Research Article

Improving the Sound Source Identification Performance of Sparsity Constrained Deconvolution Beamforming Utilizing SFISTA

Linbang Shen,[1,2] Zhigang Chu [ID],[1] Long Tan,[2] Debing Chen,[2] and Fangbiao Ye[2]

[1]*School of Automotive Engineering, Chongqing University, Chongqing 400044, China*
[2]*Chongqing Vehicle Test & Research Institute Co., Ltd., Chongqing 401120, China*

Correspondence should be addressed to Zhigang Chu; zgchu@cqu.edu.cn

In this paper, an alternative sparsity constrained deconvolution beamforming utilizing the smoothing fast iterative shrinkage-thresholding algorithm (SFISTA) is proposed for sound source identification. Theoretical background and solving procedures are introduced. The influence of SFISTA regularization and smoothing parameters on the sound source identification performance is analyzed, and the recommended values of the parameters are obtained for the presented cases. Compared with the sparsity constrained deconvolution approach for the mapping of acoustic sources (SC-DAMAS) and the fast iterative shrinkage-thresholding algorithm (FISTA), the proposed SFISTA with appropriate regularization and smoothing parameters has faster convergence speed, higher quantification accuracy and computational efficiency, and more insensitivity to measurement noise.

1. Introduction

Beamforming [1–3] based on a microphone array has become a popular sound source identification technology for aircraft [4], express train [5], wind turbine [6], automobile [7], etc. Conventional beamforming (CB) suffers from a poor spatial resolution at low frequency and plenty of spurious sources at high frequency [8–10]. To overcome these issues, various deconvolution beamforming techniques with different solving algorithms were developed, such as deconvolution approach for the mapping of acoustic sources (DAMAS) [11], nonnegative least square (NNLS) [12], and Richardson–Lucy (RL) [12] and their corresponding fast Fourier transform- (FFT-) based variants: DAMAS2 [13], FFT-NNLS [12], and FFT-RL [12]. In 2015, on the basis of the iterative shrinkage-thresholding algorithm (ISTA) [14] and the fast iterative shrinkage-thresholding algorithm (FISTA) [15], which are used to solve the inverse problem in the image processing, Lylloff et al. [16] proposed FFT-FISTA deconvolution beamforming for sound source identification. Compared to FFT-NNLS, FFT-

FISTA has higher computational efficiency and a better convergence rate. In addition, in the discussion of Ref. [16], it was suggested to extend the capabilities of FFT-FISTA to include a sparsity constraint on the solution to see whether the efficiency could be further improved. However, the proximal operator in FISTA does not have a closed-form solution when solving the sparse recovery problem. This makes it difficult for FISTA to introduce sparse constraints directly and explicitly. To overcome this difficulty, Zhao et al. [17] recently proposed the smoothing fast iterative shrinkage-threshold algorithm (SFISTA), which enjoys the advantage of quickly processing the large-scale problems in the compressive sensing framework. To the authors' knowledge, SFISTA has not yet been successfully adapted to deconvolution beamforming to enhance sound source identification performance so far. In addition, several similar deconvolution beamforming techniques successfully include the sparse distribution constraint of sound source, such as sparsity constrained DAMAS (SC-DAMAS) [18], robust super-resolution approach with sparsity constraint (SC-RDAMAS) [19], and orthogonal matching pursuit DAMAS

(OMP-DAMAS) [20]. SC-DAMAS and SC-RDAMAS are solved by the CVX toolbox [21], and the calculation speed is slow. OMP-DAMAS usually requires a priori information about the number of sound sources to obtain good sound source identification performance.

Inspired by Refs. [16, 17], this paper proposes a SFISTA deconvolution beamforming, which includes the sparsity constraint that the main sound sources are usually sparsely distributed. The proposed approach bypasses the priori information about the number of sound sources. Compared to the SC-DAMAS and FISTA, the proposed approach enjoys faster convergence speed, higher quantification accuracy and computational efficiency, and more insensitivity to measurement noise.

The remainder of this paper is organized as follows. Section 2 establishes the theory of SFISTA deconvolution beamforming for sound source identification. Sections 3 and 4 compare the performance of deconvolution beamforming utilizing SC-DAMAS, FISTA, and SFISTA by simulation and experiment, respectively. Section 5 concludes this paper.

2. Theory

The beamforming based on cross-spectral imaging function is a very common method for sound source identification, and it is as follows [22]:

$$b(\mathbf{r}) = \frac{1}{M} \frac{\mathbf{v}^{\mathrm{T}}(\mathbf{r})\mathbf{C}\mathbf{v}^*(\mathbf{r})}{\sqrt{\mathbf{w}^{\mathrm{T}}(\mathbf{r})\mathbf{1}\mathbf{w}^*(\mathbf{r})}}, \tag{1}$$

where $\mathbf{r}$ indicates the position of the focus point where the assuming acoustic source is positioned, $\mathscr{C} \in \mathbb{C}^{M \times M}$ is the cross-spectral matrix of the sound pressure signals perceived by array microphones, M is the number of microphones, $\mathbf{1} \in \mathscr{R}^{M \times M}$ is a matrix with all elements equal to 1, $\mathbf{v}(\mathbf{r}) = [v_1(\mathbf{r}), v_2(\mathbf{r}), \ldots, v_m(\mathbf{r}), \ldots, v_M(\mathbf{r})]^{\mathrm{T}}$ is the steering vector, $\mathbf{w}(\mathbf{r}) \equiv [|v_1(\mathbf{r})|^2, |v_2(\mathbf{r})|^2, \ldots, |v_m(\mathbf{r})|^2, \ldots, |v_M(\mathbf{r})|^2]^{\mathrm{T}}$, and the superscript "T" and "∗" represent the transpose and the conjugate operator, respectively. $v_m(\mathbf{r})$ is defined as

$$v_m(\mathbf{r}) = \frac{e^{-ik|\mathbf{r}-\mathbf{r}_m|}}{|\mathbf{r}-\mathbf{r}_m|}, \tag{2}$$

where $k = 2\pi f/c$ is the wave number, f is the frequency, c is the sound speed, $\mathrm{i} = \sqrt{-1}$, and $\mathbf{r}_m$ indicates the position of the mth microphone, $m = 1, 2, \ldots, M$ is the index number of microphones.

