**OPEN**

# A panoramic continuous compressive beamformer with cuboid microphone arrays

Yang Yang[1], Zhigang Chu[2], Yong-Xin Yang[2], Zhongming Xu[2] & Yongxiang Zhang[2]

Compressive beamforming is a powerful approach for the direction-of-arrival (DOA) estimation and strength quantification of acoustic sources. The conventional grid-based discrete compressive beamformer suffers from the basis mismatch conundrum. Its result degrades under the situation that sources fall off the grid. The existing continuous compressive beamformer with linear or planar microphone arrays can circumvent the conundrum, but work well only for sources in a local region. Here we develop a panoramic continuous compressive beamformer with cuboid microphone arrays based on an atomic norm minimization (ANM) and a matrix pencil and paring method. To solve the positive semidefinite programming equivalent to the ANM efficiently, we formulate a solving algorithm based on the alternating direction method of multipliers. We also present an iterative reweighted ANM to enhance sparsity and resolution. The beamformer is capable of estimating the DOAs and quantifying the strengths of acoustic sources panoramically and accurately, whether a standard uniform or a sparse cuboid microphone array is utilized.

Compressive sensing[1–3] based beamforming is an emerging and powerful approach for the direction-of-arrival (DOA) estimation and strength quantification of acoustic sources, which is also simply called compressive beamforming[4,5]. It measures signals with an array of microphones, and then processes the signals to retrieve the direction and strength information of sources by exploiting the source sparsity. Numerous applications can be encountered, for example, noise source identification in environmental protection, target detection in military surveillance, fault diagnosis in equipment maintenance, speaker tracking in video conference, etc.

In conventional compressive beamformer, the DOA domain is gridded/discretized into a finite set of look directions and all the sources are assumed to fall in these look directions. An underdetermined linear system of equations is established relating the signals measured by microphones to the unknown source distribution. It is solved by imposing a sparsity constraint. Under a single snapshot, that is minimizing the $\ell_1$ norm of the vector composed by the source strengths in all the look directions. Under the multiple snapshots, the $\ell_1$ norm of the vector becomes the $\ell_{2,1}$ norm of a matrix. Its results become inaccurate when the DOAs of sources do not conform with these look directions. The problem is termed as basis mismatch[4,6] and can be often encountered in practical applications. Using finer grids mitigates the problem but increases the complexity of computation. More seriously, grid refinement makes the coherence of the measuring process increase, which can cause offset in the estimates[4].

Inspired by the continuous methods in the field of frequency retrieval[7–14], some scholars have recently developed the continuous compressive beamformer for the DOA estimation and strength quantification of acoustic sources. Based on the minimization of the atomic norm of source strength and the polynomial rooting method, Xenaki et al.[15] developed a one-dimensional single-snapshot continuous compressive beamformer for the measurement with linear microphone arrays. Park et al.[16] extended Xenaki et al's beamformer to the multiple-snapshot case via the group atomic norm. The authors[17–20] developed a two-dimensional single-snapshot continuous compressive beamformer for the measurement with rectangular microphone arrays based on the minimization of the atomic norm of microphone signal induced by sources and the matrix enhancement and matrix pencil method[21], and extended it to the multiple-snapshot case via the group atomic norm and the matrix pencil and pairing (MaPP) method[13,14]. Treating the DOAs of sources as a continuum, these beamformer can sidestep the basis mismatch conundrum fundamentally. Moreover, the multiple-snapshot data help to obtain more accurate and robust

[1]Faculty of Vehicle Engineering, Chongqing Industry Polytechnic College, Chongqing, 401120, China. [2]State Key Laboratory of Mechanical Transmissions, Chongqing University, Chongqing, 400044, China. Correspondence and requests for materials should be addressed to Z.C. (email: zgchu@cqu.edu.cn)

results compared with the single-snapshot ones[16,20]. In the linear array, there are only a line of microphones. In the rectangular array, there are only a plane of microphones. When two sources distribute symmetrically about the line or the plane, they induce the same signals at each microphone, and thus will not be distinguished. Consequently, the measurement with linear microphone arrays requires that the sources must be in front of the array and coplanar with the microphones. The measurement with rectangular microphone arrays requires that the sources must be in front of the array. That means the existing continuous compressive beamformer cannot yet panoramically estimate the DOAs and quantify the strengths of acoustic sources. This paper is dedicated to realizing this function with a cuboid microphone array. It is impossible for two sources to induce the same signals at each microphone in a cuboid microphone array. Therefore, the panoramic estimation will be allowed. The work is of great significance because the sources scatter throughout the whole three-dimensional space in a mass of practical situations.

The key contributions are as follows: (1) we develop a panoramic continuous compressive beamformer with cuboid microphone arrays under the multiple-snapshot data model. Four steps are involved. First, an atomic norm minimization (ANM) is defined to denoise the measured signal and get the microphone signal induced by sources. Then, a positive semidefinite programming equivalent to the ANM is formulated and solved. Subsequently, the MaPP method is utilized to process the result of the positive semidefinite programming and estimate DOAs. Finally, the source strengths are quantified based on the estimated DOAs and the obtained microphone signal from sources. (2) Based on alternating direction method of multipliers (ADMM)[22–24], we formulate a reasonably fast algorithm to solve the positive semidefinite programming. (3) We enhance sparsity and resolution via an iterative reweighted ANM (IRANM). (4) We investigate the applicability of the beamformer to the sparse cuboid microphone arrays, which are constructed via randomly retaining microphones from the standard uniform cuboid microphone arrays.

This paper is directly inspired by refs[13,14], where a multidimensional super-resolution frequency retrieval approach was proposed. The main connections and differences between them are as follows: (1) refs[13,14] focus on the problem of frequency retrieval, whereas this paper solves the problem of DOA estimation and source strength quantification. (2) The ANM and IRANM in this paper are enlightened by the convex relaxation method and the reweighted trace minimization method in refs[13,14]. Differently, the methods in refs[13,14] are based on the single-snapshot data model, whereas the methods in this paper are based on the multiple-snapshot one. As proved by the literatures in one- and two-dimensional fields[5,10,12,16,20], the multiple-snapshot method performs better. (3) In this paper, we deduce and develop an algorithm based on ADMM to solve the positive semidefinite programming, and demonstrate its advantage over the IPM based SDPT3 solver. References[13,14] did not do this work. (4) Reference[14] only introduces the concept of permutation matrices abstractly when describing the MaPP method, whereas this paper presents the concrete expressions. (5) In this paper, we investigate the effect rule of the estimated value of noise level on the results and give out advices. References[13,14] did not do this work.

## Results

### Accurate DOA estimation and strength quantification function of the panoramic continuous compressive beamformer.

Here, we show the function of the developed beamformer via a simulation example. Assume six sources. Their DOAs, expressed by $(\theta, \phi)$ with $\theta \in [0°, 180°]$ being the elevation angle and $\phi \in [0°, 360°]$ being the azimuth angle, are $(45°, 90°)$, $(45°, 120°)$, $(90°, 180°)$, $(120°, 180°)$, $(135°, 270°)$ and $(155°, 290°)$, in turn. Their root mean square strengths, expressed by $s_{rms} = \|\mathbf{s}\|_2 / \sqrt{L}$ with $\mathbf{s} \in C^{1 \times L}$ being the row vector composed by the source strength under each snapshot, C being the set of complex numbers, $L$ being the total number of snapshots and $\|\cdot\|_2$ being the $\ell_2$ norm, are 100 dB, 98 dB, 96 dB, 94 dB, 92 dB and 90 dB (referring to $2 \times 10^{-5}$ Pa). The frequency of emitted signal is 4000 Hz. Figure 1 presents the utilized microphone arrays and the reconstructed source distributions. A standard uniform cuboid array with 343 microphones (Fig. 1a) is utilized to obtain Fig. 1c,e,g, while a sparse cuboid array with 170 microphones (Fig. 1b) is utilized to obtain Fig. 1d,f,h. Figure 1a,b also presents the panoramic scene of sources. Figure 1c,d corresponds to the conventional compressive beamformer. The grid is $[0°: 6°: 180°] \times [0°: 6°: 360°]$. Obviously, only the third and fourth sources that lie on the grid are accurately identified. For the other sources that lie off the grid, leakage occurs, which leads to inaccurate DOA estimations and peaks seriously deviating from the true source strengths. In contrast, as shown in Fig. 1e–h, the developed beamformer estimates the DOA and quantifies the strength of each source accurately. To sum up, the developed panoramic continuous compressive beamformer can conquer the basis mismatch conundrum, allowing accurate DOA estimation and strength quantification, whether a standard uniform or a sparse cuboid microphone array is utilized. Besides, Fig. 2 presents the result reconstructed by the two-dimensional continuous compressive beamformer with a rectangular microphone array. Only sources whose elevation angels lie in $[0°, 90°]$ are accurately identified. This demonstrates the necessity of developing the panoramic continuous compressive beamformer to realize the panoramic DOA estimation and strength quantification of acoustic sources. To obtain Figs 1 and 2, multiple snapshots are adopted. An example is also exhibited in Supplementary Note 1 to again demonstrate the advantages of the multiple-snapshot data model over the single-snapshot one, which has been demonstrated and explained in one- and two-dimensional fields[5,10,12,16,20].

### Efficiency advantage of our ADMM based algorithm.

To obtain Fig. 1e,f, we solve the positive semidefinite programming equivalent to the ANM by the off-the-peg SDPT3 solver in CVX toolbox[25], which uses the interior point method (IPM)[26]. To obtain Fig. 1g,h, we solve the positive semidefinite programming by our ADMM based algorithm. The accurate DOA estimation and strength quantification shown in Fig. 1e–h demonstrates that both the solver and our algorithm are effective. Employ $\|\hat{\mathbf{P}} - \mathbf{P}\|_F / \|\mathbf{P}\|_F$, $\|[\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}] - [\boldsymbol{\theta}, \boldsymbol{\phi}]\|_F / 2I$ and $\|\hat{\mathbf{s}}_{rms} - \mathbf{s}_{rms}\|_2 / \|\mathbf{s}_{rms}\|_2$ to measure the microphone signal reconstruction error, the DOA estimation error and the strength quantification error. Thereinto, $\|\cdot\|_F$ is the Frobenius norm, $\hat{\mathbf{P}} \in C^{ABC \times L}$ and $\mathbf{P} \in C^{ABC \times L}$ are the matrices of the reconstructed and the true microphone signals respectively, $A$, $B$ and $C$ are the row, column and layer
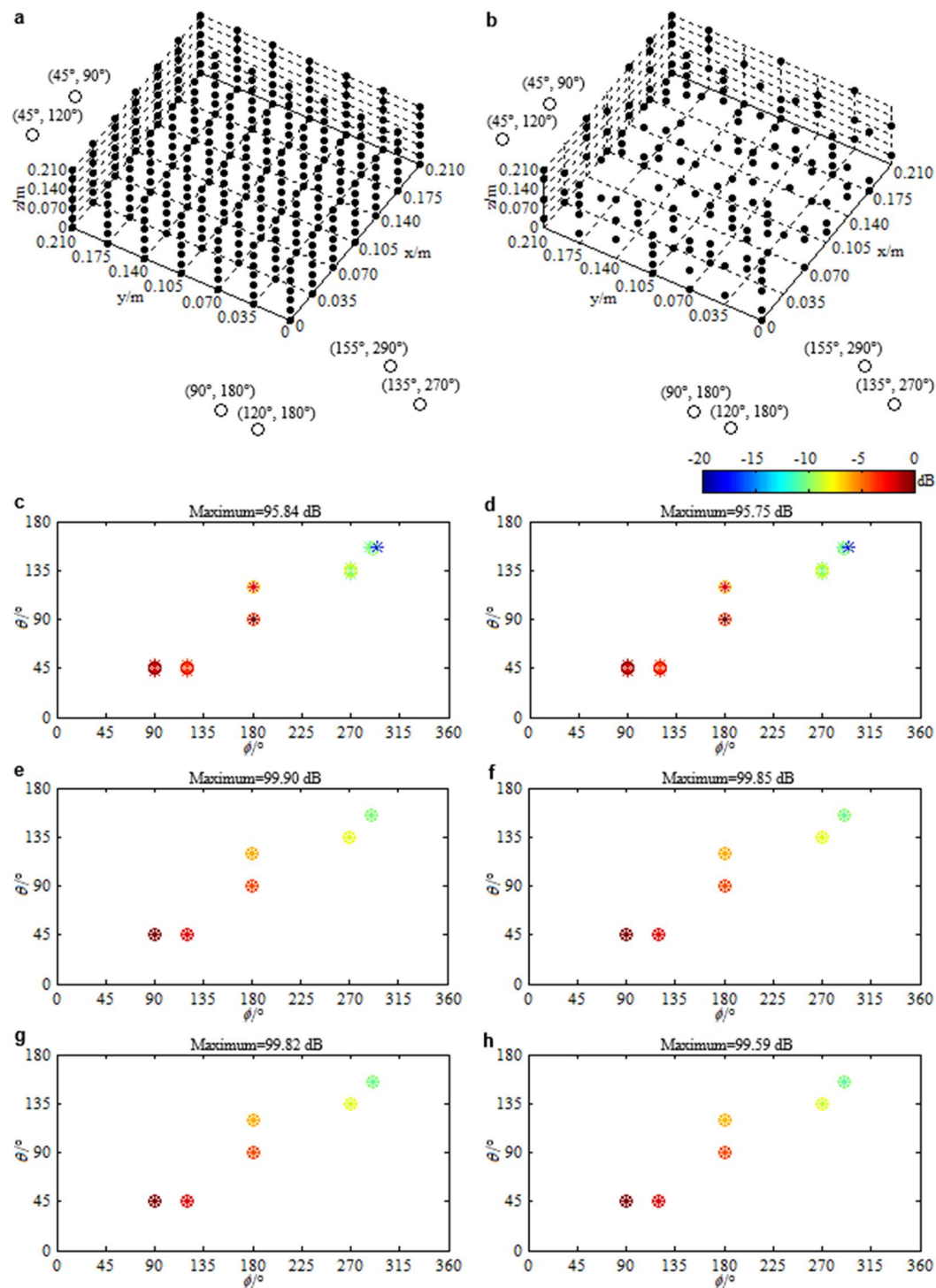
**Figure 1.** Microphone arrays and reconstructed source distributions. (**a**) The standard uniform cuboid microphone array for (**c,e** and **g**). (**b**) The sparse cuboid microphone array for (**d,f** and **h**). ● Represents the microphone and ○ represents the source. Source distributions reconstructed by (**c,d**) the conventional and (**e–h**) the developed panoramic continuous compressive beamformer. The positive semidefinite programming equivalent to the ANM is solved by (**e,f**) the SDPT3 solver in CVX toolbox and (**g,h**) our ADMM based algorithm. In (**c–h**), the reconstructed (*) and the true (○) outputs are scaled to dB via referring to their respective maximum, and at the same time, referring to $2 \times 10^{-5}$ Pa, the reconstructed maximum is labeled on the top, so do in the subsequent maps.

