Small Size Array Underwater Acoustic DOA Estimation Based on Direction-Dependent Transmission Response

Jiaheng Li, Feng Tong, *Member, IEEE*, Yuehai Zhou, Yi Yang, and Zhiqiang Hu

***Abstract*—It has been recognized that the performance of the traditional direction of arrival (DOA) estimation algorithms is highly dependent on array geometry parameters, especially the array size and element number. Unfortunately, significant performance degradation may be caused under unideal array geometry such as small size array onboard small underwater vehicle. Among all the possible directions, the presence of underwater target at a certain direction constitutes a type of spatial sparsity. Thus, spatial sparse recovery provides a potential way to address this challenging problem. As each DOA will correspond to a different transmission response pattern with respect to each array element, thus constitutes a direction-dependent transmission response (DDTR) that can be explored. However, random contribution of underwater acoustic (UWA) channel contained in DDTR makes it extremely difficult to directly construct dictionary matrix with DDTR. In this paper, a novel DDTR based small size array DOA estimation method is proposed. First, DDTRs are decomposed into time delay (TD) term and random channel response (RCR) term to decouple the adverse impact of UWA channel. Second, TD is conveniently constructed as dictionary matrix in terms of geometry relationship. Then, by accumulating the constructed DOA at frequency bins selected with energy or correlation coefficient, a selective frequency accumulation compressed sensing reconstruction (SFA-CS) algorithm is designed to mitigate the negative effect of UWA channel. Compared with the conventional DOA estimation algorithms, simulation results demonstrate that the proposed method can achieve small array DOA estimation with a lower root mean squared error (RMSE) and higher success rate (SR).**

***Index Terms*—DOA estimation, small size array, transmission response, RMSE.**

I. Introduction

DIRECTION of arrival (DOA) estimation of underwater targets is widely used in ocean observation technology, such as underwater sensor networks, navigation, and marine engineering [1], [2], [3]. Traditional DOA estimation methods such as multiple signal classification (MUSIC), minimum variance distortionless response (MVDR), and their variants [4], [5], [6] generally require a dense and large array to yield high precision, indicating a precision constraint imposed by the size and element number of the array.

Recently, rapid development of underwater technology witnessed actively evolution of small size underwater vehicles such as autonomous underwater vehicle (AUV), unmanned underwater vehicle (UUV), etc. that enable dense and low-cost deployment [7], [8]. The physical size of those small vehicles is usually limited to m-level, resulting in the size of the onboard array being inevitably limited to m-level, as well as the limited number of acoustic sensors, and steering blind region. In addition, the circumstances of low signal-to-noise ratio (SNR) and few snapshots, etc., will also significantly degrade the performance of the DOA estimation achieved by a small size array.

To overcome the above limitations, Yang et al. established an accurate acoustic localization model with a small sensor array and proved that this model is quite suitable for 4 and 5-elements spatial square arrays [9]. Su et al. proposed a time delay-based DOA estimation method based on a small size array. Using the generalized correlation method, the time delays between the signal and each element of the array are estimated and used to localize the sources [10]. Taking a tiny dimension three sensor line array, Sun et al. proposed a new signal phase matching-based DOA estimation method. The DOA estimation accuracy of this method is close to that of MUSIC and can furtherly estimate the signal waveform [11].

Different from conventional DOA estimation algorithm that directly explores spatial information contained in signal from each array element, compressed sensing (CS) theory provides a new way to address the challenging DOA estimation problem by converting the problem of DOA into that of spatial sparse recovery [12], [13], [14]. Malioutov et al. utilized the sparse signal representation methodology for source localization and imposed a penalty on the lack of sparsity of the spatial spectrum [15]. Liang et al. proposed a DOA estimation method based on sparse Bayesian learning to overcome the spatially colored noise [16]. Gurbuz et al. assumed sparsity in the bearing domain and used a basis pursuit strategy to reconstruct the bearing domain vector.

To reduce the complexity and computation, Chang et al. proposed a CS-based array processing method. This method

Manuscript received 20 February 2022; revised 11 June 2022; accepted 5 August 2022. Date of publication 10 August 2022; date of current version 19 December 2022. This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFE0110000, and in part by the National Natural Science Foundation of China under Grants 11274259 and 11574258. The review of this article was coordinated by Prof. Daniel Benevides da Costa. *(Corresponding author: Feng Tong.)*

Jiaheng Li, Feng Tong, and Yuehai Zhou are with the Key Laboratory of Underwater Acoustic Communication and Marine Information Technology of the Minister of Education, and the College of Ocean and Earth Sciences, Xiamen University, Xiamen 361001, China (e-mail: lijiaheng1009@163.com; ftong@xmu.edu.cn; zhouyuehai@xmu.edu.cn).

Yi Yang and Zhiqiang Hu are with the Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110169, China (e-mail: yangyi@sia.cn; hzq@sia.cn).

Digital Object Identifier 10.1109/TVT.2022.3197922

multiples the spatial and temporal compressed matrices with array signals to get a small size array system [17]. For a low SNR and few snapshots, Li et al. proposed a new DOA estimation method based on sparse signal recovery. This method constructs a new weighted matrix to penalize the l_1-norm constrained model, effectively suppressing the false peaks [18]. In fact, the above methods took random projections of the received signals at the sensors rather than taking random projections of the bearing vector directly [19].

However, due to the highly complicated time-frequency selectivity of underwater acoustic (UWA) channel [20], acoustic signal emitted by target source will be refracted by nonhomogeneous water medium and reflected by sea surface and seafloor. Therefore, the received signals corresponding to the same DOA may be the superposition of multipath from different sources, or the received signals corresponding to different DOAs may be the superposition of multipath from the same source [21]. Thus, besides the limitation caused by the array geometry parameter, the adverse multipath effect poses further challenges to the small size array DOA estimation.

Referring to localization examples in biological organisms, such as humans, bats, etc., sound localization or DOA estimation can be performed even just by one single ear [22], [23]. This is because, before the sound arrives at the eardrum, it will be scattered by the head, torso, and ear pinna. This type of scatters is constructed by the superposition of different multipath and changed as a function of the location or DOA of the source, referred to as the head-related transfer function (HRTF). Jia et al. designed a multisource listening system, the metamaterial enclosure (ME) was adopted to construct an artificial frequency-dependent transmission characteristic, thus enabling a single microphone 3D DOA via sparse recovery [23].

Similarly, underwater acoustic signals generated by underwater target experience different multipath pattern associated with different DOA. Therefore, for a certain direction with respect to the array, the transmission response between the target and the array constitutes the direction-dependent transmission response (DDTR). Meanwhile, among all the possible DOAs, the presence of the target in a certain direction constitutes a type of spatial sparsity, thus the DOA estimation can be achieved by sparse reconstruction algorithms, such as orthogonal matching pursuit (OMP), stagewise weak OMP (SWOMP), compressive sampling matching pursuit (CoSaMP) algorithms [24], [25], [26].

However, in [23], while the measurement matrix determined by the ME and signals can be experimentally constructed in advance, the room impulse response is not considered in this model and thus still poses a difficulty under an unknown environment. In this sense, for underwater acoustic DOA estimation, a similar problem exists as the DDTR also contains a random contribution of UWA channel, which also makes it extremely difficult to directly construct a dictionary matrix.

To address the negative impact of UWA channel, a novel CS-based underwater DOA estimation method is proposed in this paper by the means of DDTR decomposition and frequency accumulation.

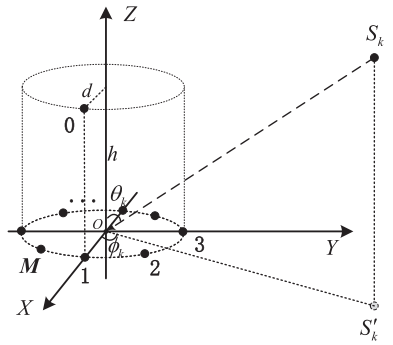

Fig. 1. Structure of the underwater acoustic signal receiver.

Specifically, DDTRs are decomposed into time delay (TD) term and random channel response (RCR) term to decouple the impact of UWA channel. While TDs associated with each DOA can be constructed as the dictionary matrix according to the array geometry, weighting of RCRs on sparse vector unavoidable affects the sparse recovery performance. Although the RCRs exhibit random frequency fading features, the target DOA remains the same at frequency-domain. Therefore, by accumulating the DOA results of frequency bins, a selective frequency accumulation compressed sensing reconstruction (SFA-CS) algorithm is designed to reconstruct DOAs. Furthermore, the energy selection and correlation coefficient selection are introduced for different selected ways, called ESFA-SC and CSFA-CS, respectively. Simulation results demonstrate that the proposed method can effectively estimate the DOAs with a lower root mean squared error (RMSE) and higher success rate (SR), compared with CS-based methods and MVDR algorithm.

The rest of this paper is organized as follows. Section II introduces the system model. Section III transforms the DOA estimation into sparse signal reconstruction and proposes the SFA-CS algorithms to estimate the DOAs. Simulations results are performed in Section IV to verify the effectiveness of the proposed algorithms. Finally, Section V concludes the paper.

