Interspeech 2022
18-22 September 2022, Incheon, Korea

Joint Estimation of Direction-of-Arrival and Distance for Arrays with Directional Sensors based on Sparse Bayesian Learning

Pengyu Wang[1,⋆] *Feifei Xiong*[2] *Zhongfu Ye*[1] *Jinwei Feng*[2]

[1]Department of Electronic Engineering and Information Science,
University of Science and Technology of China, Hefei, China
[2]Hummingbird Audio Lab, Alibaba Group, Hangzhou, China
{wpyustc,yezf}@mail.ustc.edu.cn, {xff242920,jinwei.feng}@alibaba-inc.com

Abstract

Source localization with sensor arrays is an active research topic in many areas, such as speaker localization and communication. The existing estimators usually assume the use of arrays with omnidirectional sensors, thus they would fail on arrays with directional sensors. In this work, a new method is proposed for locating the near-field sources based on sparse Bayesian learning (SBL), which is capable of integrating the near-field signal model to jointly estimate direction-of-arrival (DOA) and distance. By further considering the directionality of sensors in the signal model which takes full advantage of the magnitude information, the proposed method can handle arrays with both omnidirectional and directional sensors. Simulation results show that the proposed method yields a sharp spatial spectrum, and performs more accurately than traditional near-field Multiple Signal Classification (MUSIC) and Steered-Response Power Phase Transform (SRP-PHAT) for arrays covering heterogeneous directional sensors.

Index Terms: Source localization, sparse Bayesian learning, directional sensor, DOA estimation, distance estimation

1. Introduction

Localization of sources using sensor arrays is an important research topic in many areas such as speech, radar, and communication. As the most common position information, the direction-of-arrival (DOA) estimators of far-field sources have been proposed in many works such as the time delay based algorithms [1, 2, 3], the subspace based algorithms [4, 5, 6, 7], and the compressed sensing based algorithms [8, 9, 10, 11].

As another important position information, the distance between the array and the sources is usually considered in the near-field signal model, in which both the magnitude and the phase of the multichannel observations should be modeled properly. This introduces more challenges to some existing methods dedicated to far-field sources on the one hand. On the other hand, such near-field signal model makes the joint estimation of the DOA and the distance possible. In [12], the authors proposed a two-dimensional (2D) version of the multiple signal classification (MUSIC) algorithm and a maximum likelihood (ML) algorithm to locate the near-field sources. The authors in [13, 14, 15] further extended the subspace-based approaches with high-order cumulants, which can enlarge the array aperture. Jiang et al. proposed a low-cost geometry-based approach with a 3-sensors array by estimating the time delays of each pair of sensors [16]. According to the idea of compressed sensing, the authors in [17] proposed an off-grid approach using the modified convex clustering method with a monopole-only dictionary. Moreover, Liu et al. modeled the array observation as a mixture of Gaussian processes and designed an off-grid estimator for the frequency diverse array multiple-input multiple-output radar based on sparse Bayesian learning (SBL) [18]. However, the related work mainly focused on one task in source localization like either the DOA estimation (majority), the distance estimation, with a single source, or only with omnidirectional sensors (but suffer when applied to directional sensors which own different magnitude responses over the position of sources).

In this paper, we aim to propose a general framework for source localization to jointly estimate the DOA and the distance of multiple sources using an array flexible with either omnidirectional or directional sensors (see [19]). In fact, directional sensors are widely utilized in practical applications such as meeting devices. Recently some comparative studies have been conducted showing that the use of directional sensors brings noticeable improvement in the key performance metrics [20, 21]. Motivated by compressed sensing [22], the SBL framework is exploited in the proposed method, which can make full use of the sparsity of sources in both DOA and distance region. The magnitude response of directional sensors, as well as the attenuation and the phase shift during propagation, are seamlessly integrated into the SBL derivation. More specifically, the noise is modeled as the Gaussian process, the original sources are modeled as unknown latent variables, and the power of the sources and noise are seen as unknown parameters. To solve the latent variables and the parameters so as to accurately obtain the DOA and the distance information, the Expectation Maximization (EM) algorithm [23] can be employed and we improve the convergence of M-step via a recursive smoothing strategy inspired by the solution of over-complete problems [24]. Further, we will show that the directionality of sensors in the signal model can be straightforwardly introduced into derivation under the proposed SBL framework, achieving the generalization to cover heterogeneous directional sensors.