numbers of the cuboid microphone array, $ABC$ is the total number of microphones, $\hat{\theta} \in \mathrm{R}^I$ and $\theta \in \mathrm{R}^I$ are the vectors of the estimated and the true elevation angles respectively, R is the set of real numbers, $I$ is the total number of sources, $\hat{\phi} \in \mathrm{R}^I$ and $\phi \in \mathrm{R}^I$ are the vectors of the estimated and the true azimuth angles respectively, and
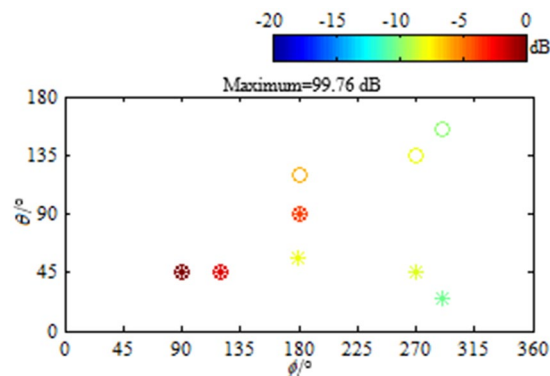
**Figure 2.** Source distribution reconstructed by two-dimensional continuous compressive beamformer.
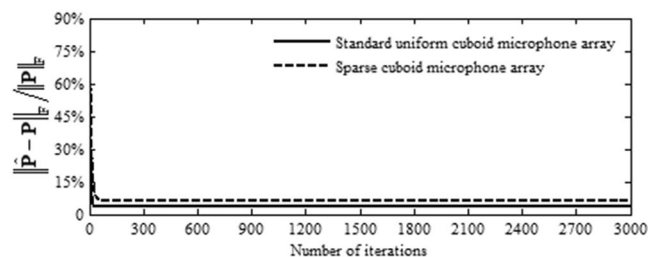


**Figure 3.** Curves of $\|\hat{\mathbf{P}} - \mathbf{P}\|_{\mathrm{F}} / \|\mathbf{P}\|_{\mathrm{F}}$ vs. number of iterations in our ADMM based algorithm.

| Microphone array | Approach to solve ANM | $\frac{\|\hat{\mathbf{P}} - \mathbf{P}\|_{\mathrm{F}}}{\|\mathbf{P}\|_{\mathrm{F}}}$ | $\frac{\|[\hat{\theta},\hat{\phi}] - [\theta,\phi]\|_{\mathrm{F}}}{2I}$ | $\frac{\|\hat{\mathbf{s}}_{rms} - \mathbf{s}_{rms}\|_2}{\|\mathbf{s}_{rms}\|_2}$ | Consuming time/s |
|---|---|---|---|---|---|
| Standard uniform cuboid array with 343 microphones | IPM based SDPT3 solver (Fig. 1e) | 2.45% | 0.06% | 1.95% | 1864 |
| | ADMM based algorithm (Fig. 1g) | 3.70% | 0.06% | 3.36% | 169 |
| Sparse cuboid array with 170 microphones | IPM based SDPT3 solver (Fig. 1f) | 3.32% | 0.04% | 2.76% | 1863 |
| | ADMM based algorithm (Fig. 1h) | 7.17% | 0.05% | 6.84% | 270 |

**Table 1.** Microphone signal reconstruction error, DOA estimation error and source strength quantification error corresponding to Fig. 1e–h, and consuming time of IPM based SDPT3 solver and ADMM based algorithm.

$\hat{\mathbf{s}}_{rms} \in \mathrm{R}^I$ and $\mathbf{s}_{rms} \in \mathrm{R}^I$ are the vectors of the quantified and the true root mean square strengths respectively. Table 1 lists the values of these errors corresponding to Fig. 1e–h and the consuming time of the IPM based SDPT3 solver and our ADMM based algorithm. Apparently, in terms of microphone signal reconstruction and strength quantification, the panoramic continuous compressive beamformer with the ADMM based algorithm has slightly lower accuracy than the one with the IPM based SDPT3 solver, whereas in term of DOA estimation, they have almost the same accuracy. The consuming time of our ADMM based algorithm is only about 1/7 of the one of the IPM based SDPT3 solver at most. Besides, when the dimensionality of the problem increases, for example, $A = B = C = 8$, the IPM based SDPT3 solver fails to obtain effective solutions, whereas our ADMM based algorithm still works well.

To test the convergence of our ADMM based algorithm, we plot the curves of the error $\|\hat{\mathbf{P}} - \mathbf{P}\|_{\mathrm{F}} / \|\mathbf{P}\|_{\mathrm{F}}$ vs. the number of iterations in Fig. 3. The parameter setup of sources keeps the same as in Fig. 1. Obviously, whether the standard uniform or the sparse cuboid microphone array is utilized, a plateau is reached after only a few dozens of iterations, which means our algorithm can converge fast. The errors after convergence are 3.38% and 6.38% respectively for the standard uniform cuboid microphone array and the sparse one, which are larger than the errors of the IPM based SDPT3 solver (2.45% and 3.32%). Changing the parameters, such as number and DOAs of sources, frequency, SNR and so on, to conduct simulations, similar phenomena can be obtained. See Supplementary Note 2. These phenomena demonstrate the inherent characteristic of ADMM that it converges fast to a moderate accuracy, but slowly even impossibly to an extremely accurate solution[11,22]. Fortunately, the moderate accuracy is typically sufficient in practical applications.

**Sparsity and resolution enhancement via IRANM in the case of small source separation.** Define the minimum separation among sources as
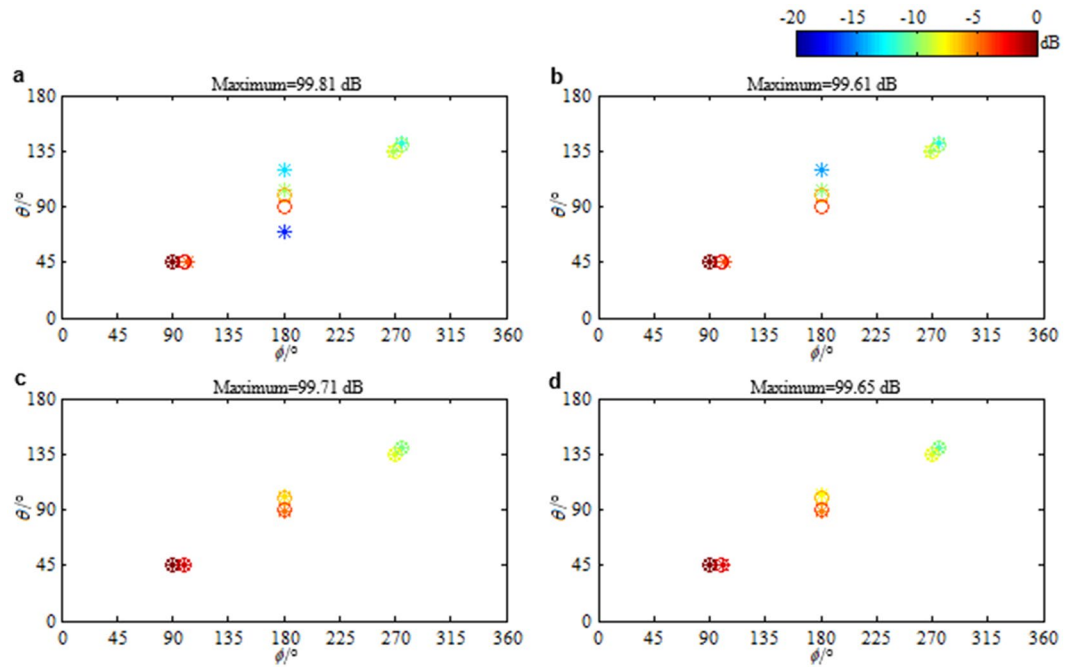
**Figure 4.** Reconstructed source distributions. (**a,b**) The ANM and (**c,d**) the IRANM based panoramic continuous compressive beamformer are utilized. (**a,c**) The standard uniform and (**b,d**) the sparse cuboid microphone array are utilized.

$$\Delta_{\min} = \min_{\substack{i,i' \in \{1,2,\cdots,I\} \\ i \neq i'}} \max\{|t_{1i} - t_{1i'}|, \ |t_{2i} - t_{2i'}|, \ |t_{3i} - t_{3i'}|\}, \tag{1}$$

where both $i$ and $i'$ are the source indices, and $t_{1i}(t_{1i'})$, $t_{2i}(t_{2i'})$ and $t_{3i}(t_{3i'})$ are the projections of the DOA of the $i$th ($i'$th) source on the $x$, $y$ and $z$ dimensions of the microphone array respectively. Denote by $\theta_i$ and $\phi_i$ the elevation and the azimuth angle of the $i$th source, $\Delta x$, $\Delta y$ and $\Delta z$ the microphone spaces, and $\lambda$ the wavelength, then $t_{1i} \equiv \sin\theta_i \cos\phi_i \Delta x/\lambda$, $t_{2i} \equiv \sin\theta_i \sin\phi_i \Delta y/\lambda$ and $t_{3i} \equiv \cos\theta_i \Delta z/\lambda$. A smaller minimum separation among sources that can be accurately identified means a higher resolution. Change the DOAs of the second, fourth and sixth sources assumed in Fig. 1 as (45°, 100°), (100°, 180°) and (140°, 275°) in turn. Namely, let the first and second sources, the third and fourth sources, and the fifth and sixth sources close to each other. $\Delta_{\min}$ decreases from 0.13 to 0.03. Figure 4 presents the reconstructed source distributions. The standard uniform cuboid array with 343 microphones (Fig. 1a) is utilized to obtain Fig. 4a,c, while the sparse cuboid array with 170 microphones (Fig. 1b) is utilized to obtain Fig. 4b,d. Figure 4a,b corresponds to the ANM based panoramic continuous compressive beamformer. It fails to identify all the sources accurately. The estimated number of sources in Fig. 4a is also more than the true one. In contrast, as shown in Fig. 4c,d, the IRANM based panoramic continuous compressive beamformer estimates the DOA and quantifies the strength of each source accurately. The phenomenon demonstrates that IRANM can enhance sparsity and resolution, whether a standard uniform or a sparse cuboid microphone array is utilized.