In the case that the acoustic source is incoherent, the output of beamforming can be expressed in the following linear equation in matrix form:

$$\mathbf{b} = \mathbf{Ax} + \mathbf{n}, \tag{3}$$

where $\mathbf{x} = [x(\mathbf{r}')] \in \mathbb{C}^{N \times 1}$ is the unknown column vector of the sound pressure at 1 m distance from the corresponding assuming point sound source, which is used to measure the sound source strength; $\mathbf{A} = [\mathrm{psf}(\mathbf{r} \,|\, \mathbf{r}')] \in \mathbb{C}^{N \times N}$ is the known PSF matrix, in which $\mathrm{psf}(\mathbf{r} \,|\, \mathbf{r}')$ expresses the beamforming contribution of the unit-amplitude point source at $\mathbf{r}'$ to the focus point at $\mathbf{r}$ and N is the total number

of the focus points; $\mathbf{b} = [b(\mathbf{r})] \in \mathbb{C}^{N \times 1}$ is the known column vector of CB outputs; $\mathbf{n} \in \mathbb{C}^{N \times 1}$ represents the noise.

Considering that acoustic sources are usually sparsely distributed, the majority of the elements in the vector $\mathbf{x}$ are zero or approximately zero. That is, the number of the nonzero elements is far less than that of zero elements. Assuming that the $\ell 2$-norm of the noise $\mathbf{n}$ is bounded by ε, equation (3) can be formulated as

$$\begin{aligned} &\min \|\mathbf{x}\|_0 \\ &\text{subject to } \|\mathbf{b} - \mathbf{Ax}\|_2 \le \varepsilon. \end{aligned} \tag{4}$$

Alternatively, in the field of acoustics, under the restricted isometry, the above nonconvex $\ell 0$-norm can be approximated by the convex $\ell 1$-norm, leading to the following relaxed problem:

$$\begin{aligned} &\min \|\mathbf{x}\|_1 \\ &\text{subject to } \|\mathbf{b} - \mathbf{Ax}\|_2 \le \varepsilon. \end{aligned} \tag{5}$$

Equation (5) is equivalent to the following unconstrained optimization [17]:

$$\min \frac{1}{2}\|\mathbf{b} - \mathbf{Ax}\|_2^2 + \lambda\|\mathbf{x}\|_1, \tag{6}$$

where λ is the regularization parameter. Let $f(\mathbf{x}) = 1/2\|\mathbf{Ax} - \mathbf{b}\|_2^2$ and $g(\mathbf{x}) = \lambda\|\mathbf{x}\|_1$.

SFISTA solves equation (6) by smoothing the sparse constraint $g(\mathbf{x})$. The nonsmoothed $g(\mathbf{x})$ is replaced approximately by the corresponding smoothed Moreau envelope $g_\mu(\mathbf{x})$. Here, $g_\mu(\mathbf{x})$ is a continuous differentiable and the gradient of $g_\mu(\mathbf{x})$ is

$$\nabla g_\mu(\mathbf{x}) = \frac{1}{\mu}\Big(\mathbf{x} - \Gamma_{\lambda\mu}(\mathbf{x})\Big), \tag{7}$$

where $\nabla(.)$ represents the gradient, $\mu > 0$ is the smoothing parameter, and $\Gamma_{\lambda\mu}$ is the soft shrinkage operator, and it is defined as

$$\Gamma_{\lambda\mu}(\mathbf{x}) = [|\mathbf{x}| - \lambda\mu]_+ \mathrm{sgn}(\mathbf{x}), \tag{8}$$

where $[|\mathbf{x}| - \lambda\mu]_+$ denotes the vector whose components are the maximum number between $|\mathbf{x}| - \lambda\mu$ and 0 and $\mathrm{sgn}(.)$ is the sign function which returns the sign of the variable in parentheses.

Initializing $\mathbf{x}^{(0)} = \mathbf{b}$, auxiliary vector $\mathbf{y}^{(0)} = \mathbf{x}^{(0)}$, and step size $t^{(0)} = 1$. The specific steps of the lth iteration are as follows:

(1) Calculating $\nabla f(\mathbf{y}^{(l-1)})$ and $\nabla g_\mu(\mathbf{x}^{(l-1)})$:

$$\begin{aligned} \nabla f\Big(\mathbf{y}^{(l-1)}\Big) &= \mathbf{A}^*\Big(\mathbf{Ay}^{(l-1)} - \mathbf{b}\Big), \\ \nabla g_\mu\Big(\mathbf{x}^{(l-1)}\Big) &= \frac{1}{\mu}\Big(\mathbf{x}^{(l-1)} - \Gamma_{\lambda\mu}\Big(\mathbf{x}^{(l-1)}\Big)\Big), \end{aligned} \tag{9}$$

(2) Calculating $\mathbf{x}^{(l)}$:

$$\mathbf{x}^{(l)} = \mathscr{P}_+\Big(\mathbf{y}^{(l-1)} - \frac{1}{L}\Big(\nabla f\Big(\mathbf{y}^{(l-1)}\Big) + \nabla g_\mu\Big(\mathbf{x}^{(l-1)}\Big)\Big)\Big), \tag{10}$$

where $\mathcal{P}_+$ is the Euclidean projection onto the nonnegative quadrant and L is the Lipschitz constant equal to the largest eigenvalue of $\mathbf{A}^\mathrm{T}\mathbf{A}$.

(3) Calculating the step size $t^{(l)}$:

$$t^{(l)} = \frac{1 + \sqrt{1 + 4\left(t^{(l-1)}\right)^2}}{2}. \tag{11}$$

(4) Calculating $\mathbf{y}^{(l)}$:

$$\mathbf{y}^{(l)} = \mathbf{x}^{(l)} + \frac{\left(t^{(l-1)} - 1\right)}{t^{(l)}}\left(\mathbf{x}^{(l)} - \mathbf{x}^{(l-1)}\right). \tag{12}$$

3. Simulation

To determine the influence of the parameters λ and μ on the sound source identification performance of SFISTA deconvolution beamforming, a 0.65 m diameter Brüel & Kjær 36-channel sector microphone array, as shown in Figure 1, is used to conduct the simulation. The calculation plane of interest is set as 1 m × 1 m with 51 × 51 focus point. The grid space of focus points is 0.02 m. The distance between the calculation plane and the array plane is 1 m.

