Augmented subspace MUSIC method for DOA estimation using acoustic vector sensor array

Aifei Liu¹,²,³, Deseng Yang¹,²,³, Shengguo Shi⁴,²,³, Zhongrui Zhu¹,²,³, Ying Li³

¹Acoustic Science and Technology Laboratory, Harbin Engineering University, Harbin, 150001, People’s Republic of China
²Key Laboratory of Marine Information Acquisition and Security (Harbin Engineering University), Ministry of Industry and Information Technology, Harbin 150001, People’s Republic of China
³College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin 150001, People’s Republic of China
E-mail: shishengguo@hrbeu.edu.cn

Abstract: In the scenario of ambient noise, the noise powers received by the pressure and velocity components in an underwater acoustic vector sensor (AVS) array are unequal. This paper proves when using the MUSIC method, this inequality causes virtual sources and thus increases the rank of the signal subspace. This fact dramatically degrades the DOA estimation performance of the MUSIC method. Then, an augmented subspace (AS) MUSIC method is proposed to take account of the virtual sources, by augmenting the number of the virtual sources into the signal subspace. Simulation results demonstrate in the case of a high signal-to-noise ratio (SNR), the performance of the AS MUSIC method and the MUSIC method is similar. However, in the case of a low SNR, the AS MUSIC method is superior to the MUSIC method in terms of spatial spectrum, estimation accuracy, and resolution. Experimental results further verify the superiority of the AS MUSIC method over the MUSIC method.

1 Introduction

An array of standard scalar sensors utilises the propagation delays between sensors for the direction of arrival (DOA) estimation of the sources impinging on the array. In contrast, an underwater acoustic vector sensor (AVS) measures the acoustic pressure as well as all three components of the acoustic particle velocity at a single point in space and thus it is superior to an acoustic pressure sensor (APS) [1–6]. Since the directional information of sources is involved in the velocity components of an AVS, an AVS array provides more information of source directions than an APS array [7–9].

The DOA estimation methods for a scalar-sensor array were extended to an AVS array, such as Capon [10], MVDR [11], MUSIC [12–14], and sparse representation methods [15]. Among them, MUSIC and sparse representation methods are significantly superior to the conventional methods such as Capon and MVDR methods in terms of resolution and estimation accuracy. However, the sparse representation method dramatically suffers from model mismatch in the practical usage. In contrast, the MUSIC method has certain robustness to model mismatch. The MUSIC method for a scalar-sensor array is well-known [16], and its performance was theoretically investigated in details by theoretical analysis and numerical results [17]. In the application of an AVS array, the quaternion-MUSIC method [14] was developed by using the quaternion formalism to elegantly model the received data of an AVS array, which reduces the computational burden and the memory space required. In [12], in order to reduce the iterative searching load, a self-initial MUSIC-based method uses ESPRIT to self-generate the coarse estimates of the DOAs of sources to start-off the MUSIC-based method. In addition, an efficient root-MUSIC-based method was proposed for a uniformly spaced AVS array by restoring the Vandermonde array manifold structure [13].

It is worth noting in practical systems, the major noise received by underwater AVSs is the ambient noise instead of the internal sensor noise. In the presence of the ambient noise, the noise power of the velocity components of an AVS array is unequal to the one of the pressure components [10], because the velocity component filters out some of the noise due to their directivity. However, the above-mentioned MUSIC-based methods did not take into account this fact. Instead, they assume that the noise powers of all the components of an AVS are the same.

For a scalar-sensor array, the noise at different sensors might have different powers due to the imperfect channel response and mutual coupling [18]. In this case, the noise is termed as non-uniform noise. In [19, 20], deterministic and stochastic ML estimators were developed by taking into account the unequal noise powers and they give essential performance improvements. However, they suffer from high computational load since they need to solve highly non-linear optimisation problems. A DOA estimator was proposed in [21], and it reduces the computational complexity by excluding the noise powers from the objective function. A subspace-separation-based estimator [22] was developed to isolate the contribution of the noise powers, based on successive array element elimination. It limits the efforts of different noise powers at the cost of the reduction in the size of the array aperture. A noise-power estimation method was proposed in [23]; however, it requires the number of array elements not <3 times the number of sources. In [24], a linear transformation method was proposed within the framework of the sparse representation in order to solve the inequality of noise powers. It faces the problems of high computational load and serious sensitivity to the model mismatch. In [25], the non-uniform noise is tackled by estimating the signal subspace and noise covariance matrix in an iterative manner. In order to avoid the iteration, [26] proposed the reduced covariance matrix (RCM) method and the rank and trace minimisation (RTM) method. The RCM method is limited to uncorrelated signals. The RTM method eliminates the limitation. However, it involves the semi-definite programming (SDP).