Figure 5 plots the curves of the above errors vs. $\Delta_{\min}$ of the ANM and the IRANM based panoramic continuous compressive beamformer. For each $\Delta_{\min}$, the errors are measured over 20 Monte Carlo runs. Two sources that are separated by $\Delta_{\min}$ are generated in each run. The standard uniform cuboid array with 343 microphones is utilized to obtain Fig. 5a–c, while a sparse cuboid array with only 170 randomly retained microphones is utilized to obtain Fig. 5d–f. In term of reconstruction of **P**, Fig. 5a,d shows that the error of ANM is distinctly higher under $\Delta_{\min} < 0.7/\sqrt[3]{ABC}$ compared to $\Delta_{\min} \geq 0.7/\sqrt[3]{ABC}$, whereas the error of IRANM varies only slightly across all $\Delta_{\min}$ and is distinctly lower than the one of ANM. In term of DOA estimation, Fig. 5b,e shows that on one hand, the IRANM based panoramic continuous compressive beamformer has very low errors across all $\Delta_{\min}$; on the other hand, compared to the former, the ANM based panoramic continuous compressive beamformer has distinctly higher errors under $\Delta_{\min} < 0.5/\sqrt[3]{ABC}$ and almost the same errors under $\Delta_{\min} \geq 0.5/\sqrt[3]{ABC}$. In term of strength quantification, Fig. 5c,f shows that the error of the ANM based panoramic continuous compressive beamformer is always higher than the one of the IRANM based panoramic continuous compressive beamformer, and that is particularly apparent under small $\Delta_{\min}$. These phenomena demonstrate that the IRANM based panoramic continuous compressive beamformer has the stronger denoising capability and the enhanced resolution compared to the ANM based one, allowing to reconstruct the microphone signal induced by sources as well as estimate the DOAs and quantify the strengths of small-separation sources more accurately, whether a standard uniform or a sparse cuboid microphone array is utilized.
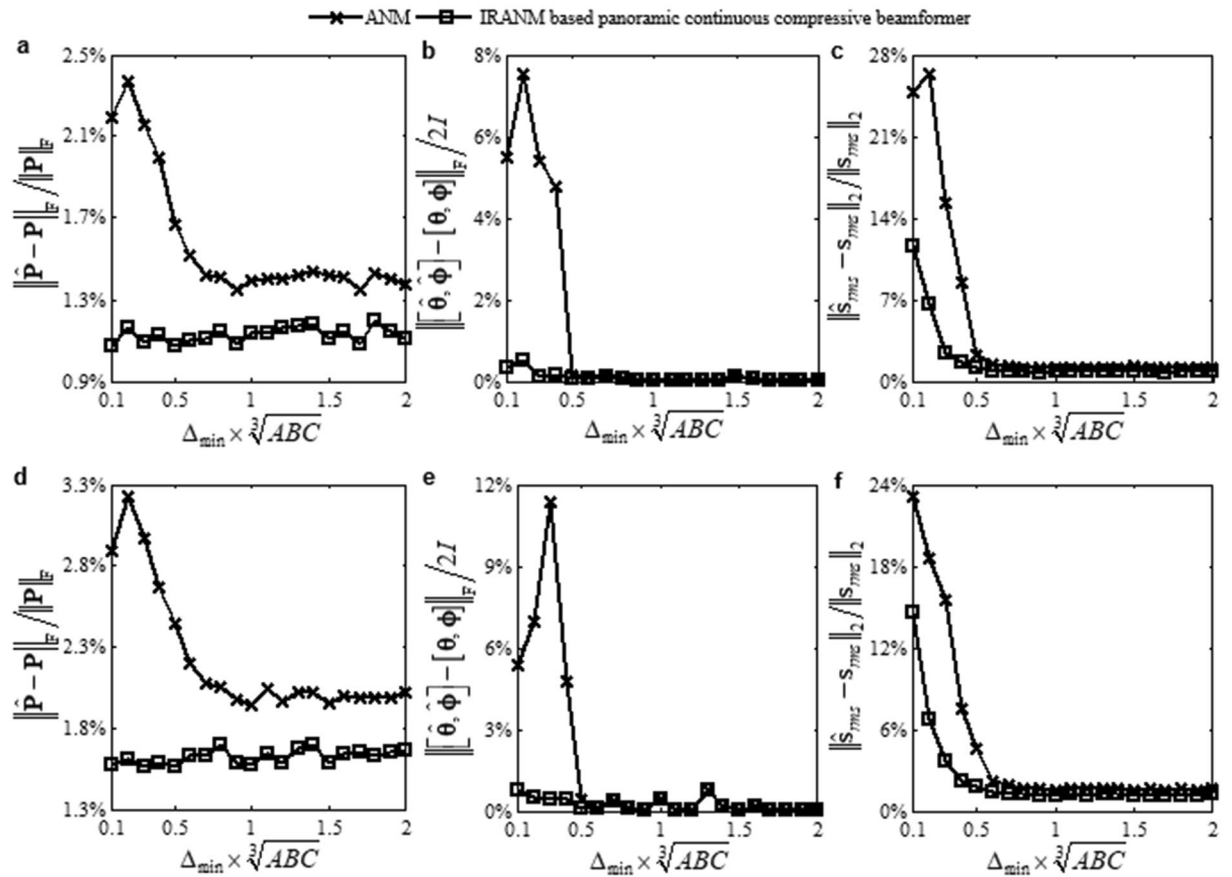
**Figure 5.** Error curves. (**a**,**d**) $\|\hat{\mathbf{P}} - \mathbf{P}\|_F / \|\mathbf{P}\|_F$, (**b**,**e**) $\|[\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}] - [\boldsymbol{\theta}, \boldsymbol{\phi}]\|_F / 2I$ and (**c**,**f**) $\|\hat{\mathbf{s}}_{rms} - \mathbf{s}_{rms}\|_2 / \|\mathbf{s}_{rms}\|_2$ vs. $\Delta_{\min}$. (**a**–**c**) The standard uniform and (**d**–**f**) the sparse cuboid microphone array are utilized.

**Sparsity and resolution enhancement via IRANM in the case of underestimated noise level.** In above results, the noise level is estimated accurately. Namely, the estimated noise level $\varepsilon$ is equal to the true one (the Frobenius norm of the true noise $\mathbf{N} \in \mathrm{C}^{ABC \times L}$). In practical applications, the noise is usually not explicitly known, and therefore, it is hard to let $\varepsilon$ be equal to the true noise level. Figure 6a–d respectively presents the distributions reconstructed by the ANM based panoramic continuous compressive beamformer for the sources assumed in Fig. 1 when $\varepsilon$ is $0.25\|\mathbf{N}\|_F$, $0.5\|\mathbf{N}\|_F$, $2\|\mathbf{N}\|_F$ and $4\|\mathbf{N}\|_F$. In Fig. 6a, the estimated number of sources is more than the true one, and the DOA estimation as well as the strength quantification is wrong. In Fig. 6b, even though the estimated sources in the display dynamic range have the same number as the true ones, the DOA estimation as well as the strength quantification is still wrong. In Fig. 6c, the DOAs are accurately estimated. The source strengths are quantified as 98.98 dB, 96.71 dB, 94.22 dB, 91.71 dB, 89.11 dB and 86.17 dB in turn, which are 1.02 dB, 1.29 dB, 1.78 dB, 2.29 dB, 2.89 dB and 3.83 dB lower than the true values. Distinctly, the weaker the source is, the more the quantified strength is lower than the true one. In Fig. 6d, the weakest source in (155°, 290°) has been lost. These phenomena demonstrate that for the ANM based panoramic continuous compressive beamformer, an underestimated noise level will lead to a less sparse solution and thus a wrong identification and a reduced resolution, whereas an overestimated noise level will cause a too sparse solution by eliminating the weak sources. This is mainly because by underestimating noise level, partial noise is identified as sources, whereas by overestimating noise level, the weak sources are removed as noise. Figure 6e,f presents the source distributions reconstructed by the IRANM based panoramic continuous compressive beamformer when $\varepsilon$ is $0.25\|\mathbf{N}\|_F$ and $0.5\|\mathbf{N}\|_F$. Apparently, all the DOAs are estimated and all the strengths are quantified accurately, which means IRANM enhances the sparsity and resolution. The standard uniform cuboid array with 343 microphones is utilized to obtain Fig. 6. Figure 7 presents the results when the sparse cuboid array with 170 microphones is utilized. It has the same rule as Fig. 6.

## Discussion

In summary, we develop a panoramic continuous compressive beamformer with cuboid microphone arrays. It first denoises the measured signal and thus obtains the signal from sources by minimizing a sparse metric of source distribution in the continuous domain, for example, the atomic norm of the microphone signal induced by sources, then retrieves the DOA information via using the MaPP method to process the result of ANM, and finally quantifies the source strength according to the DOA information and the microphone signal from sources.
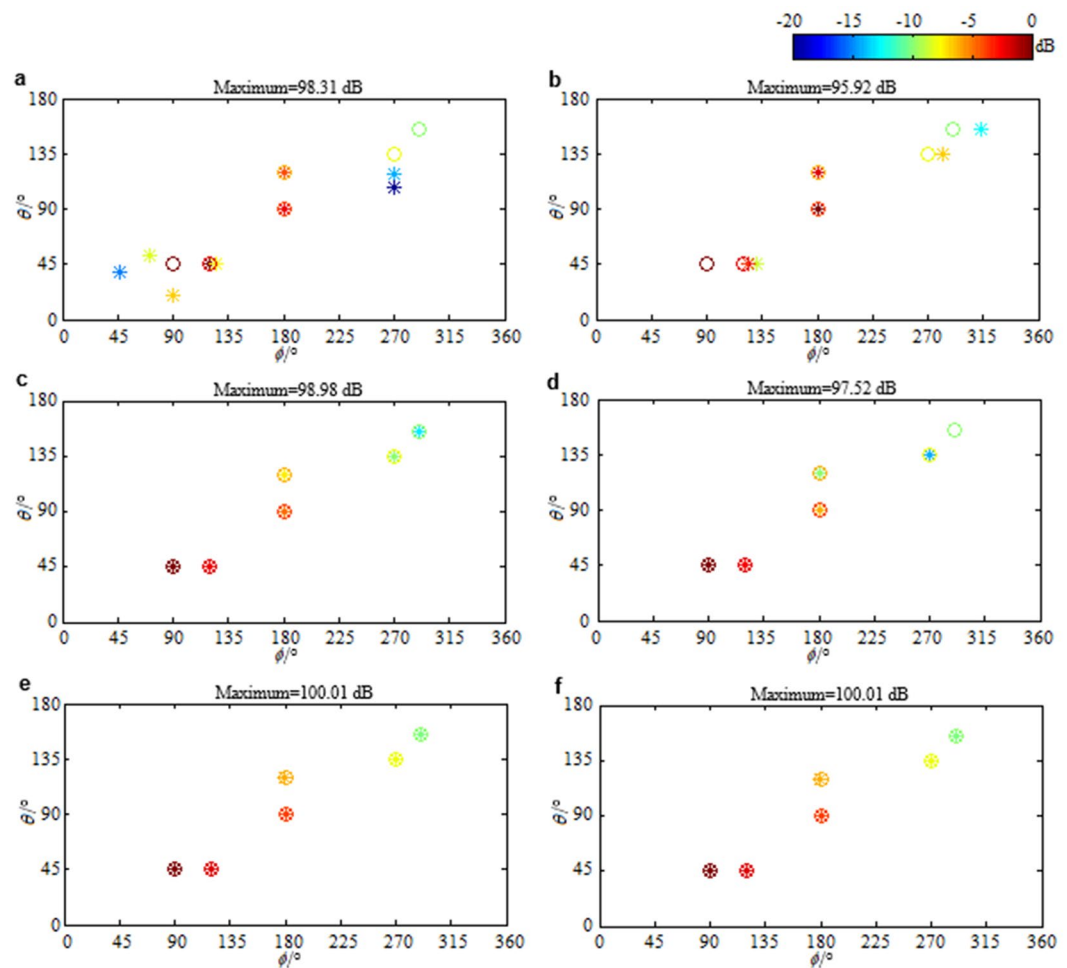
**Figure 6.** Reconstructed source distributions. (**a**–**d**) The ANM and (**e**,**f**) the IRANM based panoramic continuous compressive beamformer are utilized. $\varepsilon$ is (**a**,**e**) $0.25\|\mathbf{N}\|_{\mathrm{F}}$, (**b**,**f**) $0.5\|\mathbf{N}\|_{\mathrm{F}}$, (**c**) $2\|\mathbf{N}\|_{\mathrm{F}}$ and (**d**) $4\|\mathbf{N}\|_{\mathrm{F}}$. The standard uniform cuboid microphone array is utilized.

The beamformer can conquer the basis mismatch conundrum of the conventional grid-based discrete compressive beamformer, allowing panoramic and accurate DOA estimation and strength quantification of acoustic sources, whether a standard uniform or a sparse cuboid microphone array is utilized.

To solve the ANM, we formulate a positive semidefinite programming. It is a convex optimization problem and can be solved by the off-the-peg IPM based SDPT3 solver in CVX toolbox. The solver trends to be slow and even ineffective for large-dimensionality problems. To overcome the limitation, we present a reasonably fast algorithm based on ADMM. By establishing another sparse metric, we also present the IRANM that can be substituted for ANM. Because the metric promotes the sparsity to a larger degree than the atomic norm, IRANM can enhance sparsity and resolution compared with ANM. For the sources with a small separation, the ANM based beamformer fails to estimate the DOAs and quantify the strengths accurately, whereas the IRANM based beamformer succeeds. The beamformer takes the estimated noise level as an input. For the ANM based beamformer, overestimating the noise level may make the solution too sparse, for instance by eliminating sources of smaller strength. On the other hand, by underestimating the noise level, the solution may be less sparse than the actual solution, which will leads to inaccurate identifications and reduced resolution. In practical applications, when the noise is not explicitly known, in order to capture all the sources, we advise using a properly underestimated noise level and enhancing sparsity and resolution by IRANM.

