BeBeC-2020-D18

A FAST AND HIGH-RESOLUTION BEAMFORMING BASED ON BAYESIAN COMPRESSED SENSING

Ning CHU[1], Yue NING[1], Liang YU[2], Dazhuan WU[1]
[1]College of Energy Engineering, Zhejiang University
Zheda road 38, 310027, Hangzhou, China
[2]State Key Laboratory of Mechanical System and Vibration
School of Mechanical Engineering, Shanghai Jiao Tong University
Dongchuan Road 800, Shanghai 200240, P.R. China

ABSTRACT

Nowadays acoustic source localization and imaging by using microphone arrays has been widely studied and applied in speech tracking, mechanical noise diagnosis and aeroacoustic quality evaluation, etc. Conventional beamforming method is a practical and robust tool that can quickly obtain the position and strength of acoustic sources based on phased microphone array, but imaging resolution is very limited for middle frequencies in low SNR. Though compressed sensing technique can significantly improve the resolution by adding sparse regularisation, but it hardly performs quite well in low Signal-to-Noise Ratio. Therefore, this paper proposes a Bayesian compressed sensing beamforming to overcome the above challenges. A student-t prior is employed to enforce the sparse distribution of acoustic monopole source. And background noise can be adaptively attenuated by modelling the conjugate prior distribution. The hyper-parameter estimation is not very time-consuming thanks to variational Bayesian approximation. Through simulated and experimental data, our proposal is validated to achieve fast and super-resolution imaging in low SNR and middle frequencies.

1 INTRODUCTION

Acoustic source localization is a key technology that is widely used in many fields, such as aeronautics, non-destructive examination and vehicle manufacture. Through the analysis of acoustic pressure signals measured by phased microphones, spatial position and energy distribution can be obtained for an acoustic map. Several localization approaches have been developed over last decades to make acoustic imaging more accuracy and robust[1, 2]. The beamforming is the most universal, robust and simplest method. Nevertheless, the result obtained by beamforming is always blurred by the sidelobes, which can be interpreted as the convolution effect by the Point Spread Function (PSF) of microphone arrays. In order to overcome the beamforming limitation caused by PSF, the deconvolution and regularization approaches are proposed to improve the resolution of acoustic imaging and accuracy of source strengths. Representative algorithms include Deconvolution Approach for the Mapping of Acoustic Sources (DAMAS) and compressive sensing beamforming [3, 4]. But usually these state-of-the-art methods require a lot of computing resources for iteratation and optimization.

Recently, Bayesian method provides new framework to solve such problems and some Bayesian regularization methods have been applied for acoustic imaging successfully [5, 6].

Inspired by the advantages of Bayesian methodology, in this paper, our aim is to propose a Bayesian sparsity regularization method based on the assumptions of non-stationary noise. Concretely, a Bayesian prior framework is proposed based on the physical background of acoustic distribution, then the Variational Bayesian Approximation (VBA) is implemented to solve the inverse problem in a fast and robust way.

The rest of this paper is organised as follows: In the section 2, the detail of the forward physical model is given, then VBA algorithm is proposed to solve the inverse problem. Section 3 shows the validation of proposed method by simulations and experiments. Finally, Section 4 concludes this paper and proposes further prospects.

2 PHYSICAL MODEL AND ITS BAYESIAN SOLUTION

2.1 Forward model of acoustic source localization

In this paper, planar arrays and monopole point source are used for theoretical analysis and experimental verification. Supposing array plane with M microphones and source plane with K monopoles respectively, and these two planes are face to face. The acoustic pressure received by M microphones can be expressed as follows:

$$\boldsymbol{Z} = \boldsymbol{A}\boldsymbol{S} + \boldsymbol{E} \tag{1}$$

where $\boldsymbol{S} = (\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_K)^T \in \mathcal{R}^{K\times L}$ is the acoustic pressure at acoustic source points, $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_K) \in \boldsymbol{\mathcal{C}}^{M\times K}$ is the propagation vector that reflect the transmission relationship between microphones and sources, whose element can be expressed by Green's function. L denotes the number of samplings and $\boldsymbol{E} \in \mathcal{R}^{M\times K}$ denotes the measurement noise. $\boldsymbol{Z} \in \mathcal{C}^{K\times L}$ are the measured pressures of array output.