The above-mentioned methods can be applied to the AVS array in order to solve the inequality of the pressure and velocity components. However, as indicated above, they either face the problem of highly non-linear optimisation, the loss of the effective aperture, or the SDP.

Here, we have proposed a simple solution to estimate the DOAs of sources in the presence of unequal noise powers of the pressure and velocity components in the AVS array, which is based on the subspace theory and avoids the estimation of the noise powers. Specifically, we have derived when using the subspace-based
methods, the unequal noise powers of the pressure and velocity components of the AVS array lead to M virtual sources, where M is the number of sensors. Following that, we propose an augmented subspace (AS) MUSIC method to take care of the unequal noise powers. In addition, we have proved the standard noise subspace is not orthogonal to the steering vector of the source in the presence of the unequal noise powers, which degrades the performance of the MUSIC method. Simulation results verify the virtual sources affect the orthogonality of the noise subspace and the steering vectors of sources. Moreover, simulation results illustrate the superiority of the AS MUSIC method over the MUSIC method in terms of spatial spectrum, estimation accuracy, and resolution under different scenarios. Experimental result further confirms the superiority of the AS MUSIC method over the MUSIC method.

In Section 2, the AVS array model is briefly introduced. Section 3 proves the unequal noise powers of the pressure and velocity components cause the virtual sources when using subspace-based methods and then proposes the AS MUSIC method. Section 4 analyses how the unequal noise powers of the pressure and velocity components affect the performance of the MUSIC method. Section 5 demonstrates the effect of the virtual sources on the orthogonality between the noise subspace and the steering vector of the source, and illustrates the improvements of the AS MUSIC method by numerical results. Section 6 verifies the superiority of the AS MUSIC method over the MUSIC method by experimental results. The conclusion is drawn in Section 7.

2 Preliminary

2.1 AVS array model [1, 10]

We consider N narrowband and far-field plane waves, travelling in an isotropic, quiescent, homogeneous fluid medium and impinging on an array of M AVSs. The direction of the incident plane wave is shown in Fig. 1, where \( \theta \in [-\pi, \pi] \) is the azimuth angle and \( \phi \in [-\pi/2, \pi/2] \) is the elevation angle.

The received signal vector \( r(t) \) is modelled as

\[
 r(t) = B s(t) + n(t) \tag{1}
\]

where \( s(t) \) is the vector of signal sources, which is expressed as \( s(t) = [s_1(t), \ldots, s_N(t)]^T \), and \( s_i(t) \) is the signal of the \( i \)th source; \( n(t) \) is the noise vector modelled as a zero-mean complex white Gaussian vector, and the noise and signal are uncorrelated; \( B = [b(\theta_1, \phi_1), \ldots, b(\theta_M, \phi_M)]^T \). Note that \( b(\theta_i, \phi_i) \) is the steering vector of the \( i \)th source and it is expressed as

\[
 b(\theta_i, \phi_i) = a(\theta_i, \phi_i) \otimes u(\theta_i, \phi_i). \tag{2}
\]

where \( \otimes \) is the Kronecker product, \( a(\theta_i, \phi_i) \) is the steering vector of the array of \( M \) pressure sensors, and \( u(\theta_i, \phi_i) \) the unit vector of one AVS. Define the dimension of the \( u(\theta_i, \phi_i) \) as \( K \). Without loss of generality, we consider that the AVS consists of one pressure component and three orthogonal velocity components. Then, we have

\[
 u(\theta, \phi) = [1, \cos(\theta)\cos(\phi), \sin(\theta)\cos(\phi), \sin(\phi)]^T. \tag{3}
\]

