly invertible. Consequently, sound sources cannot be reliably reconstructed unless additional information is incorporated into the inverse problem in order to reduce the space of possible solutions. So far, Tikhonov regularization has been the prevalent method to address this issue, by selecting solutions with minimal energy. However, this is often at the expense of accuracy of the reconstruction, and especially in term of spatial resolution. When the sources of interest are known to have a sparse representation, much better strategies than Tikhonov regularization are actually feasible.

There are few reports to date of the application of sparse inversion to (near-field) acoustical holography (NAH). Reference [6] was probably first to address the subject. The objective therein is to demonstrate that a substantial reduction in the number of microphones can be achieved (by more than a factor 10) without affecting the reconstruction performance by exploiting the sparsity of the acoustical field. The two fundamental assumptions are that the acoustical field can be represented by a few plane waves (sparsity in the wavenumber domain) and that the microphones are distributed randomly.

Steered by somewhat different objectives, Ref. [11] recently introduced the concept of wideband acoustical holography (WBH). The motivation in this work is for a method that can bridge the gap between NAH and classical beamforming in order to reconstruct sources over a wide frequency range. On the one hand, NAH can cover the low frequencies but requires the array to be placed at a small distance. On the other hand, beamforming has good performance in a high frequency range but requires the array to be moved at a larger distance. By enforcing sparsity in the Equivalent Source Method (ESM) (i.e. representation by a few point sources), the proposed WBH method is able to reconstruct sources with a remarkable quantification of the sound power over a wide frequency range. A numerical experiment seems to indicate that WBH is also able to consider extended sources such as produced by the vibration of a plate. The algorithm proposed in Ref. [11] is named "fast" because it uses an Iterative Hard Thresholding algorithm instead of a direct ℓ_1-norm minimization, yet it still depends on several user-dependent parameters.

Following a similar same idea, Ref. [7] also proposed a sparse version of ESM, yet based on direct ℓ_1-norm minimization. In accordance with the results of Ref. [6], the authors also stress the importance of using a sampling scheme that leads to low column coherence of the sensing matrix, a condition which is better met when the distance between the measurement plane and the source plane is decreased. As in Ref. [11], the applicability to the reconstruction of extended sources is also demonstrated by placing the equivalent sparse sources behind the actual radiating surface. The investigated frequency range is limited to that of classical NAH, where moderate improvement is observed in the reconstruction of source spatial distribution as compared to ℓ_2 regularization. Reference [1] introduced independently a sparse wave superposition method (with sparsity placed on the so-called "charge points" which actually coincide with the monopoles generating the acoustic field), yet its validation was limited to a numerical simulation.

It is noteworthy that Ref. [17] introduced a sparse ESM before the former publications, which was later named Iterative ESM (iESM). It was compared with Generalized Inverse Beamforming in Ref. [16] and also with deconvolution approaches in Ref. [14].

The main contribution of the present paper is to demonstrate that sparse NAH is naturally devised by iterating the Bayesian Focusing (BF) method introduced in Ref. [2]. BF solves the inverse acoustical problem by considering the sound sources as random variables. They are first vaguely described by a prior probability density function (PDF) which reflects the user expectation before the experiment is run. Next, the posterior PDF is found as the product of the likelihood function and the prior PDF, which provides the probability of the sound sources after taking the measurements. A typical point estimate is then returned by the maximum value of the posterior PDF – the maximum aposteriori (MAP) – i.e. the source configuration with highest probability of occurrence given the observed data. In BF, the variance of the prior PDF – coined the aperture function (AF) – appears as a key quantity; intuitively, it will take high values where the sound sources are expected to radiate from and nil values elsewhere. Therefore, by shrinking the space of solutions to a confined region, the AF improves the quality of the reconstruction in terms of spatial resolution and of source strength quantification over a larger frequency interval than allowed by conventional methods. This is all the more remarkable as the AF is allowed to become very narrow, as in the case of sparse sources. By analogy with optics, it plays the role of a lens that focuses the light onto a point.

In Ref. [2], the AF was manually tuned by the user. A natural idea is to automate this process by using the current estimate of the source distribution as the AF to apply in the next iteration. It is shown in this paper that iterating BF generally leads to the recovery of sparse solutions. Not only is the common ℓ_1-norm penalty recovered as a particular case, but many other sparsity enforcements – possibly stronger than the ℓ_1-norm – can be devised depending on how the AF is updated. The reason is that when the AF is considered as a random quantity (in the Bayesian setting all unknowns are described by random variables), it then assigns the sound sources with a prior PDF in the form of a "scale mixture of Gaussians" (SMoG), which necessarily promotes sparsity. Based on this finding, the convergence of the iterations can be proved by establishing a formal equivalence with the Expectation-Maximization algorithm. This algorithm will be referred as Iterating Bayesian Focusing (IBF).

One advantage of IBF is to provide some physical insight into the mechanism of sparsity enforcement. Through the role played by the AF, it explains why promoting sparse solutions not only increases the spatial resolution, but also improves the estimation of the source levels and the source directivity over a larger frequency range than allowed by conventional methods.

Another prime advantage of the approach is to inherit the automated regularization of BF [18]. Contrary to Tikhonov regularization for which well established algorithms exist (e.g. the L-curve, GCV), this point remains an issue in compressive beamforming. Here, by taking a Bayesian perspective, the regularization parameter is jointly inferred in IBF with proved convergence.

The Bayesian framework also easily allows the processing of multiple snapshots and multiple frequencies, which both involve the consideration of group sparsity [12]. The multi-snapshot case is addressed here by means of the cross-spectral matrix (CSM), which turns out a sufficient statistics while, at the same time, often being the only quantity returned by commercial data acquisition systems (the approach differs in this aspect from that of Ref. [9]). The multi-frequency case is of interest when the sound sources have a constant position in space, independently of their frequency content, thus leading to broadband processing. Both the multi-snapshot and the multi-frequency considerations are shown to improve the estimation of the sound sources.