Supposing that acoustic source plane is divided into N grids, and each grid may represent where the acoustic source possibly appears. A great deal of literature [1, 2, 5] prove that conventional beamforming results can be expressed as:

$$y_n = \boldsymbol{a}_n'\overline{\boldsymbol{Z}\boldsymbol{Z}'}\boldsymbol{a}_n/\|\boldsymbol{a}_n\|_2^4 \tag{2}$$

where $(\cdot)'$ denotes conjugate transpose and $\overline{(\cdot)}$ denotes average value, n is the index of N.

The imaging of conventional beamforming is a contaminated result because it seldom takes account of the relationship between different beams, which is the PSF of microphone arrays. The deconvolution and regularization methods take this into consideration and get a clearer image. After computing the PSF of different scanning grids and mapping PSF into a measurement matrix $\boldsymbol{H} \in \mathcal{R}^{N\times N}$, then the forward model of convolution can be expressed as compressed sensing model as follows:

$$\boldsymbol{y} = \boldsymbol{H}\,\boldsymbol{x} + \boldsymbol{e} \tag{3}$$

where measurement matrix $\boldsymbol{H}$ has the item $h_{i,j} = (\boldsymbol{a}_i'\boldsymbol{a}_j)^2/\|\boldsymbol{a}_i\|_2^4$; $\boldsymbol{y} = (y_1, y_2, \cdots, y_N)^T \in \mathcal{R}^{N\times 1}$ is the result of conventional beamforming, and $\boldsymbol{e} \in \mathcal{R}^{N\times 1}$ is measured noise. Hereafter, the research target is to solve an inverse problem ill-posed by Eq. (3) so as to get accuracy imaging solution $\boldsymbol{x} = (x_1, x_2, \cdots, x_N)^T \in \mathcal{R}^{N\times 1}$.

2.2 VBA approach for the solution of inverse problem

2.2.1 Conjugate prior settings

Different from deconvolution and regularization methods, Bayesian approach assigns proper prior distributions for the unknown values based on physical characteristics of research objects. Considering the non-stationary and sparsity characteristics of acoustic monopole distrubtion, we assign Student-t prior on $\boldsymbol{e}$ and $\boldsymbol{x}$ [7]. Thanks to the Infinite Gaussian Scaled Mixture (IGSM) property of Student-t distribution as follows:

$$\mathcal{St}(x_j|\nu) = \int_0^\infty \mathcal{N}(x_j|0, 1/u_j)\mathcal{G}(u_j|\nu/2, \nu/2)d\,u_j \tag{4}$$

where $\mathcal{N}$ and $\mathcal{G}$ denote Gaussian and Gamma distribution respectively. u_j is the inverse of variance of Gaussian distribution. ν is degrees of freedom for Student-t distribution. $\mathcal{IG}(v|\alpha,\beta)$ denotes inverse Gamma distribution, whose Probability Density Function (PDF) is $\mathcal{IG}(v|\alpha,\beta) = \beta^\alpha v^{-\alpha-1}e^{-\beta/v}/\Gamma(\alpha)$. $\Gamma(\alpha)$ denotes Gamma function. Therefore, the following hierarchical model is proposed:

$$p(\boldsymbol{x}|v_f) = \mathcal{N}(\boldsymbol{x}|0, \boldsymbol{v}_x),\ \ p(\boldsymbol{v}_x) = \prod_{j=1}^{N} p\left(v_{x_j}\right) = \prod_{j=1}^{N} \mathcal{IG}\left(v_{x_j}\middle|\alpha_{x_0}, \beta_{x_0}\right) \tag{5}$$

Similar to $\boldsymbol{x}$, noise $\boldsymbol{e}$ has the similar prior structure as follows:

$$p(\boldsymbol{e}|v_e) = \mathcal{N}(\boldsymbol{e}|0, \boldsymbol{v}_e),\ \ p(\boldsymbol{v}_e) = \prod_{i=1}^{N} p(v_{e_i}) = \prod_{i=1}^{N} \mathcal{IG}(v_{e_i}|\alpha_{e_0}, \beta_{e_0}) \tag{6}$$

Under such prior conditions, the joint posterior probability becomes as:

$$p(\boldsymbol{x}, \boldsymbol{v}_x, \boldsymbol{v}_e|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{v}_e)\, p(\boldsymbol{x}|\boldsymbol{v}_x)\, p(\boldsymbol{v}_x|\alpha_{x_0}, \beta_{x_0})\, p(\boldsymbol{v}_e|\alpha_{e_0}, \beta_{e_0}) \tag{7}$$

Then the joint posterior probability in Eq.(7) can be solved by proposed VBA approach in the following subsections.