The point source at each focus point is considered, and its frequency varies from 2000 to 6000 Hz with a step size of 100 Hz (i.e., 2000 Hz, 2100 Hz, ..., 6000 Hz). The 1 m sound pressure level (SPL) of the point source is 100 dB, signal-noise ratio (SNR) is 20 dB, and the iteration number is 1000. The average deviation between the output of SFISTA and the theoretical value is acquired by all the source positions and frequencies, as shown in equation (13). Therein, N_f represents the number of frequency, $x(\mathbf{r}, \mathbf{r}', f)$ represents the reconstructed SPL (in dB) of focus point $\mathbf{r}$ for a certain frequency f and a certain point source at $\mathbf{r}'$, $x_e(\mathbf{r}, \mathbf{r}', f)$ represents the exact one. The deviation result is shown in Figure 2. Obviously, the smallest deviation occurs in the region where λ is less than 1 and μ is close to 1. Therefore, $\mu = 1$ and $\mu = 1000\lambda$ (corresponding to the red marker "+" in Figure 2) are used in this paper:

$$\sigma = \frac{1}{N_f}\frac{1}{N}\sum_f\sum_{\mathbf{r}'}\sqrt{\frac{1}{N}\sum_{\mathbf{r}}\left[x(\mathbf{r}, \mathbf{r}', f) - x_e(\mathbf{r}, \mathbf{r}', f)\right]^2}. \tag{13}$$

Simulations with two known uncorrelated point sources located at (−0.2, 0.2, 1) m and (0.2, 0.2, 1) m are demonstrated. Figures 3–5 are the acoustic source identification results in the SNR of 10 dB, 20 dB, and 40 dB, respectively. In each figure, submaps (a)–(d) and (e)–(h) show the mapping at 2000 Hz and 6000 Hz based on CB, SC-DAMAS, FISTA, and SFISTA, respectively. The iteration number for FISTA and SFISTA is 1000. The outputs are normalized to dB by referring to the respective maximum value, and the display range is 15 dB, i.e., [−15, 0] dB. Both submaps (a) and (e) in Figures 3–5 indicate that the sound sources are accurately located as the hot spots at (−0.2, 0.2, 1) m and (0.2, 0.2, 1) m. For CB, sound sources are fused with each other at 2000 Hz due to poor spatial resolution at relatively low frequency.
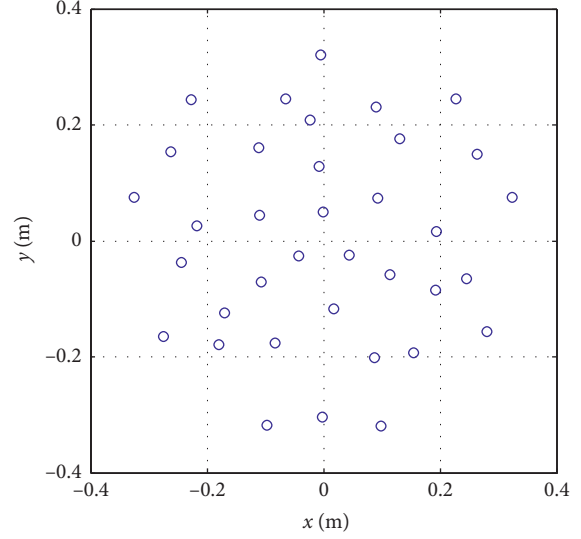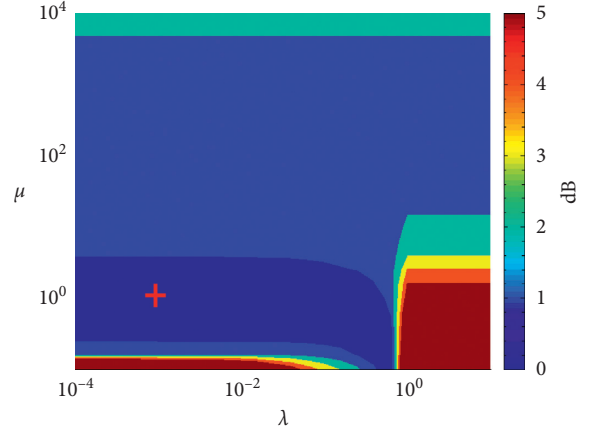

FIGURE 1: Array microphone distribution.

FIGURE 2: The map of average SPL deviation between the reconstructed and theoretical value vs. λ and μ.

The mainlobe widths of the sources reduce at 6000 Hz, and the two sources are separated. However, many spurious sources appear which leads to a blurred result. Comparing to CB, other three deconvolution algorithms can effectively narrow the mainlobe width, enhance the spatial resolution, and eliminate the spurious sources. Comparing the 2000 Hz results of three deconvolution algorithms in Figures 3–5, it can be generally seen that the lower the SNR, the more irregular the mainlobe. Comparing the submaps (b) to (d) in Figures 3–5, respectively, the mainlobe width of SFISTA is the narrowest, followed by SC-DAMAS and FISTA. Comparing the 6000 Hz results of three deconvolution algorithms in Figures 3–5, there is almost no difference among them due to the high spatial resolution of CB itself.