II. System Model

Without loss of generality, a small size quasi-cylindrical array (QCA) as shown in Fig. 1 is considered here. The number of omni-directional acoustic sensors is $M+1$, where $M>2$. $\mathbf{r_m}=[\mathrm{x}_m,\mathrm{y}_m,\mathrm{z}_m]$ and $(\widehat{\theta}_m,\widehat{\phi}_m)$ denote the position vector and the elevation and azimuth angles of the m-th element, respectively. $\mathrm{x}_m,\mathrm{y}_m,\mathrm{z}_m$ are less than the distance between the target and array. The radius of the quasi-cylindrical is $d\leq 0.1$ m, and the height is $h\leq 1$ m, referred to as a small size array. Then, the propagation vector $\boldsymbol{v}(\theta,\phi)$ is expressed as

$$\boldsymbol{v}(\theta,\phi)=[\sin(\theta)\cos(\phi),\sin(\theta)\sin(\phi),\cos(\phi)]^T, \quad (1)$$

and the $(M+1)\times 1$ steering vector $\boldsymbol{\alpha}(\theta,\phi)$ is shown as

$$\begin{aligned}\boldsymbol{\alpha}(\theta,\phi)&=[\alpha_0(\theta,\phi),\alpha_1(\theta,\phi),\ldots,\alpha_M(\theta,\phi)]^T\\&=[\exp(-j\omega\mathrm{r}_{\mathbf{0}}\mathbf{v}(\theta,\phi)),\ldots,\exp(-j\omega\mathrm{r}_{\mathbf{M}}\mathbf{v}(\theta,\phi))]^T,\end{aligned} \quad (2)$$

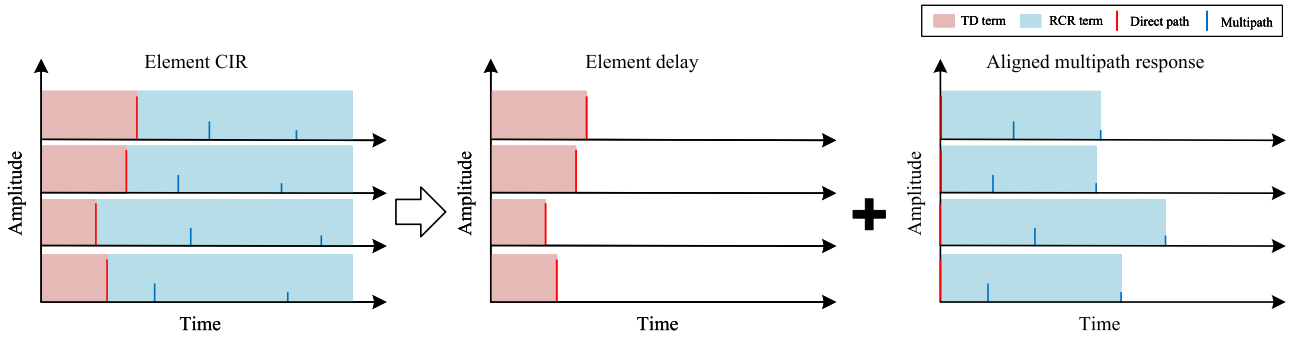

Fig. 2. Illustration of decomposing the CIR of array element into element delay and aligned multipath response.

where (θ, ϕ) are the elevation and azimuth angles of the target, $\omega = 2\pi/\lambda$. λ is the wavelength of the signal, and $(\cdot)^T$ denotes the transpose.

Generally, the target signal $s(t)$ can be described here in its digital form, as a set of discrete samples

$$s(n) = Ae^{j2\pi f_c n}, n = 1, 2, \ldots, N, \tag{3}$$

where A is the signal strength, and f_c is the central frequency. N is the number of samples.

Take the m-th element as an example, the signal received at the time instant n can be expressed as

$$\begin{aligned} y_m(n) &= h_m(\mathrm{r_k}, n) * s(n) + e_m(n), \\ m &= 0, \ldots, M, K = 1, \ldots, K, \end{aligned} \tag{4}$$

where "*" denotes the convolution, K is the number of target sources. $\boldsymbol{r}_k = [x_k, y_k, z_k]$ is the position vector of the k-th target, (θ_k, ϕ_k) are the corresponding elevation and azimuth angles, and $e_m(n)$ is the additive noise. $h_m(\boldsymbol{r}_k, n)$ is the channel impulse response (CIR) and can be expressed as

$$h_m(\boldsymbol{r}_k, n) = \sum_{p=1}^{P_k} a_{m,k}(nT)\delta(\tau - \tau_{m,k}(nT)), \tag{5}$$

where T is the sampling period satisfying the Nyquist rate, P_k, $a_{m,k}(nT)$, and $\tau_{m,k}(nT)$ are the number of multipath, path amplitudes and path delays, respectively. Moreover, when there is a relative movement of the source, array or water medium, $h_m(\boldsymbol{r}_k, n)$ is affected by the Doppler spread, as the doubly selective characteristics of UWA channel. Which means that $\tau_{m,k}(nT)$ is time-varying and can be expressed as an approximate first-order polynomial: $\tau_{m,k}(nT) \approx \tau_{m,k} - \gamma_{m,k}nT$, where $\gamma_{m,k}$ denotes the factor of Doppler spread.

As shown in Fig. 2, $h_m(\boldsymbol{r}_k, n)$ can be decomposed into an element delay and aligned multipath response. While the former is determined by array geometry, the latter is the aligned channel response obtained by aligning CIR of all elements. Namely, CIR can be expressed as

$$h_m(\boldsymbol{r}_k, n) = h_{td,m}(\boldsymbol{r}_k, n) * h_{ca,m}(\boldsymbol{r}_k, n), \tag{6}$$

where $h_{td,m}(\boldsymbol{r}_k, n)$ denotes the element time delay (TD) term between the m-th element and k-th target, and $h_{ca,m}(\boldsymbol{r}_k, n)$ denotes the aligned random multipath channel response (RCR) term caused by random UWA channel.

Therefore, (4) can be rewritten as

$$y_m(n) = h_{td,m}(\boldsymbol{r}_k, n) * h_{ca,m}(\boldsymbol{r}_k, n) * s(n) + e_m(n). \tag{7}$$

Take the coordinate origin as the reference, the time delay $\tau_{m,k}$ in $h_{td,m}(\boldsymbol{r}_k, n)$ can be written as

$$\tau_{m,k} = (\|\boldsymbol{r}_k\| - \|\boldsymbol{r}_k - \mathrm{r_m}\|)/c, \tag{8}$$

where $\|\cdot\|$ denotes the modulus length, and c is the sound velocity.

Supposed that R_m is the distance between the m-th element and the coordinate origin and $R_m \approx 1$ m as the defined small size array. R_k is the distance between the k-th target and the coordinate origin and $R_k > 50$ m as a far-field target. Then, (8) can be rewritten as (9), shown at the bottom of this page.

Due to $\frac{R_m}{R_k} < 0.02 \ll 1$, $\frac{R_m^2}{R_k^2} \approx 0$. And set $\Xi = \sin(\theta_k)\sin(\widehat{\theta}_m)\cos(\theta_k - \widehat{\theta}_m) + \cos(\widehat{\phi}_m)\cos(\phi_k)$, then

$$\begin{aligned} \tau_{m,k} &\approx \left(R_k - R_k\sqrt{1 - \frac{2R_m}{R_k}\Xi}\right)/c \\ &= \left(R_k - R_k\left(1 - \frac{2R_m}{R_k}\Xi\right)^{1/2}\right)/c. \end{aligned} \tag{10}$$

It is obvious that $\Xi \leq 2$, thus, $\frac{2R_m}{R_k}\Xi < 0.08 \ll 1$. Therefore, based on generalized binomial expansion, $(1 - \frac{2R_m}{R_k}\Xi)^{1/2}$ in

$$\begin{aligned} \tau_{m,k} &= \left(R_k - \sqrt{\left(x_k - R_m\sin\left(\widehat{\theta}_m\right)\cos\left(\widehat{\phi}_m\right)\right)^2 + \left(y_k - R_m\sin\left(\widehat{\theta}_m\right)\sin\left(\widehat{\phi}_m\right)\right)^2 + \left(z_k - R_m\cos\left(\widehat{\phi}_m\right)\right)^2}\right)/c \\ &= \left(R_k - R_k\sqrt{1 - \frac{2R_m}{R_k}\left(\sin\left(\theta_k\right)\sin\left(\widehat{\theta}_m\right)\cos\left(\theta_k - \widehat{\theta}_m\right) + \cos\left(\widehat{\phi}_m\right)\cos\left(\phi_k\right)\right) + \frac{R_m^2}{R_k^2}}\right)/c. \end{aligned} \tag{9}$$

(10) can be approximated to

$$\begin{aligned}&\left(1-\frac{2R_m}{R_k}\Xi\right)^{1/2}\\&=\sum_{i=0}^{\infty}\frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\cdots\left(\frac{1}{2}-i\right)!}{i!\left(\frac{1}{2}-i\right)!}\left(-\frac{2R_m}{R_k}\Xi\right)^i\\&=1-\frac{R_m}{R_k}\Xi+\frac{1}{2}\left(\frac{R_m}{R_k}\Xi\right)^2-\frac{1}{2}\left(\frac{R_m}{R_k}\Xi\right)^3\cdots\\&\approx 1-\frac{R_m}{R_k}\Xi,\end{aligned}\tag{11}$$

where "!" denotes the factorial. Then, (8) is simplified as

$$\begin{aligned}\tau_{m,k}&\approx\left(R_k-R_k\left(1-\frac{2R_m}{R_k}\Xi\right)^{1/2}\right)/c\\&\approx\left(R_k-R_k\left(1-\frac{R_m}{R_k}\Xi\right)\right)/c\\&=\frac{R_m}{c}\Xi.\end{aligned}\tag{12}$$

Therefore, the TD term $h_{td,m}(\boldsymbol{r}_k,n)$ is

$$h_{td,m}(\boldsymbol{r}_k,n)\approx\delta\left(n-\frac{R_m f_s}{c}\Xi\right),\tag{13}$$

where f_s is the sampling frequency.