2.2.2 Posterior solution by VBA

Normally the formula deduced by Eq. (7) can be solved in two general methods at least, Joint Maximum A posterior (JMAP) and VBA. However, it's not easy to apply JMAP to solve Eq. (7) because of large number of hyper-parameters and non-linear inversion. Considering the universality of the solution, the VBA is used for approximating posterior solution under the Kullback-Leibler (KL) divergence. The core of VBA is to separate the joint posterior into the products of separate probability distributions for each type of variables:

$$p(\boldsymbol{x}, \boldsymbol{v}_x, \boldsymbol{v}_e|\boldsymbol{y}) \propto q_1(\boldsymbol{x})q_2(\boldsymbol{v}_x)q_3(\boldsymbol{v}_e) \tag{8}$$

The most important process is to minimize the KL divergence $KL(q:p)$ to optimize the combination of $q_1(\boldsymbol{x})$, $q_2(\boldsymbol{v}_x)$ and $q_3(\boldsymbol{v}_e)$. To make full use of the analytical properities of conjugate prior, Gaussian distribution and Gamma distribution are preferred to choose:

$$\begin{cases} q_1(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_x, \boldsymbol{v}_x) \\ q_2(\boldsymbol{v}_x) = \prod_j \mathcal{IG}\left(v_{x_j}\middle|\alpha_{x_j}, \beta_{x_j}\right) \\ q_3(\boldsymbol{v}_e) = \prod_i \mathcal{IG}(v_{e_i}|\alpha_{e_i}, \beta_{e_i}) \end{cases} \tag{9}$$

In the above formula, the index of covariance matrix $\boldsymbol{v}_x$ represents the diagonal elements, so is $\boldsymbol{v}_e$. The covariance $\boldsymbol{v}_e$ and $\boldsymbol{v}_x$ are initialized as unit diagonal matrices, other hyperparameters can be initialized with random numbers. Owing to the conjugate prior, the joint posterior has the same style of probability distribution as the prior distribution. The iterative process can be expressed as follows:

$$\begin{cases} \alpha_{x_j} = \alpha_{x_0} + 0.5 \\ \beta_{x_j} = \beta_{x_0} + 0.5 * \boldsymbol{x}_j{}^2 \\ \|\boldsymbol{x}\|^2 = \|\boldsymbol{\mu}_x\|^2 + tr(\boldsymbol{v}_x) \\ v_{x_j} = \dfrac{\beta_{x_j}}{\alpha_{x_j} - 1} \\ \alpha_{e_i} = \alpha_{e_0} + 0.5 \\ \beta_{e_i} = \beta_{e_0} + 0.5 * \boldsymbol{e}_j{}^2 \\ \|\boldsymbol{e}\|^2 = \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{\mu}_x\|^2 + tr(\boldsymbol{H}\boldsymbol{v}_x\boldsymbol{H}') \\ v_{e_i} = \dfrac{\beta_{e_i}}{\alpha_{e_i} - 1} \end{cases} \tag{10}$$

Symbol $tr(\cdot)$ denotes the trace of matrix. After alternate optimization, VBA solution turns out as:

$$\begin{cases} \boldsymbol{v}_x = (\boldsymbol{H}'\boldsymbol{v}_e{}^{-1}\boldsymbol{H} + \boldsymbol{v}_x{}^{-1})^{-1} \\ \boldsymbol{\mu}_x = \boldsymbol{v}_x\boldsymbol{H}'\boldsymbol{v}_e{}^{-1}\boldsymbol{y} \end{cases} \tag{11}$$

where the super-parameters decide the form of Student-t distribution, especially control the tail thickness. The longer the tail is, the more sparsity-enforcing the Student-t prior behaves, and the larger the dynamic range of reconstructed sources performs

3 VALIDATION ON SIMULATION AND EXPERIMENTAL DATA

In this section, numerical simulation and experimental data are both implemented to validate the proposed VBA approach. The spiral array with 56 channels is used to measure acoustic source at 2500Hz. The experimental conditions are kept close to the simulation, which are specified in Table 1.

Table 1. Parameter settings of numerical simulation and real experiment.