The IRANM is solved only by the slow SDPT3 solver in this paper. We have not yet successfully developed a fast algorithm for it, even based on ADMM. The reason why IRANM cannot be solved accurately and robustly by ADMM can be explained as follows. IRANM is an iterative strategy. The weighting matrix utilized in current iteration is calculated based on the results in the previous iteration. In each iteration, if the semidefinite positive programming is solved via ADMM, the inherent characteristic that ADMM cannot fast converge to an accurate solution[11,22] will lead to a result with moderate accuracy. The error will affect the weighting matrix and thus burden the result with a larger error in next iteration. Consequently, as the iteration increases, the actually solved result deviates from the theoretical one more and more. In the future work, it is of significance to handle the problem. Besides, no experimental data are provided to support the claims in this paper. The accuracy of the
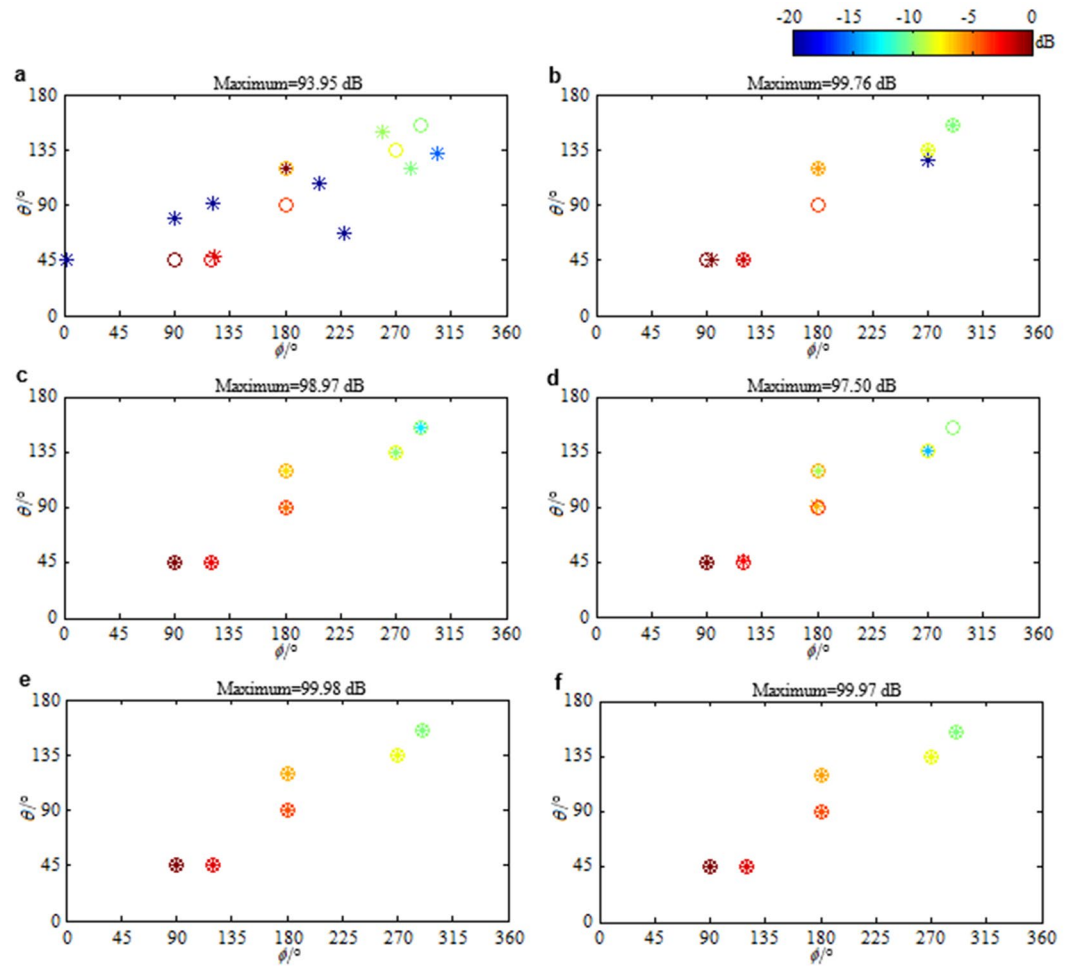
**Figure 7.** Reconstructed source distributions. (**a–d**) The ANM and (**e,f**) the IRANM based panoramic continuous compressive beamformer are utilized. $\varepsilon$ is (**a,e**) $0.25\|\mathbf{N}\|_F$, (**b,f**) $0.5\|\mathbf{N}\|_F$, (**c**) $2\|\mathbf{N}\|_F$ and (**d**) $4\|\mathbf{N}\|_F$. The sparse cuboid microphone array is utilized.

method becomes lower as the SNR decreases. It is also of significance to conduct the experimental investigation and explore a new method that can still enjoy high accuracy even under very low, for example, negative, SNRs. Finally, it is also of interest to develop the panoramic continuous compressive beamformer with other microphone arrays, for example, spherical ones, to realize the panoramic and accurate DOA estimation and source strength quantification.

## Methods

**Problem formulation.**　Figure 8 depicts the measurement layout with a cuboid microphone array. $(a, b, c)$ with $a = 0, 1, \ldots, A - 1$, $b = 0, 1, \ldots, B - 1$ and $c = 0, 1, \ldots, C - 1$ indexes the microphone. $(\theta_i, \phi_i)$ expresses the DOA of the $i$th source. Under the assumption of plane wave, the row vector $\mathbf{p}_{a,b,c} \in \mathrm{C}^{1 \times L}$ of the signal induced by sources at $(a, b, c)$th microphone under each snapshot can be modeled as

$$\mathbf{p}_{a,b,c} = \sum_{i=1}^{I} \mathbf{s}_i e^{j2\pi(t_{1i}a + t_{2i}b + t_{3i}c)},\tag{2}$$

where $j = \sqrt{-1}$ is the imaginary unit, $\mathbf{s}_i = [s_{i,1}, s_{i,2}, \cdots, s_{i,L}] \in \mathrm{C}^{1 \times L}$, $s_{i,l}$ expresses the strength of the $i$th source under the $l$th snapshot, which is the signal induced by the source at the $(0, 0, 0)$th microphone.

After forming the matrix $\mathbf{P} = [\mathbf{p}_{0,0,0}^T, \mathbf{p}_{0,0,1}^T, \cdots, \mathbf{p}_{0,0,C-1}^T, \mathbf{p}_{0,1,0}^T, \mathbf{p}_{0,1,1}^T, \cdots, \mathbf{p}_{0,1,C-1}^T, \cdots, \mathbf{p}_{0,B-1,0}^T, \mathbf{p}_{0,B-1,1}^T, \cdots,$ $\mathbf{p}_{0,B-1,C-1}^T, \mathbf{p}_{1,0,0}^T, \mathbf{p}_{1,0,1}^T, \cdots, \mathbf{p}_{1,0,C-1}^T, \mathbf{p}_{1,1,0}^T, \mathbf{p}_{1,1,1}^T, \cdots, \mathbf{p}_{1,1,C-1}^T, \cdots, \mathbf{p}_{1,B-1,0}^T, \mathbf{p}_{1,B-1,1}^T, \cdots, \mathbf{p}_{1,B-1,C-1}^T, \cdots,$ $\mathbf{p}_{A-1,0,0}^T, \mathbf{p}_{A-1,0,1}^T, \cdots, \mathbf{p}_{A-1,0,C-1}^T, \mathbf{p}_{A-1,1,0}^T, \mathbf{p}_{A-1,1,1}^T, \cdots, \mathbf{p}_{A-1,1,C-1}^T, \cdots, \mathbf{p}_{A-1,B-1,0}^T, \mathbf{p}_{A-1,B-1,1}^T, \cdots,$ $\mathbf{p}_{A-1,B-1,C-1}^T]^T \in \mathrm{C}^{ABC \times L}$ and the column vector $\mathbf{d}(t_{1i}, t_{2i}, t_{3i}) = [1, e^{j2\pi t_{1i}}, \cdots, e^{j2\pi t_{1i}(A-1)}]^T \otimes [1, e^{j2\pi t_{2i}}, \cdots,$ $e^{j2\pi t_{2i}(B-1)}]^T \otimes [1, e^{j2\pi t_{3i}}, \cdots, e^{j2\pi t_{3i}(C-1)}]^T \in \mathrm{C}^{ABC}$, where $(\cdot)^T$ and $\otimes$ denote the transpose and the Kronecker product operator respectively, we obtain
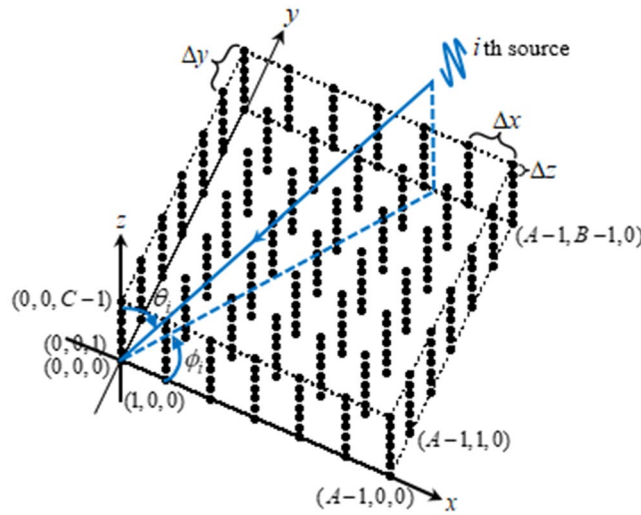
**Figure 8.** Measurement layout.

$$\mathbf{P} = \sum_{i=1}^{I} \mathbf{d}(t_{1i},\ t_{2i},\ t_{3i})\mathbf{s}_i. \tag{3}$$

Denoting by $\mathbf{N} \in \mathrm{C}^{ABC \times L}$ the additive noise, the measurement matrix $\mathbf{P}^{\star} \in \mathrm{C}^{ABC \times L}$ is described by

$$\mathbf{P}^{\star} = \mathbf{P} + \mathbf{N}. \tag{4}$$

The primary focus of this paper is to retrieve the DOA and strength information of sources with $\mathbf{P}^{\star}$ and the estimated noise level as inputs and without discretizing the target region into a grid. Equation (4) becomes $\mathbf{P}_{\Omega}^{\star} = \mathbf{P}_{\Omega} + \mathbf{N}_{\Omega}$ for sparse cuboid microphone arrays, where $\Omega$ denotes the set of the indices of the retained microphones, $\mathbf{P}_{\Omega}^{\star} \in \mathrm{C}^{|\Omega| \times L}$ denotes the matrix of the signals measured by the retained microphones, $\mathbf{P}_{\Omega} \in \mathrm{C}^{|\Omega| \times L}$ denotes the matrix of the signals induced by sources at the retained microphones, $\mathbf{N}_{\Omega} \in \mathrm{C}^{|\Omega| \times L}$ denotes the matrix of the noises born by the retained microphones, and $|\Omega|$ denotes the cardinality of $\Omega$. This paper also focuses on the applicability of the developed beamformer to the sparse cuboid microphone arrays.

**ANM.** Let $s_i = \|\mathbf{s}_i\|_2 \in \mathrm{R}^+$ and $\psi_i = \mathbf{s}_i / s_i \in \mathrm{C}^{1 \times L}$, where $\mathrm{R}^+$ is the set of positive real numbers and $\|\psi_i\|_2 = 1$. We reformulate Eq. (3) as

$$\mathbf{P} = \sum_{i=1}^{I} s_i \mathbf{d}(t_{1i},\ t_{2i},\ t_{3i})\psi_i. \tag{5}$$

Under the continuous setting, $t_1 \equiv \sin\theta \cos\phi \Delta x / \lambda$, $t_2 \equiv \sin\theta \sin\phi \Delta y / \lambda$, $t_3 \equiv \cos\theta \Delta z / \lambda$ and the elements in $\psi$ all can be regarded as the continuous functions of $\theta$ and $\phi$. In accordance to ref.[27], $\mathbf{d}(t_1, t_2, t_3)\psi$ is the atom of the signal model in Eq. (5). The infinite atomic set is expressed as

$$\mathrm{A} = \left\{ \mathbf{d}(t_1,\ t_2,\ t_3)\psi \middle| \begin{array}{l} t_1 \equiv \sin\theta \cos\phi \Delta x / \lambda,\ t_2 \equiv \sin\theta \sin\phi \Delta y / \lambda,\ t_3 \equiv \cos\theta \Delta z / \lambda, \\ \theta \in [0°,\ 180°],\ \phi \in [0°,\ 360°],\ \psi \in \mathrm{C}^{1 \times L},\ \|\psi\|_2 = 1 \end{array} \right\}. \tag{6}$$

The atomic $\ell_0$ norm and the atomic norm of $\mathbf{P}$ are respectively defined as

$$\|\mathbf{P}\|_{\mathrm{A},0} = \inf_{\substack{\mathbf{d}(t_{1i}, t_{2i}, t_{3i})\psi_i \in \mathrm{A} \\ s_i \in \mathrm{R}^+}} \left\{ I \middle| \mathbf{P} = \sum_{i=1}^{I} s_i \mathbf{d}(t_{1i},\ t_{2i},\ t_{3i})\psi_i \right\}. \tag{7}$$

and

$$\|\mathbf{P}\|_{\mathrm{A}} = \inf_{\substack{\mathbf{d}(t_{1i}, t_{2i}, t_{3i})\psi_i \in \mathrm{A} \\ s_i \in \mathrm{R}^+}} \left\{ \sum_{i} s_i \middle| \mathbf{P} = \sum_{i} s_i \mathbf{d}(t_{1i},\ t_{2i},\ t_{3i})\psi_i \right\}, \tag{8}$$

where inf denotes the infimum. $\|\mathbf{P}\|_{\mathrm{A}}$ is a convex relaxation of $\|\mathbf{P}\|_{\mathrm{A},0}$[11,18,20].