With the assumption that the signal and noise are uncorrelated, the expected covariance matrix of the received signal is then expressed as

\[
 R = \mathbb{E}[r(t)r^H(t)] = BRB^H + R_n \tag{4}
\]

where \( \mathbb{E}[\cdot] \) is the expectation operation, \( R = \mathbb{E}[s(t)s(t)^H] \), and the noise covariance matrix \( R_n \) is expressed as

\[
 R_n = \mathbb{E}[n(t)n^H(t)] = I_M \otimes R, \tag{5}
\]

where \( R \) is the noise covariance matrix of a single AVS, given as

\[
 R = \text{diag}(\sigma_n^2, \sigma_n^2, \sigma_n^2, \sigma_n^2) \tag{6}
\]

because the directivity of the velocity components as shown in (3) filters out some of the unwanted noise [10].

In practice, \( R \) can be estimated by the average sample covariance matrix of the received signal. Through this paper, \( \hat{\centerdot} \) will denote the estimate of the quantity over which it appears. This estimate is a result of replacing \( R \) with the average sample covariance matrix.

3 Augmented subspace (AS) MUSIC method

As shown in (5), the noise covariance matrix of an AVS array is not an identity matrix. According to (5)–(7), the matrix \( R_n \) can be rewritten as

\[
 R_n = \sigma_n I_K + (\sigma_n^p - \sigma_n^v) Z Z^H \tag{8}
\]

where \( I_K \) is an \( K \times K \) identity matrix, and \( Z = [1, 0, \ldots, 0]^T \) (the length of \( Z \) is equal to \( K \)).

Then, from (5) and (8), we have

\[
 R_n = \sigma_n I_K + \left( \sigma_n^p - \sigma_n^v \right) Z Z^H = \sigma_n I_K + \left( \sigma_n^p - \sigma_n^v \right) Z Z^H \tag{9}
\]

where \( \sigma_n \) is a \( K \)-dimensional vector with the \( (Km - K + 1) \)th element equal to 1 and other elements equal to zero, and \( Z = [z_1, \ldots, z_M] \).

Consequently, (4) can be rewritten as

\[
 R = \tilde{B} \tilde{R} \tilde{B}^H + \sigma_n I_K \tag{10}
\]

where

\[
 \tilde{B} = [B, Z] \tag{11}
\]

and

\[
 \tilde{R} = \mathbb{E}[\tilde{r}(t)\tilde{r}(t)^H] = \mathbb{E}[\tilde{r}(t)\tilde{r}(t)^H] = I_M \otimes \tilde{R}, \tag{12}
\]

where \( \tilde{r}(t) \) is the augmented signal vector of all AVSs.
where $\sigma_M$ is an $M \times M$ matrix with all elements equal to zero.

It is noted that $\hat{R}_s$ has a size of $(N+M) \times (N+M)$ and it is full-rank when the sources are not fully correlated. Thus, the eigen decomposition of the matrix $\hat{R}$ can be expressed as

$$R = \sum_{m=1}^{N+M} \gamma_m v_m v_m^H + \sum_{m=N+M+1}^{KM} \sigma_m v_m v_m^H$$

(13)

where the first term in the right-hand side of (13) is the eigen decomposition of $\hat{B} \hat{R} \hat{B}^H$ and the second term is the eigen decomposition of $\sigma I_{KM}$.

Equation (13) can be rewritten as

$$R = \sum_{m=1}^{N+M} (\gamma_m + \sigma_m^2) v_m v_m^H + \sum_{m=N+M+1}^{KM} \sigma_m v_m v_m^H$$

(14)

According to (10), we obtain

$$v_m^H \gamma_m = \sum_{m=1}^{N+M} v_m^H \gamma_m + \sigma_m^2, \text{ for } m = N + M + 1, \ldots, KM.$$ 

(15)

According to (14), we obtain

$$v_m^H \gamma_m = \sigma_m^2, \text{ for } m = N + M + 1, \ldots, KM.$$ 

(16)

As a sequence, we obtain $\{v_m, m = N + M + 1, \ldots, KM\}$ span the noise subspace, and the columns of $\hat{B}$ span the signal subspace. Therefore, the number of the effective sources is equal to $N + M$ instead of the number of actual sources $N$. This implies the inequality of the noise powers leads to $M$ virtual sources with the steering vectors of $Z$.