To verify the quantification performance of SFISTA, taking the result of 20 dB SNR as an example, the quantification accuracy of each approach is described. Table 1 lists mainlobe integral values and mainlobe peak values for each approach. At 2000 Hz, the mainlobe integral values of each
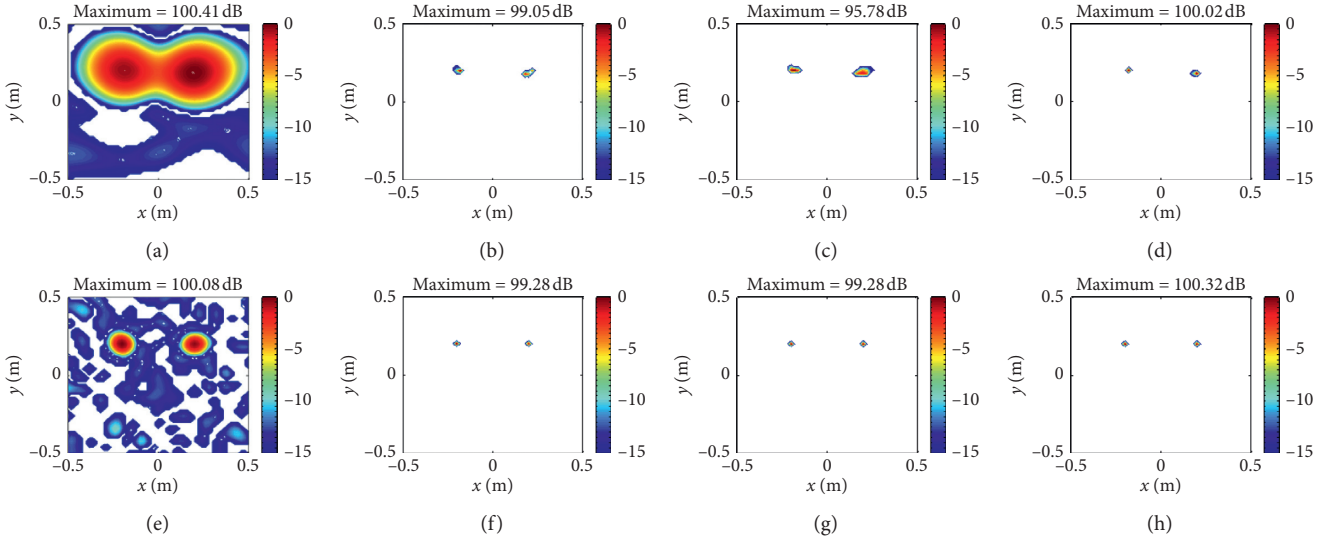
FIGURE 3: Contour maps of SPLs at 1 m distance from the source using different algorithms at a 10 dB SNR. (a) CB at 2000 Hz; (b) SC-DAMAS at 2000 Hz; (c) FISTA at 2000 Hz; (d) SFISTA at 2000 Hz; (e) CB at 6000 Hz; (f) SC-DAMAS at 6000 Hz; (g) FISTA at 6000 Hz; (h) SFISTA at 6000 Hz.
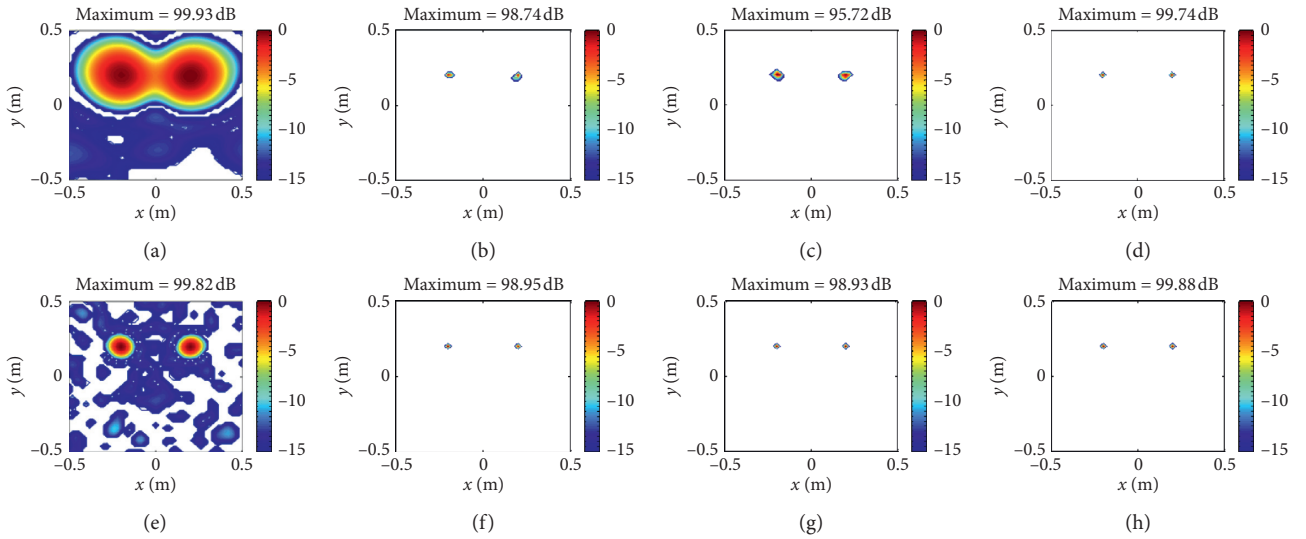

FIGURE 4: Contour maps of SPLs at 1 m distance from the source using different algorithms at a 20 dB SNR. (a) CB at 2000 Hz; (b) SC-DAMAS at 2000 Hz; (c) FISTA at 2000 Hz; (d) SFISTA at 2000 Hz; (e) CB at 6000 Hz; (f) SC-DAMAS at 6000 Hz; (g) FISTA at 6000 Hz; (h) SFISTA at 6000 Hz.

deconvolution beamforming are close to the preset 1 m SPL of the source. It indicates that the sound source can be accurately quantified by the integral value of the mainlobe. Then, the difference between the mainlobe integral value and the corresponding peak value of each deconvolution beamforming approach is compared, and the difference of SFISTA is smaller (about 0.1 dB) than SC-DAMAS (about 1.3 dB) and FISTA (about 4.2 dB). Namely, both the mainlobe integral value and mainlobe peak value of SFISTA are close to the peak value of CB. This indicates that the mainlobe width of SFISTA is the narrowest and the convergence is the best at relatively low frequency. At 6000 Hz,

due to the high spatial resolution of CB itself, the mainlobe converges to a grid point after deconvolution, and the difference between the mainlobe peak value and the corresponding mainlobe integral value of each deconvolution beamforming is zero. Further, the mainlobe integral values of each deconvolution beamforming and corresponding peak values of CB are compared with the preset 1 m SPL of the source. SFISTA and CB are almost the same and closer to the true value than the other two, and the deviation between the true value and the other two deconvolution beamforming is also less than 1 dB. It indicates that all algorithms can accurately quantify the sound source strength, and
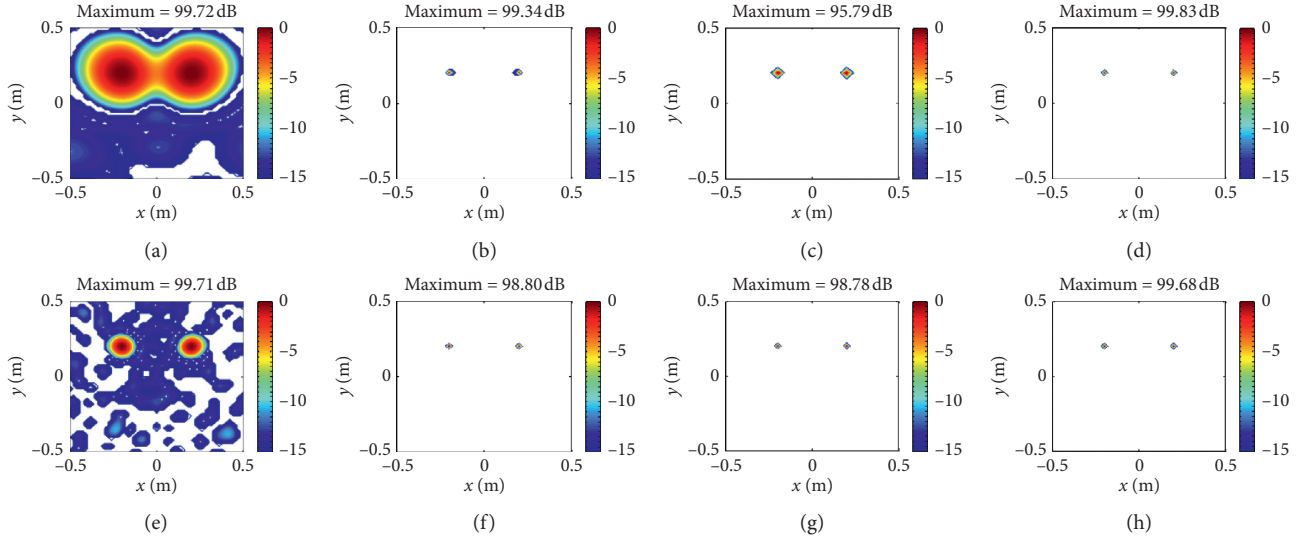
FIGURE 5: Contour maps of SPLs at 1 m distance from the source using different algorithms at 40 dB SNR. (a) CB at 2000 Hz; (b) SC-DAMAS at 2000 Hz; (c) FISTA at 2000 Hz; (d) SFISTA at 2000 Hz; (e) CB at 6000 Hz; (f) SC-DAMAS at 6000 Hz; (g) FISTA at 6000 Hz; (h) SFISTA at 6000 Hz.