In the remainder of this paper, we firstly introduce the model description of the near-field signals and the arrays with directional sensors in Section 2. Secondly in Section 3, the joint estimator of DOA and distance for arrays with directional sensors under the SBL framework as well as the corresponding derivation is described in detail. Simulation results are then presented in Section 4 to show the superiority of the proposed method compared to two classic algorithms, and Section 5 concludes the paper.

⋆The author completed this work as an intern at Alibaba Group, and this work is also partially supported by the funding from National Natural Science Foundation of China (No. 61671418).

10.21437/Interspeech.2022-497

2. Signal Model

We consider I narrowband sources placed at positions (near-field here) $(\theta_i, r_i), i \in [1, I]$, where θ and r denote DOA and distance between the source i and the reference sensor, respectively. These sources impinge a uniform linear array (ULA) with $2M+1$ either omnidirectional or directional sensors separated by a distance d in a noisy environment. The geometry of

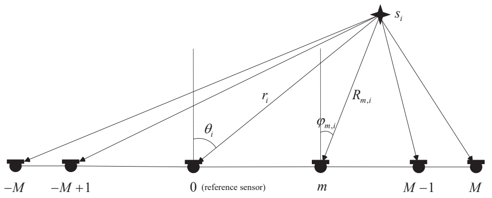


Figure 1: *Geometry of the uniform linear array with $2M+1$ amount of sensors and the ith target source.*

the array sensors and the ith source is plotted in Figure 1. Then, the N-snapshots array observation can be represented as

$$\begin{aligned} x_m(n) = \sum_i h_{m,i} s_i(n) + w_m(n), \\ i \in [1, I], m \in [-M, M], n \in [1, N], \end{aligned} \quad (1)$$

where h includes the magnitude response of directional sensors as well as the attenuation and the phase shift during propagation, s denotes the source, and w represents the additive noise. Note that the reverberation effect is not yet considered. Herein, $h_{m,i}$ can be decomposed into two portions as

$$h_{m,i} = g_{m,i} a_{m,i}, \quad (2)$$

where $g_{m,i}$ denotes the magnitude response of the directional sensor (named directional portion), and $a_{m,i}$ denotes the attenuation and the phase shift during signal propagation (named propagating portion).

According to [25], the directional portion can be expressed in a first-order form as

$$g_{m,i} = \alpha + (1-\alpha)\cos\varphi_{m,i}, \quad (3)$$

where α is the shape parameter with $\alpha = 1$ denoting omnidirectional sensors, while $0 < \alpha < 1$ for directional sensors, and

$$\varphi_{m,i} = \arctan\left(\tan\theta_i - \frac{md}{r_i \cos\theta_i}\right), \quad (4)$$

represents the angle between the ith source direction and the mth sensor as illustrated in Figure 1.

Next, according to [12], the propagating portion $a_{m,i}$ in (2) can be expressed as

$$a_{m,i} \propto \frac{1}{R_{m,i}} \exp\left(-j2\pi f \frac{R_{m,i}}{c}\right), \quad (5)$$

where

$$R_{m,i} = \frac{r_i \cos\theta_i}{\cos\varphi_{m,i}} \quad (6)$$

is the distance between the ith source and the mth sensor according to the spherical wave model [26].

After discretizing the near-field range [15, 27] into K potential source locations as $\left[\left(\bar{\theta}_1, \bar{r}_1\right), \left(\bar{\theta}_2, \bar{r}_2\right), \cdots, \left(\bar{\theta}_K, \bar{r}_K\right)\right]$, the extended array manifold and signal matrix can be expressed respectively as

$$\mathbf{H} = \begin{bmatrix} h_{-M,1} & h_{-M,2} & \cdots & h_{-M,K} \\ h_{-M+1,1} & h_{-M+1,2} & \cdots & h_{-M+1,K} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M,1} & h_{M,2} & \cdots & h_{M,K} \end{bmatrix} \quad (7)$$