As shown in (13), the TD term $h_{td,m}(\boldsymbol{r}_k,n)$ is only related to (θ_k,ϕ_k) and can be approximated as a function of the DOA of target. Thus, $h_{td,m}(\boldsymbol{r}_k,n)$ can be rewritten as $h_{td,m}((\theta_k,\phi_k),n)$. On the other hand, due to the random frequency fading, $h_{ca,m}(\boldsymbol{r}_k,n)$ also can be rewritten as $h_{ca,m}((\theta_k,\phi_k),n)$, as the impact on $h_{ca,m}$ of position $\boldsymbol{r}_k$ can be moved to the random characteristics of $h_{ca,m}$.

Suppose that the frequency-domain received signal is $\mathbf{Y}(f)=[Y_0(f),\ldots,Y_M(f)]^T, f=1,\ldots,F$, corresponding to the time-domain received signal $\boldsymbol{y}(n)=[y_0(n),\ldots,y_M(n)]^T, n=1,2,\ldots,N$. Then, $\mathbf{Y}(f)$ can be equivalently expressed in matrix form as

$$\begin{aligned}\mathbf{Y}(f)&=\mathbf{H}(\boldsymbol{r}_k,f)S(f)+\mathbf{E}(f)\\&=\mathbf{H}_{td}((\theta_k,\phi_k),f)\mathbf{H}_{ca}((\theta_k,\phi_k),f)S(f)+\mathbf{E}(f),\end{aligned}\tag{14}$$

where $S(f)$ is the discrete Fourier transform (DFT) of the transmitted signal. For position $\boldsymbol{r}_k$, $\mathbf{H}(\boldsymbol{r}_k,f)$, $\mathbf{H}_{td}((\theta_k,\phi_k),f)$, and $\mathbf{H}_{ca}((\theta_k,\phi_k),f)$ are the DFT of CIR, TD, and RCR, respectively. And which are shown as

$$\begin{aligned}&\mathbf{H}(\mathbf{r}_k,f)=[H_0(\mathbf{r}_k,f),\ldots,H_M(\mathbf{r}_k,f)]^T,\\&\mathbf{H}_{td}((\theta_k,\phi_k),f)=[H_{td,0}(\theta_k,\phi_k),f),\ldots,H_{td,M}(\theta_k,\phi_k),f)]^T,\\&\mathbf{H}_{ca}((\theta_k,\phi_k),f)=[H_{ca,0}(\theta_k,\phi_k),f),\ldots,H_{ca,M}(\theta_k,\phi_k),f)]^T.\end{aligned}\tag{15}$$

$\mathbf{E}(f)=[E_0(f),\ldots,E_M(f)]^T$ is the DFT of the additive stationary zero-mean white Gaussian noise $\boldsymbol{e}(n)=[e_0(n),\ldots,e_M(n)]^T$ satisfied $E[e_m(n)]=0$ and $E[e_m(n)s(n)]=0$. F is the length of DFT and "$E[\cdot]$" denotes the expectation.

III. DOA Estimation Method

A. Sparse Signal Reconstruction

For all the possible directions of the target sources, (θ,ϕ) are uniformly discretized to a $q_1\times q_2$ spatial direction matrix $\mathbf{Doa}=\begin{Bmatrix}(\theta_1,\phi_1)&\cdots&(\theta_1,\phi_{q_2})\\\vdots&\ddots&\vdots\\(\theta_{q_1},\phi_1)&\cdots&(\theta_{q_1},\phi_{q_2})\end{Bmatrix}$, which can be further constructed as an $1\times q_1q_2$ spatial direction vector $\boldsymbol{Doa}=\{(\theta_1,\phi_1),\ldots,(\theta_1,\phi_{q_2}),(\theta_2,\phi_1),\ldots,(\theta_{q_1},\phi_{q_2})\}$ by expanding the spatial direction matrix according to column. Define the DDTR matrix $\mathbf{D}((\theta,\phi),f)=[\mathbf{H}((\theta_1,\phi_1),f),\ldots,\mathbf{H}((\theta_D,\phi_D),f)]$, $D=q_1q_2$ denotes the number of the elements in $\boldsymbol{Doa}$, or the total number of the discretized spatial directions. Then, (14) can be written as

$$\begin{aligned}\mathbf{Y}(f)&=\mathbf{D}((\theta,\phi),f)\mathbf{S}((\theta,\phi),f)+\mathbf{E}(f)\\&=\mathbf{D}_{td}((\theta,\phi),f)\mathbf{D}_{ca}((\theta,\phi),f)\mathbf{S}((\theta,\phi),f)+\mathbf{E}(f),\end{aligned}\tag{16}$$

where $\mathbf{D}_{td}((\theta,\phi),f)=[\mathbf{H}_{td}((\theta_1,\phi_1),f),\ldots,\mathbf{H}_{td}((\theta_D,\phi_D,f)]$ and $\mathbf{D}_{ca}((\theta,\phi),f)=[\mathbf{H}_{ca}((\theta_1,\phi_1),f),\ldots,\mathbf{H}_{ca}((\theta_D,\phi_D),f)]$ are the redundant TD and RCR matrix corresponding to $\mathbf{H}_{td}((\theta,\phi),f)$ and $\mathbf{H}_{ca}((\theta,\phi),f)$. $\mathbf{S}((\theta,\phi),f)=[S((\theta_1,\phi_1),f),\ldots,S((\theta_D,\phi_D),f)]^T$ is the frequency-domain vector corresponding to $\boldsymbol{Doa}$. It is easy to find that the DOAs of targets can be estimated according to the indexes of the non-zero elements in $\mathbf{S}((\theta,\phi),f)$.

Generally, the number of target sources is much less than the number of the discretized spatial directions, i.e., $K\ll D$. Thus, $\mathbf{S}((\theta,\phi),f)$ is said to be sparse, and according to CS theory, $\mathbf{S}((\theta,\phi),f)$, $\mathbf{Y}(f)$ and $\mathbf{D}((\theta,\phi),f)$ can be regarded as the sparse vector, observation vector and dictionary matrix, respectively.

Then, the estimation of DOA can be transformed into a sparse recovery problem shown in (17) with the sparsity factor K.

$$\begin{aligned}&\min\ \|\mathbf{S}((\theta,\phi),f)\|_0\\&\text{s.t.}\ \ \|\mathbf{D}((\theta,\phi),f)\mathbf{S}((\theta,\phi),f)-\mathbf{Y}(f)\|\le\varepsilon,\end{aligned}\tag{17}$$

where ε is the allowable reconstruction error. Appendix proves that $\mathbf{D}((\theta,\phi),f)$ satisfies the restricted isometry property (RIP):

$$\begin{aligned}(1-\delta_{2K})\|\mathbf{S}((\theta,\phi),f)\|_2^2&\le\|\mathbf{D}((\theta,\phi),f)\mathbf{S}((\theta,\phi),f)\|_2^2\\&\le(1+\delta_{2K})\|\mathbf{S}((\theta,\phi),f)\|_2^2\end{aligned},\tag{18}$$

where $\delta_{2K}\ll 1$ is a static related to the sparsity of $2K$. Therefore $\mathbf{S}((\theta,\phi),f)$ can be reconstructed losslessly with high probability from $\mathbf{D}((\theta,\phi),f)$ and $\mathbf{Y}(f)$ by using an appropriate reconstruction algorithm, such as OMP algorithm [27]. And then, the DOAs of the target sources can be further estimated through the vector $\mathbf{S}((\theta,\phi),f)$.

Algorithm 1: ESFA-CS/CSFA-CS algorithm.

Input: $\boldsymbol{y}(n)$, $\boldsymbol{Doa}$, K, F_η.

1: **Initialization** : Construct the dictionary matrix $\mathbf{D}_{td}((\theta,\phi),f)$ with (13), and the frequency-domain vector $\mathbf{Y}(f), f=1,\ldots,F$ using DFT. Set the estimated DOA set $\Theta \in \emptyset$.

2: For ESFA-CS algorithm, select the top F_η frequency bins with the largest $P_{\mathbf{Y}}(f)$ as the set $\Omega_E = \{\underset{f}{max}(P_{\mathbf{Y}}(f), F_\eta)\}$. Or for CSFA-CS algorithm, select the top F_η frequency bins with the largest $\mathbf{R}(f)$ as the set $\Omega_R = \{\underset{f}{max}(\mathbf{R}(f), F_\eta)\}$.

3: For each $f \in \Omega_E$ or $f \in \Omega_R$, set $\mathbf{Y}(f)$ and $\mathbf{D}_{td}((\theta,\phi),f)$ as the observation vector and dictionary matrix, then, using CS algorithm to reconstrust $\tilde{\mathbf{S}}((\theta,\phi),f)$.

4: Calculate the estimated sparse vector $\widetilde{\tilde{\mathbf{S}}}$ using (23) or (25) for ESFA-CS or CSFA-CS algorithms.

5: Sort $\widetilde{\tilde{\mathbf{S}}}$ in descending order and obtain the corresponding index: $[\widetilde{\tilde{\mathbf{S}}}', \Theta] = sort(\widetilde{\tilde{\mathbf{S}}}, \text{'descend'})$.

6: Select the top K indexes: $\Theta = \Theta(1:K)$.

7: End.

Output: Estimated DOAs: $\{(\theta_1,\phi_1),\ldots,(\theta_K,\phi_K)\} = \boldsymbol{Doa}_\Theta$.