2 Bayesian Focusing

The array is composed of M microphones indexed by the lower-case letter m, located at positions $\mathbf{r}_m$, $m = 1,...,M$. All measurements are considered at a given frequency f, after application of the Fourier transform on a series of N snapshots indexed by the lower-case letter i ($i = 1,...,N$). The Fourier coefficients of the sound pressure measured by microphone m and assigned to snapshot i then reads $p_{m,i}$, where explicit dependence on frequency f is dropped to simplify the notation. The M Fourier coefficients $p_{m,i}$, $m = 1,...,M$ are collected in the column vector $\mathbf{p}_i$. The cross-spectral matrix (CSM) of the measurements averaged over N snapshots is defined as

$$\mathbf{S}_{pp} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{p}_i \mathbf{p}_i^H. \tag{1}$$

In the following, formulations in either continuous or discrete space will be used alternatively, the former because it provides deeper physical insight and the latter because it corresponds to the numerical resolution of the problem. In the continuous formulation, the source distribution is assumed to be a scalar field – e.g. parietal pressure, normal velocity – that adheres on the source surface Γ_s. In a functional notation, it is noted $s_i(\mathbf{r})$ where $\mathbf{r} \in \Gamma_s$ stands for the position vector. In the discrete formulation of the problem, the spatial samples of the source distribution taken at positions $\mathbf{r}_q$, $q = 1,...,Q$ are noted $s_{q,i} \doteq s_i(\mathbf{r}_q)$. They are referred to as the "source coefficients" and are collected in the column vector $\mathbf{s}_i$, with dimension Q. The power spectrum of the source coefficients at position $\mathbf{r}_q$, averaged over N snapshots, is defined as

$$S_q^2 = \frac{1}{N}\sum_{i=1}^{N} |s_{q,i}|^2. \tag{2}$$

The matricial relationship between the source distribution and the radiated pressures measured by the microphones reads, after discretization form,

$$\mathbf{p}_i = \mathbf{G}\mathbf{s}_i + \mathbf{n}_i, \quad i = 1,...,N, \tag{3}$$

where $\mathbf{G}$ stands for an $M \times Q$ matrix whose entry (i,q) is fed with $G(\mathbf{r}_m|\mathbf{r}_q)\Delta\Gamma(\mathbf{r}_q)$ – with $\Delta\Gamma(\mathbf{r}_q)$ a small surface element at position $\mathbf{r}_q$ – and where vector $\mathbf{n}_i$ stacks the noise terms $n_{m,i}$.

The aim of the inverse problem is to recover an estimate of the source distribution $s_i(\mathbf{r})$ from the observation of the measured pressures $\mathbf{P} \doteq \{\mathbf{p}_i\}_{i=1}^N$ returned by the array of microphone. BF provides a solution to this problem that makes intensive use of prior information about the spatial structure of the source field to be recovered.

The philosophy in the Bayesian approach is to see all unknowns in the inverse problem as random variables and to infer them through their PDFs. Of interest here is the posterior PDF $[\mathbf{s}_i|\mathbf{p}_i]$ of the vector $\mathbf{s}_i$ of coefficients coefficients given the measurement $\mathbf{p}_i$. The knowledge of $[\mathbf{s}_i|\mathbf{p}_i]$ completely characterizes the information that can be gained on $\mathbf{s}_i$ once the data have been measured. A popular point estimate used in this paper is the maximum aposteriori (MAP) estimate,

$$\hat{\mathbf{s}}_i = \underset{\mathbf{s}_i}{\text{Argmax}}[\mathbf{s}_i|\mathbf{p}_i], \tag{4}$$

to be interpreted as the value of $\mathbf{s}_i$ with maximum probability, yet another plausible estimate is the posterior mean $\mathbb{E}\{\mathbf{s}_i|\mathbf{p}_i\} = \int \mathbf{s}_i d[\mathbf{s}_i|\mathbf{p}_i]$.

Bayes's rule then makes it possible to express the "inverse probability" $[\mathbf{s}_i|\mathbf{p}_i]$ in terms of the probabilities assigned to the direct problem, that is

$$[\mathbf{s}_i|\mathbf{p}_i] = \frac{[\mathbf{p}_i|\mathbf{s}_i][\mathbf{s}_i]}{[\mathbf{p}_i]} \tag{5}$$

where $[\mathbf{p}_i|\mathbf{s}_i]$ is the PDF of observing the data given an instance of $\mathbf{s}_i$ (the so-called likelihood function), $[\mathbf{s}_i]$ is the PDF of the values possibly taken by $\mathbf{s}_i$ before the data are observed (the so-called prior) and $[\mathbf{p}_i] = \int [\mathbf{p}_i|\mathbf{s}_i] d[\mathbf{s}_i]$ is the "evidence". The Bayesian program is to introduce the PDFs $[\mathbf{p}_i|\mathbf{s}_i]$ and $[\mathbf{s}_i]$ according to the specificity of the problem and to deduce the posterior PDF $[\mathbf{s}_i|\mathbf{p}_i]$ by using Eq. (5).

Since the data are processed in the Fourier domain, it holds from the Central Limit Theorem applied to the Fourier transform that the elements of the noise vector $\mathbf{n}_i$ rapidly converge to a complex Gaussian [4]. Assuming for simplicity a zero mean and a covariance matrix proportional to the identity (this assumption reflects a homogeneous field; it is without loss of generality since the data can always be centered and standardized beforehand),

$$[\mathbf{p}_i|\mathbf{s}_i] = \frac{e^{-\beta^{-2}(\mathbf{p}_i - \mathbf{G}\mathbf{s}_i)^H(\mathbf{p}_i - \mathbf{G}\mathbf{s}_i)}}{\beta^{2M}\pi^M} = \mathcal{N}_{\mathbb{C}}(\mathbf{p}_i; \mathbf{0}, \beta^2\mathbf{I}) \tag{6}$$

where β^2 stands for the value of the noise power spectrum at frequency f (independently of space according to the previous assumption).