and

$$\bar{\mathbf{S}} = \begin{bmatrix} \bar{s}_1(1) & \bar{s}_1(2) & \cdots & \bar{s}_1(N) \\ \bar{s}_2(1) & \bar{s}_2(2) & \cdots & \bar{s}_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ \bar{s}_K(1) & \bar{s}_K(2) & \cdots & \bar{s}_K(N) \end{bmatrix}, \quad (8)$$

where $\bar{\mathbf{S}}$ is a row I sparse matrix whose non-zero rows are $\bar{\mathbf{s}}_k = [s_i(1), s_i(2), \cdots, s_i(N)]$ for each $\left(\bar{\theta}_k, \bar{r}_k\right) = (\theta_i, r_i), \forall k \in [1, K], i \in [1, I]$. Then the multi-snapshot array observation can be expressed as

$$\mathbf{X} = \mathbf{H}\bar{\mathbf{S}} + \mathbf{W}, \quad (9)$$

where

$$\mathbf{X} = \begin{bmatrix} x_{-M}(1) & x_{-M}(2) & \cdots & x_{-M}(N) \\ x_{-M+1}(1) & x_{-M+1}(2) & \cdots & x_{-M+1}(N) \\ \vdots & \vdots & \ddots & \vdots \\ x_M(1) & x_M(2) & \cdots & x_M(N) \end{bmatrix} \quad (10)$$

and

$$\mathbf{W} = \begin{bmatrix} w_{-M}(1) & w_{-M}(2) & \cdots & w_{-M}(N) \\ w_{-M+1}(1) & w_{-M+1}(2) & \cdots & w_{-M+1}(N) \\ \vdots & \vdots & \ddots & \vdots \\ w_M(1) & w_M(2) & \cdots & w_M(N) \end{bmatrix}. \quad (11)$$

3. The Proposed Method

3.1. Bayesian Probabilistic Model

In this work, a SBL based method is proposed to locate the sources. The probabilistic graphical model is illustrated in Figure 2.

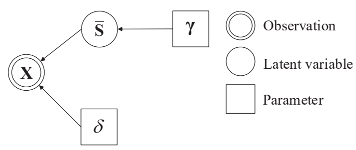


Figure 2: *Probabilistic graphical model of the SBL framework.*

Based on Figure 2, the candidate sources are seen as latent variables, whose prior follows a zero-mean complex Gaussian distribution with location-dependent precision parameters $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \cdots, \gamma_K]$, written as

$$p\left(\bar{\mathbf{S}}; \boldsymbol{\gamma}\right) = \prod_{k,n} \mathcal{CN}\left(\bar{s}_k(n); 0, \gamma_k^{-1}\right), \quad (12)$$

Similarly, the prior of the noise follows a zero-mean complex Gaussian distribution with precision parameter δ as

$$p(\mathbf{W}; \delta) = \prod_{m,n} \mathcal{CN}\left(w_m(n); 0, \delta^{-1}\right). \quad (13)$$

Then the marginalized probability density function (PDF) can be expressed as

$$p\left(\mathbf{X}|\bar{\mathbf{S}};\delta\right)=\prod_{n}\mathcal{CN}\left(\mathbf{H}\bar{\mathbf{s}}\left(n\right);\mathbf{0},\delta^{-1}\mathbf{I}_{2M+1}\right),\qquad(14)$$

where $\bar{\mathbf{s}}\left(n\right)=\left[\bar{s}_1\left(n\right),\bar{s}_2\left(n\right),\cdots,\bar{s}_K\left(n\right)\right]^T$ (see equation (8)). The EM algorithm is utilized to estimate the parameters $\boldsymbol{\gamma}$ and δ iteratively as follows.

