

Received October 18, 2021, accepted October 26, 2021, date of publication November 1, 2021, date of current version November 8, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3124509

# Reverberation Suppression Method for Active Sonar Systems Using Non-Negative Matrix Factorization With Pre-Trained Frequency Basis Matrix

GEUNHWAN KIM<sup>1</sup> AND SEOKJIN LEE<sup>1</sup>, (Member, IEEE)

Electrical and Electronics Engineering, College of IT Engineering, Kyungpook National University, Daegu 41566, Republic of Korea

Corresponding author: Seokjin Lee (sjlee6@knu.ac.kr)

This work was supported by the Agency for Defense Development (ADD), South Korea, under Grant UD190005DD.

**ABSTRACT** In active sonar systems, detection always suffers from reverberation interference from multiple scatterers in oceanic environments; therefore, numerous studies have been conducted on reverberation suppression. Recently, a non-negative matrix factorization (NMF)-based method was proposed and successfully applied to reverberation suppression. However, the conventional NMF-based method makes convergence challenging because the frequency basis matrix is initialized without considering reverberation characteristic information from oceanic environments. To solve these problems, We propose an improved NMF-based reverberation suppression method adopting a pre-trained reverberation basis matrix and modified sparse update rule. The proposed method is evaluated by analyzing simulation and sea experiment data and the study confirmed that the detection performance was improved compared to the conventional method under various signal-to-reverberation ratio conditions. Several topics are also discussed to analyze the proposed method in detail.

**INDEX TERMS** Active sonar, non-negative matrix factorization, reverberation suppression, pre-trained basis matrix, sparseness.

## I. INTRODUCTION

Active sonar is an underwater surveillance system to detect, track, and localize targets. An active sonar system transmits pulses to investigate underwater situations and analyze received signals. During propagation, both the target as well as the unwanted scatterers reflect the transmitted pulse. Therefore, interference signals, called reverberation, corrupt the received signal due to unwanted scatters. Because reverberation strongly correlates with the signal reflected by the target, it deteriorates target detection performance. Accordingly, research on how to suppress reverberation is essential for active sonar systems [1]–[6].

Autoregressive (AR) prewhitener for continuous wave (CW) reverberation is an early approach suppressing the reverberation spectrum using inverse filters, designed by estimated AR coefficients [7]. An AR prewhitener uses the

adjacent signal with reverberation as a reference signal to estimate the reverberation spectrum and design an inverse filter. The designed inverse filter is used to suppress reverberation from the current signal with both the target and reverberation.

Because a linear frequency modulated (LFM) pulse has a wideband spectrum, the AR prewhitener for CW reverberation works incorrectly. Therefore, Choi *et al.* [8] proposed an AR prewhitener using dechirping transformation as pre-processing. This pre-processing makes the received LFM signal tonal spectrum a CW case. Thus, like CW case, an inverse filter can be designed using AR modeling.

Alternatively, the principal component inversion (PCI) algorithm has been proposed [9]. It models the reverberation signal as the sum of multiple reflected signals and estimates a low-rank approximation of the forward signal matrix using singular value decomposition. As a follow-up paper, the signal subspace extraction algorithm [10] modeling reverberation as the sum of the higher and lower part, and a PCI-support

The associate editor coordinating the review of this manuscript and approving it for publication was Chengpeng Hao<sup>1</sup>.

vector machine (SVM) algorithm [11] combining PCI and SVM, were proposed.

Although AR and PCI-based reverberation suppression methods have been widely used, pulse characteristics are interpreted and used only from a time or frequency domain perspective. Recently, Lee and Lim [12] proposed a non-negative matrix factorization (NMF)-based CW reverberation suppression method that simultaneously uses pulse time and frequency characteristics. The received signal is transformed into the time-frequency domain, called a spectrogram, and the spectrogram is decomposed with frequency and time basis matrices. Because a CW target echo is represented as a straight line in the spectrogram, ideally, only a few frequency and time bases are needed to represent the target component. If a new spectrogram is reconstructed using only the target echo basis, the reverberation is removed, and only the target echo signal appears. That is, the reverberation has been suppressed.

As with the AR prewhitener, the NMF-based CW reverberation suppression method has been extended to be applied to LFM reverberation [13]. The modified NMF-based reverberation suppression method adopted two pre-processing techniques, namely, dechirping transformation and modulo operation. The dechirping transformation presents the LFM echo signal as a tonal-like spectrum, as with Choi's method. The modulo operation prevents the disconnected effect of block signal processing.

Although the usefulness of the conventional NMF-based reverberation suppression method has been verified, a problem exists that convergence might become challenging, depending on environmental fluctuation because parts of the  $W$  matrix are forced to be fixed, and others are randomly initialized [12]. Scaling between the target and reverberation frequency basis component also creates problems.

In this paper, we propose an improved NMF-based reverberation suppression method with a pre-trained reverberation frequency basis matrix to solve the convergence problem of the conventional NMF-based method. Pre-training the reverberation frequency basis matrix is performed using adjacent signals considering the Itakura–Saito distance, which measures the spectrum's similarity. Because the proposed method learned the frequency basis matrix in advance, its NMF update process converges stably. To solve the scaling problem, the power ratio parameter was adopted to balance the imbalance between the target and reverberation basis. A modified sparse update rule was also adopted to improve the NMF update process.

The rest of this article is organized as follows. In Section II, we present the problem statement, signal model, and conventional NMF-based reverberation suppression method. In Section III, we present the proposed NMF-based reverberation suppression method. In Section IV, we present the computer simulation and sea experiment analysis, respectively, to evaluate the proposed method. In Section V, several topics for the proposed method are discussed. Finally, in Section VI, we conclude this study.

## II. PRELIMINARIES

### A. PROBLEM STATEMENT

The CW pulse used in active sonars can be expressed as

$$s_p(t) = \exp(j2\pi f_c t), \quad 0 \leq t \leq T, \quad (1)$$

where  $f_c$  is the center frequency and  $T$  is the pulse length.