The specification of the prior PDF $[\mathbf{s}_i]$ allows much more flexibility and is actually central in this paper. The assumption initially used in Ref. [2] is to resort to a complex Gaussian, mainly because it makes analytic calculation tractable and also because it coincides with classical Tikhonov regularization. Namely,

$$[\mathbf{s}_i] = \frac{e^{-\alpha^{-2}\mathbf{s}_i^H \Sigma_0^{-2} \mathbf{s}_i}}{\alpha^{2K}\pi^K |\Sigma_0^2|} = \mathcal{N}_{\mathbb{C}}(\mathbf{s}_i; \mathbf{0}, \alpha^2\Sigma_0^2) \tag{7}$$

where $\alpha^2\Sigma_0^2 = \mathbb{E}\{\mathbf{s}_i^H \mathbf{s}_i\}$ stands for the prior covariance matrix of the coefficients, α^2 for their mean power and Σ_0^2 an unscaled matrix of correlation coefficients conventionally normalized such that $\text{trace}\{\Sigma_0^2\} = K$. The fact that $\alpha^2\Sigma_0^2$ does not depend on index i reflects the assumption of a stationary acoustic field, i.e. with constant statistical properties during the acquisition of the N snapshots. Using Eq. (5), one then finds the posterior PDF in the form of a complex Gaussian, whose posterior mean $\hat{\mathbf{s}}_i$ coincides with the MAP estimate as well as with the Tikhonov regularized least-square solution (LS) returned by

$$\hat{\mathbf{s}}_i = \underset{\mathbf{s}_i}{\text{Argmax}} [\mathbf{p}_i|\mathbf{s}_i][\mathbf{s}_i] = \underset{\mathbf{s}_i}{\text{Argmin}} \left(||\mathbf{p}_i - \mathbf{G}\mathbf{s}_i||_2^2 + \eta^2 ||\Sigma_0^{-1}\mathbf{s}_i||_2^2 \right) \tag{8}$$

with $\|\Sigma_0^{-1}\mathbf{s}_i\|_2^2 \doteq \mathbf{s}_i^H \Sigma_0^{-2} \mathbf{s}_i$, $\eta^2 = \frac{\beta^2}{\alpha^2}$ and wherein all terms not depending on $\mathbf{s}_i$ have been dropped out. The ratio $\eta^2 = \beta^2/\alpha^2$ in the above equations acts as the Tikhonov regularization parameter which controls the stability of the solution. One benefit of the Bayesian framework is to provide

a solution to automatically tune η^2, as described in Ref. [19].

In the context of this paper, the sources coefficients are assumed to be sparsely distributed in space; this means that only a very elements in vector $\mathbf{s}_i$ (always the same independently of snapshot i) are expected to have significant values, the other ones being close (or ideally equal) to zero. One way to enforce this property is to assign to $\mathbf{s}_i$ a prior PDF with heavy tails and very peaked at zero as compared to the Gaussian, which is denoted hereafter as leptokurtic (i.e. with a kurtosis greater than the Gaussian). The use of leptokurtic prior PDFs (sometimes referred to as sparse priors, sparsity enforcing priors or shrinkage prior) within the Bayesian framework for enforcing sparse solutions has been extensively studied in the recent statistical literature (see e.g. [3, 13]).

A physical interpretation is given hereafter in terms of SMoGs. Let the source coefficients have the following probabilistic model

$$s_{q,i} = \alpha \cdot \sigma_{0,q} \cdot \tau_q \cdot \varepsilon_{q,i}, \quad \varepsilon_{q,i} \sim \mathcal{N}_{\mathbb{C}}(0,1), \tag{9}$$

where the source power α^2 and the AF $\sigma^2_{0,q}$ are as previously defined, $\varepsilon_{i,q}$ denotes a standardized (i.e. zero-mean and unit variance) complex Gaussian random variable, and τ_q is now an additional non-negative random variable which reflects the "relative intensity" at the spatial position $\mathbf{r}_q$ (see Fig. 1) (if τ_q was deterministic, then model (9) would be equivalent to having $\mathbf{s}_i$ distributed as in Eq. (7)). It is noted at this juncture that α cannot be absorbed into τ_q because it acts as a scaling factor, whereas the latter will be shortly seen to act as a "sparifying" variable whose scale must be left free – α has the same unit as the source whereas τ_q is dimensionless. As compared to the uniform relative intensity used BF, the resulting spatial distribution of the source coefficients is thus sparser.

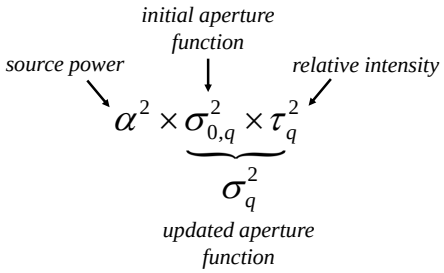

Figure 1: Structure of the sparse spatial prior, where index q relates to spatial position $\mathbf{r}_q$. α^2 is the global source power and the initial aperture function $\sigma^2_{0,q}$ reflects its expected spatial distribution before the experiment is run. The spatial variance $\alpha^2\sigma^2_{0,q}$ is used in the Gaussian prior PDF of BF; its multiplication by the relative intensity τ^2_q returns the spatial variance $\alpha^2\sigma^2_{0,q}\tau^2_q$ used in the SMoG prior PDF of Iterated Bayesian Focusing. The method updates the aperture function as $\sigma^2_q \leftarrow \sigma^2_{0,q}\tau^2_q$.

Although it is difficult to specify the values of τ^2_q before the inverse problem is solved, it is

possible to assign them a hyperprior PDF, $\left[\tau_q^2\right]$, that allows local variations in space. Given a spatial position $\mathbf{r}_q$ and considering all snapshots $i = 1,...,N$ observed at this position, the prior PDF of the source coefficients $\mathbf{S} \doteq \{s_{q,i}\}_{i=1}^N$ then reads

$$[\mathbf{S}] = \int_0^\infty \prod_{i=1}^N \left[s_{q,i}|\tau_q^2\right] d\left[\tau_q^2\right] = \mathop{\mathbb{E}}_{\tau_q^2}\left\{\prod_{i=1}^N \left[s_{q,i}|\tau_q^2\right]\right\} = \mathop{\mathbb{E}}_{\tau_q^2}\left\{\prod_{i=1}^N \mathcal{N}_{\mathbb{C}}\left(s_{q,i};0,\alpha^2\sigma_{0,q}^2\tau_q^2\right)\right\} \tag{10}$$

where $\mathbb{E}_{\tau_q^2}$ means that the expected value is taken with respect to the random variable τ_q^2. Inspection of Eq. (10) shows that the prior PDF of the source coefficients is a continuous mixture of Gaussians $\mathcal{N}_{\mathbb{C}}\left(s_{q,i};0,\alpha^2\sigma_{0,q}^2\tau_q^2\right)$ weighted by the "scale" $\left[\tau_q^2\right]$ – the so-called SMoG. This is in general no longer a Gaussian, but a leptokurtic PDF.