3.2. E-step

In E-step, the expectation of the log likelihood of the complete data is required to determine. Because both (12) and (14) are Gaussian, based on the Bayes rule, the posterior PDF of the candidate sources can be written as

$$p\left(\bar{\mathbf{S}}|\mathbf{X};\boldsymbol{\gamma},\delta\right)=\frac{p\left(\mathbf{X}|\bar{\mathbf{S}};\delta\right)p\left(\bar{\mathbf{S}};\boldsymbol{\gamma}\right)}{\int p\left(\mathbf{X}|\bar{\mathbf{S}};\delta\right)p\left(\bar{\mathbf{S}};\boldsymbol{\gamma}\right)\mathrm{d}\bar{\mathbf{S}}}=\mathcal{CN}\left(\boldsymbol{\mu},\boldsymbol{\Sigma}\right),\qquad(15)$$

where $\boldsymbol{\mu}=\delta\boldsymbol{\Sigma}\mathbf{H}^H\mathbf{X}$, $\boldsymbol{\Sigma}=\left(\delta\mathbf{H}^H\mathbf{H}+\boldsymbol{\Gamma}\right)^{-1}$ and $\boldsymbol{\Gamma}=\mathrm{diag}\left(\boldsymbol{\gamma}\right)$. Then the desired expectation can be calculated as

$$\begin{aligned}&E_{p\left(\bar{\mathbf{S}}|\mathbf{X};\boldsymbol{\gamma},\delta\right)}\left[\ln p\left(\mathbf{X},\bar{\mathbf{S}};\boldsymbol{\gamma},\delta\right)\right]\\&=E_{p\left(\bar{\mathbf{S}}|\mathbf{X};\boldsymbol{\gamma},\delta\right)}\left[\ln p\left(\mathbf{X}|\bar{\mathbf{S}};\delta\right)+\ln p\left(\bar{\mathbf{S}};\boldsymbol{\gamma}\right)\right].\end{aligned}\qquad(16)$$

3.3. M-step

In M-step, the parameters $\boldsymbol{\gamma}$ and δ are estimated by maximizing the above equation (16) respectively as

$$\begin{aligned}\hat{\gamma}_k&=\arg\max_{\gamma_k}E\left[\ln p\left(\mathbf{X},\bar{\mathbf{S}};\boldsymbol{\gamma},\delta\right)\right]\\&=\arg\max_{\gamma_k}E\left[\ln p\left(\bar{\mathbf{S}};\boldsymbol{\gamma}\right)\right]\\&=\frac{N}{\|\mu_k\|_2^2+N\Sigma_{kk}}\end{aligned}\qquad(17)$$

and

$$\begin{aligned}\hat{\delta}&=\arg\max_{\delta}E\left[\ln p\left(\mathbf{X},\bar{\mathbf{S}};\boldsymbol{\gamma},\delta\right)\right]\\&=\arg\max_{\delta}E\left[\ln p\left(\mathbf{X}|\bar{\mathbf{S}};\delta\right)\right]\\&=\frac{(2M+1)N}{\|\mathbf{X}-\mathbf{H}\boldsymbol{\mu}\|_2^2+N\,\mathrm{tr}\left(\mathbf{H}^H\mathbf{H}\boldsymbol{\Sigma}\right)},\end{aligned}\qquad(18)$$

where μ_k and Σ_{kk} denote the elements of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, respectively. Note that we alternatively calculate (17) using

$$\hat{\gamma}_k=\frac{N\left(1-\gamma_k\Sigma_{kk}\right)}{\|\mu_k\|_2^2},\qquad(19)$$

which leads to a faster convergence in large-scale and highly overcomplete problems as referred in [24, 28], and is also proved to be more efficient in our preliminary experiments.

3.4. Spatial Spectrum

Considering that γ_k is the precision of the candidate source at position $\left(\bar{\theta}_k,\bar{r}_k\right)$, the 2D spatial spectrum is defined as

$$\mathbf{P}=\{p_k\}_{k=1}^{K}\qquad(20)$$

with

$$p_k=\frac{1}{\gamma_k},\qquad(21)$$

and $\{\cdot\}^K$ represents a set with K elements. The estimated DOAs and distances are determined by the largest I peaks (corresponding to multiple sources) in the spatial spectrum $\mathbf{P}$. For clarity, the proposed method is summarized in Algorithm 1.

Algorithm 1: The proposed method.

Input: The multi-snapshot observation $\mathbf{X}$;
The hyperparameters T and ϵ.

```
Initialize γ and δ; Set the iteration counter t = 0;
while t < T and |δ̂ − δ̂^old| / δ̂^old > ε do
    t = t + 1, δ̂^old = δ;
    Calculate the expectation of the log likelihood of
      the candidate sources according to (16);
    Estimate the precision parameter of the candidate
      sources according to (19);
    Estimate the precision parameter of the additional
      noise according to (18);
end
Calculate the 2D spatial spectrum according to (20).
Find the peaks.
```

Output: The DOA and distance estimates.

