IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 7 May 2024, accepted 25 May 2024, date of publication 6 June 2024, date of current version 13 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3410276

RESEARCH ARTICLE

Blind Source Separation and Denoising of Underwater Acoustic Signals

RUBA ZAHEER, IFTEKHAR AHMAD, (Member, IEEE), QUOC VIET PHUNG, (Member, IEEE), AND DARYOUSH HABIBI, (Senior Member, IEEE)
School of Engineering, Edith Cowan University, Perth, WA 6027, Australia

Corresponding author: Ruba Zaheer (r.zaheer@ecu.edu.au)

This work was supported in part by the Department of Jobs, Tourism, Science and Innovation-Defence Science Center, Australia, under Grant G1006608.

ABSTRACT Due to the addition of new underwater vessels and other natural noise contributors, the underwater environment is becoming congested and noisy. Undersea monitoring sonobuoys receive multiple mixed acoustic signals from different vessels that need to be separated and identified in the presence of underwater noise (UWN). It is extremely challenging to separate highly correlated acoustic signals from a noisy mixture without prior knowledge of mixing process and propagation channel. Also, in many cases, the separated signals from the noisy mixture doesn't accurately describe the correct signal. This study proposes a novel multi-stage method to separate underwater acoustic source signals from noisy mixture with suppression in noise. The first stage applies multivariate blind source separation (BSS) technique known as non-negative matrix factorization (NMF) that extracts the source signals from the noisy signal mixture. In the second stage, Minimum mean square error (MMSE) estimator is used to reduce noise in separated/reconstructed source signals by minimizing the mean square error (MSE) between reconstructed acoustic signal and original clean signal, enhancing signal reconstruction quality. The results of this study indicate the effectiveness of the proposed method in terms of MSE, signal-to-distortion ratio (SDR) and cepstral distance (CD) and compare it with existing techniques. Based on simulation outcomes our proposed method demonstrates superior separation performance by reducing MSE upto 47% and improving SDR of reconstructed acoustic signals upto 28% compared to existing solutions.

INDEX TERMS Blind source separation, underwater noise, NMF, minimum mean square error (MMSE), underwater acoustic signal, denoising, target detection.

I. INTRODUCTION

The oceanic environment is becoming competitive and getting contested day by day. The importance to explore the undersea environment is increasing and is directly related to a variety of applications such as ocean environment monitoring [1], inspection [2] and surveillance [3], target recognition [4] and identification of any other marine vessels [5], The acoustic signature radiated by a particular marine vessel has unique information that can be utilized for identification, classification and detection purposes for a diverse range of military and defense applications [6], [7].

The associate editor coordinating the review of this manuscript and approving it for publication was Yougan Chen.

The harsh hydrological characteristics of the ocean and the presence of various Gaussian/non-Gaussian impulsive noises make it hard for existing detection systems to identify underwater vehicles. Figure 1 illustrates the underwater noisy environment and complexity of vessel detection.

Underwater vessels and surface vessels generate acoustic signals which are received by a passive sonar buoy [8] for the identification and further post processing. However, these acoustic mixtures are corrupted by complex marine noise making it extremely challenging for traditional detectors to separate and reconstruct them. Furthermore, due to randomness of the underwater mixing process and propagation channel, the detection system needs to be independent of these prior conditions. Blind Source Separation (BSS) techniques

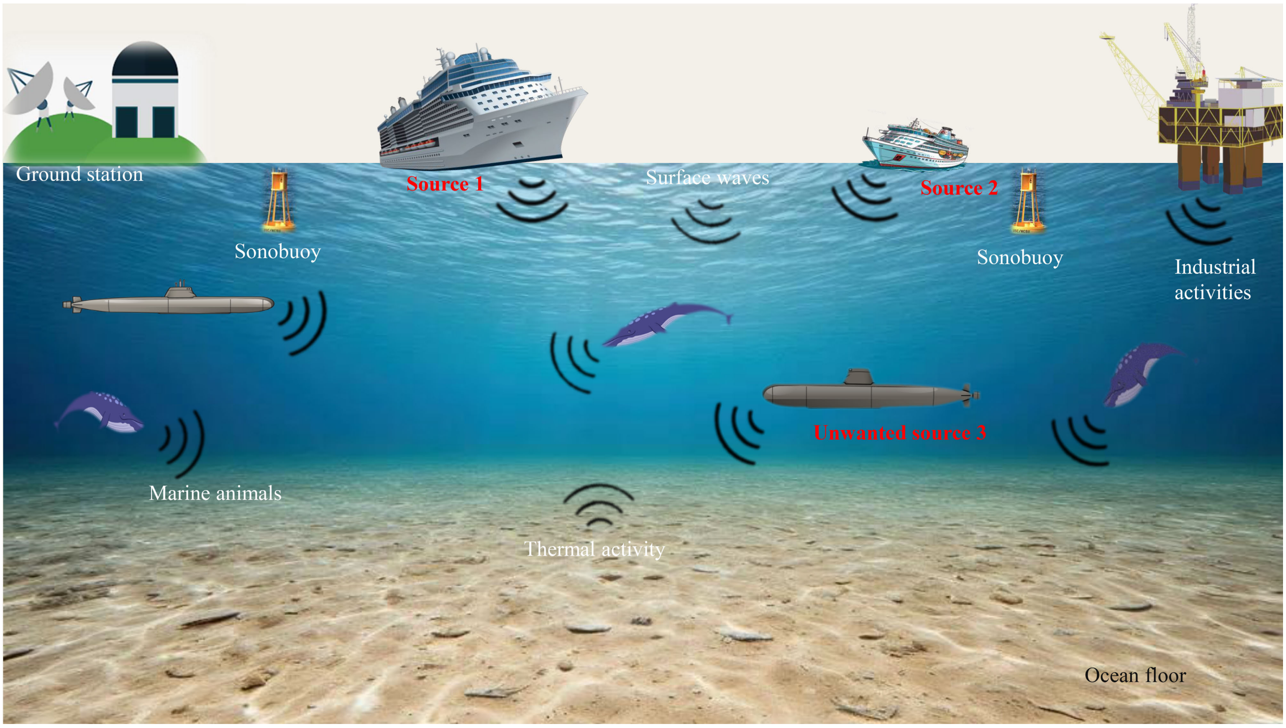

FIGURE 1. Illustration of an ocean noise sources.

serves as a promising solution in for such unpredictable environments [9]. However, the performance and efficiency of BSS is further degraded in noisy environments due to underwater background noise.

Several approaches in the BSS domain are proposed in speech recognition and enhancement systems that justifies the problem in underwater context. Previously, Kalman filtering is used in conjunction with BSS techniques to mitigate the noise problem in a sound mixture [10]. Moreover, authors in [11] improved the performance of BSS for multiple audio sources by using the Wiener filter coefficients generated from the BSS process and applying them to estimate speech signals. Another approach proposed by authors in [12] using a two-channel adaptive BSS built on the Normalized Least-Mean Square (NLMS) technique that uses variable step lengths for steady-state scenarios. Additionally, a gradient-based approach AdaGrade is proposed by authors in [13] for BSS. A combinational approach is proposed in [14] to improve traditional BSS that exploits Eigen filtering for determining the highest frequencies of a signal before applying wavelet denoising. Wavelet denoising suppresses noise components and retains acoustic signals despite their frequency composition.

Further, several statistical signal processing methods are widely applied to noisy mixtures for various applications such as Independent Component Analysis (ICA) [15], Sparse Component Analysis (SCA) [16] and NMF [17]. ICA is based on the concept that each signal source should have a minimum statistical relationship with the other. FastICA is one of the prominently used ICA algorithms for optimization [18]. Fast ICA relies on the assumption that the source signals are statistically independent and non-Gaussian. In complex underwater environments where acoustic signals can exhibit nonlinear mixing and dependencies due to multipath propagation, this assumption may not always hold [19]. Underdetermined systems can be separated using SCA, but sparse signal sources do not apply to all mixed cases [20]. The powerful NMF method solves the problems for a positive multi-source signal matrix [21]. It is applicable to non-sparse cases as well as statistically independent signals and has no non-Gaussian restrictions. Traditional NMF offers a parts-based representation that is useful for identifying sources in mixed signals based on their additive components. However, it does not directly address the issue of minimizing reconstruction error or noise. Therefore, the proposed method in this study combines the reconstruction error minimization of MMSE, which aims to reduce the MSE between the separated signals and the original signals, with the sparsity and part-based representation of NMF. This property allows for a more flexible approach that does not strictly require statistical independence, making it potentially more effective in handling the correlated and complex nature of underwater signals. Moreover, the proposed approach provides an adaptive framework that can better handle the variability and unpredictability of underwater noise sources thus offering a robust solution to aforementioned problem.

Various BSS approaches have also been studied in the maritime domain to resolve different problems, but to our knowledge, none of these studies have addressed the problem of enhancing signal quality post BSS of noisy acoustic mixtures which is a prevalent problem in marine vessel identification and classification. Authors in [22] proposed a negentropy-based FastICA acoustic signals separation technique. However, their proposed method did not consider UWN noise in acoustic signals. Then, another underwater acoustic source separation method is proposed in [23] based on the time-frequency masking in which bidirectional Long Short-Term Memory (Bi-LSTM) is used to estimate the amplitude mask. Mean Square Error (MSE) is computed between the labeled masks and estimated masks as the cost function. However, the ocean's background noise is considered negligible. Moreover, authors in [24] separated various underwater acoustic sources in the presence of multipath and coherence of their power spectral density (PSD). Authors in [25] introduced a method to identify unknown noise sources in 3D underwater environments with multipath effects using Blind Source Separation (BSS). The system applies pre-and postprocessing tailored for underwater conditions to detect dominant sources per frequency subband, characterize them using magnitude squared coherence estimate (MSCE) and maximum cross correlation (MCC), and classify based on the coherence between their estimated power spectral density (PSD) and that of known sources. Therefore, their study is tailored to classify noise sources using BSS model addressing underwater multipath problem only.

Further, a single-channel underwater BSS mechanism is developed by authors in [26] using optimized NMF and FastICA together. The spatial and spectral correlation constraints of underwater sound signals are added to NMF to improve its performance in terms of non-convexity and feature correlation. Recently, authors in [27] presented a denoising method, SVMD-FuDE-WPD, for ship-radiated noise (SN) in the marine environment. It combines successive variational mode decomposition (SVMD), fuzzy dispersion entropy (FuDE)-based dual-threshold analysis, and wavelet packet denoising (WPD) to effectively extract SN from contaminated signals. SVMD decomposes SN into intrinsic mode functions (IMFs), while FuDE classifies IMFs into signal, noise-signal, and noise categories. WPD then denoises the signal IMFs and noise-signal IMFs for final signal reconstruction. The same author in his recent study [28] introduces improved Ensemble Dispersion Lempel-Ziv complexity (EDLZC) and multi-scale (MEDLZC) by incorporating multiple effective mapping methods and coarse granulation, MEDLZC reflects complexity information at different scales, offering improved anti-noise performance, sensitivity to dynamic changes, and effective feature extraction for various types of measured ship-radiated noise (SN) signals. However, in both studies, the denoising is performed for original SN signals.