Different prior PDFs are obtained depending on the choice of the hyperprior $\left[\tau_q^2\right]$. Only two are investigated here. They are expressed in terms of a sufficient statistics, the normalized sum of squares

$$\chi_q^2 = N\frac{S_q^2}{\alpha^2\sigma_{0,q}^2} \tag{11}$$

which involves the ratio of the source power spectrum S_q^2 (see Eq. (2)) to the prior spatial variance $\alpha^2\sigma_{0,q}^2$ at the same position.

- **Multivariate complex Student-*t* (MCS).** Let us assume that τ^2 is distributed like an *inverse Gamma*, with shape and rate parameters $a > 0$ and $b > 0$, respectively:

$$\left[\tau^2\right] = \frac{b^a}{\Gamma(a)}\frac{e^{-b\tau^{-2}}}{\tau^{2(a+1)}}. \tag{12}$$

  It then holds that

$$\left[\chi_q^2\right] = \frac{\Gamma(N+a)b^a}{(\pi\alpha^2\sigma_{0,q}^2)^N\Gamma(a)\left(\chi_q^2+b\right)^{(N+a)}}, \tag{13}$$

  which is recognized as a multivariate complex version of the Student-*t* distribution. This PDF can be made very leptokurtic by setting the value of a small. By substituting the MCS for the prior PDF into the MAP, one arrives at an LS formulation with a logarithmic penalty, viz

$$\begin{aligned}\hat{\mathbf{s}}_i &= \mathop{\text{Argmax}}_{\mathbf{s}_i}[\mathbf{P}|\mathbf{S}][\mathbf{S}] = \mathop{\text{Argmin}}_{\mathbf{s}_i}\left(-\ln[\mathbf{p}_i|\mathbf{s}_i][\mathbf{S}]\right)\\ &= \mathop{\text{Argmin}}_{\mathbf{s}_i}\left(||\mathbf{p}_i-\mathbf{G}\mathbf{s}_i||_2^2+\beta^2(N+a)\ln\left(\alpha^{-2}\sum_{j=1}^N \mathbf{s}_j^H\Sigma_0^{-2}\mathbf{s}_j+b\right)\right).\end{aligned} \tag{14}$$

  Note that the logarithmic penalty is here naturally endowed with a regularization term $b > 0$ that prevents it from diverging.

- **Generalized multivariate complex Gaussian (GMCG).** The multivariate complex ver-

sion of a generalized Gaussian PDF reads

$$\left[\chi_q^2\right] = \frac{p/2}{(\pi\alpha^2\sigma_{0,q}^2)^N}\frac{\Gamma(N)}{\Gamma(2N/p)}e^{-\left(\chi_q^2\right)^{\frac{p}{2}}}, \quad 0 < p \leq 2. \tag{15}$$

Since the generalized Gaussian (also known as the exponential power distribution in the statistical literature) is an SMoG, the GMCG is also an SMoG (actually for $p \leq 2$). The expression of $\left[\tau^2\right]$ happens to be quite intricate, yet only the expression of the marginal PDF (15) will be needed in the following. Formula (15) extends the result of Ref. [15] to the multivariate case.

Sparsity is enforced by setting small values of p, strictly smaller than 2 and typically smaller than or equal to 1. As special cases, $p = 2$ corresponds to the multivariate complex Gaussian (ℓ_2-norm penalization) and $p = 1$ to the ℓ_1-norm penalization.

The GMCG prior gives rise to the penalized LS problem,

$$\hat{\mathbf{s}}_i = \underset{\mathbf{s}_i}{\text{Argmin}} \left(||\mathbf{p}_i - \mathbf{G}\mathbf{s}_i||_2^2 + \frac{\beta^2}{\alpha^p}\left(\sum_{j=1}^{N} \mathbf{s}_j^H \Sigma_0^{-2} \mathbf{s}_j\right)^{\frac{p}{2}} \right), \tag{16}$$

where the penalty function mixes the ℓ_2 and the ℓ_p norms.

3 The Iterated Bayesian Focusing algorithm

An algorithm is proposed hereafter to reach the MAP, described in terms of the normalized sum of squares

$$\chi_{q,[k-1]}^2 = N\frac{S_{q,[k-1]}^2}{\alpha_{[k-1]}^2 \sigma_{0,q,[k-1]}^2} \tag{17}$$

which involves the power spectrum $S_{q,[k-1]}^2 = \frac{1}{N}\sum_{i=1}^{N} |\hat{s}_{q,i,[k-1]}|^2$ of the source coefficients estimated at iteration $k-1$.

Algorithm 1

- **Step 0**: Design an initial AF $\sigma_{0,q}^2$ and construct the diagonal matrix Σ_0^2 whose q-th diagonal entry is $\sigma_{0,q}^2$.
- **Step 1**: Set $k = 0$. Estimate the source power $\alpha_{[0]}^2$ and the regularization parameter $\eta_{[0]}^2$ with the method of ref. ([19]). Initialize the algorithm with the MAP solution returned from a Gaussian prior.
- **Step 2**: Do $k \longleftarrow k+1$.
  Estimate the relative intensity as

$$\hat{\tau}_{q,[k]}^2 = -\left(\frac{\partial}{\partial \chi_{q,[k-1]}^2} \ln\left(\left[\chi_{q,[k-1]}^2\right]\right)\right)^{-1}, \tag{18}$$

where $[\chi^2_{q,[k-1]}]$ is the prior PDF expressed as a function of the normalized sum of squares $\chi^2_{q,[k-1]}$ at iteration $k-1$.