4. Simulation Results

4.1. Experimental Setup

Numerical simulations under different conditions are presented to evaluate the performance of the proposed method. Narrowband Gaussian sources are simulated with the center frequency $f=2000$ Hz and the wave velocity $c=343$ m/s. Three uniform linear arrays (ULAs) with nine sensors, half-wavelength spacing and $\alpha\in\{0.5,0.75,1\}$ in equation (3) (representing three types of sensor directionalities) are considered. For comparison, classic near-field MUSIC [12] and SRP-PHAT [29] are employed. The searching grids for DOA and distance are uniformly distributed in the area of $[-80^\circ,80^\circ]$ and $[0.85\text{m},5.49\text{m}]$ with spacing of 5° and 0.23m, respectively. The signal-to-noise ratio (SNR) according to (1) is defined as

$$\mathrm{SNR}=10\log_{10}\left[\frac{\sum_m E\left[\left|x_m\left(n\right)-w_m\left(n\right)\right|^2\right]}{\sum_m E\left[\left|w_m\left(n\right)\right|^2\right]}\right].$$

The root mean square error (RMSE) is selected as the evaluation metric, calculated as

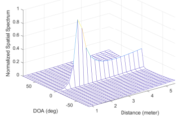
$$\mathrm{RMSE}_z=\sqrt{\frac{1}{M_tI}\sum_{m_t,i}\left|\hat{z}_i^{m_t}-z_i^{m_t}\right|^2}$$

where $M_t=200$ is the number of Monte-Carlo runs, $\hat{z}_i^{m_t}$ and $z_i^{m_t}$ represent the ith estimated value and ground truth such as θ_i and r_i for the m_tth Monte-Carlo trial.

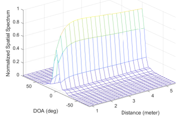
4.2. Performance with Single Source

Firstly, a single Gaussian narrowband source is considered and placed at position $(0^\circ,2\text{m})$, with SNR $=15$ dB and $N=10$. The corresponding 2D spatial spectra w.r.t. the proposed method and the two competitive algorithms in a single Monte Carlo run are shown in Figure 3.

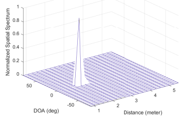
It can be found that the peak in the spatial spectrum of the proposed method is more accurate and sharper than that of the other methods in both omnidirectional ($\alpha=1$) and different directional ($\alpha=0.5,0.75$) cases, while both MUSIC and SRP-PHAT show pseudo large values in their spatial spectra. For instance, for MUSIC, these pseudo values appear at the same DOA but at different distances from the real source location.



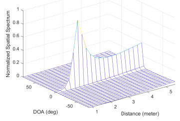
(a) MUSIC ($\alpha = 0.5$)



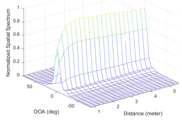
(b) SRP-PHAT ($\alpha = 0.5$)



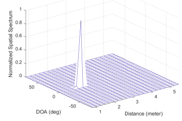
(c) Proposed method ($\alpha = 0.5$)



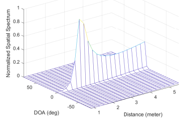
(d) MUSIC ($\alpha = 0.75$)



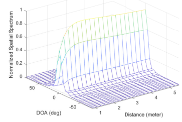
(e) SRP-PHAT ($\alpha = 0.75$)



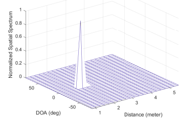
(f) Proposed method ($\alpha = 0.75$)



(g) MUSIC ($\alpha = 1$)



(h) SRP-PHAT ($\alpha = 1$)



(i) Proposed method ($\alpha = 1$)

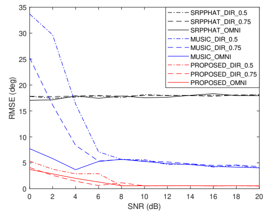
Figure 3: *The normalized spatial spectra with a single source at position* $(0^\circ, 2\text{m})$ *w.r.t. three different algorithms and three types of sensor directionalities.*

This can be explained by the fact that due to the magnitude information considered in the array manifold, the orthogonality between the noise subspace and the real steering vector in MUSIC is actually not strictly satisfied anymore. Therefore, the spatial spectrum value increases monotonically as the distance increases in the far field, and this phenomenon becomes more apparent to arrays with directional sensors. As for SRP-PHAT which is employed straightforwardly here since PHAT function only retains the phase information of the observation, the spatial spectra of SRP-PHAT are not sharp enough, instead, with many pseudo large values nearby, resulting in the peak-finding tending to be incorrect (even with a high SNR).