To the best of our knowledge, existing BSS methods in the underwater domain have primarily focused on separating underwater acoustic signals without addressing noise filtering of reconstructed/separated signals. This practice degrades the detection performance due to residual noise after separation process in realistic scenarios. Therefore, it is very important to enhance the separated signal's quality by minimizing the effect of noise after BSS. While studies in speech [29] and medical [30] domains have explored BSS under diverse noise conditions, there exists a significant gap in addressing the unique challenges posed by various forms of underwater noise on the separated signals after the BSS process for improved detection performance. This study introduces a novel approach that considers the mitigation of underwater Gaussian noise from the separated underwater acoustic signals resulting from the BSS process. Thus, our research is the first of its kind to perform BSS of real underwater signals while effectively mitigating the effects of noise in the reconstructed signals to enhance post BSS signal quality. The main contributions to this paper are as follows:

- An enhanced NMF-based BSS solution for real underwater acoustic signals corrupted by underwater Gaussian noise is proposed.
- We present an underwater noise model based on Gaussian distribution to be added to the acoustic mixture from underwater vessels.
- Then, we perform the source separation from the acoustic mixture by reconstructing the sound sources using our proposed NMF under the noise model.
- Further, the adaptive Minimum mean square error (MMSE) method is implemented to suppress the effect of noise by minimizing the MSE between the estimated and clean acoustic signal.
- Finally, the reconstructed underwater acoustic source signals from our proposed method are compared with the ones obtained from existing source separation techniques, thus, in terms of signal estimation performance, our method outperforms iFastICA and traditional NMF, as indicated by lower MSE, higher SDR, and lower CD.

The rest of the article is organized into the following sections. Section II presents an underwater noise model based on the Gaussian distribution. Section III discusses the concept of the classical NMF separation algorithm and NMF source separation of noisy mixed acoustic signals. Section IV provides the concept of the proposed enhanced MMSE-NMF separation of mixed noisy acoustic signals. In section V, the performance of our proposed MMSE-NMF algorithm is evaluated and compared with existing techniques. Section 6 concludes our paper with some future suggestions.

II. UNDERWATER NOISE MODEL

An Underwater Noise (UWN) model can be developed by employing Sound Pressure Levels (SPL) gathered through transducers deployed in a specific aquatic region or by analyzing recordings captured by hydrophones, to characterize noise statistically. Typically, to model noise in a fixed water area, noise characteristics are described using parameters such as mean, variance, probability density function (PDF), and frequency spectrum. Noise frequency components are represented as average intensities using the power spectrum function. Most of the applications in underwater domain is made using Gaussian noise assumption

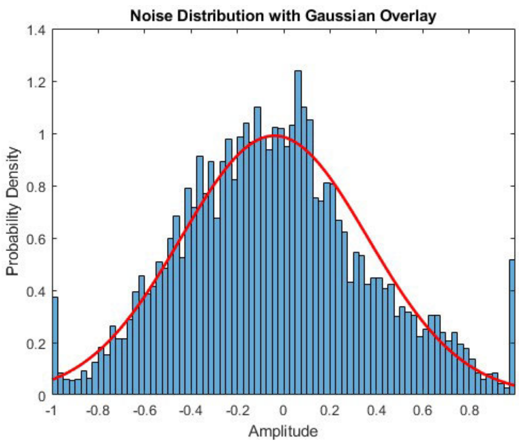

FIGURE 2. Probability distribution of noise data fitted to the Gaussian distribution model.

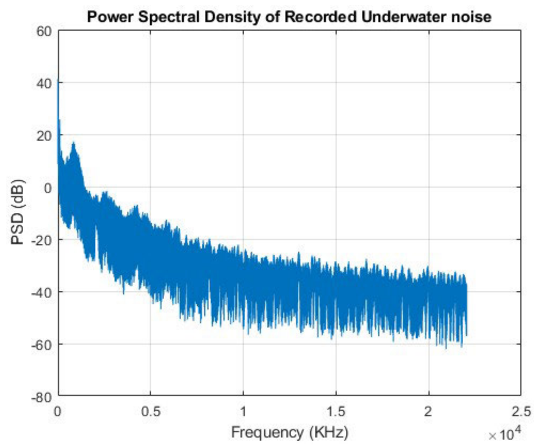

FIGURE 3. Power spectral density of the recorded noise data.

which is the cumulative effect of underwater noises. However, the underwater noise can be non-Gaussian impulsive noise as well [31]. In the study presented in this paper, the Gaussian noise model is considered only. The following equations can be used to determine an amplitude distribution that obeys $X(u, \sigma^2)$ of Gaussian distribution:

$$\phi(x) = \frac{1}{\sigma\sqrt{(2\pi)}} exp\left[-\frac{(x-u)^2}{2\sigma^2}\right] \tag{1}$$

The mean and variance are expressed as:

$$u = \langle x \rangle = \int_{-\infty}^{+\infty} \phi(x)\, pdx \tag{2}$$

$$\sigma^2 = \langle (x-u)^2 \rangle = \int_{-\infty}^{+\infty} (x-u)^2 \phi(x)\, xdx \tag{3}$$

An experiment has been conducted within a controlled environment to verify if the underwater noise follows the Gaussian distribution within a designated aquatic region. The underwater noise is recorded in laboratory shallow water field at 4 m depth using a Snap acoustic recorder with a HTI-96-min hydrophone with sensitivity of -170dBV ref:1uPa kept at one-meter distance from the noise source. The recording procedure is conducted continuously for two hours (7200 secs) at a 44.1KHz sampling rate. Afterward, the acquired data is fitted to various distribution models. Fitting results show that the above Gaussian distribution model can statistically describe noise in the trial water field as shown in Figure 2. This data fitting approach is being used in literature widely to confirm the theoretical distribution of the noise data in the experimental field [32]. The Power Spectral Density (PSD) of the recorded noise data is demonstrated in Figure 3. The noise's frequency spectrum spans from few Hz to more than 2KHz indicating that the noise components across both low and high frequencies. There is a notable decrease in power at frequencies higher than the prominent peak indicating the sound absorption characteristic of water at high frequencies.

III. PROPOSED NMF-MMSE BASED BSS MODEL

Let's assume there are n dimension sources and m dimension mixed entries. Consider the unknown acoustic source signals matrix or coefficient matrix is represented by $U \in R^{n\times r}$ the non-negative signal data matrix is $W \in R^{m\times r}$ the mixed matrix is represented by $Y \in R^{m\times n}$; where r denotes the length of the signal such that, $r \gg m, n$. Then, the signal mixture model can be written as

$$\mathbf{W}_{m\times r} = \mathbf{Y}_{m\times n}\mathbf{U}_{n\times r} \tag{4}$$

As with NMF, the hybrid model decomposes any non-negative matrix $\mathbf{W}$ into the product of two non-negative matrices $\mathbf{Y}$ and $\mathbf{U}$. Thus, NMF can be used to extract the targetted source signal from the mixed signal. This method has a simple algorithm and concise implementation that makes it easy to understand and apply multivariate statistical analysis [21]. The objective function of the NMF algorithm must be optimized and the rules of the factorization matrix must be established in an iterative process to obtain the final required factorization matrix [33]. This method is only suitable for matrix factorization. A signal can be separated based on some parameters or based on direct signal estimation depending on the conditions. By optimizing these parameters, the objective function or cost function of the algorithm is minimized or maximized. This process is known as optimization. Usually, the optimization is carried out under certain constraints. Hence, parameters can be solved or signals can be estimated in order to establish target functions. In order to achieve the factorization of non-negative matrices, it is imperative to find an objective function that can be used to measure the extent of correlation between the matrices. During the whole process, all the matrices must be non-negative. The mostly used NMF objective functions in literature are the objective function under maximum likelihood estimation, Euclidean distance and Kullback-Leibler (KL) divergence [34] which are represented mathematically as follows from (7) to (8) respectively:

$$F(\mathbf{W}|\mathbf{YU}) = \sum_{i=1}^{m}\sum_{j=1}^{n}\Big((\mathbf{W})_{ij} log_2((\mathbf{YU})_{ij} - (\mathbf{YU})_{ij})\Big) \quad (5)$$

$$E(\mathbf{W}|\mathbf{YU}) = \frac{1}{2}||\mathbf{W} - \mathbf{YU}||^2 = \frac{1}{2}\sum_{i,j}\Big((\mathbf{W})_{ij} - (\mathbf{YU})_{ij}\Big)^2 \quad (6)$$

$$KL(\mathbf{W}|\mathbf{YU}) = \sum_{i,j}\Big((\mathbf{W})_{ij}) ln\frac{(\mathbf{W})_{ij}}{(\mathbf{YU})_{ij}} - (\mathbf{W}_{ij}) + (\mathbf{YU})_{ij}\Big) \quad (7)$$

Due to the higher sensitivity of KL divergence compared to Euclidean distance and the potential of KL divergence for acoustic signal estimation, the objective function used in this paper is KL divergence. Thus, the problem is converted into an optimization in (8):

$$\begin{aligned} &[\mathbf{Y}, \mathbf{U}] = arg\, min\, KL(\mathbf{W}||\mathbf{YU}) \\ &s.t. \quad (\mathbf{Y})_{ij} \geq 0, \qquad (\mathbf{U})_{ij} \geq 0 \end{aligned} \quad (8)$$

The incremental iterations for U and Y are written mathematically as follows. These iterations follow the gradient descent method.

$$(\mathbf{U})_{au} \longleftarrow (\mathbf{U})_{au} + \eta_{au}\Big[\sum_i (\mathbf{Y})_{ia}\frac{(\mathbf{W})_{iu}}{(\mathbf{YU})_{iu}} - \sum_i (\mathbf{Y})_{ia}\Big] \quad (9)$$

$$(\mathbf{Y})_{ia} \longleftarrow (\mathbf{Y})_{ia} + \eta_{ia}\Big[\sum_u (\mathbf{U})_{au}\frac{(\mathbf{W})_{iu}}{(\mathbf{YU})_{iu}} - \sum_u (\mathbf{U})_{au}\Big] \quad (10)$$

As a result of gradient descent, the results of an unconstrained optimization problem cannot be guaranteed to be non-negative. However, the multiplication algorithm can be written using the gradient descent method as follows.

$$(\mathbf{U})_{kj} \longleftarrow (\mathbf{U})_{kj}\frac{\sum_i (\mathbf{Y})_{ik}\, (\mathbf{W})_{ij}/(\mathbf{YU})_{ij}}{\sum_i (\mathbf{Y})_{ik}} \quad (11)$$

$$(\mathbf{Y})_{ik} \longleftarrow (\mathbf{Y})_{ik}\frac{\sum_j (\mathbf{U})_{kj}\, (\mathbf{W})_{ij}/(\mathbf{YU})_{ij}}{\sum_j (\mathbf{U})_{kj}} \quad (12)$$

As discussed here, the algorithm is a classical NMF algorithm without any noise involved. A noise component can then be added to the signal mixture, and the influence of noise on the NMF algorithm will be examined.