- **Step 3**: Update the AF as
$$\sigma^2_{q,[k]} = \sigma^2_{q,0}\hat{\tau}^2_{q,[k]} \tag{19}$$
and construct the diagonal matrix $\Sigma^2_{[k]}$ whose q-th diagonal entry is $\sigma^2_{q,[k]}$.

- **Step 4**: Estimate the source power and the regularization parameter as ([19])
$$\left(\alpha^2_{[k]}, \eta^2_{[k]}\right) = \text{Argmax}\left[\alpha^2, \eta^2 | \mathbf{S}_{pp}, \Sigma^2_{[k]}\right] \tag{20}$$

- **Step 5**: Estimate the power spectrum of the the source coefficients at position $\mathbf{r}_q$ as
$$S^2_{q,[k-1]} = \sigma^4_{q,[k]}\mathbf{g}^*_q(\mathbf{G}\Sigma^2_{[k]}\mathbf{G}^H + \eta^2_{[k]}\mathbf{I})^{-1}\mathbf{S}_{pp}(\mathbf{G}\Sigma^2_{[k]}\mathbf{G}^H + \eta^2_{[k]}\mathbf{I})^{-1}\mathbf{g}_q \tag{21}$$
where $\mathbf{g}_q$ is the q-th column of matrix $\mathbf{G}$.

  Stop iterations when $||\hat{\mathbf{s}}_{i,[k]} - \hat{\mathbf{s}}_{i,[k-1]}||/||\hat{\mathbf{s}}_{i,[k-1]}||$ is smaller than a given threshold, $0 < \varepsilon_2 < 1$.

IBF belongs to the family Iterative Re-weighted Least Squares (IRLS) algorithms [10], since at each iteration it solves a weighted least square problem whose solution is given by Eq. (21) with Σ^2_0 replaced by it updated version $\Sigma^2_{[k]}$ as obtained from Eq. (19). Yet, it is more rigorously derived as an Expected-Maximization (EM) algorithm, which makes possible to prove its convergence to the MAP solution – insightful connections between IRLS and EM have been investigated in Ref. [5].

The regularization parameter η^2 is iteratively updated together with the source coefficients with proved convergence. This is to be contrasted with other approaches found in the literature where the regularization parameter is either arbitrarily imposed or tuned afterwards, typically by using cross-validation or the L-curve. The latter approaches require running the optimization as many times as values of the regularization parameter are to be tested, and are therefore computationally more demanding – if not prohibitive in some applications such as that illustrated in subsection 4. In addition, the estimation of the regularization parameter by means of the Bayesian criterion used in Step 4 has been found more efficient than the cross-validation or the L-curve in many inverse acoustic scenarios [19].

A key step in the IBF algorithm is the update of the AF in Eq. (19) rooted on the current estimate of the relative intensity $\hat{\tau}^2_{q,[k]}$ given by Eq. (18).

- **Multivariate complex Student-*t* (MCS).** The expected relative intensity related to the MCS prior PDF is easily found by application of Eq. (18) to Eq. (13):
$$\hat{\tau}^2_{q,[k]} = \frac{\chi^2_{q,[k-1]} + b}{N + a}. \tag{22}$$

- **Generalized multivariate complex Gaussian (GMCG).** The expected relative intensity related to the GMCG is readily found as

$$\hat{\tau}^2_{q,[k]} = \frac{2}{p}\left(\chi^2_{q,[k-1]}\right)^{1-\frac{p}{2}}, \quad 0 < p \leq 2. \tag{23}$$

  This is again a monotonic power function of $\chi^2_{q,[k-1]}$, where high sparsity can be reached by setting p small.

The updating rule of the GMCG is particularly intuitive since it reads, in logarithmic scale,

$$\begin{aligned} \ln \sigma^2_{q,[k]} &= \ln\left(\sigma^2_{q,0}\right) + \ln\left(\hat{\tau}^2_{q,[k]}\right) \\ &= \frac{p}{2}\ln\left(\sigma^2_{q,0}\right) + \left(1-\frac{p}{2}\right)\ln\left(S^2_{q,[k-1]}\right) + C \end{aligned} \tag{24}$$

with C a global scaling that does not depend on the local position $\mathbf{r}_q$. Therefore, the updating is interpreted as the weighted geometric mean between the initial AF $\sigma^2_{q,0}$ and the previous estimation of the source power spectrum $S^2_{q,[k-1]}$. The value of p specifies the balance between the two terms. For $p = 1$, uniform weights of value $1/2$ are assigned to the two terms, whereas for $0 < p < 1$, more weight is given to the data-dependent term.

Experimental comparison of the above priors is carried on in a numerical simulation. The range distance is $R = 0.5$ m and an ULA with 20 microphones and spacing 0.1 m is used. The SNR is set to 20 dB and the number N of snapshots to 100. The source domain $\Gamma_s = \{r : -1 \leq r \leq 1\}$ (in meters) is discretized with a spatial step of 0.01 m. Initialization of the IBF algorithm is done by using a Hann window for the AF.

The first experiment investigates the reconstruction of two incoherent monopoles with 14 dB difference in power level, located at $r_1 = -0.05$ m and $r_2 = 0.1$ m, at the very low frequency $f = 100$ Hz. It is seen that both beamforming and BF fail to detect the smallest source due to a poor spatial resolution (see Fig. 3). Although the MMCG with $p = 1$ (i.e. ℓ_1-norm penalization) can correctly identify the two sources, better results are achieved by the sparser priors MMCG with $p = 0$ and the MCS (note a small bias on the localization of the small source). The second experiment investigates another extreme case with a very high frequency $f = 10$ kHz and the same two unequal source strengths. Beamforming and BF both fail in this case (see Fig. 3). The MMCG prior with $p = 1$ clearly improves the situation, yet GMCG ($p = 0$) and MCS still return the best estimates with the deepest dynamic range achieved by the latter.

4 Broadband Bayesian Focusing (BBF)

So far, a different AF has been assumed at each frequency f. Nevertheless, it is meaningful to consider instances where the AF is independent of frequency, as would typically occur when the sound source is attached to a specific device (e.g. an obstruction, an opening, etc.). This situation has been considered for example in acoustic imaging in Refs. [8, 20] and is reminiscent of the concept of "multi-frequency" group sparsity, where the same sparse structure is enforced to the source distribution for a group of frequencies.