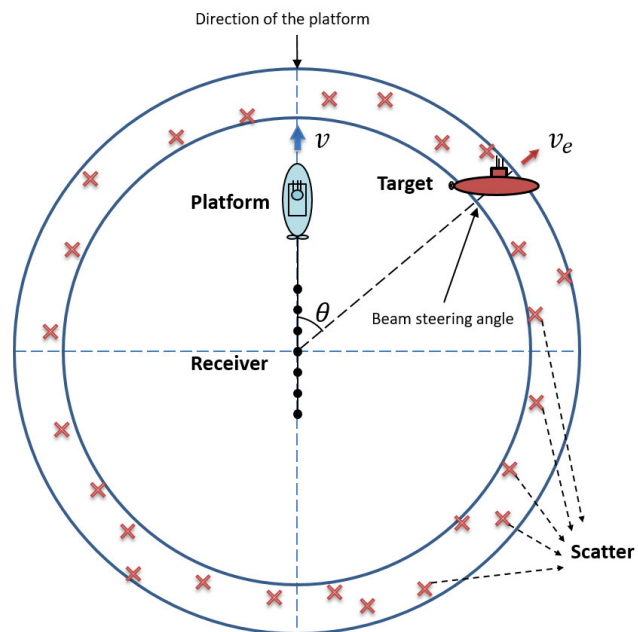


FIGURE 1. The concept of receiving a signal in an active sonar environment.

Fig. 1 shows the concept of receiving a signal in an active sonar environment. The sonar platform moves forward with speed  $v$  and the beam is steered to direction  $\theta$ . The received beam signal at target direction  $\theta$  can be expressed as

$$x(t) = s_e(t) + s_r(t) + n(t), \quad (2)$$

where  $s_e(t)$  and  $s_r(t)$  are the received target echo and reverberation signals, respectively, and  $n(t)$  is the noise signal, modeled as a zero-mean Gaussian distribution.

The target echo signal can be expressed as

$$s_e(t) = a_e s_p(t - \tau_e) \exp(j2\pi f_e t), \quad (3)$$

where  $a_e$  is the attenuation factor and  $\tau_e$  is the time delay of the target echo signal. The target signal's frequency spectrum can be expressed as

$$S_e(f) = a_e S_p(f - f_e) \exp(-j2\pi(f - f_e)\tau_e) \quad (4)$$

where  $S_p(f)$  is the Fourier transform of the CW pulse in (1). The Doppler shift frequency of echo signal  $f_e$  can be expressed as

$$\begin{aligned} f_e &= f_c \left( 1 + \frac{2}{c} (v \cos \theta + v_e) \right) \\ &= f_c \left( 1 + \frac{2v}{c} \cos \theta \right) + f_c \left( \frac{2v_e}{c} \right), \end{aligned} \quad (5)$$

where  $v$ ,  $\theta$ ,  $c$ , and  $v_e$  are the platform speed, beam steering angle, the speed of sound in underwater, and the relative target speed with the platform, respectively. From (5), the target spectrum is around  $f_c \left(1 + \frac{2v}{c} \cos \theta\right)$ , but decided by its motion.

The reverberation signal can be modeled as the sum of replica signals of the transmitted pulse from multiple scatters [6], [14]. Because the beam pattern attenuates the reverberation effect from other directions, the reverberation signal in the received beam signal at direction of  $\theta$  can be expressed as follows:

$$s_r(t) = \sum_i a_i s_p(t - \tau_i) \exp(j2\pi f_{d_i} t), \quad (6)$$

where  $a_i$  is the attenuation factor affected by the beam pattern.  $\tau_i$  is the time delay modeled as a uniform distribution over the interval. The frequency spectrum of the reverberation signal can be expressed as

$$S_r(f) = a_i \sum_i S_p(f - f_{d_i}) \exp(-j2\pi(f - f_{d_i})\tau_i). \quad (7)$$

Typically, scatters are assumed immobile; therefore, the Doppler shift frequency of  $i$ th scatter  $f_{d_i}$  is determined by platform speed  $v$  and beam steering angle  $\theta$  and can be expressed as

$$f_{d_i} = f_c \left(1 + \frac{2v}{c} \cos \theta_i\right), \quad (8)$$

where  $\theta_i$  is the scatter angle. From (8), the reverberation spectrum is distributed in range of  $f_c \left(1 - \frac{2v}{c}\right) \leq f_{d_i} \leq f_c \left(1 + \frac{2v}{c}\right)$  in the frequency domain. Note that because the beam pattern's sidelobe attenuates the reverberation from other directions, reverberation spectrum appears in the form of spreading around  $f_c \left(1 + \frac{2v}{c} \cos \theta\right)$ . Fig. 2 shows the concept of the frequency spectrum of the received beam signal described (2) to (8). From this figure, the received CW target echo signal will be buried in the reverberation spectrum [1], [14].

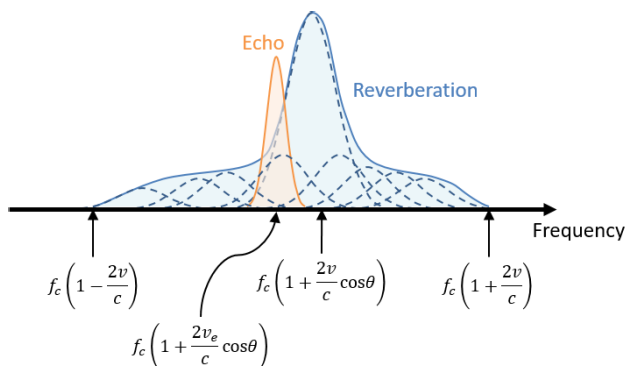


FIGURE 2. Concept of frequency spectrum of received beam signal.

The goal of this study is extracting the target echo signal from the received signal corrupted by reverberation.