A. NMF SOURCE SEPARATION UNDER NOISE

When the noise is added to the signal mixture, the signal mixing model with added noise can be represented mathematically as follows.

$$\mathbf{W} = \mathbf{YU} + \mathbf{N} \quad (13)$$

where N is the underwater Gaussian noise, correspondingly, the expanded KL divergence is:

$$KL = \sum_{i,j}\Big((\mathbf{W})_{ij}) ln\frac{(\mathbf{W})_{ij}}{(\mathbf{YU})_{\mathbf{ij}} + (\mathbf{N})_{ij}} - (\mathbf{W})_{ij} + (\mathbf{YU})_{ij} + (\mathbf{N})_{ij}\Big) \quad (14)$$

When the condition of non-negative constraint is applied.

$$\begin{aligned} \sum_{i,j}\Big(\sum_z (\mathbf{Y}_{iz})(\mathbf{U})_{zj} + (\mathbf{N}_{ij})\Big) &= \sum_{z,j}(\mathbf{U})_{zj}\sum_i (\mathbf{Y})_{iz} + \sum_{i,j}(\mathbf{N})_{ij} \\ &= \sum_{z,j}(\mathbf{U})_{zj} + \sum_{i,j}(\mathbf{N})_{ij} \\ &= \sum_{i,j}(\mathbf{W})_{ij} \end{aligned} \quad (15)$$

the optimization model can be written as:

$$\begin{aligned} &[\mathbf{Y}, \mathbf{U}, \mathbf{N}] = arg\, min\, KL(\mathbf{W}||\mathbf{YU}) \\ &s.t. \quad \sum_i (\mathbf{Y})_{ij} = 1 \\ &(\mathbf{Y})_{ij} \geq 0, \qquad (\mathbf{U})_{ij} \geq 0, \qquad (\mathbf{N})_{ij} \geq 0, \end{aligned} \quad (16)$$

The multiplication algorithm is used for iteration as shown below.

$$(\mathbf{U})_{kj} \longleftarrow (\mathbf{U})_{kj}\frac{\sum_i (\mathbf{Y})_{ik}\, (\mathbf{W})_{ij}/\Big((\mathbf{YU})_{ij} + (\mathbf{N})_{ij}\Big)}{\sum_i (\mathbf{Y})_{ik}} \quad (17)$$

$$(\mathbf{Y})_{ik} \longleftarrow (\mathbf{Y})_{ik}\frac{\sum_j (\mathbf{U})_{kj}\, (\mathbf{W})_{ij}/\Big((\mathbf{YU})_{ij} + (\mathbf{N})_{ij}\Big)}{\sum_j (\mathbf{U})_{kj}} \quad (18)$$

B. NMF SOURCE SEPARATION OF NOISY MIXTURE WITH MMSE

Let's consider N number of source signals are linearly mixed together and the mixture of these source signals are observed at M number of sensors/observed mixture. In this study, we have used an underdetermined BSS system which implies that $N > M$ as described in [16]. In this study model, only one receiving sensor is assumed therefore, there is no cross-coupling effect of multiple channels in the received signal mixture. If $s(t)$ denotes an acoustic source signal for $i = 1, \ldots, N$, $M = 1$ then the received underwater acoustic mixture $y(t)$ is an instantaneous linear combination of N number of acoustic sources contaminated by underwater noise $u(t)$, then it can be mathematically modeled as:

$$y(t) = \sum_{i=1}^{N}[h_i(t)s_i(t)] + u(t) \quad (19)$$

where $h_i(t)$ is the impulse response of the unknown underwater channel that is from the i^{th} source to the sensor. The core idea is to separate received mixed signals into their source components without any prior information on clean signals or channel. Our current model assumes that

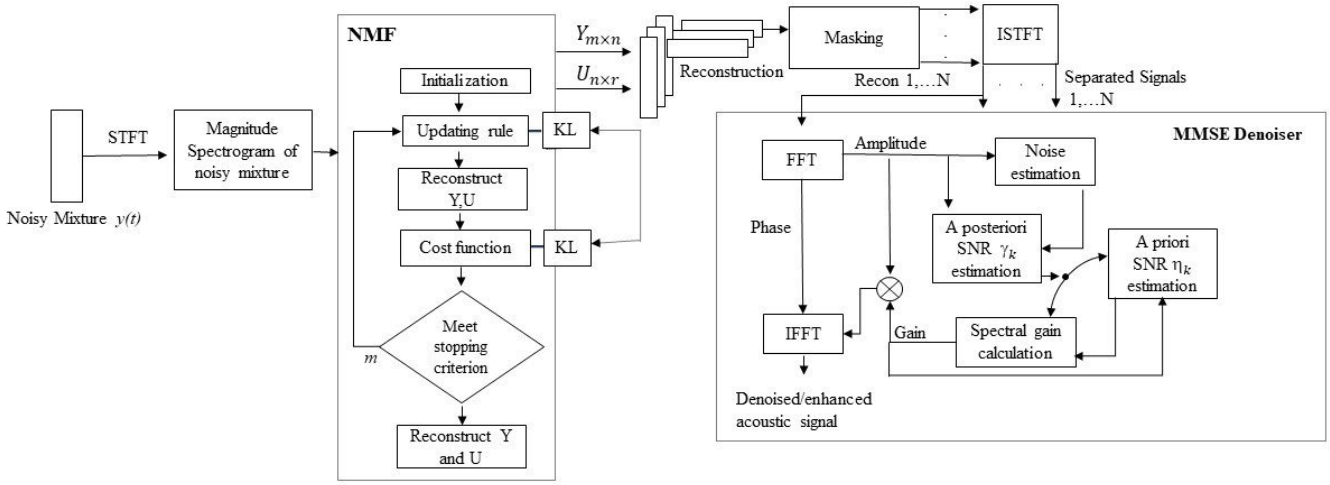

FIGURE 4. **Block diagram of proposed MMSE-NMF system.**

the UWA channel, like the source signals, is unknown and not explicitly modeled. This approach is common in BSS scenarios where the goal is to develop methods robust enough to handle variabilities in the channel without requiring explicit modeling or prior knowledge of channel characteristics. The mixing process and the proposed model of this study are shown in Figure 4. Generally, an acoustic sensor can receive sounds from N number of sources in the presence of an interfering noise. Then the noisy mixture $y(t)$ is fed to the MMSE-NMF blind source separator and estimator unit which separates and estimates the sources $\hat{s}_1(t)$, $\hat{s}_2(t),\ldots,\hat{s}_N(t)$.

Algorithm 1 NMF Separation from Noisy Mixture

Input: W,Y^0, U^0, srcnumber, iter, stopcriterion
Output: Y and U matrices, optimized cost function
Define function *divergence* of KL cost function in Eq. (7)
Set *count* $\leftarrow$ 0;
Set *beta* $\leftarrow$ 1
Initialize cost function as an empty list
Random initialization of matrices Y and U
while *beta* $\geq$ *stopcriterion* and *count* $\leq$ *iter* **do**
  Update Y and U matrices based on gradient descent according to Eq. (9), (10), (11), (12).
  Compute cost function
  beta=*divergence*(W,Y,U)
  costfunction.append(*beta*)
  count=*count*+1
end
if *count* $\leftarrow$ *iter* **then**
  stop the iterative process
else
  convergence after *count* iterations
end
return **Y**, **U**, cost function

First, the goal is to separate all the acoustic source signals from the mixed signal using NMF separation modelled in (16), (17) and (18). Then, well known MMSE estimator is employed to minimize the noise in all of its separated components. MMSE denoising reduces noise while preserving the important features of the observed signal and minimizing signal degradation. In this method, the original signal is estimated by minimizing the MSE between the separated and the clean signal. Its demonstrated ability to effectively denoise signal components corrupted by Gaussian noise makes it statistically optimal for linear signal models with Gaussian noise distributions [35]. Also, MMSE is a highly adaptive technique that has the ability to adjust to the observed acoustic mixture accordingly. Adaptive MMSE filtering dynamically adjusts its filter coefficients based on the characteristics of the input signal and noise. This adaptability allows the filter to effectively track changes in the noise statistics over time, ensuring promising denoising performance in non-stationary environments, such as underwater acoustic channels with varying noise levels and characteristics [36], [37].

If underwater acoustic source signals, noise and received noisy acoustic mixture are converted into the frequency domain, then these are denoted by $S_i(k)$, $N(k)$, $Y(k)$. Here, k denotes the position of the coefficients in the transform domain. With the assumption, that transformed coefficients are statistically independent. The design criteria for the estimator for the observation to minimize the MSE can be written as follows:

$$E[S_i(k) - \hat{S}_i(k)] \tag{20}$$

$E[.]$ represents the expectation operator and $\hat{S}(k)$ is the estimated acoustic source signal. The minimization of mean square error (MSE) in Equ. (20) can be achieved using the MMSE filter. In the presence of underwater noise $y(t)$; $0 \leq t \leq T$ with received signal mixture $Y(k)$. We can obtain estimated $\hat{S}_i(k)$ using [38] model.

$$\hat{S}_i(k) = E[S_i(k)|Y(k)] \tag{21}$$

Then, Equ. (21) can be further expanded by using Bayes' theorem [38]:

$$\hat{S}_i(k) = \frac{\int_{-\infty}^{\infty} a_k p\Big(Y(k)/a_k\Big) p(a_k) da_k}{\int_{-\infty}^{\infty} p\Big(Y(k)/a_k\Big) p(a_k) da_k} \tag{22}$$