B. SFA-CS Algorithm

However, practically $\mathbf{D}_{ca}((\theta,\phi),f)$ exhibits random frequency selective fading, and cannot be experimentally obtained. Thus, $\mathbf{D}((\theta,\phi),f)$ is hard to be constructed to recover $\mathbf{S}((\theta,\phi),f)$ from (16). Considering that the redundant TD matrix $\mathbf{D}_{td}((\theta,\phi),f)$ is directly related to the array geometry and the target DOAs, $\mathbf{D}_{td}((\theta,\phi),f)$ can be conveniently constructed by calculating delay according to the geometry relationship. Therefore, by rearranging $\mathbf{S}((\theta,\phi),f)$ and $\mathbf{D}_{ca}((\theta,\phi),f)$, (16) can be rewritten as

$$\mathbf{Y}(f) = \mathbf{D}_{td}((\theta,\phi),f)\tilde{\mathbf{S}}((\theta,\phi),f) + \mathbf{E}(f), \tag{19}$$

where $\tilde{\mathbf{S}}((\theta,\phi),f)$ denotes $\mathbf{S}((\theta,\phi),f)$ weighted by $\mathbf{H}_{ca}((\theta,\phi),f)$, which is expressed as

$$\widetilde{\mathbf{S}}((\theta,\phi),f) = \begin{bmatrix} \mathbf{H}_{ca}((\theta_1,\phi_1),f)S((\theta_1,\phi_1),f) \\ \vdots \\ \mathbf{H}_{ca}((\theta_D,\phi_D),f)S((\theta_D,\phi_D),f) \end{bmatrix}. \tag{20}$$

And the sparse recovery problem can be further described as

$$\begin{aligned} \min \quad & \left\| \tilde{\mathbf{S}}((\theta,\phi),f) \right\|_0 \\ \text{s.t.} \quad & \left\| \mathbf{D}_{td}((\theta,\phi),f)\tilde{\mathbf{S}}((\theta,\phi),f) - \mathbf{Y}(f) \right\| \le \varepsilon. \end{aligned} \tag{21}$$

As expressed in (19), the random frequency fading characteristic of $\mathbf{H}_{ca}((\theta,\phi),f)$ unavoidably renders the uncertain spatial sparsity of $\tilde{\mathbf{S}}((\theta,\phi),f)$ and performance degradation of the CS reconstruction algorithm. Considering that in each frequency bin within the target source bandwidth, while the $\mathbf{H}_{ca}((\theta,\phi),f)$ exhibits a random fading feature, the $\tilde{\mathbf{S}}((\theta,\phi),f)$ remains a similar sparse pattern as the target DOA is the same. Thus, the negative impact of $\mathbf{H}_{ca}((\theta,\phi),f)$ on reconstructing $\tilde{\mathbf{S}}((\theta,\phi),f)$ can be suppressed by accumulation $\tilde{\mathbf{S}}((\theta,\phi),f)$ from selected frequency bins, i.e., selective frequency accumulation (SFA). Under the framework of SFA compressed sensing (SFA-CS), introduced in the next paragraph, two types of frequency bins selection methods, energy selection and correlation coefficient selection are adopted to derive ESFA-CS (Energy SFA-CS) and CSFA-CS (Correlation SFA-CS) algorithm, respectively. For moving underwater vehicles, the inherent Doppler spread will inevitably cause signal mismatch and thus degrade the conventional DOA estimation methods. Note that, because the proposed SFA-CS algorithms estimate the DOA by accumulating selected frequency bins, theoretically it can work normally if the Doppler frequency shift is within a certain range.

1) ESFA-CS Algorithm: Define the total frequency energy of the received signals from different array elements is

$$P_{\mathbf{Y}}(f) = \sum_{i=0}^{M} |Y_i(f)|, f = 1,\ldots,F. \tag{22}$$

where "$|\cdot|$" denotes absolute value. Then, suppose that Ω_E is the set of the top $F_\eta = \text{ceil}(\eta * F)$ frequency bins containing the largest $P_{\mathbf{Y}}(f)$, where η denotes the proportion of all frequency bins, and "ceil$(\cdot)$" denotes rounding up. For each energy selected frequency bin $f \in \Omega_E$, ESFA-CS algorithm reconstrucsts the sparse vector $\tilde{\mathbf{S}}((\theta,\phi),f)$ using CS reconstruction algorithm, and calculates the estimated sparse vector $\widetilde{\tilde{\mathbf{S}}}$ according to (23).

$$\widetilde{\tilde{\mathbf{S}}} = \sum_{f\in\Omega_E} \widetilde{\mathbf{S}}((\theta,\phi),f). \tag{23}$$

Then, the indexes of the K largest elements in $\widetilde{\tilde{\mathbf{S}}}$ correspond to the indexes of K DOAs in $\boldsymbol{Doa}$, and the estimated DOAs are obtained.

2) CSFA-CS Algorithm: Note that, as revealed in (23), by energy selection accumulation of $\tilde{\mathbf{S}}((\theta,\phi),f)$ from energetic bins, the performance of ESFA-CS algorithm is still subject to random fading caused by $\mathbf{H}_{ca}((\theta,\phi),f)$, as the energy distribution of $\tilde{\mathbf{S}}((\theta,\phi),f)$ has been weighted by $\mathbf{H}_{ca}((\theta,\phi),f)$.

However, among different frequency bins f, while the random $\mathbf{H}_{ca}((\theta,\phi),f)$ are uncorrelated, the sparse vector $\tilde{\mathbf{S}}((\theta,\phi),f)$ tends to exhibit similar dominant component at (θ,ϕ) associated with target direction. This indicates that the cross-correlation coefficient of $\tilde{\mathbf{S}}((\theta,\phi),f)$ along different f will contribute to suppressing the random fading impact of $\mathbf{H}_{ca}((\theta,\phi),f)$ on $\tilde{\mathbf{S}}((\theta,\phi),f)$.

Thus, a correlation selection strategy is designed under the SFA-CS framework as CSFA-CS. Specifically, suppose that f_{Pmax} is the frequency bin corresponding to the largest $P_{\mathbf{Y}}(f)$, the cross-correlation coefficient between $\tilde{\mathbf{S}}((\theta,\phi),f_{Pmax})$ and $\tilde{\mathbf{S}}((\theta,\phi),f)$ can be written as

$$\mathbf{R}(f) = \frac{|\langle \mathbf{Y}(f_{Pmax}), \mathbf{Y}(f)\rangle|}{\langle |\mathbf{Y}(f_{Pmax})|, |\mathbf{Y}(f)|\rangle}, f = 1,\ldots,F, \tag{24}$$

where "$\langle\cdot,\cdot\rangle$" denotes inner product. Define Ω_R is the set of the top $F_\eta = \text{ceil}(\eta * F)$ frequency bins containing the largest $\mathbf{R}(f)$. Different from ESFA-CS algorithm, CSFA-CS algorithm calculates the estimated sparse vector $\widehat{\widetilde{\mathbf{S}}}$ according to (25).

$$\widehat{\widetilde{\mathbf{S}}} = \sum_{f\in\Omega_R} \widetilde{\mathbf{S}}((\theta,\phi),f). \tag{25}$$

The proposed ESFA-CS/CSFA-CS algorithms are described in Algorithm 1. Here, function "$max\{\mathbf{a}, b\}$" returns the indexes of the top b columns with the largest value in $\mathbf{a}$. "$[\mathbf{a},\mathbf{b}] = sort(\mathbf{c},$'descend'$)$" is a sorting function, sorts $\mathbf{c}$ in descending order and returns the corresponding elements and indexes as separate output variables: $\mathbf{a}$ and $\mathbf{b}$. $\boldsymbol{Doa}_\Theta$ returns the elements in $\boldsymbol{Doa}$ whose indexes are the elements of the set Θ.

In particular, the CS algorithm in Step 3 can be replaced by any other CS reconstruction algorithms, such as convex optimization methods or greedy schemes. However, considered practical underwater engineering applications, greedy schemes such as OMP, SWOMP, CoSaMP are simple to implement, and have low computational complexity and fast convergence speed. Therefore, the ESFA-CS and CSFA-CS in this paper are proposed as ESFA-OMP, ESFA-SWOMP, ESFA-CoSaMP, and CSFA-OMP. And, in Step 2, ESFA-CS and CSFA-CS select the frequency bins as the sets of Ω_E and Ω_R according to energy selection and correlation coefficient selection, respectively. Meanwhile, in Step 4, ESFA-CS and CSFA-CS calculate $\widehat{\widetilde{\mathbf{S}}}$ using (23) and (25), respectively.

IV. Simulation Results

In this section, numerical simulations are carried out to verify the performance of the proposed SFA-CS algorithm, including ESFA-OMP, ESFA-SWOMP, ESFA-CoSaMP, and CSFA-OMP with OMP, SWOMP, CoSaMP, and MVDR algorithms adopted for comparison reference. The simulations were performed using MATLAB R2016a running on a computer with Intel i7-10700 processor and 16 G memory. The root mean squared error (RMSE) and success rate (SR) are used to assess the performance of DOA estimation, which are defined as

$$\text{RMSE} = \sqrt{\frac{1}{2N_{total}}\sum_{n=1}^{N_{total}}\left(\left(\theta_n-\widehat{\theta}_n\right)^2+\left(\phi_n-\widehat{\phi}_n\right)^2\right)}, \tag{26}$$

$$\text{SR} = \frac{N|_{\theta_n=\widehat{\theta}_n\phi_n=\widehat{\phi}_n}}{N_{total}}, \tag{27}$$

where N_{total} is the total number of the spatial location of the target sources. θ_n and ϕ_n are the true value of the elevation and azimuth angles of the n-th source, where $\widehat{\theta}_n$ and $\widehat{\phi}_n$ are the estimation of θ_n and ϕ_n, respectively. $N|_{\theta_n=\widehat{\theta}_n\phi_n=\widehat{\phi}_n}$ denotes the number where $\theta_n=\widehat{\theta}_n$ and $\phi_n=\widehat{\phi}_n$. Consider a small size 4-elements QCA with a radius $d=0.1$ m and height $h=1$ m under one single target. The omnidirectional space is divided into 30 different discretized spatial directions, that is, there are 30 different DOAs in the spatial direction matrix, whose elevation angles θ range from 45° to 135° with a step size of 45°, and azimuth angles ϕ range from 0° to 180° with a step size of 20°. The number of Monte Carlo trials is 200, each of which is obtained by averaging results among all the 30 possible directions.