To accomplish this goal, the received signal's spectrogram is analyzed using NMF, and the components of the target echo and reverberation are decomposed. Note that we focus on CW reverberation suppression because it is easy to extend the NMF-based CW reverberation suppression method to the LFM reverberation case using the pre-processing method proposed by Kim *et al.* [13].

### B. NON-NEGATIVE TIME-FREQUENCY MODEL OF THE RECEIVED SIGNAL

Lee and Seung originally developed the NMF method [15], inspired by psychological and physiological evidence for parts-based representations of the brain. The NMF method limits data to non-negative leading to a parts-based representation because they allow only additive combinations. Therefore, the NMF method has the advantage of representing the sparse matrix.

To apply the NMF method to the received signal, it must be expressed as a non-negative matrix. Therefore, the time-frequency analysis method is applied to the received signal and converted into spectrogram data by taking only the magnitude component. For this purpose, a short-time Fourier transform (STFT) is generally used and can be expressed as [16]

$$\begin{aligned} [\mathbf{X}]_{(k,n)} &= \text{STFT}[x(m)] \\ &= \sum_{m=-\infty}^{\infty} x(m)w(m - n\Delta h)e^{-j2\pi km/l_w}, \end{aligned} \quad (9)$$

where  $x(m)$  is the sampled received beam signal,  $w(m)$  is the window function of  $l_w$  samples,  $m$  is the index of each sample, and  $\Delta h$  is the hop size.  $k$  and  $n$  are the frequency bin and time frame bin indices, respectively. The magnitude square of (9) yields a spectrogram  $\mathbf{V}$  as follows:

$$[\mathbf{V}]_{(k,n)} = |[\mathbf{X}]_{(k,n)}|^2, \quad (10)$$

where  $[\mathbf{V}]_{(k,n)}$  is the  $(k, n)$  element of matrix  $\mathbf{V}$ . Furthermore, the phase information is defined as

$$\angle \mathbf{X}_{(k,n)} = \tan^{-1} \left( \frac{\text{Im}([\mathbf{X}]_{(k,n)})}{\text{Re}([\mathbf{X}]_{(k,n)})} \right), \quad (11)$$

where  $\text{Re}(a)$  and  $\text{Im}(a)$  are the real and imaginary parts of  $a$ , respectively.

### C. CONVENTIONAL NMF-BASED REVERBERATION SUPPRESSION METHOD AND ITS LIMITATIONS

From the non-negative spectrogram in (10), we could develop the NMF method. NMF decomposes the non-negative matrix into a multiplication of two non-negative matrices [15]. One advantage of the NMF method is the ability of part-based representation, meaning that NMF can analyze the sparse components in the matrix. In the CW active sonar, the CW echo signal is shown as a straight line, sparse components in the spectrogram. NMF can extract the sparse component by analyzing the frequency and time information simultaneously; therefore, NMF is suitable for sonar reverberation

suppression. The NMF model is expressed as

$$\mathbf{V} = \mathbf{W}\mathbf{H} + \mathbf{E}, \quad (12)$$

where  $\mathbf{V} \in \mathbb{R}_{K \times N}^+$ ,  $\mathbf{W} \in \mathbb{R}_{K \times L}^+$ ,  $\mathbf{H} \in \mathbb{R}_{L \times N}^+$ , and  $\mathbf{E} \in \mathbb{R}_{K \times N}^+$ . The matrix  $\mathbf{V}$  is the spectrogram of the input signals with  $K$  frequency bins and  $N$  time frames, and  $\mathbf{W}$  and  $\mathbf{H}$  are the frequency and time basis matrices, respectively. Further,  $L$  is the rank of NMF. The matrices  $\mathbf{W}$  and  $\mathbf{H}$  can be estimated from the input spectrogram  $\mathbf{V}$  using the multiplication update (MU) rule using a cost function with the Kullback–Leibler (KL) divergence [17].

Recently, Lee and Lim [12] proposed an NMF-based CW reverberation suppression algorithm. To establish the frame of the algorithm, they first divided the input spectrogram into the target echo and reverberation parts and applied different constraints to each basis. The frequency basis matrix  $\mathbf{W}$  and time basis matrix  $\mathbf{H}$  are, respectively, expressed as

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_P \\ \mathbf{W}_R \end{bmatrix}, \quad (13)$$

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_P \\ \mathbf{H}_R \end{bmatrix}, \quad (14)$$

where  $\mathbf{W}_P \in \mathbb{R}_{K \times L_P}^+$  and  $\mathbf{W}_R \in \mathbb{R}_{K \times L_R}^+$  are, respectively, the target echo and reverberation parts of the frequency basis matrix;  $\mathbf{H}_P \in \mathbb{R}_{L_P \times N}^+$  and  $\mathbf{H}_R \in \mathbb{R}_{L_R \times N}^+$  are, respectively, the target echo and reverberation parts of the time basis matrix;  $L_P$  and  $L_R$  are the basis numbers of the target echo and reverberation, respectively, and  $L = L_P + L_R$ . Then, the NMF model in (12) can be rearranged as

$$\mathbf{V} = \mathbf{W}_P \mathbf{H}_P + \mathbf{W}_R \mathbf{H}_R + \mathbf{E}. \quad (15)$$

Because of the division, NMF could estimate information on the target echo and reverberation separately. In (15), the additivity assumption of the target echo part  $\mathbf{W}_P \mathbf{H}_P$  and reverberation part  $\mathbf{W}_R \mathbf{H}_R$  in the NMF model (15) is reasonable because it agrees with the signal model (2).