Here, $p(.)$ denotes probability density function (PDF) and a_k is a dummy variable that represents all values of $S_i(k)$. In this paper, Gaussian distribution model is presented, then $p\Big(Y(k)/a_k\Big)$ and $p(a_k)$ can be mathematically represented in terms of variance of Gaussian noise as follows:

$$p\Big(Y(k)/a_k\Big) = \frac{1}{\sqrt{(2\pi\lambda_n(k))}} exp\Big(-\frac{(Y(k)-a_k)^2}{2\lambda_n(k)}\Big) \tag{23}$$

$$p(a_k) = \frac{1}{\sqrt{(2\pi\lambda_x(k))}} exp\Big(-\frac{(a_k)^2}{2\lambda_s(k)}\Big) \tag{24}$$

where $\lambda_n(k) = E[|N(k)^2]$ and $\lambda_s(k) = E[|S(k)^2]$ represent the variance of noise and clean acoustic source signal respectively. If $\eta(k)$ denotes a priori SNR, then it can be mathematically explained as the ratio of signal variance to noise variance as follows:

$$\eta(k) = \frac{\lambda_s(k)}{\lambda_n(k)}. \tag{25}$$

Then, by inserting (23) and (24) in (22), the estimated acoustic signal $\hat{S}(k)$ in terms of a priori SNR can now be written mathematically as:

$$\hat{S}(k) = \frac{\eta(k)}{\eta(k)+1}\gamma_k \tag{26}$$

The detailed method for estimating λ_s is mentioned in [39]. The parameters $\hat{\lambda}_s$ and λ_s can be estimated by using method in [38]. For each frame, the a priori SNR estimate is updated by using the following equation:

$$\hat{\lambda}_s = \alpha\hat{\lambda}_s(k)_p + (1-\alpha)max(Y(k)^2 - \lambda_s(k), 0) \tag{27}$$

Here, the max function determines the non-negative values. The estimated value of λ_s in the previous frame is $\hat{\lambda}_s(k)_p$ and α is a adjustable tuning constant to improve estimation results that ranges from 0 to 1. In this study, the α is set to 0.98 that exhibited best estimated results in our case. The complete implementation of our proposed method is detailed in the Algorithm 1 and Algorithm 2.

Algorithm 2 Denoising and Enhancement of Separated Signals

Function: MMSE denoiser
Input data: Separated audio **U** {from algorithm 1}
Result: Denoised audio (.wav)
Calculate frame size *len* from sampling rate of audio
Partition **U** into frames[k] of size *len* using Hamming window with 50% overlap
Initialize noise statistics using the first segment of frames (approximately 0.5 seconds)
Define firstseg[k, f] as the magnitude of the first segment of frames.
Calculate the mean noise level for each frequency index y: NoiseMean[f] = Mean(firstseg[:, f])
Calculate the mean noise variance for each frequency index y: NoiseVar[f] = Mean(firstseg[:, f]^2)
Define Nframes[k, f] as the magnitude of all subsequent frames.
{Set threshold for acoustic activity detection (AAD) alp=0.15}
{Set smoothing parameter alpha=0.98 }
for k in (Nframes) **do**
  for f in (len) **do**
    Perform Fast Fourier Transform (FFT) on each frame to compute Signal variance sig2 ($\lambda_s(k)$) in {Eq. (25)}
    sig[k,f] = |Coeffs[k, f]|
    sig2[k,f] = |Coeffs[k, f]|^2
    $gammak = min(sig2[k,f]/NoiseVar[f], 40)$
    $ksi = alpha + (1-alpha) * max(gammak - 1, 0)$ {from Eq. (27)}
    $logsigmak = gammak \times ksi/(1+ksi) - log(1+ksi)$ {from Eq. 26}
    $aad = sum(logsigmak)/len$ {Acoustic activity detection}
    if Aad < alp **then**
      {update NoiseMean[f], NoiseVar[f]}
    end
    $A = ksi/1 + ksi$ {MMSE estimation}
    $Svk = A \times gammak$ Calculate the spectral gain factor for each frequency bin f in frame k: spectralgain[k, f] for Nframes[k, f]
    Apply the spectral gain to denoise: deframes[k, p] = Nframes[k, f] x spectralgain[k, p]
  end
  InverseFFT(deframes) {Inverse FFT to time domain on all frames}
end

C. COMPLEXITY OF MMSE-NMF AND COMPARISION

Our proposed method combines two distinct approaches: NMF-based Blind Source Separation (BSS) and MMSE-based denoising and enhancement of separated acoustic signals. The computational complexity of each method is evaluated separately. In addition, the complexities of other statistical BSS methods are also evaluated and compared with NMF which is used in this study. Since, NMF is an iterative process to decompose a non -negative matrix W of size $m \times r$ into two non-negative matrices Y (size $m \times n$) and U (size $n \times r$) and multiplicative updates on every iteration, its overall complexity can be written as $O(i.mrk)$ for i number of operations. The solution of NMF based model $W = YU$ refer to the multiplication operations of multiple $n \times n$ matrices, therefore their time complexities are all above $O(n^2)$. The optimization problem associated with NMF involves determining the non-negative rank and and computing the corresponding factorization is more complex than its with respect to determining the non-negative rank and computing the associated factorization, is more difficult than its unconstrained counterpart. It is considered as NP-hard that require both the dimension and the factorization rank of W to increase, which was proved via relating it to NP-hard intermediate simplex problem by a study [40].

The computational complexity of FastICA is roughly $O(mn^2)$ for estimation of n components from m samples involving iterations. The number of iterations can vary significantly based on the data and convergence criteria. Subsequently, Singular Value decomposition (SVD) of $m \times n$ matrix generaly runs at $O(mn^2)$ when using standard algorithms. While, principal component analysis (PCA) typically involves computing the covariance matrix of the dataset which is $O(mn^2)$ (assuming m samples and n features). The overall complexity is dominated by the eigen decomposition step, therefore offering $O(n^3)$. Among these methods, PCA and SVD have the highest computational complexity due to the $O(n^3)$ term involved in eigen decomposition or matrix factorization. FastICA's complexity, while also significant, may be preferable in scenarios where number of components is much smaller than the number of samples. NMF is potentially costly due to the need for multiple iterations and can be efficient if the rank of decomposition is kept low and convergence is reached quickly.

The adaptive MMSE used in this study offers low computational complexity when it comes to Gaussian noise compared to state of the art denoising methods. The computational complexity of the MMSE estimator is mainly determined by the Fast Fourier Transform (FFT) and Inverse Fast Fourier Transform (FFT) operations each contributing $O(N \times LlogL)$. N represents the number of frames is proportional to M/L with 50% overlap, it would be closer to 2M/L and scales to $O(MlogL)$, where M denotes the signal length and L is the frame size.

IV. SIMULATION RESULTS AND PERFORMANCE DISCUSSION

In this section, we discuss the performance evaluation of the proposed method using the MATLAB (R2021b) software on Intel(R) Core(TM) i7-8700 CPU @ 3.2GHz with 16GB RAM system. First, we use underwater source signals from the Shipsear database [41] to generate a mixture of underwater sound sources. Then, we add underwater Gaussian noise to the acoustic mixture to see the effectiveness of our proposed MMSE-NMF separation.

A. SIMULATION SETUP

The BSS problem used in this study is underdetermined, indicating that the number of sources exceeds the observed

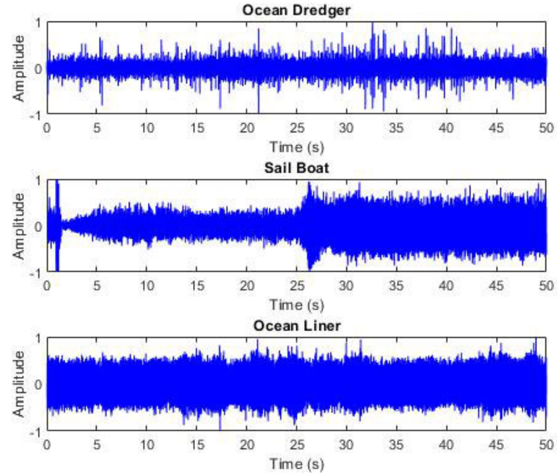

FIGURE 5. Normalized time evolution of source signals.

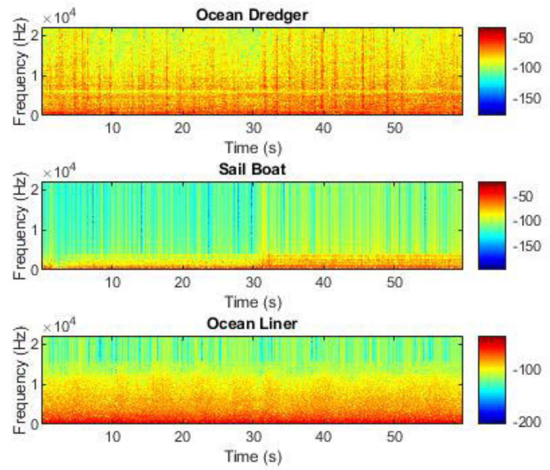

FIGURE 6. Spectrogram of underwater source signals.

mixtures. Theoretically, N number of sources has been considered in the proposed BSS model. A practical approach to address real-world acoustic mixtures is to impose constraints by assuming the number of sources for the mixture. It is a common practice in the literature to make assumptions about the number of sources involved [42], [43]. This assumption allows for a quantitative evaluation of the effectiveness of our proposed algorithm in separating sources under controlled conditions.

Without the loss of generality, we assumed three underwater acoustic source signals ocean dredger, sailboat and ocean liner. All the source acoustic signals have the same length and sampling rate of 52734Hz for simulation purposes as this assumption is being used in studies in the literatures [42] and [30]. However, considering stochastic signals may need enhanced statistical source separation that may increase the complexity of the work. The hamming window is applied with 75% overlap. The three acoustic source signals and their spectrograms are shown in Figure 5 and Figure 6. The spectrogram of three acoustic sources displays the distribution

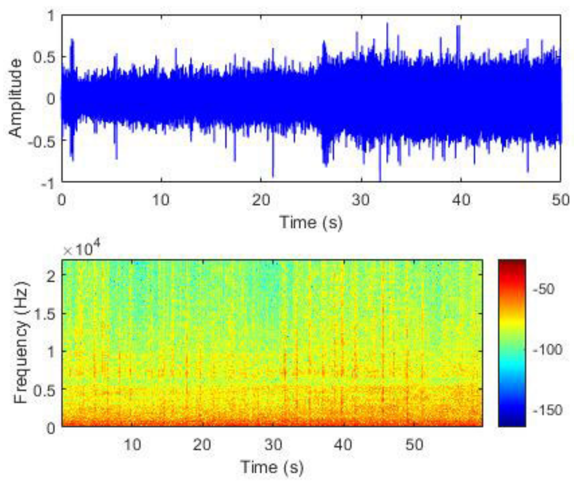

FIGURE 7. Time evolution and spectrogram of mixed acoustic signal.