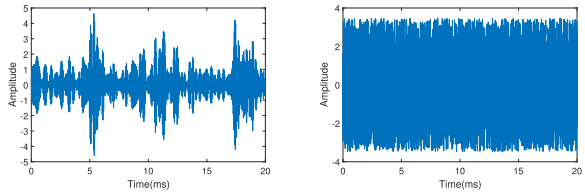

Fig. 3. The waveform of the two types of target signals. (a) Motor-boat noise. (b) White Gaussian noise.

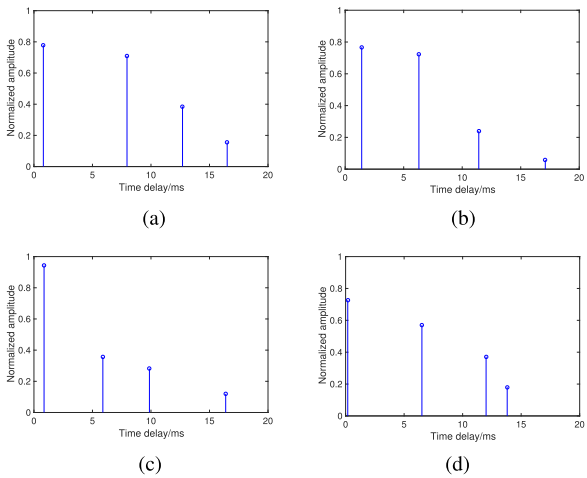

Fig. 4. The underwater channel impulse responses between different array elements and target sources in different directions. (a) Direction 1 of element 0. (b) Direction 2 of element 1. (c) Direction 3 of element 2. (d) Direction 4 of element 3.

To evaluate the influence of different target signals on DOA estimation performance, motor-boat noise experimentally collected in Wuyuan Bay, Xiamen, China and white Gaussian noise artificially generated is adopted as the target signals, respectively. As shown in Fig. 3(a) and Fig. 3(b), these two types of noises exhibit different temporal characteristics. With a sampling rate of 250 kHz, the length of processing window and DFT are both 4096. Different level of additive white Gaussian noise (AWGN) is used as the additive noise to generate an input SNR range of -10 dB to 30 dB with a step size of 5 dB.

An artificial underwater acoustic multipath channel is generated by Bellhop model through Matlab [28]. With a water depth of 50 m, the distance between the array and target source is 200 m, and the depth of array and target source are both 10 m. Fig. 4(a)–4(d) show the underwater channel impulse responses between different array elements and target sources in different directions. It is shown that the TD terms and RCR terms are different for different directions and array elements. Such difference constitutes DDTR which can be explored for spatial sparsity recovery DOA estimation.

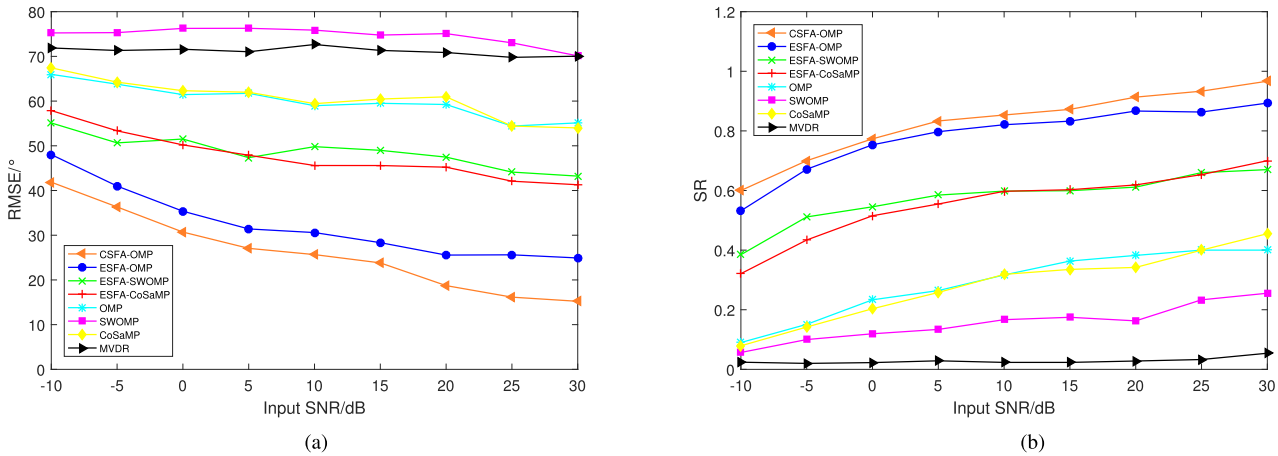

Fig. 5. Comparison of RMSEs and SRs of eight DOA estimation methods as a function of the input SNR with motor-boat noise as the target signal. (a) RMSE. (b) SR.

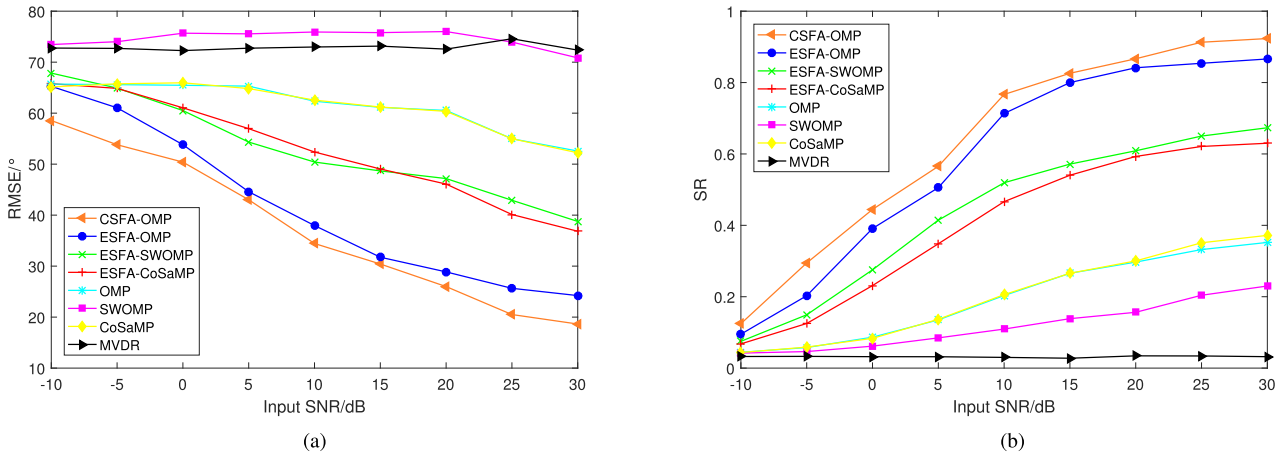

Fig. 6. Comparison of RMSEs and SRs of eight DOA estimation methods as a function of the input SNR with white Gaussian noise as the target signal. (a) RMSE. (b) SR.

A. *Effects of UWA Channel on DOA Estimation*

This section shows three sets of simulations to verify the effects of UWA channel on DOA estimation, including the comparison of RMSEs and SRs under different input SNR using two types of target signals, and the comparison of RMSEs, and SRs under different Doppler spread.

Fig. 5(a) and Fig. 5(b) show the RMSEs and SRs versus the input SNR with motor-boat noise as the target signal for eight different methods, where four SFA-CS algorithms, three CS algorithms, and one traditional algorithm are selected as the comparisons. The proportion for SFA-CS algorithms is set to be $\eta = 20\%$. It can be seen that with an increase in the input SNR, the DOA estimation performances in RMSE and SR of the SFA-CS algorithms and CS algorithms increase. However, MVDR algorithm shows the poor DOA estimation performance for all input SNR. The reason is that the traditional DOA estimation algorithms are subject to significant performance degradation as the array size is seriously limited. On the other hand, compared with the three CS algorithms, SFA-CS algorithms show obvious advantages in RMSE and SR. This is because that the superposition of multi-frequency bins suppresses the influence of $\mathbf{H}_{ca}((\theta, \phi), f)$ on sparsity. Then, the four SFA-CS algorithms can reconstruct the sparse vectors more efficiently and obtain satisfactory DOA estimation performances.

Furthermore, as shown in Fig. 5(a) and Fig. 5(b), CSFA-OMP algorithm achieves a smaller RMSE and higher SR than ESFA-OMP algorithm does, indicating that the correlation selection strategy outperforms the energy selection in mitigating the adverse impact of randomly fading channel response.

Fig. 6(a) and Fig. 6(b) show the RMSEs and SRs versus the input SNR with white Gaussian noise as the target signal, respectively. The proportion for SFA-CS algorithms is set to be $\eta = 20\%$. It can be seen that with an increase in the input SNR, the DOA estimation performances also increase. Even if the RMSEs of SWOMP are larger than that of MVDR, the SRs of SWOMP are higher than that of MVDR, which means that SWOMP could estimate the true DOAs with a greater probability. For MVDR, the performances are still poor even if the input SNR $= 30$ dB.