Thus, we take $\{v_m, m = N + M + 1, \ldots, KM\}$ to generate the projection matrix of the modified noise subspace $\hat{P}_n$, that is

$$\hat{P}_n = \sum_{m=N+M+1}^{KM} v_m v_m^H.$$ 

(18)

As of (17), we have

$$b^H(\theta, \phi) \hat{P}_n b(\theta, \phi) = 0.$$ 

(19)

Then, we propose the AS MUSIC method to estimate the DOAs of sources by searching the peaks of the following spatial spectrum

$$\left(\theta_i, \phi_i\right)^{\text{AS}} = \arg \max_{(\theta, \phi)} \frac{1}{b^H(\theta, \phi) \hat{P}_n b(\theta, \phi)}$$

(20)

4 Performance degradation of MUSIC method

It is worth noting that the conventional MUSIC method does not consider the virtual sources. It treats the dimension of the signal subspace as $N$ and estimates the DOAs of sources by searching the peaks of the following spatial spectrum

$$\left(\theta_i, \phi_i\right)^{\text{MUSIC}} = \arg \max_{(\theta, \phi)} \frac{1}{b^H(\theta, \phi) P_n b(\theta, \phi)}$$

(21)

where $\hat{P}_n$ is the estimate of $P_n$, and

$$P_n = \frac{1}{N+1} \sum_{m=0}^{N} v_m v_m^H.$$ 

(22)

It is noted that $P_n$ is the projection matrix of the standard noise subspace.

Similarly to (15) and (16), we obtain

$$v_m^H \gamma_m \hat{B}^H \hat{B}^H \gamma_m = \gamma_m, \text{ for } m = N + 1, \ldots, N + M.$$ 

(23)

Assume there is only one single source and denote $\sigma_M^2$ as the power of the pressure component of the signal source. Then, from (11), (12), and (23), we have (24).

$$\sigma_M^2 b(\theta_i, \phi_i) b^H(\theta_i, \phi_i) v_m + (\sigma_m^2 - \sigma_M^2) v_m v_m^H = \gamma_m, \text{ for } m = N + 1, \ldots, N + M.$$ 

(24)

Let $v_m(n)$ denote the $n$th element of the $v_m$ vector, from (24), we obtain (25).

$$v_m^H b(\theta_i, \phi_i) b^H(\theta_i, \phi_i) v_m = \frac{\gamma_m}{\sigma_m^2 - \sigma_M^2} \sum_{n=1}^{N} v_m(n)^2 \| v_m(n') \|_2, \text{ where } \| . \|_2 = \text{norm};$$

(25)

It is noted that $\sigma_M^2 > \sigma_m^2$ as shown in (7). By simulation, we find the right-hand side of (25) is larger than zero, thus

$$v_m^H b(\theta_i, \phi_i) b^H(\theta_i, \phi_i) v_m > 0, \text{ for } m = N + 1, \ldots, N + M.$$ 

(26)

By observing (21) and (22), we can see due to (26), we have

$$b^H(\theta_i, \phi_i) \sum_{n=1}^{N+M} v_m v_m^H b(\theta_i, \phi_i) > 0.$$ 

(27)

Therefore, in contrast to (19), we have

$$b^H(\theta_i, \phi_i) P_n b(\theta_i, \phi_i) > 0.$$ 

(28)

As of (27), $b^H(\theta_i, \phi_i) \hat{P}_n b(\theta_i, \phi_i)$ is always larger than $b^H(\theta_i, \phi_i) P_n b(\theta_i, \phi_i)$. This fact is verified in Section 5.1. Therefore, we can expect the performance of the MUSIC method behaves worse than the AS MUSIC method because of the virtual sources.

5 Numerical results

In the following, we compare the AS MUSIC method, the MUSIC method, and the RTM method in different scenarios. We consider a horizontal uniform linear array with $M = 3$ AVSs and the inter-element spacing equal to 0.2 m, where each AVS is composed of a pressure component and three-dimensional particle velocity components (i.e., $K = 4$). We consider the far-field slowly moving targets and set the number of samples to be 5000. The noise is generated by a bandlimited Gaussian random process with a frequency range from 2 KHz to 4 KHz. It is noted that we consider the noise is spherically isotropic in an ideal case, which implies $(\delta^2_{\phi} / \delta^2_{\theta}) = 3$ [10]. The SNR of the source is defined as $\text{SNR} = 10 \log_{10} (\hat{\sigma}_b^2 / \sigma_s^2)$. We provide the following results based on 100 Monte–Carlo simulation trials.