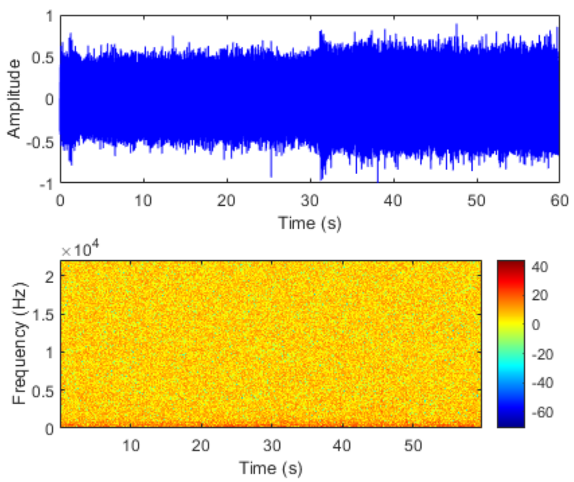

FIGURE 8. Time evolution and spectrogram of noisy mixture.

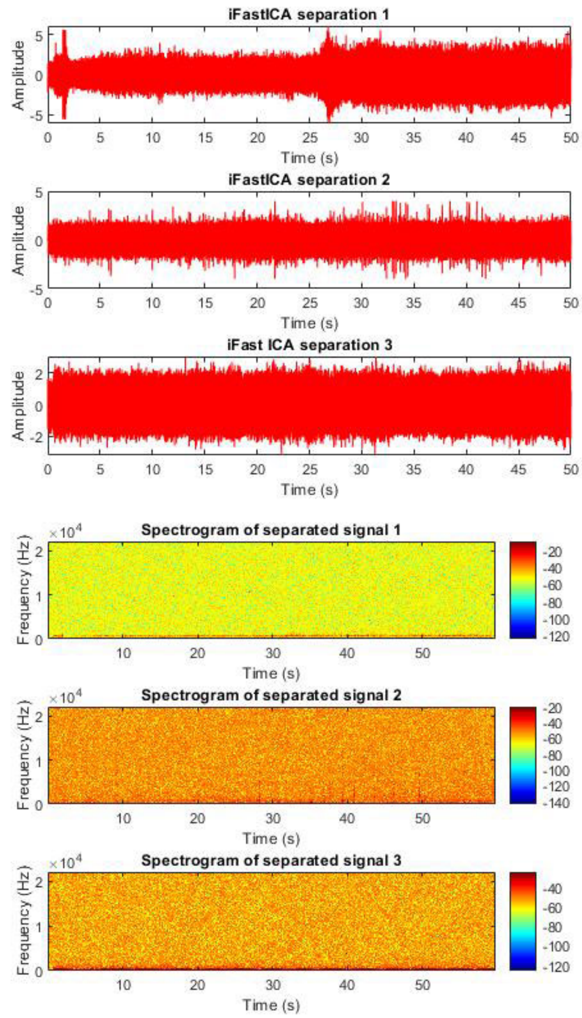

FIGURE 9. Separated/reconstructed signals from noisy acoustic mixture using iFastICA method.

of energies across a wide frequency range. Especially, the sound from the ocean dredger has dense concentration of energy across a wide frequency range which is consistent over time. The consistent yellow-orange color suggests a steady presence of loud sounds across the entire duration. Whereas, the spectrogram of sailboat is mostly blue indicating low energy and quieter sounds across the frequency spectrum. In the ocean liner's spectrogram, there is a noticeable red area that is consistent over time at lower frequency range possibly from the propeller's noise. The mixed signal matrix is generated by mixing all three acoustic source signals. The time domain waveform and spectrogram of the mixing signal are shown graphically in Figure 7. Then, the modelled Gaussian noise in Section II is added to the acoustic signal mixture at SNR 0dB to the acoustic signal mixture as shown in Figure 8. However, to evaluate the performance of the proposed method, Gaussian noise generated at various values of SNR is added to the acoustic signal mixture for generating results and performance analysis.

B. RESULTS AND ANALYSIS

This section provides the performance evaluation of our proposed solution. Firstly, we provide the time evolution and spectrogram evaluation of output signals. This study aims to assess and contrast the effectiveness of our approach by comparing it against established methods such as improved FastICA [44] and traditional NMF [45]. iFastICA and NMF are both unsupervised learning algorithms designed for linear mixture models, assuming statistically independent source signals. They are easily applicable to underdetermined BSS problems. The primary objective of both methods is to extract the underlying components or sources from mixed signals. iFastICA achieves this goal by maximizing the non-Gaussianity of the components, while NMF accomplishes it through the factorization of the observed matrix into non-negative matrices that represent the sources and the mixing coefficients.We implement iFastICA and traditional NMF-based BSS approaches on the noisy mixture and then compare it with the proposed MMSE-NMF technique. iFastICA algorithm solves the problem of source separation in the presence of noise. The algorithm can cancel the effect of noise to a certain extent but in the case of low SNR the noise reduction is limited. The time evolution and spectrogram of separated source signals from the noisy mixture at SNR 10dB are shown in Figure 9.

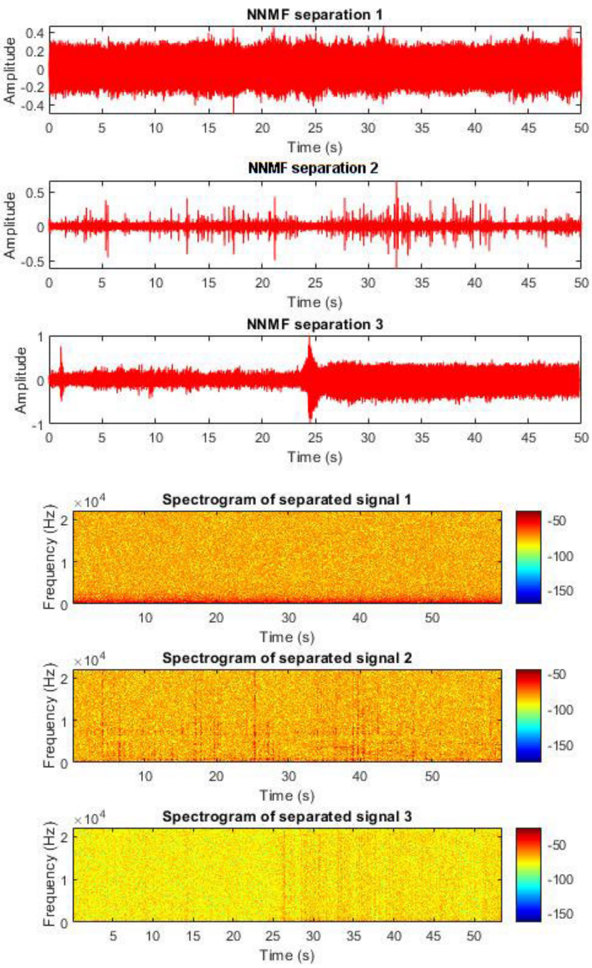

FIGURE 10. Separated/reconstructed signals from noisy acoustic mixture using classical NMF method.

The performance of the two-stage proposed method is validated by implementing source separation first followed by denoising of acoustic signals. Initially, the traditional optimized NMF algorithm-based BSS is implemented only on the noisy signal mixture so that the performance of traditional NMF-BSS can be compared with our MMSE-NMF method. The performance of classical NMF-BSS in the presence of underwater noise at SNR 10dB, is shown in Figure 10. These figures present the time evolution and spectrogram representation of separated acoustic signals. The visual analysis of the figures reveals an enhancement in the time evolution of the separated signals compared to the iFastICA separation results. Nevertheless, the influence of interfering noise at a signal-to-noise ratio (SNR) of 10dB is dominant, particularly evident in the spectrogram. The separated signal 1 has a dense energy presence across the entire frequency range. However, the energy levels are not as intense, suggesting some level of attenuation in the separation process. The separated signal 2 shows a clear pattern of energy presence, with less energy spread across the frequency range than the first separated signal thus resembling both ocean dredger and ocean liner. The separated signal 3 has the lowest energy presence across the frequency spectrum, which may suggest a degree of signal loss during the separation. The separated signal indicates the pronounced signal content of ocean dredger. The signal quality needs improvement by mitigating the impact of noise on the signal. Subsequently, the separated source signals undergo the process of MMSE filtering further to enhance the quality of the signal by minimizing the estimation error as detailed in the previous section.

The separated time evolution acoustic source signals and their spectrogram after MMSE filtering are shown in Figure 11.

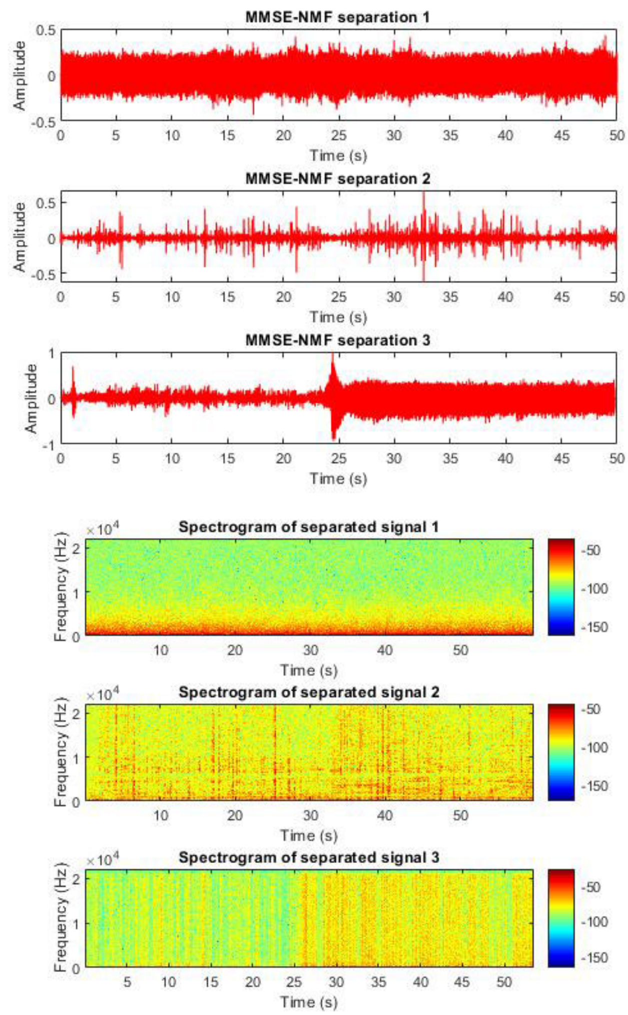

FIGURE 11. Separated/reconstructed signals from noisy acoustic mixture using MMSE-NMF method.