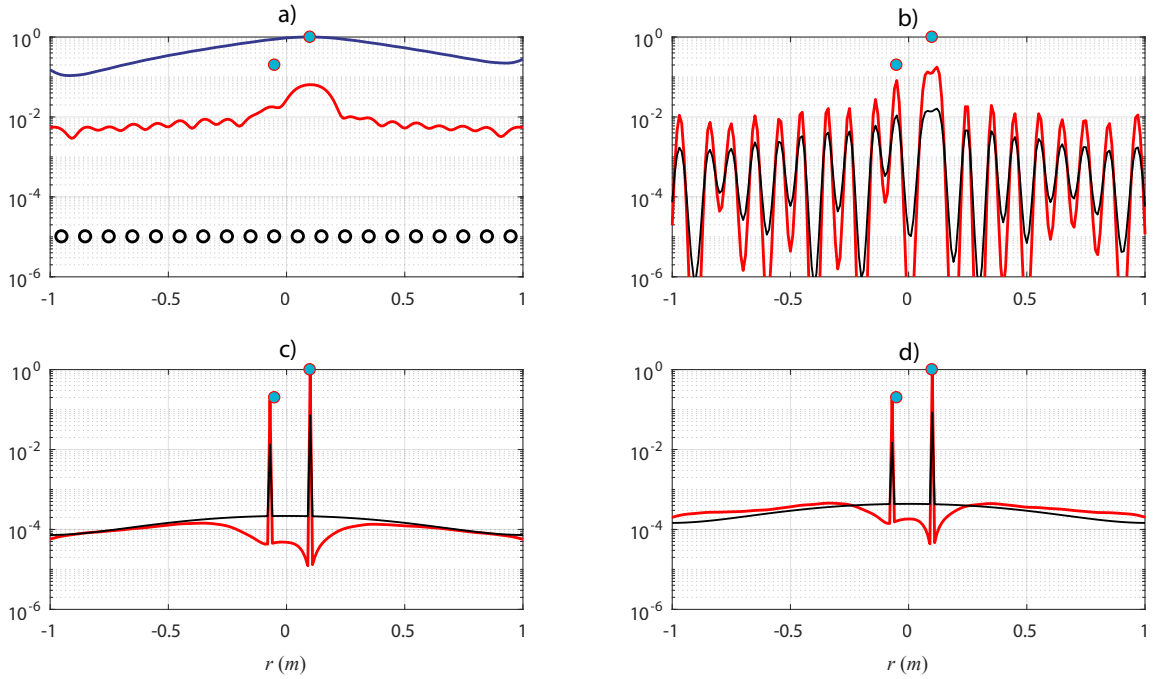

Figure 2: Estimated source power spectrum S_q^2 at at $f = 100$ Hz with $N = 100$ snapshots in the case of two incoherent monopoles with 14 dB difference in levels. The positions of the microphones is marked by black circles in (a). a) Results returned by conventional beamforming (blue) and Bayesian Focusing (red). Results returned by Iterated Bayesian Focusing with b) an GMCG prior with $p = 1$, c) an GMCG prior with $p = 0$ and d) an MCS prior ($a = b = 0.01$). The red curves in (b-d) relate to the source power spectrum S_q^2 and the black curves to the updated aperture function σ_q^2.

Forcing a constant AF in a frequency band is straightforward within the proposed IBF framework. It is also found to yield improved results. The algorithm proceeds as follows.

Algorithm 3

- Define a frequency band $\mathscr{B} = \{f_1, ..., f_B\}$ as a collection of frequency bins (as typically returned by the discrete Fourier transform) where the self-tunning AF is invariant.
- Apply Algorithm 1 where
  - steps 0 and 3 are unchanged,
  - step 2 is modified with

$$\overline{\chi^2}_{q,i,[k-1]} = \frac{\sum_{k=1}^{B} |\hat{s}_{q,i,[k-1]}(f_k)|^2}{\alpha^2_{[k-1]} \sigma^2_{0,q,[k-1]}} \tag{25}$$

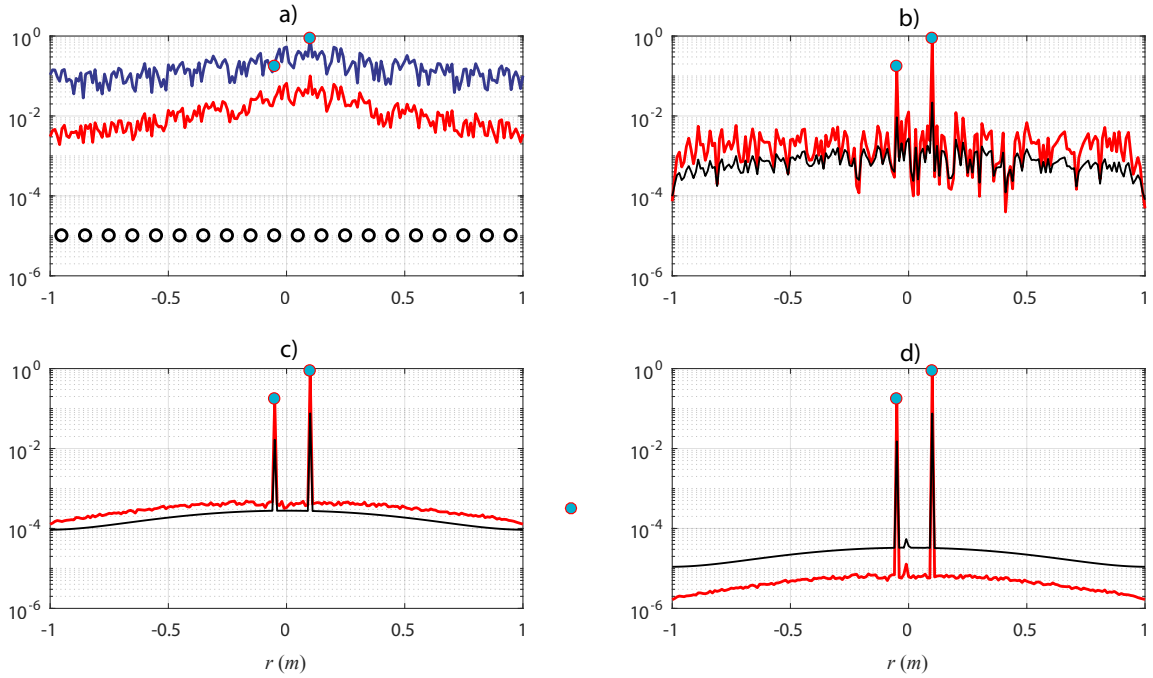

Figure 3: Same configuration as in Fig. 3 at $f = 10$ kHz.