5.1 Orthogonality verification

We consider a pure-tone source at a frequency of 3 KHz is impinging on the array from $\theta_1 = 50^\circ$ and $\phi_1 = 10^\circ$. 

IET Radar Sonar Navig., 2019, Vol. 13 Iss. 6, pp. 969-975
© The Institution of Engineering and Technology 2019
When the expected covariance matrix $\mathbf{R}$ is employed, Fig. 2 provides $\mathbf{b}^H(\theta_1, \phi_1) \mathbf{P}_n \mathbf{b}(\theta_1, \phi_1)$ used in the AS MUSIC method and $\mathbf{b}^H(\theta_1, \phi_1) \mathbf{P}_n \mathbf{b}(\theta_1, \phi_1)$ used in the MUSIC method. As shown in Fig. 2, the orthogonality in the AS MUSIC is perfect. This result coincides with (19). In contrast, $\mathbf{b}^H(\theta_1, \phi_1) \mathbf{P}_n \mathbf{b}(\theta_1, \phi_1)$ is far from zero, even it is small in the case of high SNRs, which is consistent with (28).

When the estimated covariance matrix $\hat{\mathbf{R}}$ is employed, Fig. 3 illustrates $\mathbf{b}^H(\theta_1, \phi_1) \hat{\mathbf{P}}_n \mathbf{b}(\theta_1, \phi_1)$ used in the AS MUSIC method and $\mathbf{b}^H(\theta_1, \phi_1) \hat{\mathbf{P}}_n \mathbf{b}(\theta_1, \phi_1)$ used in the MUSIC method. In Fig. 3, $\mathbf{b}^H(\theta_1, \phi_1) \hat{\mathbf{P}}_n \mathbf{b}(\theta_1, \phi_1)$ is obviously smaller than $\mathbf{b}^H(\theta_1, \phi_1) \hat{\mathbf{P}}_n \mathbf{b}(\theta_1, \phi_1)$, especially in low SNRs. Therefore, we can expect the AS MUSIC method performs better than the MUSIC method in low SNRs. It is noted the orthogonality is not perfect for the AS MUSIC method in low SNRs such as $-15$ dB. This is due to the fact that the practical covariance matrix is estimated by a limited number of samples instead of the expected one, resulting in the subspace leakage [27].

5.2 Spatial spectrum

5.2.1 One source: The source is the same as that in Section 5.1. The two-dimensional spatial spectrums of the AS MUSIC and the MUSIC when the SNRs are equal to 10 dB and $-5$ dB are shown in Figs. 4 and 5, respectively. From Fig. 4, we observe when SNR = 10 dB, the aforementioned methods have a sharp peak around $(50^\circ, 10^\circ)$. Among them, the RTM method has the highest sidelobes. This might be because of the estimation error of the noise non-uniformity in the RTM method. In the case of a low SNR such as $-5$ dB, as shown in Fig. 5, both the AS MUSIC method and the RTM method have narrower mainlobes and much lower sidelobes than the MUSIC method.

Since the estimation performance of the azimuth and elevation angles is similar, we only give the results of the azimuth angle estimation in the following.

5.2.2 Two sources: Consider two pure-tone sources at the frequencies of 3 and 3.5 KHz impinging on the AVS array from $(20^\circ, 0^\circ)$ and $(60^\circ, 0^\circ)$, respectively. The spatial spectrums of the AS MUSIC, MUSIC, and RTM are provided in Fig. 6 for different
SNRs. From Fig. 6a, we observe the aforementioned methods are able to identify the two sources when SNR = 10 dB. From Fig. 6b, we can see the MUSIC method fails to identify the two sources in the case of −8 dB. In contrast, both the AS MUSIC method and the RTM method successfully identify the two sources.