The CW target echo components were shown as horizontal lines in the spectrogram. In other words, the CW target echo information is concentrated in a narrow or single frequency bin and is continuous along several time bins. Fig. 3 shows the meaning of this assumption. In the figure, matrix  $\mathbf{V}$  is decomposed as  $\mathbf{W}_P$ ,  $\mathbf{W}_R$ ,  $\mathbf{H}_P$ , and  $\mathbf{H}_R$ . As  $\mathbf{W}_P$  is fixed in advance, after decomposition, information on the target (green cell) in  $\mathbf{V}$  appears in a form in which a specific region of the  $\mathbf{H}_P$  matrix is activated on a specific frequency basis in  $\mathbf{W}_P$ . Therefore, we can reconstruct the target echo signal from  $\mathbf{W}_P$  and  $\mathbf{H}_P$ , shown inside the red rectangles.

To use the CW pulse information, in the conventional NMF-based method,  $\mathbf{W}_P$  is initialized with a set of one-hot encoded vectors, which represent the specific Doppler target echoes and are fixed during the update. The target echo frequency basis matrix  $\mathbf{W}_P$  can be expressed as follows:

$$\mathbf{W}_P = [\mathbf{w}_{P,1}, \dots, \mathbf{w}_{P,l}, \dots, \mathbf{w}_{P,L_P}], \quad (16)$$

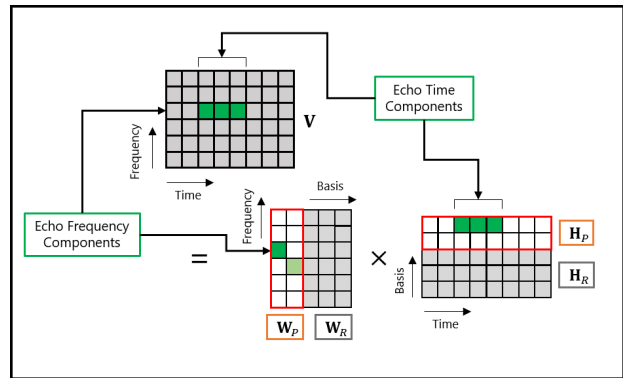


FIGURE 3. NMF scheme of two parts, target echo ( $\mathbf{W}_P$ ,  $\mathbf{H}_P$ ) and reverberation ( $\mathbf{W}_R$ ,  $\mathbf{H}_R$ ).

where  $\mathbf{w}_{P,l} \in \mathbb{R}_{K \times 1}^+$  is the  $l^{\text{th}}$  target echo frequency basis.  $\mathbf{w}_{P,l}$  is expressed as

$$\mathbf{w}_{P,l} = [0, 0, \dots, 1, \dots, 0, 0]^T. \quad (17)$$

The value one is placed at the frequency bin corresponding to the target echo-Doppler shift. This structure is similar to the Doppler replica of the matched filter, verifying all hypotheses of the target Doppler to be detected. Then, other matrices are updated using the MU rule with several constraints until convergence is achieved. Finally, echo information is reconstructed from target echo basis matrices.

Although the conventional NMF-based reverberation suppression method has been well used [12], [13], room for performance improvement exists. The conventional NMF-based reverberation suppression method's limitations can be explained as follows. First, the NMF update process of the conventional method is sensitive to environmental conditions because parts of the  $\mathbf{W}$  matrix are forced to be fixed, and others are randomly initialized. Therefore, hyperparameters must be set well for NMF updates to converge stably. Second, the conventional method is sensitive to the imbalance between the target and reverberation signals because the level scaling between the target and reverberation frequency basis components is excluded in the initial  $\mathbf{W}$  matrix. Accordingly, a problem might occur in that the time basis power is unfairly estimated and can be concentrated on either the target or reverberation. Consequently, convergence might not be performed well, or even if convergence occurs, reverberation suppression might fail because non-target components are emphasized. These problems are related to setting the initial NMF value, and similar cases have been reported in studies [18], [19].

### III. THE PROPOSED NMF-BASED REVERBERATION SUPPRESSION METHOD WITH THE PRE-TRAINED FREQUENCY BASIS MATRIX

#### A. CONTRIBUTION OF THE PROPOSED METHOD

We modified the conventional NMF-based reverberation suppression method to overcome its limitations. The main contributions of this article are as follows. First, the reverberation