TABLE 1: Simulation quantification results using different approach.

			CB	SC-DAMAS	FISTA	SFISTA
2000 Hz	Integral value	Left source	—	99.66	99.74	99.74
		Right source	—	99.88	99.86	99.86
	Peak value	Left source	99.76	98.14	95.72	99.74
		Right source	99.93	98.74	95.50	99.65
6000 Hz	Integral value	Left source	—	98.95	98.93	99.88
		Right source	—	98.80	98.80	99.66
	Peak value	Left source	99.82	98.95	98.93	99.88
		Right source	99.72	98.80	98.80	99.66

SFISTA slightly outperforms SC-DAMAS and FISTA at relatively high frequency.

Assuming a known point source with 100 dB is located at (0, 0, 1) m. In Figure 6, the quantification accuracy, convergence performance, and computational efficiency are further compared for the three deconvolution beamforming algorithms. The black dotted line, the red solid line, and the blue dashed line represent SC-DAMAS, SFISTA, and FISTA, respectively. The iteration number of FISTA and SFISTA is 1000. Since SC-DAMAS is solved by the convex optimization MATLAB toolbox, the iteration number cannot be set and the default terminal condition is applied.

Figure 6(a) shows the 1 m SPL deviation between the mainlobe integral and the true value at each frequency. When the frequency is lower than 3000 Hz, the deviations of the three algorithms are similar. When the frequency is higher than 3000 Hz, the deviation of SFISTA is smaller than that of SC-DAMAS and FISTA. In summary, the quantification accuracy of SFISTA is superior to the others.

The standard deviation, which is used to measure the convergence, is defined as [12]

$$\sigma^{(l)} = \frac{1}{\sqrt{N}} \left\| \mathbf{x}^{(l)} - \mathbf{x}_{e2} \right\|, \qquad (14)$$

where $\mathbf{x}^{(l)}$ is the reconstructed SPL at the lth iteration and $\mathbf{x}_e$ is the true one. Since the stopping criteria of SC-DAMAS does not depend on the iteration number and its computational efficiency is low, only the standard deviation curves of FISTA and SFISTA at 2000 Hz are given in Figures 6(b) and 6(c). Figure 6(b) shows the curves of the standard deviation vs. the iteration number and, Figure 6(c) shows the curves of the standard deviation vs. computational time.

As shown in Figure 6(b), standard deviation of SFISTA decreases rapidly and tends to be stable after about 1000 iterations. Standard deviation of FISTA decreases slower and tends to be stable after about 2500 iterations. In addition, the stable standard deviation of SFISTA is less than that of FISTA, which indicates that SFISTA enjoys better quantification accuracy. Furthermore, Figure 6(c) shows that SFISTA takes less computational time than FISTA to achieve the same standard deviation. This more intuitively shows that SFISTA converges faster than FISTA.
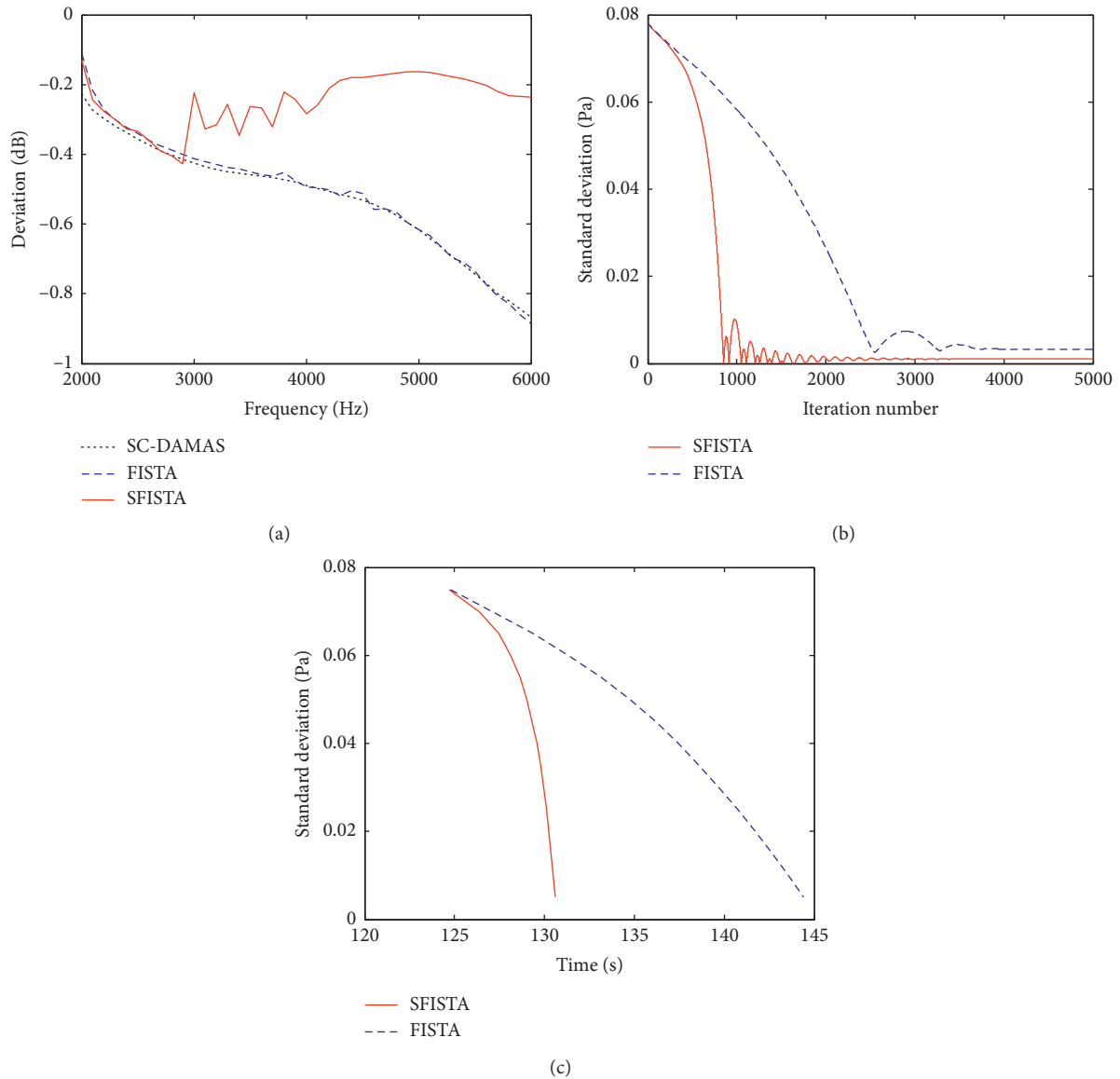
(a)