Exploiting the source sparsity to solve Eq. (4) can denoise $\mathbf{P}^\star$ and thus obtain $\mathbf{P}$. $\|\mathbf{P}\|_{A,0}$ is the direct metric of source sparsity[11,18,20]. However, it is non-deterministic polynomial-time hard to solve Eq. (4) with the minimization of $\|\mathbf{P}\|_{A,0}$ as a constraint. Replacing $\|\mathbf{P}\|_{A,0}$ with $\|\mathbf{P}\|_A$, we write the reconstruction problem of $\mathbf{P}$ as

$$\hat{\mathbf{P}} = \underset{\mathbf{P} \in C^{ABC \times L}}{\arg\min} \|\mathbf{P}\|_A \qquad \text{subject to} \|\mathbf{P}^\star - \mathbf{P}\|_F \leq \varepsilon, \tag{9}$$

where $\varepsilon$ is the estimated noise level. Theoretically, let $\varepsilon = \|\mathbf{N}\|_F$. Equation (9) is a convex optimization problem[26]. For sparse cuboid microphone arrays, $\mathbf{P}^\star - \mathbf{P}$ in Eq. (9) becomes $\mathbf{P}^\star_\Omega - \mathbf{P}_\Omega$. In this case, the signals induced by sources at full microphones, $\mathbf{P}$, are reconstructed from the measurements of partial microphones, $\mathbf{P}^\star_\Omega$.

## Positive semidefinite programming to solve ANM.

For a given column vector $\mathbf{u} = \left[u_{\alpha_1,\alpha_2,\alpha_3} | (\alpha_1, \alpha_2, \alpha_3) \in \mathrm{H}\right] \in C^{N_u}$ with $\mathrm{H} = (\{0\} \times \{0\} \times \{0, 1, \ldots, C-1\}) \cup (((\{0\} \times \{1, 2, \cdots, B-1\}) \cup (\{1, 2, \ldots, A-1\} \times \{1-B, 2-B, \ldots, B-1\})) \times \{1-C, 2-C, \ldots, C-1\})$ being the halfspace[28] of $(A-1, B-1, C-1)$ and $N_u = C + (B-1 + (A-1)(2B-1))(2C-1)$, we define a three-fold Toeplitz operator $T_{bb}(\cdot)$ to map $\mathbf{u}$ into a Hermitian $A \times A$ block Toeplitz matrix:

$$T_{bb}(\mathbf{u}) = \begin{bmatrix} \mathbf{T}_0 & \mathbf{T}_1^H & \cdots & \mathbf{T}_{A-1}^H \\ \mathbf{T}_1 & \mathbf{T}_0 & \cdots & \mathbf{T}_{A-2}^H \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{T}_{A-1} & \mathbf{T}_{A-2} & \cdots & \mathbf{T}_0 \end{bmatrix}, \tag{10}$$

where $(\cdot)^H$ denotes the Hermitian operator, each block $\mathbf{T}_{\alpha_1}(0 \leq \alpha_1 \leq A-1)$ is a $B \times B$ block Toeplitz matrix:

$$\mathbf{T}_{\alpha_1} = \begin{bmatrix} \mathbf{T}_{\alpha_1,0} & \mathbf{T}_{\alpha_1,-1} & \cdots & \mathbf{T}_{\alpha_1,1-B} \\ \mathbf{T}_{\alpha_1,1} & \mathbf{T}_{\alpha_1,0} & \cdots & \mathbf{T}_{\alpha_1,2-B} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{T}_{\alpha_1,B-1} & \mathbf{T}_{\alpha_1,B-2} & \cdots & \mathbf{T}_{\alpha_1,0} \end{bmatrix}, \tag{11}$$

and $\mathbf{T}_0$ is Hermitian, i.e., $\mathbf{T}_{0,-\alpha_2} = \mathbf{T}_{0,\alpha_2}^H (0 \leq \alpha_2 \leq B-1)$. Each block $\mathbf{T}_{\alpha_1,\alpha_2}(1-B \leq \alpha_2 \leq B-1)$ in Eq. (11) is a $C \times C$ Toeplitz matrix:

$$\mathbf{T}_{\alpha_1,\alpha_2} = \begin{bmatrix} u_{\alpha_1,\alpha_2,0} & u_{\alpha_1,\alpha_2,-1} & \cdots & u_{\alpha_1,\alpha_2,1-C} \\ u_{\alpha_1,\alpha_2,1} & u_{\alpha_1,\alpha_2,0} & \cdots & u_{\alpha_1,\alpha_2,2-C} \\ \vdots & \vdots & \ddots & \vdots \\ u_{\alpha_1,\alpha_2,C-1} & u_{\alpha_1,\alpha_2,C-2} & \cdots & u_{\alpha_1,\alpha_2,0} \end{bmatrix}, \tag{12}$$

and $\mathbf{T}_{0,0}$ is Hermitian, i.e., $u_{0,0,-\alpha_3} = u_{0,0,\alpha_3}^*(0 \leq \alpha_3 \leq C-1)$.

We propose the following proposition.

*Proposition:* Denote

$$\{\hat{\mathbf{u}}, \hat{\mathbf{E}}\} = \underset{\mathbf{u} \in C^{N_u}, \mathbf{E} \in C^{L \times L}}{\arg\min} \frac{1}{2\sqrt{ABC}}(\text{tr}(T_{bb}(\mathbf{u})) + \text{tr}(\mathbf{E})) \quad \text{subject to} \begin{bmatrix} T_{bb}(\mathbf{u}) & \mathbf{P} \\ \mathbf{P}^H & \mathbf{E} \end{bmatrix} \geq 0. \tag{13}$$

and

$$\|\mathbf{P}\|_T = \frac{1}{2\sqrt{ABC}}(\text{tr}(T_{bb}(\hat{\mathbf{u}})) + \text{tr}(\hat{\mathbf{E}})), \tag{14}$$

where $\text{tr}(\cdot)$ represents the trace and $\geq 0$ means the matrix is positive semidefinite. If $T_{bb}(\hat{\mathbf{u}})$ admits a Vandermonde decomposition[13,14], i.e.,

$$T_{bb}(\hat{\mathbf{u}}) = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^H, \tag{15}$$

where $\mathbf{V} = [\mathbf{d}(t_{11}, t_{21}, t_{31}), \mathbf{d}(t_{12}, t_{22}, t_{32}), \ldots, \mathbf{d}(t_{1r}, t_{2r}, t_{3r})], \mathbf{\Sigma} = \text{diag}([\sigma_1, \sigma_2, \ldots, \sigma_r])$, $\text{diag}(\cdot)$ forms a diagonal matrix with diagonal being the vector in the brackets, $\sigma_i(i = 1, 2, \ldots, r) \in R^+$, and $r$ is the rank of $T_{bb}(\hat{\mathbf{u}})$, then $\|\mathbf{P}\|_T = \|\mathbf{P}\|_A$.

The proof about the proposition can be found in Supplementary Note 3. Based on the proposition, the following positive semidefinite programming can be used to characterize the ANM in Eq. (9):

$$\{\hat{\mathbf{u}}, \hat{\mathbf{P}}, \hat{\mathbf{E}}\} = \underset{\mathbf{u} \in C^{N_u}, \mathbf{P} \in C^{ABC \times L}, \mathbf{E} \in C^{L \times L}}{\arg\min} \frac{1}{2\sqrt{ABC}} \times (\text{tr}(T_{bb}(\mathbf{u})) + \text{tr}(\mathbf{E}))$$

$$\text{subject to} \begin{bmatrix} T_{bb}(\mathbf{u}) & \mathbf{P} \\ \mathbf{P}^H & \mathbf{E} \end{bmatrix} \geq 0, \quad \left\|\mathbf{P}^\star - \mathbf{P}\right\|_F \leq \varepsilon \tag{16}$$

The Vandermonde decomposition of $T_{bb}(\hat{\mathbf{u}})$ is the precondition to make Eqs (16) and (9) strictly equivalent. References[13,14] has proved that $r \leq \min\{A, B, C\}$ is the sufficient condition for $T_{bb}(\hat{\mathbf{u}})$ to admit a Vandermonde decomposition. For sparse cuboid microphone arrays, $\mathbf{P}^{\star} - \mathbf{P}$ in Eq. (16) becomes $\mathbf{P}_{\Omega}^{\star} - \mathbf{P}_{\Omega}$.

**ADMM to solve positive semidefinite programming.** Because $\mathbf{P}^{\star} - \mathbf{P}$ is a special case of $\mathbf{P}_{\Omega}^{\star} - \mathbf{P}_{\Omega}$ with $\Omega$ including the indices of the full microphones, we conduct the derivation on $\mathbf{P}_{\Omega}^{\star} - \mathbf{P}_{\Omega}$. We reformulate Eq. (16) as

$$\{\hat{\mathbf{u}}, \hat{\mathbf{P}}, \hat{\mathbf{E}}, \hat{\mathbf{Z}}\} = \underset{\substack{\mathbf{u}\in C^{N_u}, \mathbf{P}\in C^{ABC\times L}, \mathbf{E}\in C^{L\times L} \\ \mathbf{Z}\in C^{(ABC+L)\times(ABC+L)}}}{\arg\min} \frac{1}{2}\|\mathbf{P}_{\Omega}^{\star} - \mathbf{P}_{\Omega}\|_{F}^{2} + \frac{\tau}{2\sqrt{ABC}}(\mathrm{tr}(T_{bb}(\mathbf{u})) + \mathrm{tr}(\mathbf{E}))$$

$$\text{subject to } \mathbf{Z} = \begin{bmatrix} T_{bb}(\mathbf{u}) & \mathbf{P} \\ \mathbf{P}^{H} & \mathbf{E} \end{bmatrix}, \mathbf{Z} \geq 0 \tag{17}$$

where $\tau$ is a regularization parameter and $\mathbf{Z}$ is an auxiliary matrix. The augmented Lagrangian function of Eq. (17) is

$$L_{\rho}(\mathbf{E}, \mathbf{u}, \mathbf{P}, \mathbf{Z}, \boldsymbol{\Lambda}) = \frac{1}{2}\|\mathbf{P}_{\Omega}^{\star} - \mathbf{P}_{\Omega}\|_{F}^{2} + \frac{\tau}{2\sqrt{ABC}}(\mathrm{tr}(T_{bb}(\mathbf{u})) + \mathrm{tr}(\mathbf{E}))$$

$$+ \left\langle \boldsymbol{\Lambda}, \mathbf{Z} - \begin{bmatrix} T_{bb}(\mathbf{u}) & \mathbf{P} \\ \mathbf{P}^{H} & \mathbf{E} \end{bmatrix} \right\rangle + \frac{\rho}{2}\left\| \mathbf{Z} - \begin{bmatrix} T_{bb}(\mathbf{u}) & \mathbf{P} \\ \mathbf{P}^{H} & \mathbf{E} \end{bmatrix} \right\|_{F}^{2}, \tag{18}$$

where the Hermitian matrix $\boldsymbol{\Lambda} \in C^{(ABC+L)\times(ABC+L)}$ is the Lagrangian multiplier, $\rho > 0$ is the penalty parameter and $\langle \cdot, \cdot \rangle$ denotes the inner product. The ADMM solves Eq. (17) iteratively. Initializing $\mathbf{Z}^0 = \boldsymbol{\Lambda}^0 = \mathbf{0}$, the updates in $(q + 1)$th iteration are as follows:

$$\{\mathbf{E}^{q+1}, \mathbf{u}^{q+1}, \mathbf{P}^{q+1}\} = \underset{\mathbf{E}\in C^{L\times L}, \mathbf{u}\in C^{N_u}, \mathbf{P}\in C^{ABC\times L}}{\arg\min} L_{\rho}(\mathbf{E}, \mathbf{u}, \mathbf{P}, \mathbf{Z}^q, \boldsymbol{\Lambda}^q), \tag{19}$$

$$\mathbf{Z}^{q+1} = \underset{\mathbf{Z}\geq 0}{\arg\min} L_{\rho}(\mathbf{E}^{q+1}, \mathbf{u}^{q+1}, \mathbf{P}^{q+1}, \mathbf{Z}, \boldsymbol{\Lambda}^q), \tag{20}$$

$$\boldsymbol{\Lambda}^{q+1} = \boldsymbol{\Lambda}^q + \rho\left(\mathbf{Z}^{q+1} - \begin{bmatrix} T_{bb}(\mathbf{u}^{q+1}) & \mathbf{P}^{q+1} \\ (\mathbf{P}^{q+1})^{H} & \mathbf{E}^{q+1} \end{bmatrix}\right). \tag{21}$$