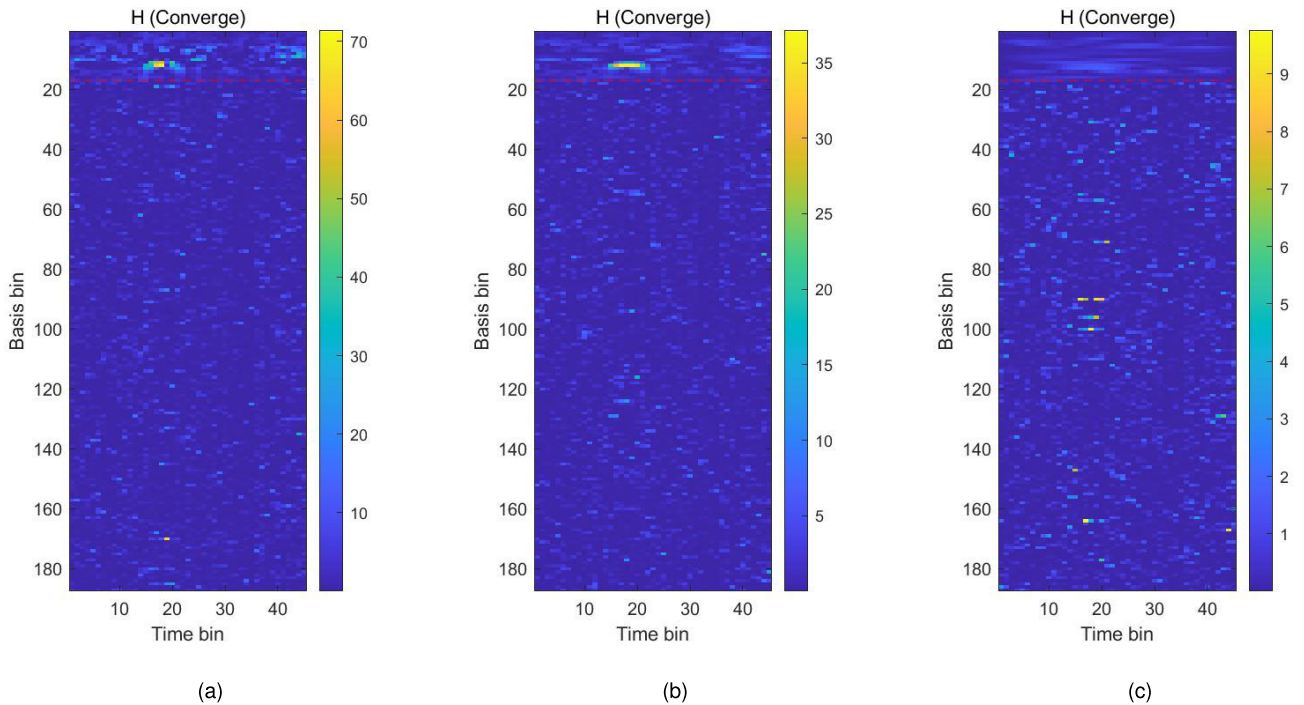


FIGURE 4. Converged  $\mathbf{H}$  matrix using conventional NMF method where parameter  $\alpha$  was set to (a) 1 (b) 10 and (c) 50.

frequency basis matrix  $\mathbf{W}_R$  was initialized (pre-trained) before NMF update to making it stable and robust with environmental fluctuations. This concept is widely employed in speech, audio, and acoustic signal processing [20]–[22], but we apply it to the sonar reverberation suppression method for the first time. Second, a power normalization scheme was adopted to compensate for the imbalance between the target echo and reverberation signal. The Frobenius norm was employed to normalize the target echo frequency basis matrix  $\mathbf{W}_P$  and reverberation frequency basis matrix  $\mathbf{W}_R$ , and a power ratio parameter was adopted to tune their relationship. Third, sophisticated sparse update rules adjusting the update rate were proposed. The proposed sparse update rules measure the sparseness of an updated matrix and applied it to the step-size parameter of the NMF update.

**B. PRE-TRAINING THE FREQUENCY BASIS MATRIX WITH POWER NORMALIZATION**

In the conventional method,  $\mathbf{W}_R$  was randomly initialized and updated using a simple MU rule. However, since the target and reverberation are received complexly through the oceanic environment, and the target is often buried in the reverberation signal. Therefore, not only the target component but also the unnecessary reverberation component might be activated during the NMF update process. Although the NMF method provides few constrained options to enhance the convergence performance, it is sensitive to the hyperparameter selection. Fig.4 (a), (b), and (c) show the

converged  $\mathbf{H}$  matrix using the conventional method, where parameter  $\alpha$  was set to 1, 10, and 50, respectively. In the figures, the red dashed line is the boundary between the target and reverberation base (based on the boundary line, the upper and lower parts mean the target and reverberation bases, respectively). For  $\alpha = 10$ , the target component appeared because NMF converged with ease. For  $\alpha = 1$ , however, unnecessary components besides the target component appeared. For  $\alpha = 50$ , the target base component was inactivated.

To overcome the limitations of the conventional method, we adopted a methodology of taking information from the reference reverberation signal in research on AR prewhiteners [7], [8] and combined it with the NMF method. This scheme can be implemented by learning the reverberation frequency basis matrix  $\mathbf{W}_R$  in advance. Notably, this scheme of pre-trained basis matrix is widely used in speech, audio, and acoustic signal processing [20]–[22].

The reverberation frequency basis matrix  $\mathbf{W}_R$  can be trained with the adjacent time or beam signals under the assumption that adjacent signals would have similar statistical characteristics of reverberation in the current block signals. We used the simple MU rule to train the  $\mathbf{W}_R$  in advance as follows:

$$\mathbf{W}_R \leftarrow \mathbf{W}_R \circ \frac{[\mathbf{V}_R / (\mathbf{W}_R \tilde{\mathbf{H}}_R)] \tilde{\mathbf{H}}_R}{\mathbf{1}_{K \times N} \tilde{\mathbf{H}}_R^T}, \tag{18}$$

$$\tilde{\mathbf{H}}_R \leftarrow \tilde{\mathbf{H}}_R \circ \frac{\mathbf{W}_R^T [\mathbf{V}_R / (\mathbf{W}_R \tilde{\mathbf{H}}_R)]}{\mathbf{W}_R^T \mathbf{1}_{K \times N}}, \tag{19}$$

where  $\mathbf{V}_R$  is the spectrogram reference signal, which containing reverberation components;  $\tilde{\mathbf{H}}_R$  is a dummy reverberation time basis matrix, used temporarily to train  $\mathbf{W}_R$ . Notably, we used the tilde symbol to distinguish  $\tilde{\mathbf{H}}_R$ , which is updated in the main NMF process. There is a possibility that the pre-trained and actual reverberation frequency bases might mismatch. However, if the reference reverberation signal is locally stationary, there is no problem using it [7], [8]. Moreover, a little mismatch can be compensated for during the NMF update process. The detailed mismatch issue will be discussed in Section V-B.

Through (16) to (19), we can initialize frequency basis matrix  $\mathbf{W}$ . However, the relationship between the level of target echo and reverberation signals is excluded in this formula. In this study, we devised a formula considering the imbalance level. To complete the formula of frequency basis matrix  $\mathbf{W}$ , we normalized both target echo part  $\mathbf{W}_P$  and reverberation part  $\mathbf{W}_R$  using the Frobenius norm and adopted a new parameter as

$$\mathbf{W} = \left[ \gamma \frac{\mathbf{W}_P}{\|\mathbf{W}_P\|_F^2} : (1 - \gamma) \frac{\mathbf{W}_R}{\|\mathbf{W}_R\|_F^2} \right], \quad (20)$$

where  $\gamma$  is the power ratio parameter and  $\|\mathbf{A}\|_F^2$  is the Frobenius norm of matrix  $\mathbf{A}$  defined as adding the square of the absolute value to all elements of  $\mathbf{A}$ . The power ratio parameter  $\gamma$  should be adjusted appropriately to ensure the performance. This issue will be discussed in Section V-A.