In addition, compared Fig. 5 with Fig. 6, one may see that, for both two types of source signals, the SFA-CS algorithms achieve

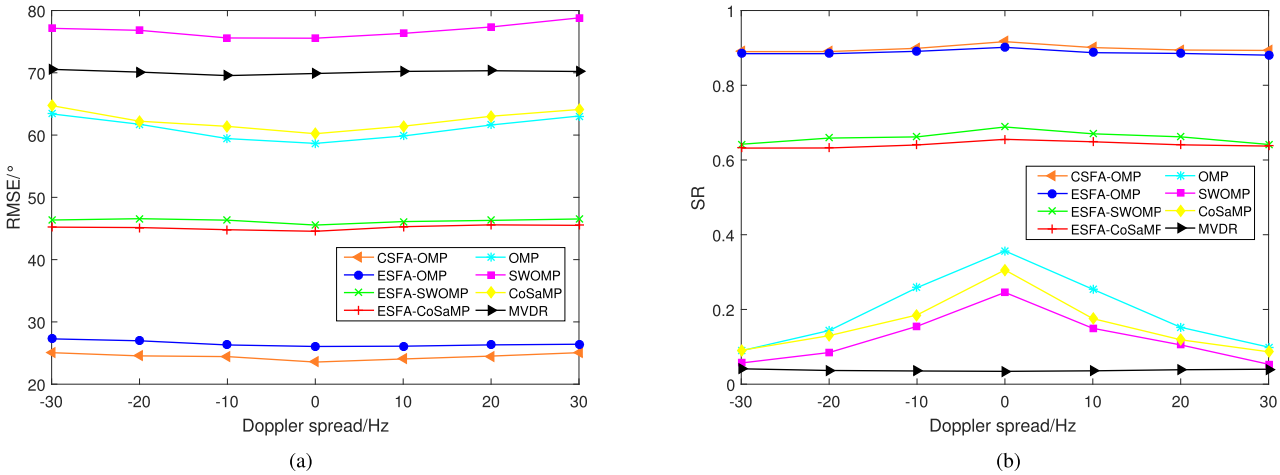

Fig. 7. Comparison of RMSEs and SRs of eight DOA estimation methods as a function of different Doppler spread with white Gaussian noise as the target signal. (a) RMSE. (b) SR.

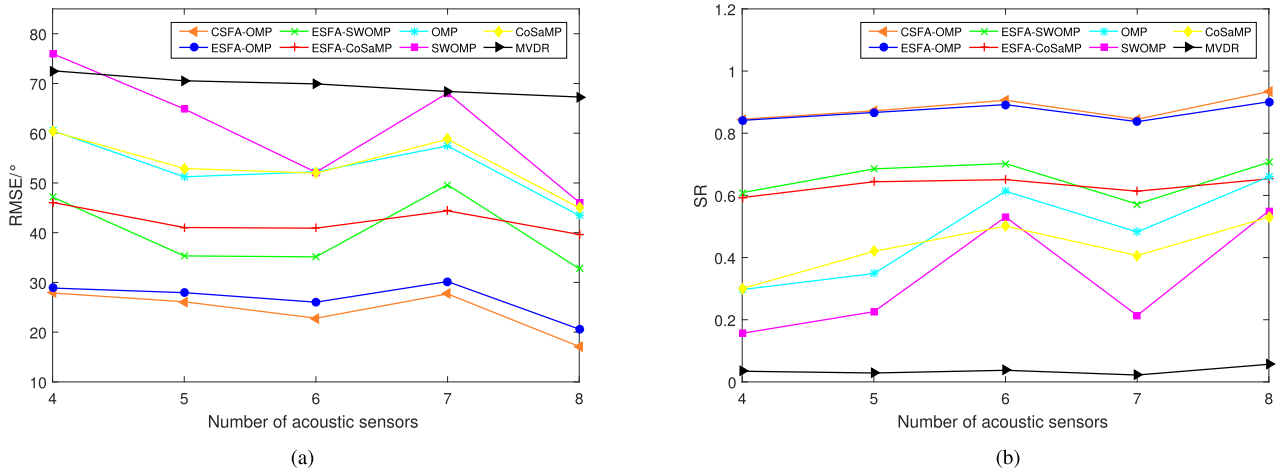

Fig. 8. Comparison of RMSEs and SRs of eight DOA estimation methods as a function of different numbers of array elements with white Gaussian noise as the target signal. (a) RMSE. (b) SR.

better DOA estimation performance than the CS type algorithms and MVDR algorithm do under different SNR. Meanwhile, from both Fig. 5 and Fig. 6, it is also obvious that the CSFA-OMP algorithm exhibits better RMSE and SR results than the ESFA-OMP does under the SFA framework. This reveals that the correlation selection strategy outperforms the energy selection in mitigating the adverse impact of randomly fading UWA channel. Considering the different target signal types, under the same SNR, the motor-boat noise corresponds to a slightly better behavior than the white Gaussian noise. For example, under the specific SNR of 5 dB, the RMESs of the former and latter are 30.7° and 44.6°, respectively.

Considering the doubly selective characteristics of UWA channel, Fig. 7(a) and Fig. 7(b) show the comparison of RMSEs and SRs of eight DOA estimation methods as a function of different Doppler spread, respectively. The input SNR is set to be 20 dB. The proportion for SFA-CS algorithms is $\eta = 20\%$. The Doppler spread is imposed on the source signal by artificial resampling and set to range from -30 Hz to 30 Hz with a step size of 10 Hz. It can be seen that the SRs of OMP, SWOMP, and CoSaMP algorithms seriously decline with an increase in Doppler spread. While the performances of the proposed SFA-CS algorithms exhibit certain robustness upon the Doppler spread. The reason is that the DDTR decomposition and frequency accumulation offer particular tolerance capability upon the Doppler spread.

B. *Effects of Array Parameters on DOA Estimation*

This section shows two sets of simulations to verify the effects of array parameters on DOA estimation, including the comparison of RMSEs and SRs under different numbers of array elements, and different sizes of the array. In addition, the DOA estimation performances of the proposed methods under different input SNR and proportion of the frequency η are also presented here.

Fig. 8(a) and Fig. 8(b) show the RMSEs and SRs of eight different methods as functions of different numbers of array elements, respectively. The input SNR is set to be 20 dB and the proportion for SFA-CS algorithms is $\eta = 20\%$. It can be seen that with an increase in the number of array elements, the RMSEs of all the methods decline slightly, meanwhile the SRs

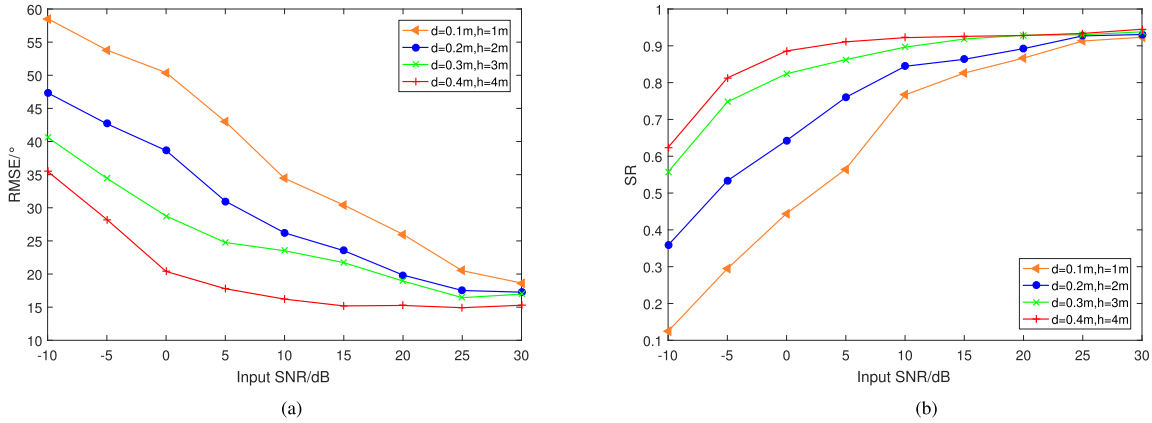

Fig. 9. Comparison of RMSEs and SRs of CSFA-OMP algorithm as a function of different sizes of the array with white Gaussian noise as the target signal. (a) RMSE. (b) SR.

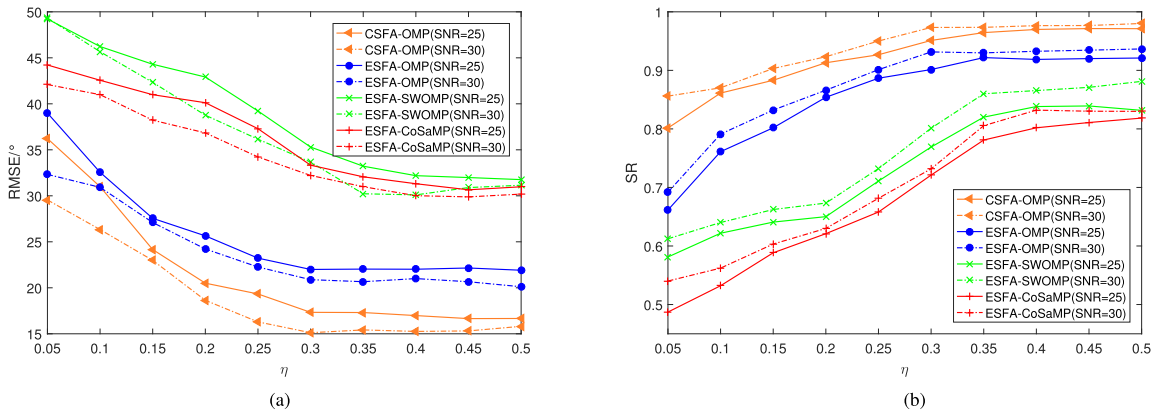

Fig. 10. Comparison of RMSEs and SRs of eight DOA estimation methods as a function of different input SNR and different η. (a) RMSE. (b) SR.

increase slightly. This means that increasing the number of array elements can improve the DOA estimation performance, though the improvement is not obvious. Besides, it is noted that when there are seven array elements, the RMSEs and SRs of most algorithms have deteriorated. According to Fig. 1, the bottom of the array is divided by 6 elements, the azimuth angle of those elements is a multiple of 30°, as well as some of the targets. Therefore, these targets may be coplanar with some array elements, resulting in performance degradation. For MVDR, even if the array elements are increased, the performance is still poor.