Introduce the partitions

$$\boldsymbol{\Lambda}^q = \begin{bmatrix} \boldsymbol{\Lambda}_0^q \in C^{ABC\times ABC} & \boldsymbol{\Lambda}_1^q \in C^{ABC\times L} \\ \boldsymbol{\Lambda}_1^{qH} \in C^{L\times ABC} & \boldsymbol{\Lambda}_2^q \in C^{L\times L} \end{bmatrix} \text{ and } \mathbf{Z}^q = \begin{bmatrix} \mathbf{Z}_0^q \in C^{ABC\times ABC} & \mathbf{Z}_1^q \in C^{ABC\times L} \\ \mathbf{Z}_1^{qH} \in C^{L\times ABC} & \mathbf{Z}_2^q \in C^{L\times L} \end{bmatrix} \tag{22}$$

Denote by $\boldsymbol{\Lambda}_{1\Omega}^q \in C^{|\Omega|\times L}$ and $\mathbf{Z}_{1\Omega}^q \in C^{|\Omega|\times L}$, respectively, the matrices of the rows in $\boldsymbol{\Lambda}_1^q$ and $\mathbf{Z}_1^q$ corresponding to the retained microphones, $\Omega^c$ the set of the indices of the unretained microphones, $|\Omega^c|$ the cardinality of $\Omega^c$, $\mathbf{P}_{\Omega^c} \in C^{|\Omega^c|\times L}$ the matrix of the signals induced by sources at the unretained microphones, $\boldsymbol{\Lambda}_{1\Omega^c}^q \in C^{|\Omega^c|\times L}$ and $\mathbf{Z}_{1\Omega^c}^q \in C^{|\Omega^c|\times L}$, respectively, the matrices of the rows in $\boldsymbol{\Lambda}_1^q$ and $\mathbf{Z}_1^q$ corresponding to the unretained microphones, and $\mathbf{I}_1 \in R^{L\times L}$ and $\mathbf{I}_2 \in R^{ABC\times ABC}$, both, the identity matrices. It can be derived that the variable updates in Eq. (19) have the following closed forms:

$$\mathbf{E}^{q+1} = \mathbf{Z}_2^q + \frac{1}{\rho}\left(\boldsymbol{\Lambda}_2^q - \frac{\tau}{2\sqrt{ABC}}\mathbf{I}_1\right), \tag{23}$$

$$T_{bb}(\mathbf{u})^{q+1} = \mathbf{Z}_0^q + \frac{1}{\rho}\left(\boldsymbol{\Lambda}_0^q - \frac{\tau}{2\sqrt{ABC}}\mathbf{I}_2\right), \tag{24}$$

$$\mathbf{P}_{\Omega}^{q+1} = \frac{1}{1+2\rho}(\mathbf{P}_{\Omega}^{\star} + 2\boldsymbol{\Lambda}_{1\Omega}^q + 2\rho\mathbf{Z}_{1\Omega}^q), \tag{25}$$

$$\mathbf{P}_{\Omega^c}^{q+1} = \mathbf{Z}_{1\Omega^c}^q + \frac{1}{\rho}\boldsymbol{\Lambda}_{1\Omega^c}^q. \tag{26}$$

The derivation can be found in Supplementary Note 4. Let $\mathbf{M} = \mathrm{diag}([ABC, AB(C-1), \cdots, AB, [A(B-1), A(B-2), \cdots, A, [A-1, A-2, \cdots, 1] \otimes [1, 2, \cdots, B, B-1, \cdots, 1]]\otimes[1, 2, \cdots, C, C-1, \cdots, 1]])\in R^{N_u\times N_u}$

---

and $T_{bb}^*(\cdot)$ be the adjoint of $T_{bb}(\cdot)$. For a given matrix $\mathbf{A} \in C^{ABC \times ABC}$, $T_{bb}^*(\mathbf{A}) = [\text{tr}\left(\left(\Theta_{\alpha_1} \otimes \Theta_{\alpha_2} \otimes \Theta_{\alpha_3}\right)\mathbf{A}\right)$ $|(\alpha_1, \alpha_2, \alpha_3) \in H] \in C^{N_u}$, where $\Theta_\alpha$ is an elementary Toeplitz matrix with ones on the $\alpha$th diagonal and zeros elsewhere. Then,

$$\mathbf{u}^{q+1} = \mathbf{M}^{-1}T_{bb}^*(T_{bb}(\mathbf{u})^{q+1}). \tag{27}$$

We rewritten the update of $\mathbf{Z}$ in Eq. (20) as

$$\mathbf{Z}^{q+1} = \arg\min_{\mathbf{Z} \geq 0} \left\| \mathbf{Z} - \begin{bmatrix} T_{bb}(\mathbf{u}^{q+1}) & \mathbf{P}^{q+1} \\ (\mathbf{P}^{q+1})^H & \mathbf{E}^{q+1} \end{bmatrix} + \frac{\Lambda^q}{\rho} \right\|_F^2, \tag{28}$$

which can be performed by conducting the eigenvalue decomposition of the Hermitian matrix $\begin{bmatrix} T_{bb}(\mathbf{u}^{q+1}) & \mathbf{P}^{q+1} \\ (\mathbf{P}^{q+1})^H & \mathbf{E}^{q+1} \end{bmatrix} - \frac{\Lambda^q}{\rho}$ and setting all negative eigenvalues to zero.

A reasonable termination criterion of ADMM is that the primal and dual residuals or infeasibilities must be small[22–24]. For the semidefinite programming, the primal and dual infeasibilities are recommended for use[23,24]. According to refs[23,24], the primal and dual infeasibilities after $q$ iterations for the current problem can be respectively defined as

$$\text{pri\_inf}^q = \frac{\left\| \mathbf{Z}^q - \begin{bmatrix} T_{bb}(\mathbf{u}^q) & \mathbf{P}^q \\ (\mathbf{P}^q)^H & \mathbf{E}^q \end{bmatrix} \right\|_F}{ABC + L + \max\left( \|\mathbf{Z}^q\|_F, \left\| \begin{bmatrix} T_{bb}(\mathbf{u}^q) & \mathbf{P}^q \\ (\mathbf{P}^q)^H & \mathbf{E}^q \end{bmatrix} \right\|_F \right)}. \tag{29}$$

and

$$\text{dual\_inf}^q = \frac{\rho\sqrt{\|T_{bb}^*(\mathbf{Z}_0^q - \mathbf{Z}_0^{q-1})\|_2^2 + \|\mathbf{Z}_1^q - \mathbf{Z}_1^{q-1}\|_F^2 + \|\mathbf{Z}_2^q - \mathbf{Z}_2^{q-1}\|_F^2}}{\sqrt{N_u + ABCL + L^2} + \rho\sqrt{\|T_{bb}^*(\Lambda_0^q - \Lambda_0^{q-1})\|_2^2 + \|\Lambda_1^q - \Lambda_1^{q-1}\|_F^2 + \|\Lambda_2^q - \Lambda_2^{q-1}\|_F^2}}. \tag{30}$$

The algorithm terminates when $\max\{\text{pri\_inf}^q, \text{dual\_inf}^q\}$ is less than a pre-set tolerance level or the pre-set maximum number of iterations is reached. Another reasonable termination criterion is that the relative change of the result at two consecutive iterations is small enough, which means the result has converged.

**DOA estimation via MaPP.** $T_{bb}(\hat{\mathbf{u}})$ contains the DOA information of sources. That can be retrieved via the MaPP method[14,20]. The concrete steps are:

1) Conduct the eigendecomposition on $T_{bb}(\hat{\mathbf{u}})$:

$$T_{bb}(\hat{\mathbf{u}}) = \mathbf{U}_1\mathbf{C}_1\mathbf{U}_1^H, \tag{31}$$

where $\mathbf{U}_1 \in C^{ABC \times ABC}$ is the unitary matrix whose columns are the eigenvectors of $T_{bb}(\hat{\mathbf{u}})$ and $\mathbf{C}_1 \in R^{ABC \times ABC}$ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues. Estimate the total number of sources as the number of eigenvalues larger than a given threshold and denote it by $\hat{I}$. Denote by $\mathbf{C}_{1e} \in R^{\hat{I} \times \hat{I}}$ the diagonal matrix with diagonal elements being the square roots of the $\hat{I}$ larger eigenvalues and $\mathbf{U}_{1e} \in C^{ABC \times \hat{I}}$ the matrix with columns being the corresponding eigenvectors. Let $\mathbf{Y}_1 = \mathbf{U}_{1e}\mathbf{C}_{1e} \in C^{ABC \times \hat{I}}$.

2) Delete the last $BC$ rows of $\mathbf{Y}_1$ to obtain $\mathbf{Y}_{1u} \in C^{(A-1)BC \times \hat{I}}$ and the first $BC$ rows to obtain $\mathbf{Y}_{1d} \in C^{(A-1)BC \times \hat{I}}$. Compute the generalized eigenvalues of the matrix pencil $(\mathbf{Y}_{1d}, \mathbf{Y}_{1u})$ to obtain $\{e^{j2\pi t_{1m}} | m = 1, 2, \cdots, \hat{I}\}$.

3) Let $\rho(\alpha) \in R^{ABC}$ be the column vector with one on the $\alpha$th position and zeros elsewhere, $\rho_{\alpha_1,\alpha_2,\alpha_3} = \rho(\alpha_1 C + \alpha_2 AC + \alpha_3 + 1)$, $\mathbf{P}_{\alpha_1,\alpha_2} = [\rho_{\alpha_1,\alpha_2,0}, \rho_{\alpha_1,\alpha_2,1}, \cdots, \rho_{\alpha_1,\alpha_2,C-1}] \in R^{ABC \times C}$, $\mathbf{P}_{\alpha_1} = [\mathbf{P}_{\alpha_1,0}, \mathbf{P}_{\alpha_1,1}, \cdots, \mathbf{P}_{\alpha_1,B-1}] \in R^{ABC \times BC}$, and $\mathbf{P} = [\mathbf{P}_0, \mathbf{P}_1, \cdots, \mathbf{P}_{A-1}] \in R^{ABC \times ABC}$. $\mathbf{P}T_{bb}(\hat{\mathbf{u}})\mathbf{P}^T$ permutes the elements of $T_{bb}(\hat{\mathbf{u}})$ to obtain a new three-fold Toeplitz matrix. The first fold contains $B \times B$ big blocks. The second fold contains $A \times A$ small blocks. The third fold is the $C \times C$ matrix. Conduct the eigendecomposition on $\mathbf{P}T_{bb}(\hat{\mathbf{u}})\mathbf{P}^T$:

$$\mathbf{P}T_{bb}(\hat{\mathbf{u}})\mathbf{P}^T = \mathbf{U}_2\mathbf{C}_2\mathbf{U}_2^H, \tag{32}$$

where $\mathbf{U}_2 \in C^{ABC \times ABC}$ is the unitary matrix whose columns are the eigenvectors of $\mathbf{P}T_{bb}(\hat{\mathbf{u}})\mathbf{P}^T$ and $\mathbf{C}_2 \in R^{ABC \times ABC}$ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues. Denote by $\mathbf{C}_{2e} \in R^{\hat{I} \times \hat{I}}$ the diagonal matrix with diagonal elements being the square roots of the $\hat{I}$ larger eigenvalues and $\mathbf{U}_{2e} \in C^{ABC \times \hat{I}}$ the matrix with columns being the corresponding eigenvectors. Let $\mathbf{Y}_2 = \mathbf{U}_{2e}\mathbf{C}_{2e} \in C^{ABC \times \hat{I}}$.

4) Delete the last $AC$ rows of $\mathbf{Y}_2$ to obtain $\mathbf{Y}_{2u} \in C^{A(B-1)C \times \hat{I}}$ and the first $AC$ rows to obtain $\mathbf{Y}_{2d} \in C^{A(B-1)C \times \hat{I}}$. Compute the generalized eigenvalues of the matrix pencil $(\mathbf{Y}_{2d}, \mathbf{Y}_{2u})$ to obtain $\{e^{j2\pi t_{2n}} | n = 1, 2, \cdots, \hat{I}\}$.