### C. ESTIMATING THE TIME BASIS MATRIX WITH THE MODIFIED SPARSE UPDATE RULE

To use the target echo information that has continuous features in the time bin,  $\mathbf{H}_P$  is estimated using a cost function expressed as [12], [13], [23]

$$C(\mathbf{W}_P, \mathbf{H}_P) = C_{RE}(\mathbf{W}_P, \mathbf{H}_P) + \alpha C_{TC}(\mathbf{H}_P) + \beta C_{TLL}(\mathbf{H}_P), \quad (21)$$

where  $C_{RE}(\mathbf{W}_P, \mathbf{H}_P)$ ,  $C_{TC}(\mathbf{H}_P)$ , and  $C_{TLL}(\mathbf{H}_P)$  are the costs of reconstruction error, temporal continuity (TC), and temporal length limitation (TLL), respectively.  $\alpha$  and  $\beta$  are the weighting parameters for TC and TLL, respectively [12], [23].

The gradient of each cost is the weighted sum of the gradients of each cost and expressed as

$$\nabla_{\mathbf{H}_P} C = \nabla_{\mathbf{H}_P} C_{RE} + \alpha \nabla_{\mathbf{H}_P} C_{TC} + \beta \nabla_{\mathbf{H}_P} C_{TLL}. \quad (22)$$

Notably, we omitted the parentheses for convenience. If we split this gradient into positive and negative terms, (22) is expressed as

$$\begin{aligned} \nabla_{\mathbf{H}_P} C = & \left( \nabla_{\mathbf{H}_P}^+ C_{RE} - \nabla_{\mathbf{H}_P}^- C_{RE} \right) \\ & + \alpha \left( \nabla_{\mathbf{H}_P}^+ C_{TC} - \nabla_{\mathbf{H}_P}^- C_{TC} \right) \\ & + \beta \left( \nabla_{\mathbf{H}_P}^+ C_{TLL} - \nabla_{\mathbf{H}_P}^- C_{TLL} \right). \end{aligned} \quad (23)$$

Then, we can express the total gradient term as

$$\nabla_{\mathbf{H}_P} C = \nabla_{\mathbf{H}_P}^+ C - \nabla_{\mathbf{H}_P}^- C, \quad (24)$$

where the positive terms  $\nabla_{\mathbf{H}_P}^+ C$  and the negative terms  $\nabla_{\mathbf{H}_P}^- C$  are, respectively, given by

$$\nabla_{\mathbf{H}_P}^+ C = \nabla_{\mathbf{H}_P}^+ C_{RE} + \alpha \nabla_{\mathbf{H}_P}^+ C_{TC} + \beta \nabla_{\mathbf{H}_P}^+ C_{TLL}, \quad (25)$$

$$\nabla_{\mathbf{H}_P}^- C = \nabla_{\mathbf{H}_P}^- C_{RE} + \alpha \nabla_{\mathbf{H}_P}^- C_{TC} + \beta \nabla_{\mathbf{H}_P}^- C_{TLL}. \quad (26)$$

Consequently,  $\mathbf{H}_P$  can be estimated using [23]

$$\mathbf{H}_P \leftarrow \mathbf{H}_P \circ \frac{\nabla_{\mathbf{H}_P}^- C}{\nabla_{\mathbf{H}_P}^+ C}. \quad (27)$$

See Appendix A for detailed definition of each cost and derivation of each gradient term.

In reference paper [13], a sparse update rule for matrix  $\mathbf{H}_P$  was proposed to improve NMF performance. Although the proposed sparse update rules work well, the rapid update rate due to sparse constraints causes a problem of highlighting local non-target components and fail to find the target information, even if the NMF algorithm converges. Therefore, we modified sparse update rules for  $\mathbf{H}_P$  as follows

$$\mathbf{H}_P \leftarrow (\mathbf{I} - \tilde{\mathbf{D}}) \mathbf{H}_P + \tilde{\mathbf{D}} \frac{\nabla_{\mathbf{H}_P}^- C}{\nabla_{\mathbf{H}_P}^+ C} \mathbf{H}_P, \quad (28)$$

where  $\mathbf{I}$  is an identity matrix.  $\nabla_{\mathbf{H}_P}^+ C$  and  $\nabla_{\mathbf{H}_P}^- C$  are the positive and negative total gradient terms in (25) and (26), respectively.  $\tilde{\mathbf{D}}$  is a modified normalized basis power matrix controlling the NMF update rate more sophisticated than the conventional method. See Appendix B for the detailed derivation process and Section V-C for a detailed discussion.

Estimating the reverberation basis is achieved using the MU rule without additional constraints as follows

$$\mathbf{W}_R \leftarrow \mathbf{W}_R \circ \frac{[\mathbf{V}/(\mathbf{W}\mathbf{H})]\mathbf{H}_R}{\mathbf{1}_{K \times N} \mathbf{H}_R^T}, \quad (29)$$

$$\mathbf{H}_R \leftarrow \mathbf{H}_R \circ \frac{\mathbf{W}_R^T [\mathbf{V}/(\mathbf{W}\mathbf{H})]}{\mathbf{W}_R^T \mathbf{1}_{K \times N}}. \quad (30)$$

### D. RECONSTRUCTING TARGET ECHO SIGNAL

The NMF method is applied iteratively using (28), (29), and (30) until convergence is attained. After convergence, the target echo information can be estimated by multiplying the matrices  $\mathbf{W}_P$  and  $\mathbf{H}_P$ . The NMF results can be expressed as

$$\hat{\mathbf{V}} = \mathbf{W}_P \mathbf{H}_P. \quad (31)$$

Because (31) has only magnitude information, phase information is needed to complete the target echo signal spectrogram. It is easily achieved using the original phase information in (11) [12]. The output time signal with reverberation removed is reconstructed by applying the inverse STFT as follows

$$\hat{x}(m) = \text{ISTFT} \left[ \hat{\mathbf{V}} \exp(\angle \mathbf{X}) \right] \quad (32)$$

where ISTFT means inverse STFT transformation, the inverse process of (9).