5.3 Estimation accuracy

Consider a pure-tone source at a frequency of 3 KHz impinging on the array from (50°, 0°). Fig. 7 shows the root mean square error (RMSE) of the azimuth angle estimates versus SNR. In addition, we use the (31) in [19] for the case of unequal noise powers to provide the CRB in Fig. 7. From Fig. 7, it is illustrated that the AS MUSIC method has significantly higher estimation accuracy than the MUSIC method when the SNR is lower than −5 dB. Therefore, the AS MUSIC method is more suitable for the DOA estimation of weak targets. However, it deviates from the CRB when the SNR is lower than −10 dB. On the other hand, as shown in Fig. 7, when the SNR is < −5 dB, the RTM method behaves better than the MUSIC method. However, in this case, its estimation accuracy is lower than that of the AS MUSIC method.

5.4 Resolution

Consider two pure-tone sources at the frequencies of 3 and 3.5 KHz impinging on the AVS array from (30°, 0°) and (30° + Δθ, 0°), respectively, where Δθ is the DOA separation between the azimuth angles of the two sources. The SNR is set to be −5 dB. The RMSE of the azimuth angle estimates and the CRB versus Δθ are shown in Fig. 8. From Fig. 8, we observe in order to successfully discriminate the two sources, the minimum DOA separation required by the AS MUSIC method and the RTM method is 30 degrees, while the one required by the MUSIC method is 40 degrees. As a result, both the AS MUSIC method and the RTM method have higher resolution than the MUSIC method.
However, their performance is far from the CRB when the DOA separation is <30°.

6 Experimental results

We test our proposed method and the MUSIC method in the presence of the ambient noise collected in the anechoic water tank with a size of 25 m × 15 m × 10 m. The experimental array is composed of two AVSs, and it is arranged horizontally in the tank with a depth of 5 m. Each of AVSs consists of one pressure component and two orthogonal velocity components in the horizontal plane, and the distance between the two AVSs is equal to 0.7 m.

The acoustic emission transducer transmits a pure-tone source with a frequency of 4 KHz and an azimuth angle of 225° with respect to the axis of the array. It is located at the same depth as the array (that is, its elevation angle is equal to 0°), and it is 16 m away from the centre of the array. The employed bandpass filter has a frequency range from 3 to 5 KHz. The received signal by the AVS array when the acoustic emission transducer does not transmit source signal is the ambient noise. We adjust the ratio of the power of the received pure source signal to that of the received ambient noise and add them together to generate the real data with different SNRs.

The spatial spectrums of the aforementioned methods are given in Fig. 9 in the cases of different SNRs. From Fig. 9a, it is illustrated that the aforementioned methods have a peak at the true target direction of 225° when SNR = 15 dB. From Figs. 9b and c, we observe, when the SNR is decreased to ~3 dB, the peaks of the MUSIC and RTM methods slightly deviate from the true target direction. In contrast, the AS MUSIC method remains the peak at the true source direction. On the other hand, the AS MUSIC method has significantly lower sidelobes than the MUSIC method.

7 Conclusion

We derive the inequality of the noise powers in an array of $M$ AVSs causes $M$ virtual sources when using the MUSIC method. Following that, we propose an AS MUSIC method, which avoids the highly non-linear optimisation involved in the ML-based methods and the SDP encountered in the RTM method. Thus, it is a simple solution to tackle the unequal noise powers. Furthermore, we prove the MUSIC method for the AVS array degrades because of the virtual sources. Simulation and experimental results demonstrate the AS MUSIC method performs better than the MUSIC method as well as the RTM method when the SNR is low. Thus, the AS MUSIC is more suitable for the DOA estimation of the real data collected in the anechoic water tank (a) SNR = 15 dB, (b) SNR = −3 dB, (c) Zoom-in version of (b) around 225 degrees.
weak underwater targets. It is worth noting the virtual sources cause the performance degradation to the other subspace-based methods as well as the MUSIC method, since it is no longer proper to separate the signal subspace and noise subspace according to the number of sources when the virtual sources are present. As a result, the methodology of the AS MUSIC method can be extended to other subspace-based methods for the AVS array.

8 Acknowledgment
This research was supported by the National Natural Science Foundation of China (grant no. 61701133), the Fundamental Research Funds for the Central Universities (grant no. HEUCFJ180502 and HEUCFJ180505), the Stable Supporting Fund of Acoustics Science and Technology Laboratory (SSJSWDZC2018003).

9 References