5) Let $\widetilde{\boldsymbol{\rho}}_{\alpha_1,\alpha_2,\alpha_3} = \boldsymbol{\rho}(1 + \alpha_1 + \alpha_2 A + \alpha_3 AB)$, $\widetilde{\mathbf{P}}_{\alpha_1,\alpha_2} = \left[\widetilde{\boldsymbol{\rho}}_{\alpha_1,\alpha_2,0}, \widetilde{\boldsymbol{\rho}}_{\alpha_1,\alpha_2,1}, \cdots, \widetilde{\boldsymbol{\rho}}_{\alpha_1,\alpha_2,C-1}\right] \in \mathrm{R}^{ABC \times C}$, $\widetilde{\mathbf{P}}_{\alpha_1} = \left[\widetilde{\mathbf{P}}_{\alpha_1,0}, \widetilde{\mathbf{P}}_{\alpha_1,1}, \cdots, \widetilde{\mathbf{P}}_{\alpha_1,B-1}\right] \in \mathrm{R}^{ABC \times BC}$, and $\widetilde{\mathbf{P}} = [\widetilde{\mathbf{P}}_0, \widetilde{\mathbf{P}}_1, \cdots, \widetilde{\mathbf{P}}_{A-1}] \in \mathrm{R}^{ABC \times ABC}$. $\widetilde{\mathbf{P}}T_{bb}(\hat{\mathbf{u}})\widetilde{\mathbf{P}}^{\mathrm{T}}$ permutes the elements of $T_{bb}(\hat{\mathbf{u}})$ to obtain another new three-fold Toeplitz matrix. The first fold contains $C \times C$ big blocks. The second fold contains $B \times B$ small blocks. The third fold is the $A \times A$ matrix. Conduct the eigendecomposition on $\widetilde{\mathbf{P}}T_{bb}(\hat{\mathbf{u}})\widetilde{\mathbf{P}}^{\mathrm{T}}$:

$$\widetilde{\mathbf{P}}T_{bb}(\hat{\mathbf{u}})\widetilde{\mathbf{P}}^{\mathrm{T}} = \mathbf{U}_3\mathbf{C}_3\mathbf{U}_3^{\mathrm{H}}, \tag{33}$$

where $\mathbf{U}_3 \in \mathrm{C}^{ABC \times ABC}$ is the unitary matrix whose columns are the eigenvectors of $\widetilde{\mathbf{P}}T_{bb}(\hat{\mathbf{u}})\widetilde{\mathbf{P}}^{\mathrm{T}}$ and $\mathbf{C}_3 \in \mathrm{R}^{ABC \times ABC}$ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues. Denote by $\mathbf{C}_{3e} \in \mathrm{R}^{\hat{I} \times \hat{I}}$ the diagonal matrix with diagonal elements being the square roots of the $\hat{I}$ larger eigenvalues and $\mathbf{U}_{3e} \in \mathrm{C}^{ABC \times \hat{I}}$ the matrix with columns being the corresponding eigenvectors. Let $\mathbf{Y}_3 = \mathbf{U}_{3e}\mathbf{C}_{3e} \in \mathrm{C}^{ABC \times \hat{I}}$.

6) Delete the last $AB$ rows of $\mathbf{Y}_3$ to obtain $\mathbf{Y}_{3u} \in \mathrm{C}^{AB(C-1) \times \hat{I}}$ and the first $AB$ rows to obtain $\mathbf{Y}_{3d} \in \mathrm{C}^{AB(C-1) \times \hat{I}}$. Compute the generalized eigenvalues of the matrix pencil $(\mathbf{Y}_{3d}, \mathbf{Y}_{3u})$ to obtain $\{e^{j2\pi t_{3o}}|o = 1, 2, \cdots, \hat{I}\}$.

7) Compute the function

$$(f(m), g(m)) = \arg\max_{(n,o)\in\{1,2,\cdots,\hat{I}\}\times\{1,2,\cdots,\hat{I}\}} \left\|\mathbf{U}_{1e}^{\mathrm{H}}\mathbf{d}(t_{1m}, t_{2n}, t_{3o})\right\|_2^2 \tag{34}$$

to pair $m$, $n$ and $o$, and then obtain the pairs $\{(e^{j2\pi t_{1m}}, e^{j2\pi t_{2f(m)}}, e^{j2\pi t_{3g(m)}})|m = 1, 2, \cdots, \hat{I}\}$ (denoted as $\{(e^{j2\pi t_{1i}}, e^{j2\pi t_{2i}}, e^{j2\pi t_{3i}})|i = 1, 2, \cdots, \hat{I}\}$ for simplicity).

8) Compute $\{(t_{1i}, t_{2i}, t_{3i})|i = 1, 2, \cdots, \hat{I}\}$ according to $t_{1i} = \mathrm{Im}(\ln(e^{j2\pi t_{1i}}))/2\pi$, $t_{2i} = \mathrm{Im}(\ln(e^{j2\pi t_{2i}}))/2\pi$ and $t_{3i} = \mathrm{Im}(\ln(e^{j2\pi t_{3i}}))/2\pi$, where $\mathrm{Im}(\cdot)$ denotes the imaginary part. Compute $\{(\theta_i, \phi_i)|i = 1, 2, \cdots, \hat{I}\}$ according to the relationship between $(t_{1i}, t_{2i}, t_{3i})$ and $(\theta_i, \phi_i)$, i.e., $t_{1i} \equiv \sin\theta_i\cos\phi_i\Delta x/\lambda$, $t_{2i} \equiv \sin\theta_i\sin\phi_i\Delta y/\lambda$ and $t_{3i} \equiv \cos\theta_i\Delta z/\lambda$.

**Source strength quantification.** After computing the sensing matrix $\hat{\mathbf{D}}=[\mathbf{d}(t_{11}, t_{21}, t_{31}), \mathbf{d}(t_{12}, t_{22}, t_{32}), \cdots, \mathbf{d}(t_{1\hat{I}}, t_{2\hat{I}}, t_{3\hat{I}})]\in\mathrm{C}^{ABC \times \hat{I}}$ according to the estimated DOAs, we quantify the matrix $\hat{\mathbf{S}} = \left[\mathbf{s}_1^{\mathrm{T}}, \mathbf{s}_2^{\mathrm{T}}, \cdots, \mathbf{s}_{\hat{I}}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathrm{C}^{\hat{I} \times L}$ composed by the strength of each source under each snapshot as

$$\hat{\mathbf{S}} = \hat{\mathbf{D}}^+\hat{\mathbf{P}}, \tag{35}$$

where $\cdot^+$ denotes the pseudo-inverse.

**IRANM.** Establish the metric

$$\mathrm{M}^{\kappa}(\mathbf{P}) = \min_{\mathbf{u}\in\mathrm{C}^{N_u},\mathbf{E}\in\mathrm{C}^{L\times L}} \frac{1}{2\sqrt{ABC}}(\ln|T_{bb}(\mathbf{u}) + \kappa\mathbf{I}_2| + \mathrm{tr}(\mathbf{E})) \text{ subject to } \begin{bmatrix} T_{bb}(\mathbf{u}) & \mathbf{P} \\ \mathbf{P}^{\mathrm{H}} & \mathbf{E} \end{bmatrix} \geq 0, \tag{36}$$

where $|\cdot|$ denotes the determinant of a matrix and $\kappa > 0$ is a regularization parameter. $\mathrm{M}^{\kappa}(\mathbf{P})$ has the following properties under a certain condition[11,14,18]: (1) $\mathrm{M}^{\kappa}(\mathbf{P})$ is on the order of $(\|\mathbf{P}\|_{\mathrm{A},0}/2\sqrt{ABC} - \sqrt{ABC}/2)\ln\kappa^{-1}$ as $\kappa$ approaches 0 if $\|\mathbf{P}\|_{\mathrm{A},0} < ABC$, i.e., $\lim_{\kappa\to 0}\mathrm{M}^{\kappa}(\mathbf{P})/((\|\mathbf{P}\|_{\mathrm{A},0}/2\sqrt{ABC} - \sqrt{ABC}/2)\ln\kappa^{-1}) = 1$; (2) $\mathrm{M}^{\kappa}(\mathbf{P}) - (\sqrt{ABC}/2)\ln\kappa$ is on the order of $\|\mathbf{P}\|_{\mathrm{A}}\kappa^{-1/2}$ as $\kappa$ approaches $+\infty$, i.e., $\lim_{\kappa\to+\infty}(\mathrm{M}^{\kappa}(\mathbf{P}) - (\sqrt{ABC}/2)\ln\kappa)/(\|\mathbf{P}\|_{\mathrm{A}}\kappa^{-1/2}) = 1$; (3) Denote by $\hat{\mathbf{u}}_{\kappa\to 0}$ the optimal variable when $\kappa$ approaches 0. The smallest $ABC - \|\mathbf{P}\|_{\mathrm{A},0}$ eigenvalues of $T_{bb}(\hat{\mathbf{u}}_{\kappa\to 0})$ are either 0 or approach 0, i.e., only $\|\mathbf{P}\|_{\mathrm{A},0}$ eigenvalues are large. According to the first property, minimizing $\mathrm{M}^{\kappa}(\mathbf{P})$ is equivalent to minimizing $\|\mathbf{P}\|_{\mathrm{A},0}$ as $\kappa$ approaches 0. According to the second property, minimizing $\mathrm{M}^{\kappa}(\mathbf{P})$ is equivalent to minimizing $\|\mathbf{P}\|_{\mathrm{A}}$ as $\kappa$ approaches $+\infty$. This means $\mathrm{M}^{\kappa}(\mathbf{P})$ serves as a bridge between $\|\mathbf{P}\|_{\mathrm{A},0}$ and $\|\mathbf{P}\|_{\mathrm{A}}$, and can enhance sparsity compared with $\|\mathbf{P}\|_{\mathrm{A}}$. According to the third property, the sufficient condition for $T_{bb}(\hat{\mathbf{u}}_{\kappa\to 0})$ to admit a Vandermonde decomposition can be guaranteed.

Replacing $\|\mathbf{P}\|_{\mathrm{A}}$ with $\mathrm{M}^{\kappa}(\mathbf{P})$, we reformulate the reconstruction problem of $\mathbf{P}$ as

$$\hat{\mathbf{P}} = \arg\min_{\mathbf{P}\in\mathrm{C}^{ABC\times L}} \mathrm{M}^{\kappa}(\mathbf{P}) \quad \text{subject to}\|\mathbf{P}^{\star} - \mathbf{P}\|_{\mathrm{F}} \leq \varepsilon. \tag{37}$$

Simultaneous Eqs (36) and (37) yield

$$\begin{aligned}\{\hat{\mathbf{u}}, \hat{\mathbf{P}}, \hat{\mathbf{E}}\} &= \arg\min_{\mathbf{u}\in\mathrm{C}^{N_u},\mathbf{P}\in\mathrm{C}^{ABC\times L},\mathbf{E}\in\mathrm{C}^{L\times L}} \frac{1}{2\sqrt{ABC}} \times (\ln|T_{bb}(\mathbf{u}) + \kappa\mathbf{I}_2| + \mathrm{tr}(\mathbf{E})) \\ &\text{subject to}\begin{bmatrix} T_{bb}(\mathbf{u}) & \mathbf{P} \\ \mathbf{P}^{\mathrm{H}} & \mathbf{E} \end{bmatrix} \geq 0, \|\mathbf{P}^{\star} - \mathbf{P}\|_{\mathrm{F}} \leq \varepsilon\end{aligned} \tag{38}$$

In Eq. (38), $\ln|T_{bb}(\mathbf{u}) + \kappa\mathbf{I}|$ is a concave function of $\mathbf{u}$, while $\mathrm{tr}(\mathbf{E})$ is a convex function of $\mathbf{E}$. To minimize such a concave + convex function, an effective algorithm is the majorization-minimization[29,30]. The algorithm solves

the optimal variable and the optimal value of the objective function iteratively. Two steps are involved in each iteration. In the first majorization step, a surrogate function is constructed to locally approximate the objective function. The surrogate function should be no smaller than the objective function, and the equality holds at the current optimal variable. Then in the minimization step, the surrogate function is minimized to obtain a new optimal variable. Let $\hat{\mathbf{u}}^k$ be the optimal variable, $\kappa^k$ be the regularization parameter, $\mathbf{W}^k \equiv (T_{bb}(\hat{\mathbf{u}}^k) + \kappa^k \mathbf{I}_2)^{-1} \in \mathbf{C}^{ABC \times ABC}$ be the weighting matrix, determined by the $k$th iteration. The surrogate function in $(k+1)$th iteration can be constructed as

$$\frac{1}{2\sqrt{ABC}}(\ln\left|T_{bb}(\hat{\mathbf{u}}^k) + \kappa^k \mathbf{I}_2\right| + \mathrm{tr}(\mathbf{W}^k T_{bb}(\mathbf{u} - \hat{\mathbf{u}}^k)) + \mathrm{tr}(\mathbf{E})) = \frac{1}{2\sqrt{ABC}}(\mathrm{tr}(\mathbf{W}^k T_{bb}(\mathbf{u})) + \mathrm{tr}(\mathbf{E})) + c^k, \quad (39)$$

where $\ln\left|T_{bb}(\hat{\mathbf{u}}^k) + \kappa^k \mathbf{I}_2\right| + \mathrm{tr}(\mathbf{W}^k T_{bb}(\mathbf{u} - \hat{\mathbf{u}}^k))$ is the tangent plane of $\ln\left|T_{bb}(\mathbf{u}) + \kappa^k \mathbf{I}_2\right|$ at $\mathbf{u} = \hat{\mathbf{u}}^k$, and $c^k$ is a constant independent of variables. Ignoring $c^k$, we write the minimization problem in $(k+1)$th iteration as

$$\{\hat{\mathbf{u}}^{k+1}, \hat{\mathbf{P}}^{k+1}, \hat{\mathbf{E}}^{k+1}\} = \underset{\mathbf{u} \in \mathbf{C}^{N_u}, \mathbf{P} \in \mathbf{C}^{ABC \times L}, \mathbf{E} \in \mathbf{C}^{L \times L}}{\arg\min} \frac{1}{2\sqrt{ABC}} \times (\mathrm{tr}(\mathbf{W}^k T_{bb}(\mathbf{u})) + \mathrm{tr}(\mathbf{E}))$$

$$\text{subject to } \begin{bmatrix} T_{bb}(\mathbf{u}) & \mathbf{P} \\ \mathbf{P}^{\mathrm{H}} & \mathbf{E} \end{bmatrix} \geq 0, \ \|\mathbf{P}^{\star} - \mathbf{P}\|_{\mathrm{F}} \leq \varepsilon \quad (40)$$

which is also a disciplined convex optimization problem. Initialize $\hat{\mathbf{u}}^0 = \mathbf{0}$ and $\kappa^0 = 1$, then $\mathbf{W}^0 = \mathbf{I}$. The first iteration in Eq. (40) agrees with the ANM in Eq. (16). For the purpose of enhancing sparsity, $\kappa$ should decrease gradually. The specific strategy is[18]

$$\kappa^k = \begin{cases} \min\{\kappa^{k-1}/2, \ \lambda_{\max}(T_{bb}(\hat{\mathbf{u}}^k))/10\}, & k = 1 \\ \kappa^{k-1}/2, & 2 \leq k \leq 10 \\ \kappa^{10}, & k > 10 \end{cases} \quad (41)$$

where $\lambda_{\max}(T_{bb}(\hat{\mathbf{u}}^k))$ is the largest eigenvalue of $T_{bb}(\hat{\mathbf{u}}^k)$.