substituted for $\chi^2_{q,[k-1]}$, where $\hat{s}_{q,i,[k-1]}(f_k)$ stands for the evaluation of $\hat{s}_{q,i,[k-1]}$ at frequency f_k and snapshot i,

– steps 1, 4 and 5 are unchanged but performed for *each* frequency f_k in band $\mathscr{B}$.

Equation (25) is the main modification to address the broadband case. The quantity $\overline{\chi^2}_{q,i,[k-1]}$ may be interpreted as the normalized energy of the source coefficients – as measured by their power spectra – summed in band $\mathscr{B}$. A resolution in terms of the CSM is also trivially implemented by considering the quantity $\sum_{i=1}^{N} \overline{\chi^2}_{q,i,[k-1]}/N$ averaged over N snapshots in Eq. (25) and proceeding as in Algorithm 2. Eventually, it is noteworthy that a different regularization parameter is reasonably assumed at each frequency.

BBF is illustrated here for the reconstruction of four uncorrelated monopole sources driven by white noises in a wide frequency interval $\mathscr{B}$ ranging from 100 Hz to 20 kHz. The sources are located at $r_1 = -0.1$, $r_2 = 0$, $r_3 = 0.1$ and $r_4 = 0.2$ m on the source domain $\Gamma_s = \{r : -1 \leq r \leq 1\}$ (in meters) and have decreasing levels in the proportion 1, 0.25, 0.04 and 0.01 (i.e. 0dB, -6dB, -14dB, -20dB). The array is the same as in Example 3 and the range distance is $R = 0.1$ m. The SNR is set to 20 dB. BBF is run with the CSM computed on 100 snapshots and a spatial resolution of 10 Hz is used (thus leading to $B = 1991$ frequency bins). The results of conventional beamforming, BF, IBF and BBF are compared in Fig. 5. It is again checked that beamforming and BF both suffer from 1) a poor spatial resolution which makes the localization of the sources difficult below 2 kHz and 2) a limited dynamic range which prevents the two smallest sources from being identified. On the one hand, whereas IBF greatly upgrades the results, it is limited

downward to 1 kHz for the reconstruction of the strongest source and to 2 kHz for the smallest source. The identification of the two small sources is actually intermittent and the smallest one is missed at most frequencies. On the other hand, BBF completely fixes these errors. The accuracy in terms of localization and quantification is excellent over the full frequency range and for all sources. This is because a frequency-independent AF is better estimated from an "average" of the map in Fig. 5(c), which in turns improves the estimation of the sources at low frequencies and low SNRs.

Concerning computational time, the BBF algorithm was actually found faster than repeating the IBF for all frequencies in the scrutinized band. As previoulsy explained, this is because accounting for all frequencies at once makes the convergence of the iterations faster (convergence was achieved after 7 iterations for BBF while each run of the IBF required about twice as many iterations on the average).

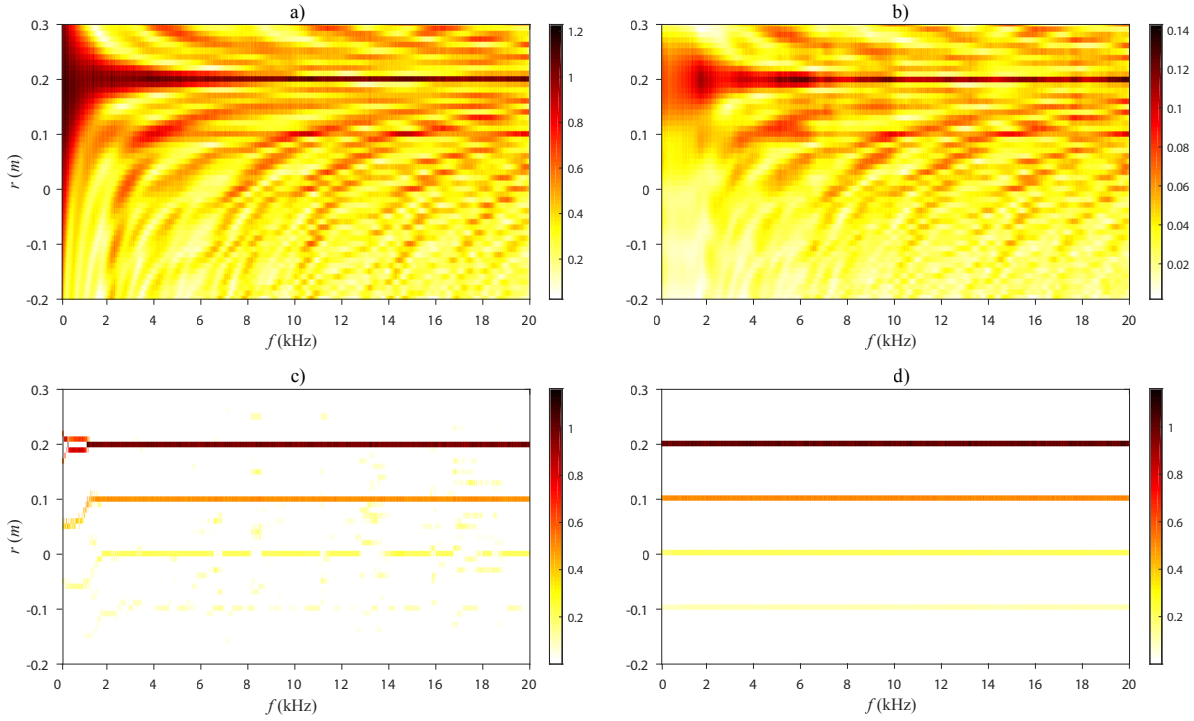

Figure 4: Estimated source distribution on the 1D domain $\Gamma_s = \{r : -0.2 \leq r \leq 0.3\}$, as a function of frequency, when the theoretical distribution comprises four monopoles placed at $r = 0.2$ m, $r = 0.1$ m, $r = 0$ m and $r = -0.1$ m, driven by uncorrelated white Gaussian signals with levels in the proportion 1, 0.25, 0.04 and 0.01, respectively, and SNR of 20 dB. Estimates returned by a) BF with a fixed uniform aperture function, b) conventional beamforming, c) IBF with an MCS prior, and d) broadband IBF with an MCS prior.