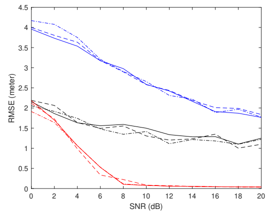
4.3. Performance with Two Sources

Without loss of generality, two Gaussian sources are placed at (θ_1, r_1) and (θ_2, r_2), where θ_1, θ_2, r_1 and r_2 are uniformly distributed in $[-5^\circ, -4^\circ]$, $[5^\circ, 6^\circ]$, $[1\text{m}, 1.1\text{m}]$ and $[1.5\text{m}, 1.6\text{m}]$, respectively. The RMSEs of DOA and distance estimation w.r.t. different SNRs and numbers of snapshots are shown in Figure 4 and Figure 5, respectively.

As clearly shown in Figure 4, the proposed method performs the best among the three algorithms. SRP-PHAT barely distinguishes the two closely spaced sources, probably due to the blunt spectrum peaks as also illustrated in Figure 3 (b) (e) (h). It seems that MUSIC behaves better for estimating DOA rather than distance. The proposed method gives consistent performance for arrays with both omnidirectional and directional sensors, indicating the *good* applicability of the proposed method to various sensor directionalities. While MUSIC performs struggling for arrays with directional sensors compared to omnidirectional case, particularly at low SNRs. As well, it can be observed that when SNR reaches 8 dB, an



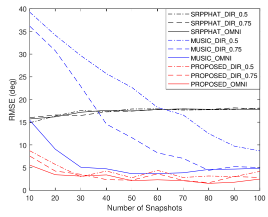
(a) DOA



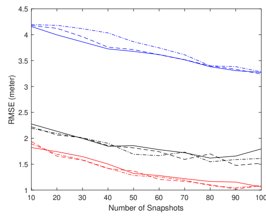
(b) Distance

Figure 4: RMSEs *of DOA and distance estimation with* $N = 50$ *w.r.t. different SNRs for the three algorithms with three types of sensor directionalities.*

accurate DOA and distance estimation for two sources can be achieved by the proposed method with RMSEs of less than 1 degree and less than 0.1 m, respectively.



(a) DOA



(b) Distance

Figure 5: RMSEs *of DOA and distance estimation at* SNR = 3 *dB w.r.t. different numbers of snapshots* N *for the three algorithms with three types of sensor directionalities.*

For real-time application, the number of snapshots N of the received signal required for an accurate estimation is also an important metric. Figure 5 illustrates that performance of all these three algorithms can be further improved with more snapshots, except that SRP-PHAT seems to perform saturated already for the DOA estimation with a low SNR. It is also observed that MUSIC seems to be more applicable to omnidirectional sensors, and SRP-PHAT can not distinguish the two closely spaced sources. Still, the proposed method gives the best performance, and it also shows that the distance estimation requires more snapshots than the DOA estimation, indicating that the distance estimation is more challenging to accomplish.

5. Conclusions

This paper proposed a novel DOA and distance joint estimation method based on sparse Bayesian learning for multiple source localization. The magnitude information (magnitude response and attention of directional sensors) and the phase information (phase shift during propagation) have been seamlessly integrated into the SBL derivation to allow the proposed approach to be applicable to arrays with various directional sensors. Simulation results showed that the proposed method yields sharp peaks in the spatial spectrum, leading to superior performance in comparison to the existing methods such as near-field MUSIC and SRP-PHAT. Accurate estimation results even with low SNRs and few snapshots further showed the potential of the proposed method for practical applications. As part of future work, the reverberation effect and non-Gaussian distributed sources will be integrated into the proposed framework.