(b)

(c)

Figure 6: Performance comparison. (a) Amplitude deviation; (b) convergence; (c) computational efficiency.

To sum up, SFISTA has faster convergence speed and higher quantification accuracy and computational efficiency compared to SC-DAMAS and FISTA. Further, the uncertainty analysis of the sound source identification performance of SFISTA is performed. In practice, the sound source positions are unknown prior to measurement. So, a statistical simulation based on the Monte Carlo approach is to be used to archive uncertainty [23–25]. A 200-time Monte Carlo simulation is performed. The monopole sound source with 100 dB is randomly placed on a 50 cm × 50 cm plane with a 1 m distance from the microphone array. The SPL at 1 m distance from the sound source is retrieved by integration over 4 segments of 0.02 cm × 0.02 cm that are defined in the map around the maximum value [26, 27]. The sound frequencies are 2000 Hz and 6000 Hz, and the SNR is 20 dB and 40 dB. Figure 7 illustrates the cumulative distribution functions (CDFs) of the proposed SFISTA at different SNRs

and different frequencies. The location error is measured by the ratio of the distance between the identified position of the maximum SPL and the preset sound source position to the grid spacing. The quantification error of strength is measured by the deviation between the mainlobe integral of the identified sound source and the preset 100 dB. The black dot dash line and blue dotted line are the results of 2000 Hz and 6000 Hz at 20 dB SNR, and the red solid line and green solid line are the results of 2000 Hz and 6000 Hz at 40 dB SNR. Figure 7(a) is the CDF of the location error, and in general, the location accuracy of 6000 Hz is higher than that of 2000 Hz, and the accuracy is higher in the case of 40 dB SNR than that of 20 dB SNR. Except that there are a few points whose location error is greater than one grid interval in the case of 2000 Hz and 20 dB SNR, the location errors of other points are less than one grid interval, which indicates that almost all the identified sound source positions fall on
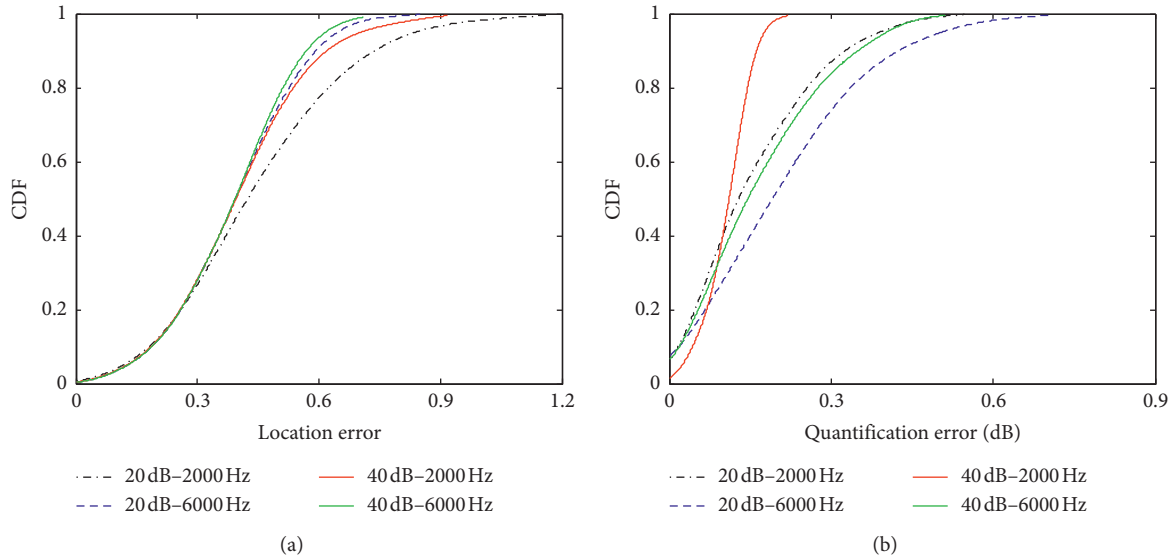
FIGURE 7: Cumulative distribution function for error of SFISTA under 200 trials of Monte Carlo simulations. (a) Location error; (b) quantification error.
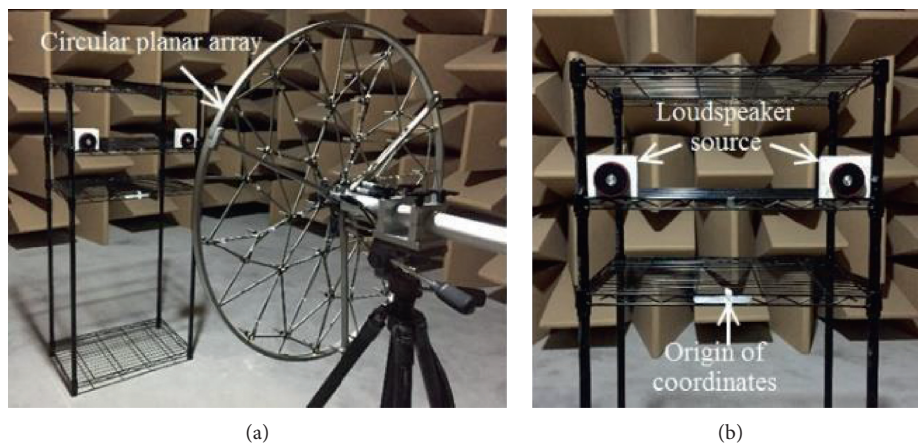