Configurations:	**Simulation / Experiment**
Signal types	cyclostationary acoustic sources / single-frequency at 2500Hz
Sampling frequency	50000Hz(fs)
Number of microphones	56(M)
Number of acoustic sources	4(K) in a cross form
Source intensities	83.52dB (top), 86.02dB (bottom), 87.96dB (right), 89.54dB (left)
Duration of sampling	1s (that is L=50000)
Resolution of acoustic imaging	25×25=625(N)
Distance between acoustic sources and array	1.5m / 0.75m
Area of single grid	0.5m×0.5m=0.25m^2
Speed of acoustic propagation	340m/s

3.1 Validation on numerical simulation

To make the simulation closer to experiments, 4 nonstationary monopole sources are arranged in a cross and non-Gaussian noise is added to measurements. Signal to Noise Ratio (SNR) is set as low as -5dB. The nonstationary source is modelled by cyclic amplitude modulation signal in Eq. (11):

$$s(t) = \left[1 + \sum b_i \cos 2\pi f_i t\right] \cdot c(t) \tag{12}$$

where $s(t)$ is a random but cyclostationary signal emitting by each monopole related to Eq.(1); $c(t)$ is the carrier which has wide spectrum, for example Gaussian process; f_i is the discrete frequency component which consists of the envelope spectrum of $s(t)$; b_i is the amplitude of discrete frequency component. The left picture of Fig. 1 is the patten of microphone array. Sound Pressure Level (SPL) and positions of acoustic sources are shown in right picture of Fig. 1.

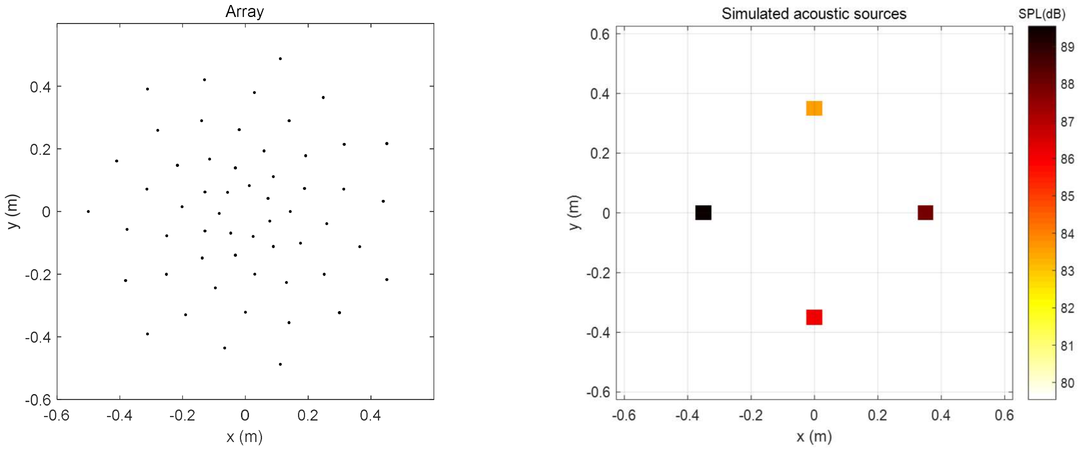

Fig. 1. Spiral array with 56 channels (left), distribution of simulated acoustic sources (right).

Then the beamforming result and proposed VBA result are shown in Fig. 2 respectively. Although both methods have uncertainty in the reconstruction of source intensity, proposed VBA method has greatly improvements in terms of positions and intensities of acoustic sources, compared to conventional beamforming. Reasonable sparse prior provides more physical information for solving inverse problem and enhances the robustness of the solution.