TABLE 1. Summary of the proposed algorithm.

**1. Pre-training frequency basis matrix  $\mathbf{W}$ :**

- 1.1. Initialize  $\mathbf{W}_P$  using (16)
- 1.2. Pre-train  $\mathbf{W}_R$  using (18) and (19) with reference reverberation signal

**2. NMF method:**

- 2.1. Initialize  $\mathbf{H}$  using a uniform distribution with non-negative values
- 2.2. Update  $\mathbf{W}_R$ ,  $\mathbf{H}_P$ , and  $\mathbf{H}_R$  using (28), (29), and (30), respectively. These updated equations are calculated iteratively until convergence.
- 2.3. Estimate the target echo signal using (31)
- 2.4. Finally, the target echo signal can be reconstructed by performing inverse STFT using (32)

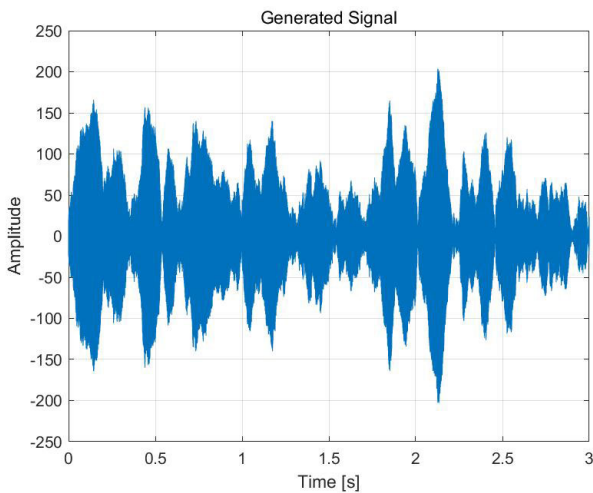


FIGURE 5. Time signal of simulated data.

We summarize the proposed reverberation suppression method in Table 1.

**IV. PERFORMANCE EVALUATION**

**A. SIMULATION ANALYSIS**

We evaluated the proposed method using simulation. The received beam signal was generated based on (2), (3), and (6) with 0.4 s CW pulse. The number of point scatterers to synthesize the reverberation signal was set to 1000. The arrival time was set to have a uniform distribution of 3 s, and the attenuation factor was set to a Gaussian distribution with a mean of 1 and a standard deviation of 0.1. The platform speed was set to 4 m/s such that  $\frac{2v}{c} = 0.0053$ .

STFT was performed using a 0.5 s Hamming window and a 75% overlap. The numbers of NMF target and reverberation echo bases ( $L_P$  and  $L_R$ ) were set to 33 and 330, respectively. The TC weighting parameter  $\alpha$ , TLL weighting parameter  $\beta$ , power normalization factor  $\gamma$ , and expected echo length  $L_T$  were set to 10, 1, 0.75, and 0.52 s, respectively.

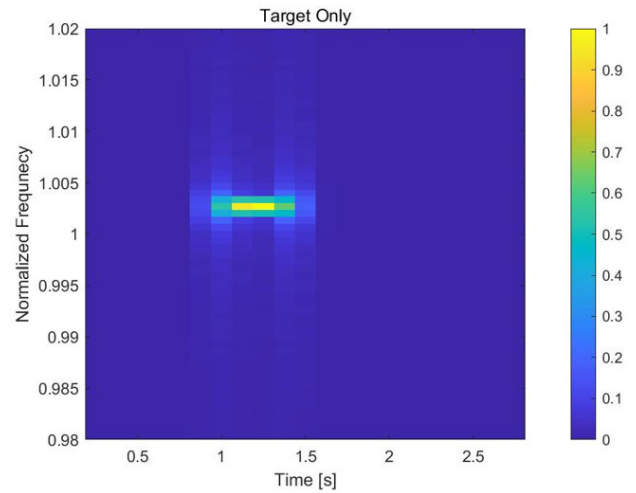


FIGURE 6. Spectrogram of target echo signal.

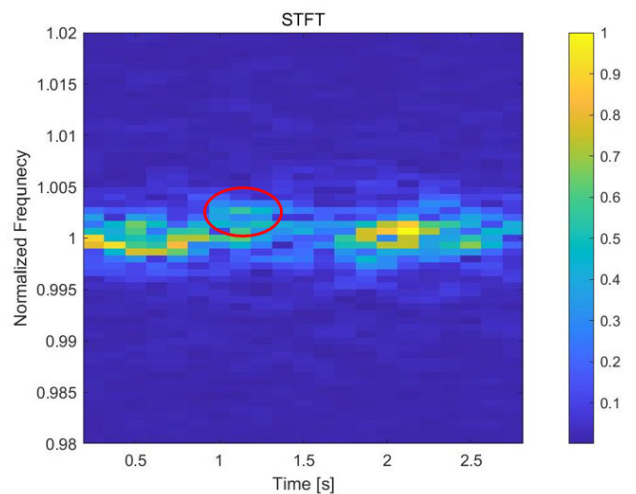
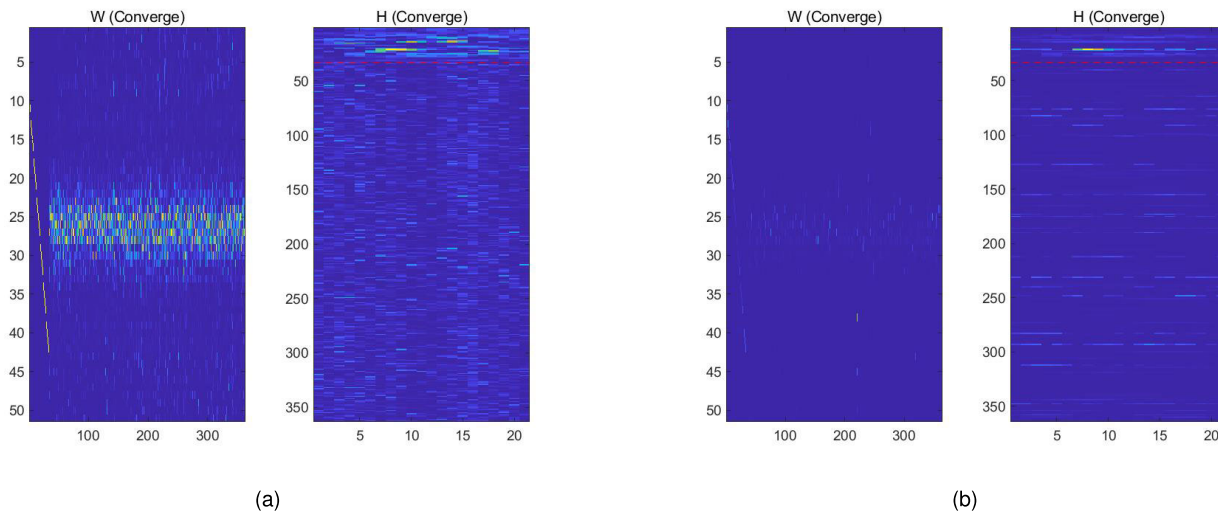