5 CONCLUSIONS

This paper has demonstrated how sparse acoustical holography is naturally achieved by simply iterating the Bayesian Focusing method previously proposed in Ref. [2]. The principle consists in using the estimated sources in the current iteration to update the aperture function in the next iteration. Formally, this is equivalent to considering the aperture function as a random quantity (endowed with a prior probability density function) that is estimated conjointly with the sources. Not only does this point of view provide physical insight into the mechanism of sparse acoustical holography – and in particular why it can considerably improve the estimation of source quantification and directivity high in frequency – but it also allows direct extension to different scenarios. For instance, the method is easily formulated in terms of the cross-spectral matrix and adapted to the broadband case (i.e. frequency group sparsity) where source positions are assumed stationary in space. The method also allows complete flexibility in the choice of the sparsity penalty. The multivariate complex Student$-t$ prior appears to be a good candidate, which returns sparser results than the typical ℓ_1-norm approach while at the time offering stable results. A last but not least advantage is that Iterated Bayesian Focusing inherits the possibility of automatic setting of the regularization parameter, an issue which has remained critical in sparse methods.

Overall, it is believed that Iterated Bayesian Focusing provides a comprehensive view of sparse acoustical holography that might not be shared by other approaches.

References

[1] N. M. Abusag and D. J. Chappell. "Near-field acoustic holography using the method of superposition." *Journal of Computational Acoustics*, 24, 1–22, 2016.

[2] J. Antoni. "A bayesian approach to sound source reconstruction: Optimal basis, regularization, and focusing." *The Journal of the Acoustical Society of America*, 131(4), 2873–2890, 2012.

[3] S. D. Babacan, R. Molina, and A. K. Katsaggelos. "Bayesian compressive sensing using laplace priors." *IEEE Transactions on Image Processing*, 19(1), 53–63, 2010. ISSN 1057-7149.

[4] D. R. Brillinger. *Time Series: Data Analysis and Theory*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001. ISBN 0-89871-501-6.

[5] F. Champagnat and J. Idier. "A connection between half-quadratic criteria and EM algorithms." *IEEE Signal Processing Letters*, 11(9), 709–712, 2004. ISSN 10709908.

[6] G. Chardon, L. Daudet, A. Peillot, F. Ollivier, N. Bertin, and R. Gribonval. "Near-field acoustic holography using sparse regularization and compressive sampling principles." *The Journal of the Acoustical Society of America*, 132(3), 1521–1534, 2012.

[7] E. Fernandez-Grande, A. Xenaki, and P. Gerstoft. "A sparse equivalent source method for near-field acoustic holography." *The Journal of the Acoustical Society of America*, 141(1), 532–542, 2017. ISSN 0001-4966.

[8] K. L. Gemba, S. Nannuru, P. Gerstoft, and W. S. Hodgkiss. "Multi-frequency sparse bayesian learning for robust matched field processing." *The Journal of the Acoustical Society of America*, 141(5), 3411–3420, 2017.

[9] P. Gerstoft, A. Xenaki, and C. F. Mecklenbräuker. "Multiple and single snapshot compressive beamforming." *The Journal of the Acoustical Society of America*, 138(4), 2003–2014, 2015.

[10] I. F. Gorodnitsky and B. D. Rao. "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm." *IEEE Transactions on Signal Processing*, 45(3), 600–616, 1997. ISSN 1053587X.

[11] J. Hald. "Fast wideband acoustical holography." *The Journal of the Acoustical Society of America*, 139(4), 1508–1517, 2016.

[12] J. Huang and T. Zhang. "The benefit of group sparsity." *Ann. Statist.*, 38(4), 1978–2004, 2010. doi:10.1214/09-AOS778.

[13] S. Ji, Y. Xue, and L. Carin. "Bayesian compressive sensing." *IEEE Transactions on Signal Processing*, 56(6), 2346–2356, 2008. ISSN 1053-587X.

[14] Q. Leclère, A. Pereira, C. Bailly, J. Antoni, and C. Picard. "A unified formalism for acoustic imaging based on microphone array measurements." *International Journal of Aeroacoustics*, 16(4-5), 431–456, 2017.

[15] M. Novey, T. Adali, and A. Roy. "A complex generalized gaussian distribution – characterization, generation, and estimation." *IEEE Transactions on Signal Processing*, 58(3), 1427–1433, 2010. ISSN 1053-587X.

[16] B. Oudompheng, A. Pereira, C. Picard, and B. Nicolas. "A theoretical and experimental comparison of the iterative equivalent source method and the generalized inverse beamforming." *Proceedings of the 5th Berlin Beamforming Conference*, pages 1–15, 2014.

[17] A. Pereira. "Acoustic imaging in closed spaces." *Ph.D. thesis, University of Lyon*, 2013.

[18] A. Pereira, J. Antoni, and Q. Leclère. "Empirical bayesian regularization of the inverse acoustic problem." *Applied Acoustics*, 97(Supplement C), 11 – 29, 2015. ISSN 0003-682X.

[19] A. Pereira and Q. Leclére. "A theoretical and experimental comparison of the equivalent source method and a bayesian approach to noise source identification." *Proceedings of Berlin Beamforming Conference, Berlin 2012*, pages 1–12, 2012.

[20] Y. Wang, J. Li, P. Stoica, M. Sheplak, and T. Nishida. "Wideband relax and wideband clean for aeroacoustic imaging." *The Journal of the Acoustical Society of America*, 115(2), 757–767, 2004.