To verify the DOA estimation performance of the proposed CSFA-OMP algorithm under different sizes of the QAC, Fig. 9(a) and Fig. 9(b) show the comparison of RMSEs and SRs, respectively. The size of the QAC is determined by the radius d and height h. Therefore, four groups of simulations are designed in this paper, that is, $d = 0.1$ m, $h = 1$ m; $d = 0.2$ m, $h = 2$ m; $d = 0.3$ m, $h = 3$ m; $d = 0.4$ m, $h = 4$ m. The proportion for CSFA-OMP algorithm is $\eta = 20\%$, and the input SNR is set to a range of -10 dB to 30 dB with a step size of 5 dB. It can be seen that when the input SNR is at a low level, the DOA estimation performance of large-sized arrays is significantly better than that of small-sized arrays, i.e. SR $= 62.4\%$ for $d = 0.4$ m, $h = 4$ m versus SR $= 12.5\%$ for $d = 0.1$ m, $h = 1$ m when SNR $= -10$ dB. With an increase in the input SNR, the difference in DOA estimation performance between different size arrays gradually decreases. However, the performance of large size arrays is still better than that of small size arrays. This is because a large size array leads to a wide distribution of acoustic sensor spacing, resulting in a more negligible correlation of the received signals of each acoustic sensor compared with the small-sized array. Therefore, the signals received by the large-sized array contain more spatial information and can be used to estimate the DOA more accurately.

Fig. 10(a) and Fig. 10(b) show the RMSEs and SRs of four SFA-CS algorithms as functions of different input SNR and proportion of the frequency η with white Gaussian noise as the target signal. The input SNRs are set to be 25 dB and 30 dB, respectively. It can be seen from the figures that with the increase of η, the performances of the four algorithms under different input SNRs increase slowly and then tend to be stable. That is, the performances of the four algorithms will not keep increasing as the increase of η. For CSFA-OMP and ESFA-OMP algorithms, the performances tend to be stable when $\eta = 0.3$. And, for ESFA-SWOMP and ESFA-CoSaMP algorithms, the performances tend to be stable when $\eta = 0.4$. The above results show that a suitable proportion of the frequency η can improve the performance of SFA-CS algorithms. However, considering

TABLE I
COMPLEXITIES AND AVERAGE RUNNING TIMES OF EIGHT DOA ESTIMATION METHODS

Algorithm	Complexity	Ave-time (s)
CSFA-OMP	$O\left(\eta FKD(M+1)+F\right)$	0.0054
ESFA-OMP	$O\left(\eta FKD(M+1)\right)$	0.0053
ESFA-SWOMP	$O\left(\eta FSD(M+1)\right)$	0.0053
ESFA-CoSaMP	$O\left(2\eta FK^2D(M+1)\right)$	0.0056
OMP	$O\left(KD(M+1)\right)$	0.0002
SWOMP	$O\left(SD(M+1)\right)$	0.0002
CoSaMP	$O\left(2K^2D(M+1)\right)$	0.0003
MVDR	$O\left(N(M+1)^2\right)$	0.0113

the complexity of the algorithms, η needs to be determined according to the actual situation.

C. Complexities

Table I lists the computational complexities and average running times of the eight DOA estimation methods, where N, $M+1$ and F are the numbers of samples, array elements and the length of DFT. D is the number of elements in $\mathbf{Doa}$, and K is the sparsity factor. The proportion for SFA-CS algorithms is η. Each algorithm was run for 50 times to calculate the average running time. OMP algorithm is mainly in the calculation of the inner product, the common computational complexity is approximately $O(KD(M+1))$. SWOMP method needs to set maximum iterations to halt the SWOMP algorithm, the computational complexity is approximately $O(SD(M+1))$. CoSaMP method selects multiple atoms in each iteration, and some of them may be discarded in the next iteration. Therefore, the computational complexity is approximately $O(2K^2D(M+1))$. SFA-CS algorithms add the processing of multi-frequency bins on the basis of CS algorithms. Thus, the complexity of SFA-CS algorithms will increase too. And, for CSFA-OMP algorithm, the process of correlation coefficient calculation also increases the computational complexity. MVDR algorithm has a computational complexity of about $O(N(M+1)^2)$, due to calculating and inverting of the covariance matrix, corresponding to the longest running time as given in Table I.

We may also observe from Tab. 1 that, with additional calculation for multi-frequency bins accumulation, the running time of SFA-CS algorithms is an order of magnitude higher than that of CS algorithms. However, considering the obvious performance enhancements yielded by the SFA-CS type algorithms, moderately increasing computational complexity is still acceptable.

V. CONCLUSION

In this paper, a novel small size array underwater DOA estimation method is proposed by exploring DDTR constituted spatial sparsity to overcome the performance constraint of traditional DOA estimation approaches under the limited array size and element number. Meanwhile, the adverse impact of randomly fading UWA channel was mitigated by decomposing the DDTR into TD term and RCR term. While TD term is conveniently related to array geometry, RCR term can be addressed by SFA-CS. Moreover, two types of selection strategies, i.e., energy selection and correlation selection, are designed under the framework of SFA-CS to enable sparse reconstruction of target DOAs.

Finally, simulation results demonstrate the superiority of the proposed algorithms in terms of RMSE and SR. And the proposed SFA-CS algorithms are also verified to be robust to Doppler spread. The proposed method provides a new way for small size array-based underwater DOA estimation and has the potential of being applied in booming small underwater vehicle scenarios.

APPENDIX

This appendix proves the RIP.

Firstly, as shown in (16), for a certain f, $\mathbf{Y}(f) \in \mathbb{C}^{(M+1)\times 1}$, $\mathbf{D}((\theta,\phi),f) \in \mathbb{C}^{(M+1)\times D}$, and $\mathbf{S}((\theta,\phi),f) \in \mathbb{C}^{D\times 1}$, where $\mathbf{D}((\theta,\phi),f)$ and $\mathbf{S}((\theta,\phi),f)$ can be expressed as

$$\mathbf{D}\left((\theta,\phi),f\right) = \begin{bmatrix} D_{01} & D_{02} & \cdots & D_{0D} \\ D_{11} & D_{12} & \cdots & D_{1D} \\ \vdots & \vdots & \ddots & \vdots \\ D_{M1} & D_{M2} & \cdots & D_{MD} \end{bmatrix}, \tag{28}$$

$$\mathbf{S}\left((\theta,\phi),f\right) = [S_1, S_2, \ldots, S_D]^T. \tag{29}$$

It is easy to find that the DOAs of targets correspond to the indexes of non-zero elements in $\mathbf{S}((\theta,\phi),f)$. Suppose that there are K targets and the DOAs are $\{(\theta_{k_1},\phi_{k_1}),(\theta_{k_2},\phi_{k_2}),\ldots,(\theta_{k_K},\phi_{k_K})\}$. Thus, $S_d \neq 0, d \in \{k_1,\ldots,k_K\}$ and $S_d = 0, d \notin \{k_1,\ldots,k_K\}$. Then,

$$(1-\delta_{2K})\left\|\mathbf{S}\left((\theta,\phi),f\right)\right\|_2^2 = (1-\delta_{2K}) \sum_{d\in\{k_1,\ldots,k_K\}} S_d^2, \tag{30}$$

and

$$(1+\delta_{2K})\left\|\mathbf{S}\left((\theta,\phi),f\right)\right\|_2^2 = (1+\delta_{2K}) \sum_{d\in\{k_1,\ldots,k_K\}} S_d^2. \tag{31}$$

Secondly, $\|\mathbf{D}((\theta,\phi),f)\mathbf{S}((\theta,\phi),f)\|_2^2$ can be written as

$$\left\|\mathbf{D}\left((\theta,\phi),f\right)\mathbf{S}\left((\theta,\phi),f\right)\right\|_2^2 = \sum_{m=0}^{M}\left(\sum_{d\in\{k_1,\ldots,k_K\}} D_{md}S_d\right)^2. \tag{32}$$

According to the CS theory, D_{md} in $\mathbf{D}((\theta,\phi),f)$ is the value after normalization. Suppose that $\mathbf{D}'((\theta,\phi),f)$ is the matrix before normalization and which is given as

$$\mathbf{D}'\left((\theta,\phi),f\right) = \begin{bmatrix} D'_{01} & D'_{02} & \cdots & D'_{0D} \\ D'_{11} & D'_{12} & \cdots & D'_{1D} \\ \vdots & \vdots & \ddots & \vdots \\ D'_{M1} & D'_{M2} & \cdots & D'_{MD} \end{bmatrix}. \tag{33}$$

Therefore,

$$D_{md} = \frac{D'_{md}}{\left\|\mathbf{D}'\left((\theta,\phi),f\right)\right\|} = \frac{D'_{md}}{\sqrt{\sum_{m=0}^{M}\sum_{d=1}^{D} D'^2_{md}}}. \tag{34}$$

And then, (32) can be rewritten as

$$\begin{aligned}&\left\|\mathbf{D}\left((\theta,\phi),f\right)\mathbf{S}\left((\theta,\phi),f\right)\right\|_2^2\\&=\sum_{m=0}^{M}\left(\sum_{d\in\{k_1,\ldots,k_K\}}\left(\frac{D'_{md}}{\sqrt{\sum_{m=0}^{M}\sum_{d=1}^{D}D'^2_{md}}}\right)S_d\right)^2\\&=\frac{\sum_{m=0}^{M}\left(\sum_{d\in\{k_1,\ldots,k_K\}}D'_{md}S_d\right)^2}{\sum_{m=0}^{M}\sum_{d=1}^{D}D'^2_{md}}.\end{aligned}\tag{35}$$

When there is one single target, that is $K=1$. Then, (35) is

$$\begin{aligned}&\left\|\mathbf{D}\left((\theta,\phi),f\right)\mathbf{S}\left((\theta,\phi),f\right)\right\|_2^2=\frac{\sum_{m=0}^{M}\left(D'_{mk_1}S_{k_1}\right)^2}{\sum_{m=0}^{M}\sum_{d=1}^{D}D'^2_{md}}\\&=\frac{\sum_{m=0}^{M}D'^2_{mk_1}}{\sum_{m=0}^{M}\sum_{d=1}^{D}D'^2_{md}}S_{k_1}^2.\end{aligned}\tag{36}$$

Finally, due to $K\ll D$, combined (36) with (30) and (31), it is obvious that there always exists a static δ_{2K} satisfied

$$\left(1-\delta_{2K}\right)S_{k_1}^2\le\frac{\sum_{m=0}^{M}D'^2_{mk_1}}{\sum_{m=0}^{M}\sum_{d=1}^{D}D'^2_{md}}S_{k_1}^2\le\left(1+\delta_{2K}\right)S_{k_1}^2\tag{37}$$

by constructing a suitable $\mathbf{D}((\theta,\phi),f)$. This completes the proof of the RIP. Similarly, when $K>1$, $\mathbf{D}((\theta,\phi),f)$ also satisfies the RIP.