FIGURE 8: Experimental configuration.

one of the four closest grid points and most of them fall on the nearest point. The maximum location errors of 6000 Hz at 40 dB SNR, 6000 Hz at 20 dB SNR, 2000 Hz at 40 dB SNR, and 2000 Hz at 20 dB SNR are about 0.71, 0.85, 0.92, and 1.17, respectively. The mean location errors of 6000 Hz at 40 dB SNR, 6000 Hz at 20 dB SNR, 2000 Hz at 40 dB SNR, and 2000 Hz at 20 dB SNR are about 0.39, 0.39, 0.40, and 0.45, respectively. In other words, SFISTA enjoys high location accuracy. Figure 7(b) is the CDF of the quantification error. Results show that the quantification accuracy of 2000 Hz is higher than that of 6000 Hz, and the accuracy is higher in the case of 40 dB SNR than that of 20 dB SNR. This is because the conventional beamforming result of 2000 Hz has less ghosts than that of 6000 Hz, and the energy is concentrated on the mainlobe. Due to the interference of noise, the CDF of the quantification error of 20 dB SNR is larger than that of 40 dB SNR. The maximum quantification errors of 2000 Hz at 40 dB SNR, 2000 Hz at 20 dB SNR,

6000 Hz at 40 dB SNR, and 6000 Hz at 20 dB SNR are about 0.20, 0.55, 0.51, and 0.71, respectively. The mean quantification errors of 2000 Hz at 40 dB SNR, 2000 Hz at 20 dB SNR, 6000 Hz at 40 dB SNR, and 6000 Hz at 20 dB SNR are about 0.11, 0.15, 0.17, and 0.21, respectively. In other words, SFISTA also enjoys high quantification accuracy.

4. Experiment

As shown in Figure 8, the same microphone array as that in Section 3 is used for the experiment. The uncorrelated speaker sources driven by the white noise are approximately located at (−0.2, 0.2, 1) m and (0.2, 0.2, 1) m.

Same as the simulation, the iteration number of FISTA and SFISTA is set as 1000. CB results at 2000 Hz and 6000 Hz are shown in Figures 9(a) and 9(e), respectively. Sources fuse together, and the spatial resolution is poor at 2000 Hz, and many spurious sources appear at 6000 Hz. Figures 9(b)–9(d)
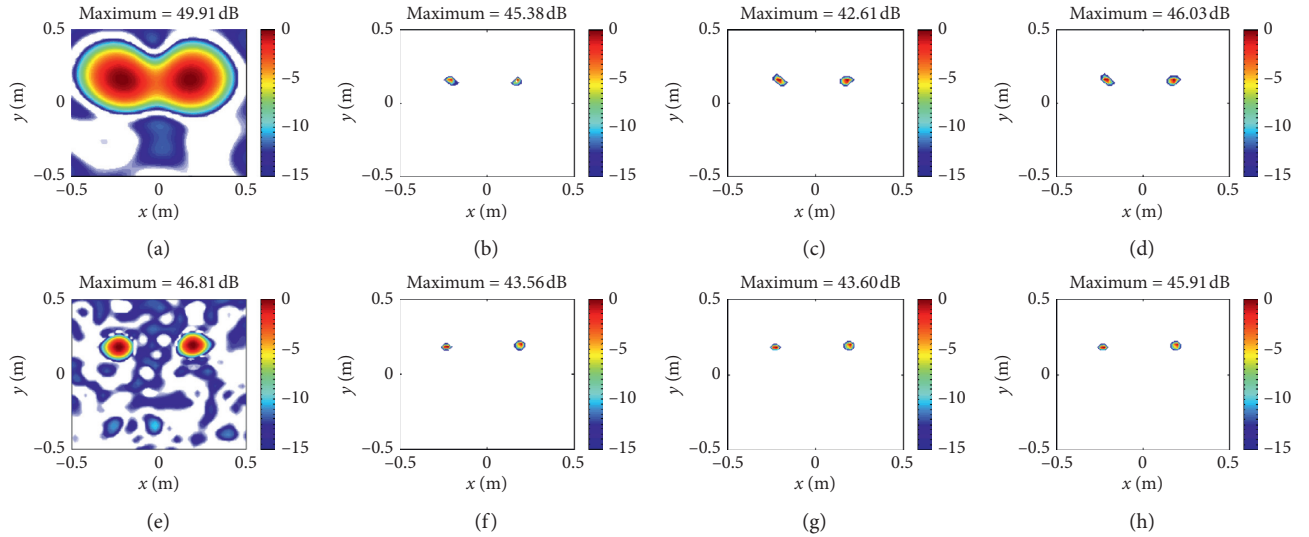
FIGURE 9: Experimental contour maps of SPLs at 1 m distance from the source using different algorithms. (a) CB at 2000 Hz; (b) SC-DAMAS at 2000 Hz; (c) FISTA at 2000 Hz; (d) SFISTA at 2000 Hz; (e) CB at 6000 Hz; (f) SC-DAMAS at 6000 Hz; (g) FISTA at 6000 Hz; (h) SFISTA at 6000 Hz.

TABLE 2: Experimental quantification results of different algorithms.

			CB	SC-DAMAS	FISTA	SFISTA
2000 Hz	Integral value	Left source	—	49.85	49.39	50.05
		Right source	—	49.39	48.94	49.61
	Peak value	Left source	49.92	43.73	42.21	45.46
		Right source	49.41	45.38	42.61	46.03
6000 Hz	Integral value	Left source	—	46.23	46.14	46.92
		Right source	—	44.60	44.52	45.30
	Peak value	Left source	46.84	43.56	43.60	45.91
		Right source	45.56	42.88	42.76	45.09

and 9(f)–9(h) are the results of SC-DAMAS, FISTA, and SFISTA at 2000 Hz and 6000 Hz, respectively. As shown, all deconvolution beamforming can locate the sound sources accurately. Table 2 gives the experimental results of amplitude quantification. Because we do not obtain the actual strengths of the loudspeaker sources, we use the mainlobe peak value of CB instead of strength to measure the quantification accuracy here. Both the results of 2000 Hz and 6000 Hz indicate that the results of the mainlobe integral value of each deconvolution algorithm are close, and the difference between the mainlobe peak value and the corresponding mainlobe integral value of SFISTA is the smallest. Besides, similar with the above simulation conclusion, compared with SC-DAMAS and FISTA, the mainlobe peak values of SFISTA are closer to those of CB. It also shows that SFISTA enjoys the better spatial convergence performance than SC-DAMAS and FISTA. In summary, SFISTA performs better than SC-DAMAS and FISTA, which is consistent with the simulation conclusion.