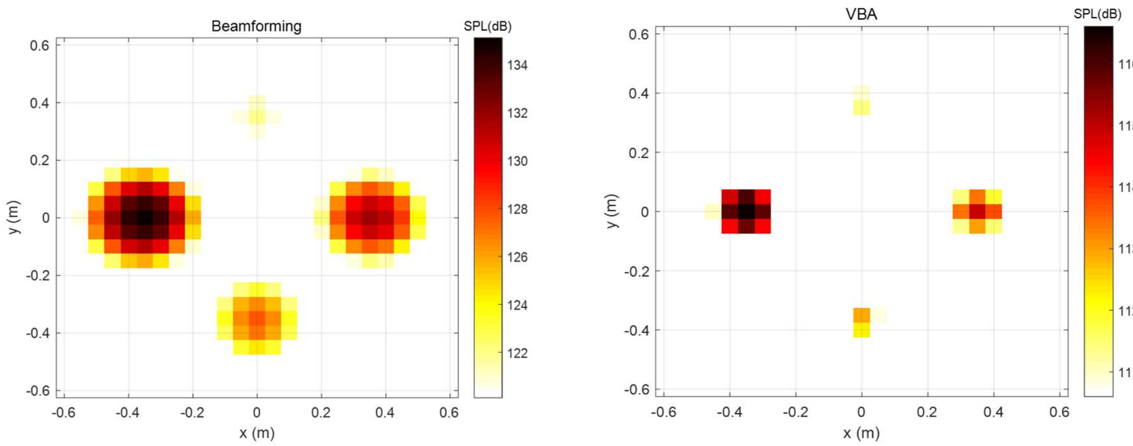

Fig. 2. Beamforming result (left) and purpose VBA result (right) obtained by simulation at 2500Hz.

Actually, energy propagation model expressed by Eq. (3) can be represented by a convolution formation. Assuming that there is a spatially invariant kernel $\boldsymbol{h}_0$ and $\boldsymbol{x}_0$ is source-power image with matrix form, then the relationship between energy propagation model and spatially invariant convolution model is expressed as follows:

$$\boldsymbol{H}\,\boldsymbol{x} \approx \boldsymbol{h}_0 * \boldsymbol{x}_0 \tag{13}$$

Where ($*$) denotes valid convolution, so that the output matrix after convolution is the same size as input matrix. ($\approx$) denotes the approximation in the sense of each item from left side (vector form) being approximated to the corresponding one from the right side (matrix form). Under the simulation conditions in Table 1, the PSF can be depicted as:

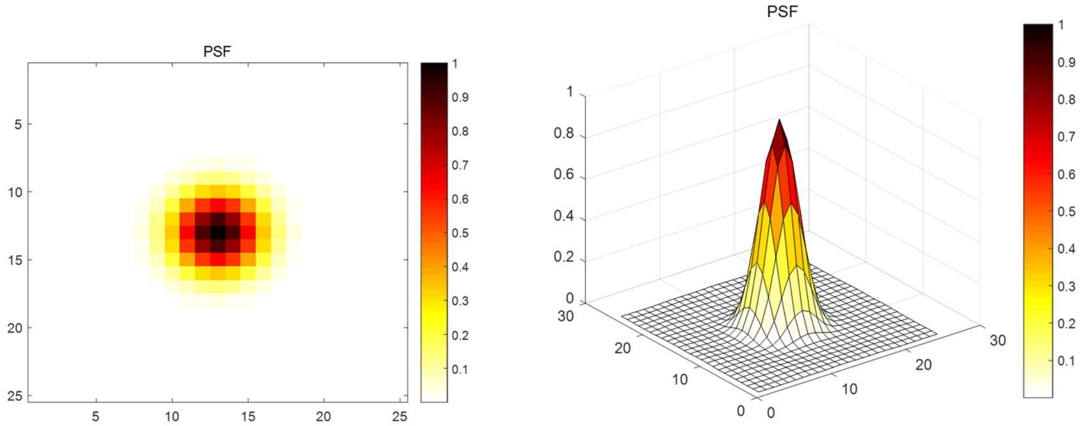

Fig. 3. 2D image and 3D stereo image of spatially invariant PSF under simulation conditions (seen in Table 1).

The proposed method in this paper is able to achieve the deconvolution of blurred map through VBA-based sparse regularization.

3.2 Validation on experimental data

Experiment is implemented in anechoic chamber under the same conditions as the simulation conditions in Table 1 for sinusoidal acoustic monopoles. The acoustic sources are distributed in a cross shape and the centre of the sources plane is aligned with the centre of the array. As shown in Fig. 4, the distance between diagonal sources is 0.75 meters.

Fig. 4. Practical experiment scenario in an anechoic chamber, the left is microphone array, the right is four loud-speakers .

Figure. 5 shows the blurred map reconstructed by beamforming and a clean-up map reconstructed by proposed VBA. Even though the intensity of 4 sources is different from each other, after adding the sparsity prior as sparse regularization, the energy characteristics of each source are enhanced by our proposal. Moreover, imaging resolution is greatly improved owing to sidelobe reduction. In addition, compared to deconvolution such as DAMAS, proposed VBA is faster and more efficient due to its analytical iteration. As for the sparse regularization approaches, there is less artifact originated by proposed VBA.