REFERENCES

[1] W. Yan, X. Fang, and J. Li, "Formation optimization for AUV localization with range-dependent measurements noise," *IEEE Commun. Lett.*, vol. 18, no. 9, pp. 1579–1582, Sep. 2014.

[2] G. Han, C. Zhang, L. Shu, and J. J. P. C. Rodrigues, "Impacts of deployment strategies on localization performance in underwater acoustic sensor networks," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1725–1733, Mar. 2015.

[3] I. Ullah, J. Chen, X. Su, C. Esposito, and C. Choi, "Localization and detection of targets in underwater wireless sensor using distance and angle based algorithms," *IEEE Access*, vol. 7, pp. 45693–45704, 2019.

[4] K. T. Wong and M. D. Zoltowski, "Root-MUSIC-based azimuth-elevation angle-of-arrival estimation with uniformly spaced but arbitrarily oriented velocity hydrophones," *IEEE Trans. Signal Process.*, vol. 47, no. 12, pp. 3250–3260, Dec. 1999.

[5] K. T. Wong and M. D. Zoltowski, "Self-initiating MUSIC-based direction finding in underwater acoustic particle velocity-field beamspace," *IEEE J. Ocean. Eng.*, vol. 25, no. 2, pp. 262–273, Apr. 2000.

[6] N. Zou and A. Nehorai, "Circular acoustic vector-sensor array for mode beamforming," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3041–3052, Aug. 2009.

[7] B. Tang and Q. Chen, "Positioning method for an AUV in arctic seawater based on polarization," *IEEE Antennas Wirel. Propag. Lett.*, vol. 18, no. 1, pp. 177–181, Jan. 2019.

[8] J. Yan, D. B. Guo, X. Y. Luo, and X. Guan, "AUV-aided localization for underwater acoustic sensor networks wWith current field estimation," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 8855–8870, Aug. 2020.

[9] Y. Yang, P. Teng, X. Li, and J. Tian, " Study of acoustic position for moving sources in low sky with a small spatial array," *Acta Acustica*, vol. 29, no. 4, pp. 346–352, Jul. 2014.

[10] S. Su X. Wang and C. Sun, " Direction of Arrival(DOA) estimation of multiple sources in the air using small-aperture arrays," *Torpedo Technol.*, vol. 16, no. 6, pp. 48–52, Dec. 2008.

[11] J. C. Sun, W. J. Zhu, and Y. Y. Sun, "DOA and waveform estimation by using small-dimension array," *J. Northw estern Polytechnical Univ.*, vol. 21, no. 2, pp. 155–158, Apr. 2003.

[12] S. Nannuru et al., "Sparse Bayesian learning with multiple dictionaries," *Signal Process.*, vol. 159, no. 1, pp. 159–170, Jun. 2019.

[13] Y. N. Liu, H. Q. Niu, S. S. Yang, and Z. l. Li, "Multiple source localization using learning-based sparse estimation in deep ocean," *J. Acoustic. Soc. Amer.*, vol. 150, no. 5, pp. 3773–3786, Nov. 2021.

[14] Z. B. Shi, G. L. Liang, L. H. Qiu, and G. M. Wan, "Vector hydrophone array design based on off-grid compressed sensing," *Sensors*, vol. 20, no. 23, pp. 6949–6962, Dec. 2020.

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

[16] G. L. Liang et al., "Sparse Bayesian learning based direction-of-arrival estimation under spatially colored noise using acoustic hydrophone arrays," *J. Mar. Sci. Eng.*, vol. 9, no. 2, pp. 127–148, Jan. 2021.

[17] C. Chang and G. Huang, "Low-complexity spatial-temporal filtering method via compressive sensing for interference mitigation in a GNSS receiver," *Int. J. Antennas Propag.*, vol. 2014, pp. 1–8, May 2014.

[18] W. T. Li, Y. J. Lei, and X. W. Shi, "DOA estimation of time-modulated linear array based on sparse signal recovery," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, pp. 2336–2340, Jun. 2017.

[19] A. C. Gurbuz, V. Cevher, and J. H. Mcclellan, "Bearing estimation via spatial sparsity using compressive sensing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1358–1369, Apr. 2012.

[20] J. H. Li, Y. V. Zakharov, and B. Henson, "Multibranch autocorrelation method for doppler estimation in underwater acoustic channels," *IEEE J. Ocean. Eng.*, vol. 43, no. 4, pp. 1099–1113, Oct. 2018.

[21] X. Y. Han, M. Q. Liu, S. Zhang, R. Zheng, and J. Lan, "A passive DOA estimation algorithm of underwater multipath signals via spatial time-frequency distributions," *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3439–3455, Apr. 2021.

[22] M. Aytekin, E. Grassi, M. Sahotam, and C. F. Moss, "The bat head-related transfer function reveals binaural cues for sound localization in azimuth and elevation," *J. Acoust. Soc. America*, vol. 116, no. 6, pp. 3594–3605, Dec. 2004.

[23] X. C. Sun et al., "Sound localization and separation in three-dimensional space using a single microphone with a metamaterial enclosure," *Adv. Sci.*, vol. 7, no. 2, pp. 1–7, Aug. 2019.

[24] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[25] T. Blumensath and M. E. Davies, "Stagewise weak gradient pursuits," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4333–4346, Nov. 2009.

[26] D. Needell and J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.

[27] T. Zhang, "Sparse recovery with orthogonal matching pursuit under RIP," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 6215–6221, Sep. 2011.

[28] M. B. Porter, *The Bellhop Manual and User's Guide: Preliminary Draft*. La Jolla, CA, USA: Heat, Light, and Sound Research, Inc., Jan. 2011.

Jiaheng Li received the master's degree in computer science and technology from the Qingdao University of Science and Technology, Qingdao, China, in 2020. He is currently working toward the Ph.D. degree with the Xiamen University, Xiamen, China. He is a Reviewer of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. His research interests include underwater acoustic communications, array signal processing, and DOA estimation.

Feng Tong (Member, IEEE) received the Ph.D. degree in underwater acoustics from Xiamen University, Xiamen, China, in 2000. From 2000 to 2002, he was a Postdoctoral Fellow with the Department of Radio Engineering, Southeast University, Nanjing, China. From January 2003 to July 2004, he was a Research Associate with the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, Hong Kong. From December 2009 to December 2010, he was a Visiting Scholar with the Department of Computer Science and Engineering, University of California San Diego, La Jolla, CA, USA. He is currently a Professor with the Department of Applied Marine Physics and Engineering, Xiamen University. His research interests include underwater acoustic communication and acoustic signal processing. Dr. Tong is also a Member of the Acoustical Society of China and China Ship Instrument Society. He is on the Editorial Board for the *Journal of Marine Science and Application*.

Yuehai Zhou received the bachelor's, master's, and Ph.D. degrees in marine sciences from the College of Ocean and Earth Sciences, Xiamen University, Xiamen, China, in 2010, 2013, and 2018, respectively. From February 2016 to January 2018, he was a Visiting Student with the Department of Electrical and Computer Engineering, The University of Alabama, Tuscaloosa, AL, USA. From November 2018 to January 2020, he was a Postdoctoral Researcher with the Acoustic and Navigation Laboratory, University of Haifa, Haifa, Israel. He is currently an Associate Professor with the College of Ocean and Earth Sciences, Xiamen University. His research interests include underwater acoustic signal processing, underwater acoustic network, and underwater acoustic modem design.

Yi Yang received the master's degree in mechanical design and theory from Northeastern University, Shenyang, China in 2011. From 2011 to 2012, he was a R&D Engineer with the Center of Automatic Instruments and Meters, Northeastern University. From 2012, He is with the Shenyang Institute of Automation, Chinese Academy of Sciences. He is currently an Associate Professor with Oceanic Underwater Robot Technology and Engineering Department, Shenyang Institute of Automation, Chinese Academy of Sciences, Beijing, China. His research interests include mini AUV and AUV swarms, air-water trans-domain unmanned vehicles, underwater sensing, and communication technologies.

Zhiqiang Hu received the Ph.D. degree in mechatronic engineering from the Shenyang Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2013. Since 2002, he has been with the Shenyang Institute of Automation, Chinese Academy of Sciences. He is currently a Professor and the Dean of Oceanic Underwater Robot Technology and Engineering Department, Shenyang Institute of Automation, Chinese Academy of Sciences. His research interests include unmanned underwater vehicles, AUV swarms, and air-water trans-domain unmanned vehicles.