These figures demonstrate that our proposed method has performed much better than iFastICA and the traditional NMF-BSS algorithm in the separation of source signal at high SNR and suppressing the effect of noise during signal estimation. Separated Signal 1 have a uniform distribution of energy across the entire frequency range, similar to the Ocean liner. This broad frequency content and the high energy level throughout suggest that this signal could be associated with the Ocean liner. Subsequently, the separated signal 2 has periodic energy spikes at specific intervals and frequencies, which is more structured similar to ocean dredger. Lastly, the separated signal 3 appears to have the energy spikes similar to sail boat but some frequency content is attenuate showing poor separation performance compared to other two sources. For numerical evaluation, we have used four reconstruction evaluation metrics to analyze the performance of our proposed method against iFastICA and NNMF. These metrics have been used extensively in blind source separation to evaluate the separation performance in the presence of noise.

1) CORRELATION COEFFICIENT EVALUATION

The correlation coefficient can determine how similar the waveforms of the two signals are; therefore, it is used as a criterion for evaluating the separation performance between the separated source signal and the original source signal [26]. Cross-correlation between two acoustic signals is often denoted by $corr(y, y')$ where y and y' represent the clean acoustic signal and estimated/separated signal respectively.

TABLE 1. Cross-correlation coefficient between separated signals and original source signals obtained by iFastICA at different SNRs.

Correlation coefficient at SNR = 0dB				
Separated signals	**Ocean dredger**	**Sailboat**	**Oceanliner**	**Mixed signal**
Separated signal 1	0.161	0.005	0.601	0.181
Separated signal 2	0.413	0.141	0.091	0.335
Separated signal 3	0.061	0.624	0.261	0.412
Correlation coefficient at SNR = 10dB				
Separated signals	**Ocean dredger**	**Sailboat**	**Oceanliner**	**Mixed signal**
Separated signal 1	0.201	0.043	0.689	0.291
Separated signal 2	0.636	0.013	0.221	0.504
Separated signal 3	0.093	0.744	0.346	0.493
Correlation coefficient at SNR = 20dB				
Separated signals	**Ocean dredger**	**Sailboat**	**Oceanliner**	**Mixed signal**
Separated signal 1	0.201	0.106	0.797	0.503
Separated signal 2	0.771	0.013	0.216	0.562
Separated signal 3	0.005	0.841	0.331	0.684

Typically, correlation coefficients span a scale from -1 to 1. A value of 1 suggests maximum similarity, 0 indicates no relationship, and -1 signifies negative similarity. The cross-correlation coefficients obtained as a result of the separation process by existing methods and proposed are shown in Table 1, 2, 3 at different values of SNR. From the table, it is apparent that the cross-correlation between the separated signal and its original source signal is notably higher compared to the other two source signals, which exhibit lower cross-correlation values. At low signal-to-noise ratio (SNR) values, the cross-correlations between the separated and original signals are low due to the interference

TABLE 2. Cross-correlation coefficient between separated signals and original source signals obtained by NMF at different SNRs.

Correlation coefficient at SNR = 0dB				
Separated signals	**Ocean dredger**	**Sailboat**	**Oceanliner**	**Mixed signal**
Separated signal 1	0.177	0.017	0.617	0.197
Separated signal 2	0.447	0.157	0.107	0.339
Separated signal 3	0.077	0.690	0.277	0.481
Correlation coefficient at SNR = 10dB				
Separated signals	**Ocean dredger**	**Sailboat**	**Oceanliner**	**Mixed signal**
Separated signal 1	0.217	0.058	0.712	0.307
Separated signal 2	0.687	0.028	0.237	0.564
Separated signal 3	0.108	0.789	0.361	0.554
Correlation coefficient at SNR = 20dB				
Separated signals	**Ocean dredger**	**Sailboat**	**Oceanliner**	**Mixed signal**
Separated signal 1	0.217	0.122	0.821	0.516
Separated signal 2	0.797	0.028	0.232	0.597
Separated signal 3	0.009	0.864	0.347	0.701

TABLE 3. Cross-correlation coefficient between separated signals and original source signals obtained by MMSE-NMF at different SNRs.

Correlation coefficient at SNR = 0dB				
Separated signals	**Ocean dredger**	**Sailboat**	**Oceanliner**	**Mixed signal**
Separated signal 1	0.191	0.026	0.635	0.212
Separated signal 2	0.562	0.173	0.121	0.445
Separated signal 3	0.094	0.791	0.291	0.563
Correlation coefficient at SNR = 10dB				
Separated signals	**Ocean dredger**	**Sailboat**	**Oceanliner**	**Mixed signal**
Separated signal 1	0.233	0.071	0.825	0.321
Separated signal 2	0.740	0.042	0.252	0.672
Separated signal 3	0.120	0.821	0.378	0.664
Correlation coefficient at SNR = 20dB				
Separated signals	**Ocean dredger**	**Sailboat**	**Oceanliner**	**Mixed signal**
Separated signal 1	0.235	0.137	0.890	0.554
Separated signal 2	0.811	0.039	0.247	0.617
Separated signal 3	0.019	0.912	0.361	0.747

introduced by underwater noise in the estimated acoustic signals at low SNR. However, there exists a direct correlation between cross-correlation and SNR. As the SNR values increase, the cross-correlation improves. This characteristic suggests that our proposed algorithm demonstrates a high level of accuracy in separating sources.

2) MEAN SQUARE ERROR (MSE) CRITERION EVALUATION

We also compare the performance of our proposed model on the MSE parameter. Using MSE, the average squared difference between the estimated source signal and the original source signal is measured. In short, it tells the error rate between the estimated signal and the clean/original signal. The MSE between the clean and estimated signal of N length can be written mathematically as:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(S_i - \hat{S}_i)^2 \tag{28}$$

This expression is being used to evaluate the performance of iFastICA, NMF and MMSE-NMF. Smaller MSE values indicate the estimated signal is closer to the original signal. We calculate the MSE of iFastICA, conventional NMF and MMSE-NMF. The results of the MSE evaluation are shown in Figure 12. Regarding MSE, our proposed approach exhibits notable enhancement, consistently reducing MSE across various SNR values. Specifically, at an SNR of 20dB, MMSE-NMF achieved an average MSE of 0.4dB,

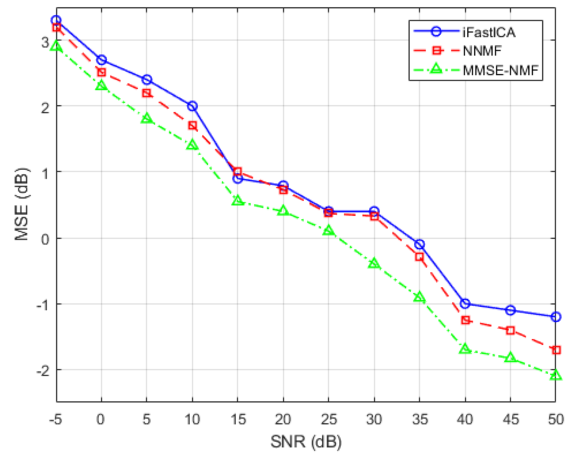

FIGURE 12. Comparison of overall MSE obtained by iFastICA, NNMF and proposed MMSE-NMF over different SNRs.

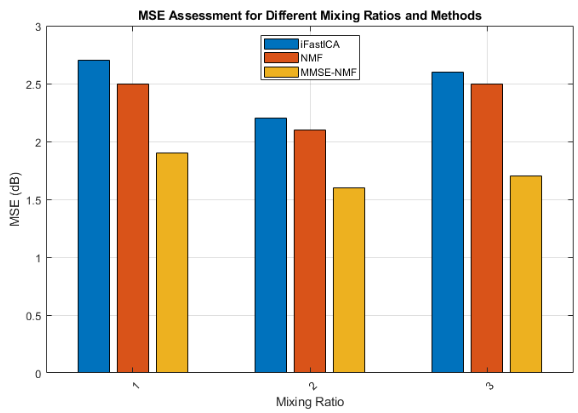

FIGURE 13. MSE evaluation at different mixing ratios for iFastICA, NMF and proposed MMSE-NMF.

surpassing iFastICA and NNMF, which indicated 0.73dB and 0.8dB, respectively. In challenging conditions, such as SNR 0dB, where noise impact is more pronounced, our method demonstrated an MSE of 2.3dB, outperforming iFastICA (2.7dB) and NMF (2.5dB). Similarly, at SNR 10dB, MMSE-NMF exhibited improved performance, minimizing MSE to 1.4dB, significantly lower than iFastICA and NMF. The effect of noise reduction and reduction in MSE is more prominent at high values of SNR. It is quite evident, NMF is robust to noise to some extent as compared to iFastICA while MMSE component in the system enhances system's performance by improving clarity and accuracy of the separated signals. Overall, our proposed method demonstrated superior performance compared to iFastICA and conventional NMF in minimizing the Mean Squared Error (MSE) between the reconstructed signal and the clean signal. On average, our method outperformed iFastICA by 47% and conventional NMF by 35% over the given SNR range.

Moreover, we further evaluate the performance in terms of MSE by computing MSE between estimated signals and clean signals by using different acoustic mixing ratios at a specific value of SNR. For this case, the MSE for three mixing ratios at SNR 5dB is simulated which are shown in Table 4. From Figure 13, it can be observed that our proposed algorithm performed better than iFastICA and NMF at different mixing ratios. Nonetheless, the influence of the mixing proportion on Mean Square Error (MSE) is evident in the figure. All algorithms resulted better outcomes, particularly in scenario where one source contribute more significantly to the mixture than the other two sources. In this specific situation, the MSE achieved by our proposed algorithm drops to 1.6 dB at 5 dB SNR, outperforming both iFastICA and NMF. This observation underscores the effectiveness of our proposed method, particularly in cases involving three sources. It also validates the suitability of this algorithm for scenarios with a limited number of sources, as an increase in the number of sources could increase system complexity, potentially leading to poorer separation performance.

TABLE 4. Mixing ratios of three acoustic sources used for simulation.