We define the weighted atomic norm of $\mathbf{P}$ as

$$\|\mathbf{P}\|_{\mathrm{A}^w} = \underset{\substack{\mathbf{d}(t_{1i}, t_{2i}, t_{3i})\psi_i \in \mathrm{A} \\ s_i \in \mathrm{R}^+}}{\inf} \left\{ \sum_i \frac{s_i}{w(t_{1i}, t_{2i}, t_{3i})} \middle| \mathbf{P} = \sum_i s_i \mathbf{d}(t_{1i}, t_{2i}, t_{3i})\psi_i \right\}, \quad (42)$$

where $w(t_{1i}, t_{2i}, t_{3i}) \geq 0$ is the weighting coefficient. When $w(t_{1i}, t_{2i}, t_{3i}) \equiv 1$, $\|\mathbf{P}\|_{\mathrm{A}^w} = \|\mathbf{P}\|_{\mathrm{A}}$. Denote by $w^k(t_{1i}, t_{2i}, t_{3i})$ and $\|\mathbf{P}\|_{\mathrm{A}^{w^k}}$ the weighting coefficient and the weighted atomic norm corresponding to the results of the $k$th iteration. If

$$w^k(t_{1i}, t_{2i}, t_{3i}) = \sqrt{\frac{ABC}{\mathbf{d}(t_{1i}, t_{2i}, t_{3i})^{\mathrm{H}} \mathbf{W}^k \mathbf{d}(t_{1i}, t_{2i}, t_{3i})}} \quad (43)$$

and $T_{bb}(\hat{\mathbf{u}}^{k+1})$ admits a Vandermonde decomposition, then

$$\|\mathbf{P}\|_{\mathrm{A}^{w^k}} = \underset{\mathbf{u} \in \mathbf{C}^{N_u}, \mathbf{E} \in \mathbf{C}^{L \times L}}{\min} \frac{1}{2\sqrt{ABC}}(\mathrm{tr}(\mathbf{W}^k T_{bb}(\mathbf{u})) + \mathrm{tr}(\mathbf{E})) \text{ subject to } \begin{bmatrix} T_{bb}(\mathbf{u}) & \mathbf{P} \\ \mathbf{P}^{\mathrm{H}} & \mathbf{E} \end{bmatrix} \geq 0. \quad (44)$$

The proof can be found in Supplementary Note 5. Simultaneous Eqs (40) and (44) yield

$$\hat{\mathbf{P}}^{k+1} = \underset{\mathbf{P} \in \mathbf{C}^{ABC \times L}}{\arg\min} \left( \underset{\mathbf{u} \in \mathbf{C}^{N_u}, \mathbf{E} \in \mathbf{C}^{L \times L}}{\min} \frac{1}{2\sqrt{ABC}} \times (\mathrm{tr}(\mathbf{W}^k T_{bb}(\mathbf{u})) + \mathrm{tr}(\mathbf{E})) \text{ subject to } \begin{bmatrix} T_{bb}(\mathbf{u}) & \mathbf{P} \\ \mathbf{P}^{\mathrm{H}} & \mathbf{E} \end{bmatrix} \geq 0 \right)$$

$$\text{subject to } \|\mathbf{P}^{\star} - \mathbf{P}\|_{\mathrm{F}} \leq \varepsilon$$

$$= \underset{\mathbf{P} \in \mathbf{C}^{ABC \times L}}{\arg\min} \|\mathbf{P}\|_{\mathrm{A}^{w^k}} \quad \text{subject to } \|\mathbf{P}^{\star} - \mathbf{P}\|_{\mathrm{F}} \leq \varepsilon \quad (45)$$

Apparently, $\mathbf{P}$ is reconstructed via minimizing its weighted atomic norm iteratively, and the weighting coefficient is updated in each iteration. Hence, the method can be named as IRANM. For sparse cuboid microphone arrays, $\mathbf{P}^{\star} - \mathbf{P}$ in Eqs (37), (38), (40) and (45) becomes $\mathbf{P}^{\star}_{\Omega} - \mathbf{P}_{\Omega}$.

**Implementation details.** The simulations are conducted in Matlab R2014a on a PC with a Windows 10 system and a 2.2 GHz Intel(R) Core(TM) i5-5200U CPU. $A = B = C = 7$ and $\Delta x = \Delta y = \Delta z = 0.035$ m. The number of snapshots is taken as 10. The independent and identically distributed complex Gaussian noise is utilized. The signal-to-noise ratio ($20\lg(\|\mathbf{P}\|_{\mathrm{F}} / \|\mathbf{N}\|_{\mathrm{F}})$) is taken as 20 dB. In the ADMM based algorithm, $\tau$ is determined according to ref.[10], $\rho$ is set to 1, and the iteration is terminated if the relative changes of $\mathbf{u}$ and $\mathbf{P}$ at two consecutive iterations, i.e., $\|\mathbf{u}^q - \mathbf{u}^{q-1}\|_2 / \|\mathbf{u}^{q-1}\|_2$ and $\|\mathbf{P}^q - \mathbf{P}^{q-1}\|_{\mathrm{F}} / \|\mathbf{P}^{q-1}\|_{\mathrm{F}}$, both are less than $10^{-3}$ or the maximum number of iterations, set to 1000, is reached. Simulations of the ADMM based algorithm with another termination criterion relating to the primal and dual infeasibilities shown in Eqs (29) and (30) are also presented

and discussed in Supplementary Note 6. In the MaPP method, the threshold is set as the maximal eigenvalue divided by 100, which means sources in 20 dB dynamic range are considered. When implementing IRANM, we terminate the iteration if the relative changes of the solution $\hat{\mathbf{P}}$ at two consecutive iterations, i.e., $\|\hat{\mathbf{P}}^k - \hat{\mathbf{P}}^{k-1}\|_F / \|\hat{\mathbf{P}}^{k-1}\|_F$, is less than $10^{-3}$ or the maximum number of iterations, set to 20, is reached.

## Data Availability

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

## References

1. Elad, M. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing* (Springer, New York, USA, 2010).
2. Foucart, S. & Rauhut, H. *A Mathematical Introduction to Compressive Sensing* (Springer, New York, USA, 2013).
3. Boche, H., Calderbank, R., Kutyniok, G. & Vybíral, J. *Compressed Sensing and its Applications* (Springer, New York, USA, 2015).
4. Xenaki, A., Gerstoft, P. & Mosegaard, K. Compressive beamforming. *J. Acoust. Soc. Am.* **136**, 260–271 (2014).
5. Gerstoft, P., Xenaki, A. & Mecklenbräuker, C. F. Multiple and single snapshot compressive beamforming. *J. Acoust. Soc. Am.* **138**, 2003–2014 (2015).
6. Chi, Y., Scharf, L. L., Pezeshki, A. & Calderbank, A. R. Sensitivity to basis mismatch in compressed sensing. *IEEE Trans. Signal Process.* **59**, 2182–2195 (2011).
7. Candès, E. J. & Fernandez-Granda, C. Towards a mathematical theory of super-resolution. *Comm. Pure Appl. Math.* **67**, 906–956 (2014).
8. Tang, G., Bhaskar, B. N., Shah, P. & Recht, B. Compressed sensing off the grid. *IEEE Trans. Inf. Theory* **59**, 7465–7490 (2013).
9. Chi, Y. & Chen, Y. Compressive two-dimensional harmonic retrieval via atomic norm minimization. *IEEE Trans. Signal Process.* **63**, 1030–1042 (2015).
10. Li, Y. & Chi, Y. Off-the-grid line spectrum denoising and estimation with multiple measurement vectors. *IEEE Trans. Signal Process.* **64**, 1257–1269 (2016).
11. Yang, Z. & Xie, L. Enhancing sparsity and resolution via reweighted atomic norm minimization. *IEEE Trans. Signal Process.* **64**, 995–1006 (2016).
12. Yang, Z. & Xie, L. Exact joint sparse frequency recovery via optimization methods. *IEEE Trans. Signal Process.* **64**, 5145–5157 (2016).
13. Yang, Z., Xie, L. & Stoica, P. Generalized Vandermonde decomposition and its use for multi-dimensional super-resolution. *IEEE Int. Symp. Inf. Theor. Proc.* **2015-June**, 2011–2015 (2015).
14. Yang, Z., Xie, L. & Stoica, P. Vandermonde decomposition of multilevel Toeplitz matrices with application to multidimensional super-resolution. *IEEE Trans. Inf. Theory* **62**, 3685–3701 (2016).
15. Xenaki, A. & Gerstoft, P. Grid-free compressive beamforming. *J. Acoust. Soc. Am.* **137**, 1923–1935 (2015).
16. Park, Y., Choo, Y. & Seong, W. Multiple snapshot grid free compressive beamforming. *J. Acoust. Soc. Am.* **143**, 3849–3859 (2018).
17. Yang, Y., Chu, Z., Xu, Z. & Ping, G. Two-dimensional grid-free compressive beamforming. *J. Acoust. Soc. Am.* **142**, 618–629 (2017).
18. Yang, Y., Chu, Z., Ping, G. & Xu, Z. Resolution enhancement of two-dimensional grid-free compressive beamforming. *J. Acoust. Soc. Am.* **143**, 3860–3872 (2018).
19. Yang, Y., Chu, Z. & Ping, G. Alternating direction method of multipliers for weighted atomic norm minimization in two-dimensional grid-free compressive beamforming. *J. Acoust. Soc. Am.* **144**, EL361–EL366 (2018).
20. Yang, Y., Chu, Z. & Ping, G. Two-dimensional multiple-snapshot grid-free compressive beamforming. *Mech. Syst. Signal Process.* **124**, 524–540 (2019).
21. Hua, Y. Estimating two-dimensional frequencies by matrix enhancement and matrix pencil. *IEEE Trans. Signal Process.* **40**, 2267–2280 (1992).
22. Boyd, S., Parikh, N., Chu, E., Peleato, B. & Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Lean.* **3**, 1–122 (2010).
23. Wen, Z., Goldfarb, D. & Yin, W. Alternating direction augmented Lagrangian methods for semidefinite programming. *Math. Program. Comput.* **2**, 203–230 (2010).
24. Yang, L. *Research on Compressed Sensing Radar Imaging Under Model Mismatch* (Doctoral dissertation, National University of Defense Technology, Changsha, China, 2016).
25. Grant, M. & Boyd, S. CVX: MATLAB software for disciplined convex programming. http://cvxr.com/cvx (2019).
26. Boyd, S. & Vandenberghe, L. *Convex Optimization* (Cambridge Univ. Press, Cambridge, UK, 2004).
27. Chandrasekaran, V., Recht, B., Parrilo, P. A. & Willsky, A. S. The convex geometry of linear inverse problems. *Found. Comput. Math.* **12**, 805–849 (2012).
28. Dumitrescu, B. *Positive Trigonometric Polynomials and Signal Processing Applications* (Springer, Dordrecht, Netherlands, 2007).
29. Hunter, D. R. & Lange, K. A tutorial on MM algorithms. *Am. Stat.* **58**, 30–37 (2004).
30. Sun, Y., Babu, P. & Palomar, D. P. Majorization-Minimization algorithms in signal processing, communications, and machine learning. *IEEE Trans. Signal Process.* **65**, 794–816 (2017).

## Acknowledgements

## Author Contributions

Y.Y. and Z.C. designed the methods and algorithm. Y.Y. and Y.X.Y. implemented the simulations. Z.X. and Y.Z. checked the simulations. Y.Y. wrote the manuscript. Z.C., Z.X. and Y.Z. reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at https://doi.org/10.1038/s41598-019-47845-7.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.