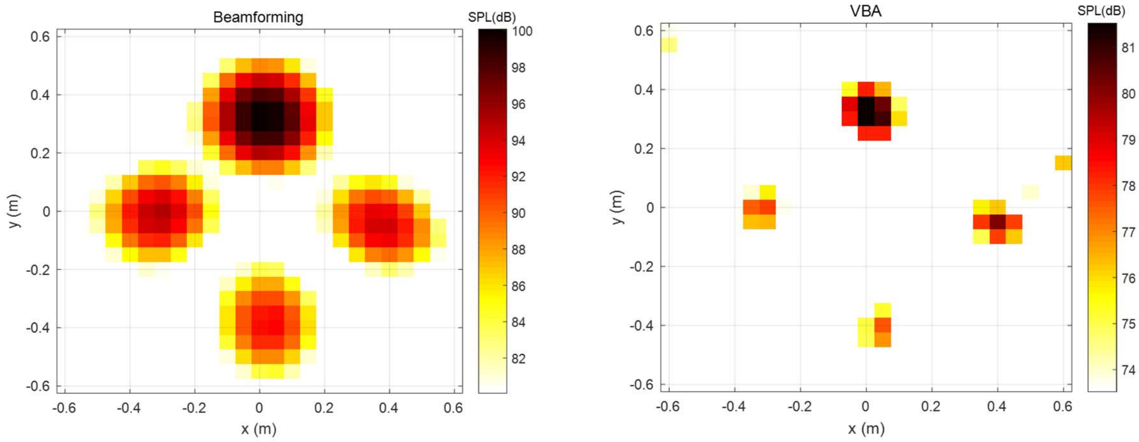

Fig. 5. Beamforming result (left) and proposed VBA (right) obtained by experimental data at 2500Hz.

4 CONCLUSIONS

This paper transforms the beamforming model into compressed sensing problem, nd proposes VBA method to solve this problem efficiently. The application of VBA provides fast and robust solution for the posterior probability problem based on Student-t prior and its conjugate pair of background noise. Numerical simulation with nonstationary sources and experiments with single-frequency acoustic sources are given to validate the proposed method.

From a profound perspective, different acoustic signals should correspond to targeted prior distributions, especially for nonstationary sources and non-Gaussian nosie, and there are expected to exploit different VBA approaches to match various prior frameworks. Therefore, it is necessary to explore the proper priors of different signals related to acoutical physics, then take full advantages of Bayesian VBA inference.

ACKNOWLEDGMENT

This research is funded by Natural Science Foundation of China (Grant No. 61701440, 11704248), National Key Lab Funding of Compressor Technology (Grant No. SKL-YSL201812), Science and Technology on Sonar Laboratory (Grant No. 6142109KF201901), State Key Laboratory of Mechanical System and Vibration (Grant No. MSV202001), Science and Technology on Underwater test and Control Laboratory.

REFERENCES

[1] Merino-Martínez R, Sijtsma P, Snellen M, et al. A review of acoustic imaging methods using phased microphone arrays[J]. CEAS Aeronautical Journal, 2019, 10(1): 197-230.

[2] Chiariotti P, Martarelli M, Castellini P. Acoustic beamforming for noise source localization–Reviews, methodology and applications[J]. Mechanical Systems and Signal Processing, 2019, 120: 422-448.

[3] Dougherty R. Extensions of DAMAS and benefits and limitations of deconvolution in beamforming[C]//11th AIAA/CEAS aeroacoustics conference. 2005: 2961.

[4] Simard P, Antoni J. Acoustic source identification: Experimenting the ℓ1 minimization approach[J]. Applied Acoustics, 2013, 74(7): 974-986.

[5] Chu N, Mohammad-Djafari A, Picheral J. Robust Bayesian super-resolution approach via sparsity enforcing a priori for near-field aeroacoustic source imaging[J]. Journal of Sound and Vibration, 2013, 332(18): 4369-4389.

[6] Pereira A, Antoni J, Leclere Q. Empirical Bayesian regularization of the inverse acoustic problem[J]. Applied Acoustics, 2015, 97: 11-29.

[7] Mohammad-Djafari A, Dumitru M. Bayesian sparse solutions to linear inverse problems with non-stationary noise with Student-t priors[J]. Digital Signal Processing, 2015, 47: 128-156.