5. Conclusions

In this paper, SFISTA deconvolution beamforming for the sparse sound source identification is proposed. Simulations and experiments indicate that the proposed SFISTA has good acoustic source identification performance. Compared with SC-DAMAS and FISTA, SFISTA enjoys better spatial resolution, convergence performance, quantification accuracy, and computational efficiency.

Data Availability

Datasets generated and analyzed in the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] P. Chiariotti, M. Martarelli, and P. Castellini, "Acoustic beamforming for noise source localization-reviews, methodology and applications," *Mechanical Systems and Signal Processing*, vol. 120, pp. 422–448, 2019.

[2] R. Merino-Martínez, P. Sijtsma, M. Snellen et al., "A review of acoustic imaging methods using phased microphone arrays," *CEAS Aeronautical Journal*, vol. 10, no. 1, pp. 197–230, 2019.

[3] Z. Chu and Y. Yang, "Comparison of deconvolution methods for the visualization of acoustic sources based on cross-spectral imaging function beamforming," *Mechanical Systems and Signal Processing*, vol. 48, no. 1-2, pp. 404–422, 2014.

[4] V. F. Kopiev, M. Y. Zaytsev, and I. V. Belyaev, "Investigation of airframe noise for a large-scale wing model with high-lift devices," *Acoustical Physics*, vol. 62, no. 1, pp. 97–107, 2016.

[5] D. Yang, J. Wen, F. Miao, Z. Wang, X. Gu, and X. Lian, "Complete de-dopplerization and acoustic holography for external noise of a high-speed train," *The Journal of the Acoustical Society of America*, vol. 140, no. 3, pp. 1739–1748, 2016.

[6] B. Zajamsek, C. J. Doolan, D. J. Moreau, J. Fischer, and Z. Prime, "Experimental investigation of trailing edge noise from stationary and rotating airfoils," *The Journal of the Acoustical Society of America*, vol. 141, no. 5, pp. 3291–3301, 2017.

[7] P. Chiariotti, G. Battista, M. Ettorre, and P. Castellini, "Average acoustic beamforming in car cabins: an automatic system for acoustic mapping over 3D surfaces," *Applied Acoustics*, vol. 129, pp. 47–63, 2018.

[8] R. Porteous, Z. Prime, C. J. Doolan, D. J. Moreau, and V. Valeau, "Three-dimensional beamforming of dipolar aeroacoustic sources," *Journal of Sound and Vibration*, vol. 355, pp. 117–134, 2015.

[9] W. Ma and X. Liu, "DAMAS with compression computational grid for acoustic source mapping," *Journal of Sound and Vibration*, vol. 410, pp. 473–484, 2017.

[10] F. Ning, J. Wei, L. Qiu, H. Shi, and X. Li, "Three-dimensional acoustic imaging with planar microphone arrays and compressive sensing," *Journal of Sound and Vibration*, vol. 380, pp. 112–128, 2016.

[11] T. F. Brooks and W. M. Humphreys, "A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays," *Journal of Sound and Vibration*, vol. 294, no. 4-5, pp. 856–879, 2006.

[12] K. Ehrenfried and L. Koop, "Comparison of iterative deconvolution algorithms for the mapping of acoustic sources," *AIAA Journal*, vol. 45, no. 7, pp. 1584–1595, 2007.

[13] R. Dougherty, *Extensions of DAMAS and Benefits and Limitations of Deconvolution in Beamforming*, Monterey, CA, USA, 2005.

[14] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[15] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *Siam Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[16] O. Lylloff, E. Fernández-Grande, F. Agerkvist, J. Hald, E. Tiana Roig, and M. S. Andersen, "Improving the efficiency of deconvolution algorithms for sound source localization," *The Journal of the Acoustical Society of America*, vol. 138, no. 1, pp. 172–180, 2015.

[17] T. Zhao, Y. C. Eldar, A. Beck et al., "Smoothing and decomposition for analysis sparse recovery," *IEEE Transactions on Signal Processing*, vol. 62, no. 7, pp. 1762–1774, 2014.

[18] T. Yardibi, J. Li, P. Stoica, and L. N. Cattafesta, "Sparsity constrained deconvolution approaches for acoustic source mapping," *The Journal of the Acoustical Society of America*, vol. 123, no. 5, pp. 2631–2642, 2008.

[19] N. Chu, J. Picheral, A. Mohammad-Djafari, and N. Gac, "A robust super-resolution approach with sparsity constraint in acoustic imaging," *Applied Acoustics*, vol. 76, pp. 197–208, 2014.

[20] T. Padois and A. Berry, "Orthogonal matching pursuit applied to the deconvolution approach for the mapping of acoustic sources inverse problem," *The Journal of the Acoustical Society of America*, vol. 138, no. 6, pp. 3678–3685, 2015.

[21] G. Grant and S. Boyd, "CVX – Matlab software for disciplined convex programming, version 2.0 beta," September 2013, http://cvxr.com/cvx.

[22] E. Sarradj, "Three-dimensional acoustic source mapping with different beamforming steering vector formulations," *Advances in Acoustics and Vibration*, vol. 2012, Article ID 292695, 12 pages, 2012.

[23] T. Yardibi, C. Bahr, N. Zawodny, F. Liu, L. N. Cattafesta III, and J. Li, "Uncertainty analysis of the standard delay-and-sum beamformer and array calibration," *Journal of Sound and Vibration*, vol. 329, no. 13, pp. 2654–2682, 2010.

[24] L. Battista, E. Schena, G. Schiavone, S. A. Sciuto, and S. Silvestri, "Calibration and uncertainty evaluation using Monte Carlo method of a simple 2D sound localization system," *IEEE Sensors Journal*, vol. 13, no. 9, pp. 3312–3318, 2013.

[25] P. Chiariotti and M. Martarelli, "Acoustic beamforming-analysis of uncertainty and metrological performances," *Mechanical Systems and Signal Processing*, vol. 22, no. 3, pp. 672–692, 2008.

[26] S. Yin, Z. Chu, Y. Zhang, and Y. Liu, "Adaptive reweighting homotopy algorithm based compressive spherical beamforming with spherical microphone arrays," *The Journal of the Acoustical Society of America*, vol. 147, no. 1, pp. 480–489, 2020.

[27] E. Sarradj, "A fast signal subspace approach for the determination of absolute levels from phased microphone array measurements," *Journal of Sound and Vibration*, vol. 329, no. 9, pp. 1553–1569, 2010.