	Sailboat	Ocean dredger	Ocean liner
Mixing 1	0.5	0.3	0.2
Mixing 2	0.8	0.1	0.1
Mixing 3	0.2	0.5	0.3

3) SIGNAL-TO-DISTORTION RATIO (SDR) CRITERION EVALUATION

We have computed the overall SDR of output source signals using iFastICA, conventional NNMF and the proposed method after running various simulations for each estimated source signal. SDR measures the purity of the signal and conveys information about the possible distortion in a signal due to added noise or other interference of convolutive signals during BSS and estimation. It is a popular performance parameter used by audio processing researchers to assess the BSS performance of convolutive mixture [46]. SDR is less sensitive to amplitude scaling issues, making it a robust metric. We have used the following expression to evaluate this criterion on output signals by computing energy ratios as explained by [47].

$$SDR = 10log_{10}\frac{||S_{output}||^2}{||e_{interf} + e_{noise} + e_{artif}||^2} \tag{29}$$

where S_{output} is the targetted estimated signal and e_{interf}, e_{noise} and e_{artif} are the interference, noise and artifacts error terms respectively. From Figure 14, the output overall SDR of all sources is computed for iFastICA, NMF and the proposed method and compared at various values of SNR. It is well observed that the output SDR is directly proportional to the SNR i.e. the output increases as a function of the input SNR. The superiority of our proposed MMSE-NMF method can easily be observed at lower values of SNR. The other two methods resulted in negative SDR when SNR is considerably low whereas, MMSE-NMF

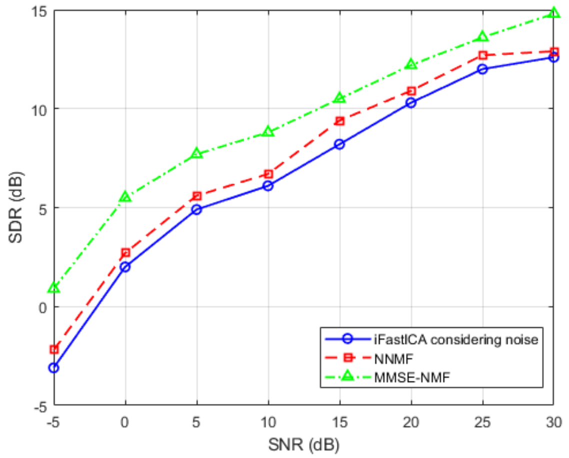

FIGURE 14. **Comparison of overall SDR obtained by iFastICA, NMF and proposed MMSE-NMF over different SNRs.**

achieved 0.9dB SDR. This shows, that our proposed method demonstrated better enhancement of signal quality at low SNR. It is noted that the performance of iFastICA and NMF is good at high SNR only when the interference from the noise is generally less.

4) CEPSTRAL DISTANCE (CD) EVALUATION

In this subsection, the performance of enhanced signal after separation is evaluated using frequency-based metric, Cepstral Distance (CD). Linear predictive coding coefficients (LPCCs) have been employed in speech systems to evaluate the measure of the extent of distortion in the output of estimation algorithms [48]. LPC coefficients can be computed efficiently and offer a compact representation of the spectral envelope of a signal. This efficiency makes LPC attractive for real-time applications, noisy conditions and scenarios where computational resources are limited. If the c and $\hat{c}$ are the LPC coefficients of clean signal s and estimated signal $\hat{s}$, then the cepstral coefficients A and $\hat{A}$ can be computed from the following relation.

$$A(n) = c_n + \sum_{m=1}^{n-1} \frac{m}{n}[A(m)c_{n-m}] \qquad for \quad 1 \leq n \leq Z \quad (30)$$

where Z denotes the order. Then, the cepstral distance between s and $\hat{s}$ can be easily calculated by using cepstral coefficients A and $\hat{A}$ using the following expression.

$$CD = 10/log_{10}\sqrt{2\sum_{i=1}^{Z}[A_s(i) - A_{\hat{s}}(i)]^2} \qquad (31)$$

We have evaluated the CD criterion for five input SNRs i.e $-5dB$, $0dB$, $5dB$, $10dB$, $15dB$ to have a clear idea of the performance of the estimated source signal of the sailboat, ocean liner and ocean dredger to their clean versions in the presence of strong noise. Figure 15 shows the performance comparisons based on the CD criterion obtained with the iFastICA, NMF and MMSE-NMF. The CD values for noisy signals are demonstrated for the relative comparison of the

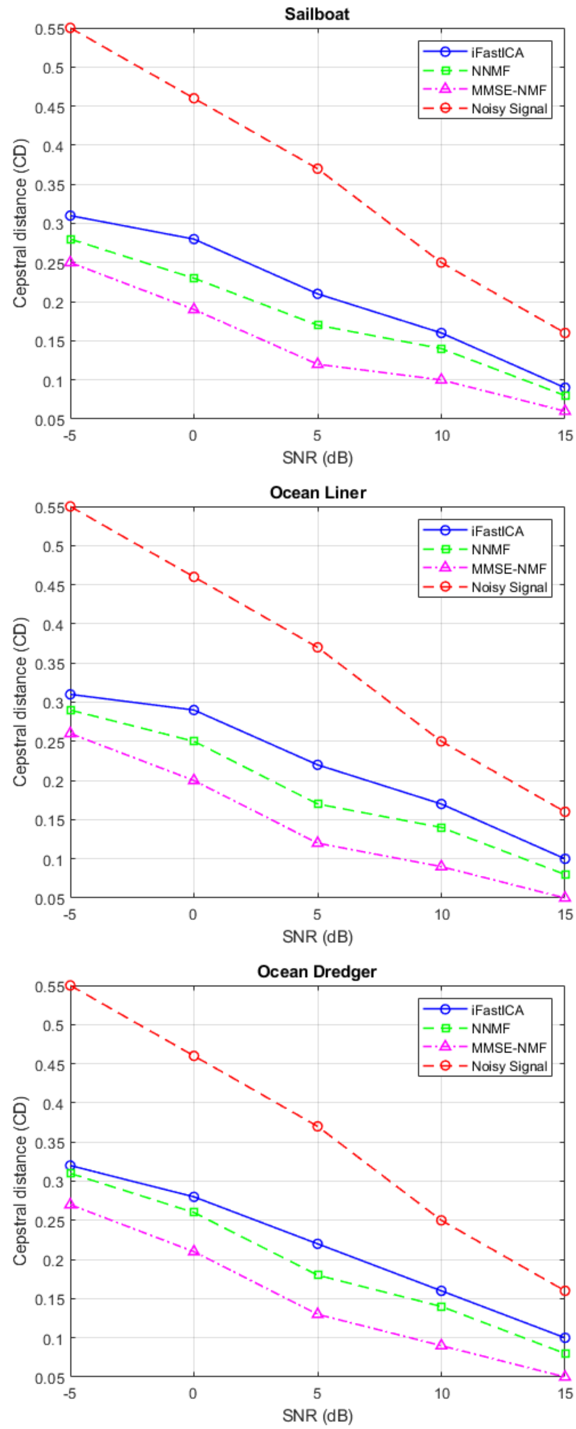

FIGURE 15. **Cepstral Distance (CD) evaluation of iFastICA, NMF, and the proposed MMSE-NMF algorithm for estimated sources: Sailboat, Ocean Liner, Ocean dredger.**

systems. It can be observed from the figure, that the proposed MMSE-NMF method has reported the lowest overall CD for all three sources as compared to iFastICA and NMF. The proposed MMSE-NMF resulted in the mean CD of 0.146, while NMF produced 0.194, and iFastICA resulted in 0.228. This shows the superiority of our proposed method on existing by reducing CD upto 36%. It is also observed that the effectiveness of BSS methods exhibits minor variations with distinct source signals. In general, the estimated signal of the sailboat consistently demonstrates the lowest CD in comparison to the ocean dredger and ocean liner. This suggests that the reconstruction of the sailboat's acoustic signal is superior to the other two acoustic signals and therefore, it is closer to its cleaner version.

V. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

As a result of the fact that water noise has a major impact on the ability to separate acoustic signals from undersea vessels from a reverberant and noisy mixture in underwater BSS, this paper proposes a novel two-stage approach to counter noise interference after source separation. Source signals are separated using single-channel NMF-based BSS, and then denoised using adaptive MMSE for effective extraction of clean signals to determine sources more accurately. The enhancement of the separated signal from the noisy mixture is performed by minimizing the MSE between the estimated signal and the true signal.

To validate the effectiveness of the proposed model, various simulations using different evaluation metrics (correlation, MSE, SDR and CD) are carried out. The results are evaluated and compared with iFastICA and classical NMF from the literature where results clearly indicate better performance in terms of separation accuracy, and residual distortion in the estimated signal. Our proposed method outperformed existing methods when it comes to minimizing MSE between reconstructed signal and clean signal upto 47%. It also validates the fact that different mixing ratios of acoustic signals in a mixture can alter separation accuracy and reconstruction error. In addition, our proposed method has consistently demonstrated superior performance both in improving SDR and lowering CD upto 36%. However, at SNR less than -5dB, the reconstruction performance of NMF system degrades but adaptive MMSE focuses on optimizing the signal quality post-separation. As the noise is dominant over the the target acoustic signal and NMF does not consider temporal or spectral correlations, therefore, NMF can mistakenly fit the noise components rather underlying signal. Thus, NMF requires improvements in further studies in initialization of noisy data and adding additional constraints during decomposition process.

In the future, further practical studies are required to model water noise in complicated scenarios because the baseline Gaussian model may not be applied to various other water environments. Underwater noise in the ocean is stratified according to its spatial distribution due to the changes in noise characteristics and reverberations. In that case, the Rayleigh distribution can be used to model pressure variations and reverberations, especially in shallow water environments where sound waves undergo random fluctuations that occur as a result of reflection, refraction, and scattering of the sound waves [49]. Moreover, we aim to explore non-Gaussian underwater noise along with advanced noise handling methods in future. Additionally, the system can be improved by upgrading to a multi-channel/sensor BSS setup. A multi-channel BSS system estimates the mixing process and reconstructs the source signals through iterative optimization, aiming to achieve the best match with the observed mixtures based on the assumption of source independence. Further studies would include spatial diversity and accounts for the cross-coupling effects of acoustic signals received by multiple sensors.

REFERENCES

[1] J. P. Mishra, K. Singh, and H. Chaudhary, "Research advancements in ocean environmental monitoring systems using wireless sensor networks: A review," *TELKOMNIKA, Telecommun. Comput. Electron. Control*, vol. 21, no. 3, p. 513, Jun. 2023.

[2] K. Y. Islam, I. Ahmad, D. Habibi, and A. Waqar, "A survey on energy efficiency in underwater wireless communications," *J. Netw. Comput. Appl.*, vol. 198, Feb. 2022, Art. no. 103295.

[3] J. McNulty, "Western and central Pacific Ocean fisheries and the opportunities for transnational organised crime: Monitoring, control and surveillance (MCS) operation kurukuru," *Austral. J. Maritime Ocean Affairs*, vol. 5, no. 4, pp. 145–152, Jan. 2013.

[4] B. Teng and H. Zhao, "Underwater target recognition methods based on the framework of deep learning: A survey," *Int. J. Adv. Robot. Syst.*, vol. 17, no. 6, Nov. 2020, Art. no. 172988142097630.