6. References

[1] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 4, pp. 320–327, 1976.

[2] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, 1996.

[3] J. D. Hector, "A high-accuracy, low-latency technique for talker localization in reverberant environments using microphone arrays," *Brown University*, 2000.

[4] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[5] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.

[6] A. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *1983 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 8. IEEE, 1983, pp. 336–339.

[7] B. Porat and B. Friedlander, "Direction finding algorithms based on high-order statistics," in *1990 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, 1990, pp. 2675–2678.

[8] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.

[9] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse bayesian inference," *IEEE Transactions on Signal Processing*, vol. 61, no. 1, pp. 38–43, 2012.

[10] P. Gerstoft, C. F. Mecklenbräuker, A. Xenaki, and S. Nannuru, "Multisnapshot sparse Bayesian learning for DOA," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1469–1473, 2016.

[11] A. Das and T. J. Sejnowski, "Narrowband and wideband off-grid direction-of-arrival estimation via sparse Bayesian learning," *IEEE Journal of Oceanic Engineering*, vol. 43, no. 1, pp. 108–118, 2017.

[12] Y.-D. Huang and M. Barkat, "Near-field multiple source localization by passive sensor array," *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 7, pp. 968–975, 1991.

[13] Y. Wu, L. Ma, C. Hou, G. Zhang, and J. Li, "Subspace-based method for joint range and DOA estimation of multiple near-field sources," *Signal Processing*, vol. 86, no. 8, pp. 2129–2133, 2006.

[14] J. Liang and D. Liu, "Passive localization of mixed near-field and far-field sources using two-stage MUSIC algorithm," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 108–120, 2009.

[15] J. Li, G. Wei, B. Ma, Y. Wang, and C. Le Bastard, "A simple way for near-field source localization with MUSIC," in *2016 IEEE International Conference on Computational Electromagnetics (ICCEM)*. IEEE, 2016, pp. 330–332.

[16] W. Jiang, Z. Cai, M. Luo, and Z. L. Yu, "A simple microphone array for source direction and distance estimation," in *2011 6th IEEE Conference on Industrial Electronics and Applications*. IEEE, 2011, pp. 1002–1005.

[17] T. Tachikawa, K. Yatabe, Y. Ikeda, and Y. Oikawa, "Sound source localization based on sparse estimation and convex clustering," in *172nd Meeting of the Acoustical Society of America*, vol. 29, no. 1. Acoustical Society of America, 2016, pp. 1–14.

[18] Q. Liu, X. Wang, M. Huang, X. Lan, and L. Sun, "DOA and range estimation for FDA-MIMO radar with sparse Bayesian learning," *Remote Sensing*, vol. 13, no. 2553, pp. 1–19, 2021.

[19] J. Eargle, *The Microphone Book: From mono to stereo to surround-a guide to microphone design and application*. Routledge, 2012.

[20] W. Huang and J. Feng, "Differential beamforming for uniform circular array with directional microphones." in *INTERSPEECH*, 2020, pp. 71–75.

[21] W. Chen, C. Xue, and X. Zhong, "Cramér-rao lower bound for doa estimation with an array of directional microphones in reverberant environments." in *INTERSPEECH*, 2021, pp. 2172–2176.

[22] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[23] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[24] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, no. Jun, pp. 211–244, 2001.

[25] A. Politis, S. Delikaris-Manias, and V. Pulkki, "Direction-of-arrival and diffuseness estimation above spatial aliasing for symmetrical directional microphone arrays," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015, pp. 6–10.

[26] H. Gazzah and J.-P. Delmas, "CRB-based design of linear antenna arrays for near-field source localization," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 4, pp. 1965–1974, 2014.

[27] N. Venkateswaran and C. Ashok, "DOA estimation of near-field sources using support vector regression," in *2016 IEEE Region 10 Conference (TENCON)*. IEEE, 2016, pp. 1570–1574.

[28] D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3704–3716, 2007.

[29] A. Johansson, G. Cook, and S. Nordholm, "Acoustic direction of arrival estimation, a comparison between ROOT-MUSIC and SRP-PHAT," in *2004 IEEE Region 10 Conference TENCON 2004*. IEEE, 2004, pp. 629–632.