[5] Y.-H. Lin, T.-L. Wu, C.-M. Yu, and I.-C. Wu, "Development of an intelligent underwater recognition system based on the deep reinforcement learning algorithm in an autonomous underwater vehicle," *Measurement*, vol. 214, Jun. 2023, Art. no. 112844.

[6] R. Zaheer, I. Ahmad, D. Habibi, K. Y. Islam, and Q. V. Phung, "A survey on artificial intelligence-based acoustic source identification," *IEEE Access*, vol. 11, pp. 60078–60108, 2023.

[7] O. Yaman, T. Tuncer, and B. Tasar, "DES-Pat: A novel DES pattern-based propeller recognition method using underwater acoustical sounds," *Appl. Acoust.*, vol. 175, Apr. 2021, Art. no. 107859.

[8] T. A. Smith and J. Rigby, "Underwater radiated noise from marine vessels: A review of noise reduction methods and technology," *Ocean Eng.*, vol. 266, Dec. 2022, Art. no. 112863.

[9] Z. Xinhua, Z. Anqing, F. Jianping, and Y. Shaoqing, "Study on blind separation of underwater acoustic signals," in *Proc. 5th Int. Conf. Signal Process., 16th World Comput. Congr. (WCC ICSP)*, 2000, pp. 1802–1805.

[10] S. K. Beack, B. Lee, M. Hahn, and S. H. Nam, "Blind source separation and Kalman filter-based speech enhancement in a car environment," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst. (ISPACS)*, 2004, pp. 520–523.

[11] D. N. Parikh and D. V. Anderson, "Blind source separation with perceptual post processing," in *Proc. Digit. Signal Process. Signal Process. Educ. Meeting (DSP/SPE)*, Jan. 2011, pp. 321–325.

[12] R. Bendoumia and M. Djendi, "Speech enhancement using backward adaptive filtering algorithm: Variable step-sizes approaches," in *Proc. 3rd Int. Conf. Control, Eng. Inf. Technol. (CEIT)*, May 2015, pp. 1–5.

[13] H. Tang and S. Wang, "Noisy blind source separation based on adaptive noise removal," in *Proc. 10th World Congr. Intell. Control Autom.*, Jul. 2012, pp. 4255–4257.

[14] A. Routray, N. Das, and P. K. Dash, "Robust preprocessing: Denoising and whitening in the context of blind source separation of instantaneous mixtures," in *Proc. 5th IEEE Int. Conf. Ind. Informat.*, Jul. 2007, pp. 377–380.

[15] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287–314, Apr. 1994.

[16] P. Bofill and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," *Signal Process.*, vol. 81, no. 11, pp. 2353–2362, Nov. 2001.

[17] D. Seung and L. Lee, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 13, 2001, pp. 556–562.

[18] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Netw.*, vol. 10, no. 3, pp. 626–634, May 1999.

[19] M. Kermit, *Independent Component Analysis: Theory and Applications*, 2000.

[20] Y. Li, S. Amari, A. Cichocki, D. W. C. Ho, and S. Xie, "Underdetermined blind source separation based on sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 423–437, Feb. 2006.
[21] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999.
[22] T. Shijie and C. Hang, "Blind source separation of underwater acoustic signal by use of negentropy-based fast ICA algorithm," in *Proc. IEEE Int. Conf. Comput. Intell. Commun. Technol.*, Feb. 2015, pp. 608–611.
[23] W. Zhang, X. Li, A. Zhou, K. Ren, and J. Song, "Underwater acoustic source separation with deep bi-LSTM networks," in *Proc. 4th Int. Conf. Inf. Commun. Signal Process. (ICICSP)*, Sep. 2021, pp. 254–258.
[24] M. Rahmati, P. Pandey, and D. Pompili, "Separation and classification of underwater acoustic sources," in *Proc. Underwater Commun. Netw. (UComms)*, Sep. 2014, pp. 1–5.
[25] M. Rahmati and D. Pompili, "UNISeC: Inspection, separation, and classification of underwater acoustic noise point sources," *IEEE J. Ocean. Eng.*, vol. 43, no. 3, pp. 777–791, Jul. 2018.
[26] D. Li, M. Wu, L. Yu, J. Han, and H. Zhang, "Single-channel blind source separation of underwater acoustic signals using improved NMF and FastICA," *Frontiers Mar. Sci.*, vol. 9, Jan. 2023, Art. no. 1097003.
[27] Y. Li, C. Zhang, and Y. Zhou, "A novel denoising method for ship-radiated noise," *J. Mar. Sci. Eng.*, vol. 11, no. 9, p. 1730, Sep. 2023.
[28] Y. Li, Y. Zhou, and S. Jiao, "Multi-scale ensemble dispersion Lempel–Ziv complexity and its application on feature extraction for ship-radiated noise," *Appl. Acoust.*, vol. 218, Mar. 2024, Art. no. 109890.
[29] M. Zoulikha and M. Djendi, "A new regularized forward blind source separation algorithm for automatic speech quality enhancement," *Appl. Acoust.*, vol. 112, pp. 192–200, Nov. 2016.
[30] X. Yan, M.-H. Boudrias, and G. D. Mitsis, "Removal of transcranial alternating current stimulation EEG artifacts using blind source separation and wavelets," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 10, pp. 3183–3192, Oct. 2022.
[31] J. Wang, J. Li, S. Yan, W. Shi, X. Yang, Y. Guo, and T. A. Gulliver, "A novel underwater acoustic signal denoising algorithm for Gaussian/non-Gaussian impulsive noise," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 429–445, Jan. 2021.
[32] M. Zuo, X. Tu, S. Yang, H. Fang, X. Wen, and F. Qu, "Channel distribution and noise characteristics of distributed acoustic sensing underwater communications," *IEEE Sensors J.*, vol. 21, no. 21, pp. 24185–24194, Nov. 2021.
[33] D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 13, 2000.
[34] T. Van Erven and P. Harremos, "Rényi divergence and Kullback–Leibler divergence," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 3797–3820, Jul. 2014.
[35] D. Guo, Y. Wu, S. S. Shitz, and S. Verdú, "Estimation in Gaussian noise: Properties of the minimum mean-square error," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2371–2385, Apr. 2011.
[36] M. Souden, S. Araki, K. Kinoshita, T. Nakatani, and H. Sawada, "A multichannel MMSE-based framework for speech source separation and noise reduction," *IEEE Trans. Audio, Speech, Languages Process.*, vol. 21, no. 9, pp. 1913–1928, Sep. 2013.
[37] G. Enzner and P. Thüne, "Robust MMSE filtering for single-microphone speech enhancement," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 4009–4013.
[38] I. Y. Soon, S. N. Koh, and C. K. Yeo, "Noisy speech enhancement using discrete cosine transform," *Speech Commun.*, vol. 24, no. 3, pp. 249–257, Jun. 1998.
[39] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 443–445, Apr. 1985.
[40] S. A. Vavasis, "On the complexity of nonnegative matrix factorization," *SIAM J. Optim.*, vol. 20, no. 3, pp. 1364–1377, Jan. 2010.
[41] D. Santos-Domínguez, S. Torres-Guijarro, A. Cardenal-López, and A. Pena-Gimenez, "ShipsEar: An underwater vessel noise database," *Appl. Acoust.*, vol. 113, pp. 64–69, Dec. 2016.
[42] T. Mei and A. Mertins, "Convolutive blind source separation based on disjointness maximization of subband signals," *IEEE Signal Process. Lett.*, vol. 15, pp. 725–728, 2008.
[43] Z. Yang, Y. Xiang, Y. Rong, and K. Xie, "A convex geometry-based blind source separation method for separating nonnegative sources," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 8, pp. 1635–1644, Aug. 2015.
[44] A. Al-Qaisi, "Blind source separation of mixed noisy audio signals using an improved FastICA," *J. Appl. Sci.*, vol. 15, no. 9, pp. 1158–1166, Aug. 2015.
[45] N. Hassan and D. A. Ramli, "A comparative study of blind source separation for bioacoustics sounds based on FastICA, PCA and NMF," *Proc. Comput. Sci.*, vol. 126, pp. 363–372, Jan. 2018.
[46] Z. Zhang and P. R. White, "A blind source separation approach for humpback whale song separation," *J. Acoust. Soc. Amer.*, vol. 141, no. 4, pp. 2705–2714, Apr. 2017.
[47] E. Vincent, R. Gribonval, and C. Fevotte, "Performance measurement in blind audio source separation," *IEEE Trans. Audio, Speech Languages Process.*, vol. 14, no. 4, pp. 1462–1469, Jul. 2006.
[48] I. McLoughlin, *Speech and Audio Processing: A MATLAB-Based Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2016.
[49] A. P. Lyons and D. A. Abraham, "Statistical characterization of high-frequency shallow-water seafloor backscatter," *J. Acoust. Soc. Amer.*, vol. 106, no. 3, pp. 1307–1315, Sep. 1999.

RUBA ZAHEER received the Bachelor of Engineering degree in telecommunication from the National University of Sciences and Technology, Pakistan, in 2015. She is currently pursuing the Ph.D. degree in electrical and electronics engineering with Edith Cowan University, Australia. She has been associated with COMSATS University, Pakistan, as a Lab Engineer. Her research interests include data analytics, machine learning, signal processing, and underwater acoustics.

IFTEKHAR AHMAD (Member, IEEE) received the Ph.D. degree in communication networks from Monash University, Australia, in 2007. He is currently an Associate Professor with the School of Engineering, Edith Cowan University, Australia. His research interests include 5G technologies, green communications, QoS in communication networks, software-defined radio, wireless sensor networks, and computational intelligence.

QUOC VIET PHUNG (Member, IEEE) received the Ph.D. degree in communication engineering from Edith Cowan University, Australia, in 2010. He is currently a Lecturer with the School of Engineering, Edith Cowan University. His research interests include smart sensors, 5G technologies, underwater communication, software-defined radio, data analytics, network security, and applied artificial intelligence.

DARYOUSH HABIBI (Senior Member, IEEE) received the Bachelor of Engineering degree (Hons.) in electrical and the Ph.D. degree from the University of Tasmania, in 1989 and 1994, respectively. His employment history includes Telstra Research Laboratories, Flinders University, Intelligent Pixels Inc., and Edith Cowan University, where he is currently a Professor, a Pro Vice-Chancellor, and the Executive Dean of the School of Engineering. He has more than 200 refereed publications in high-impact journals, conference proceedings and book chapters. His research interests include engineering design for sustainable development, renewable and smart energy systems, environmental monitoring technologies, and reliability and quality of service in engineering systems and networks. He is a Fellow of Engineers Australia and the Institute of Marine Engineering, Science and Technology.