FIGURE 7. Spectrogram of generated signal (NMF input).

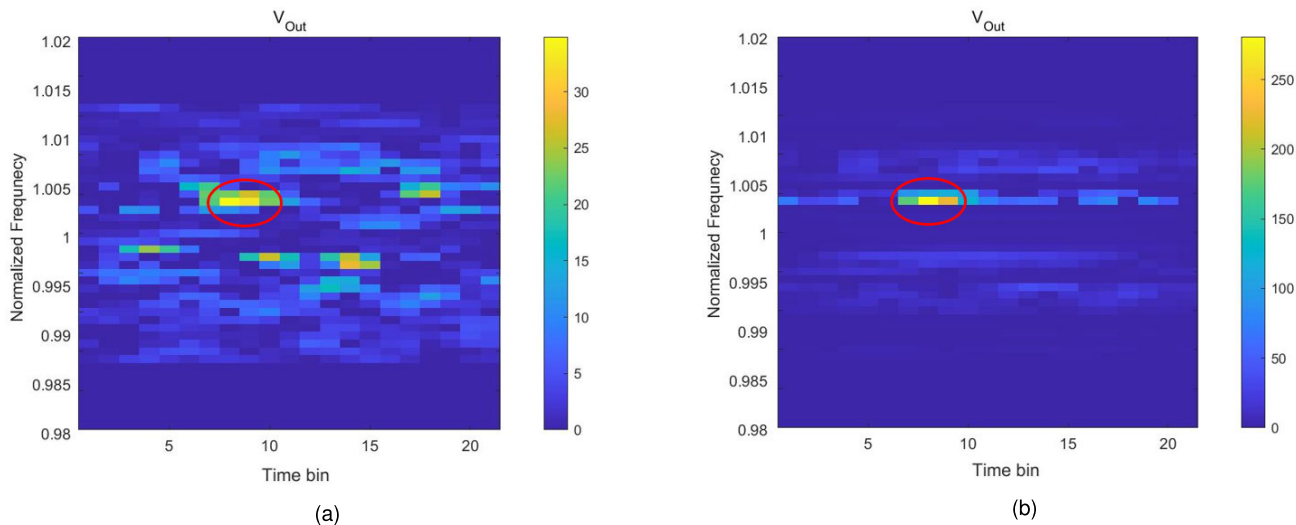
The NMF method's matrices were initialized to perform the update process. The time basis matrix  $\mathbf{H}$  was initialized with a non-negative value, and the target echo frequency basis matrix  $\mathbf{W}_P$  was initialized with one-hot encoding (16) and (17). For the conventional method, the reverberation frequency basis matrix  $\mathbf{W}_R$  was initialized with a non-negative value. However, for the proposed method, the MU rule learned  $\mathbf{W}_R$  using the reference reverberation signal.

Fig. 5 shows the time signal of the generated beam signal. Fig.6 and Fig.7 show the spectrograms of the target echo signal and generated beam signal, respectively. The target echo signal was set to arrive at 1 s with a 1.0026 normalized Doppler. The signal-to-reverberation ratio (SRR) was set to  $-19$  dB. In Fig.7, the target echo signal, highlighted with a red ellipse, is buried in the reverberation; thus, challenging to identify.

In Fig. 8, we present the estimated frequency and time basis matrices ( $\mathbf{W}$  and  $\mathbf{H}$ , respectively) of the conventional



**FIGURE 8.** W matrix and converged H matrix of (a) the conventional method and (b) the proposed method with simulated data. The red line represents the boundary between the target and reverberation bases.



**FIGURE 9.** NMF output of (a) the conventional method and (b) the proposed method with simulated data.

and proposed methods, respectively. The figure shows that each basis of the reverberation frequency basis matrix of the conventional method does not occupy a specific frequency bin. However, each basis of the reverberation frequency basis matrix of the proposed method occupied a specific frequency bin, and time bases were activated in specific regions. This phenomenon occurred because the reverberation basis matrix of the proposed method acquired reverberation information from the reference reverberation signal; therefore, it could accurately estimate the reverberation components.

Fig. 9 (a) and (b) depict the NMF output of conventional and proposed methods, respectively. From Fig. 9 (a), energy occurred in components other than the target, but the result of the proposed method in Fig. 9 (b) showed only the target component; most reverberation components were removed. Notably, near the normalized frequency 1, components were

strongly suppressed because the NMF algorithm accurately estimated the reverberation components.

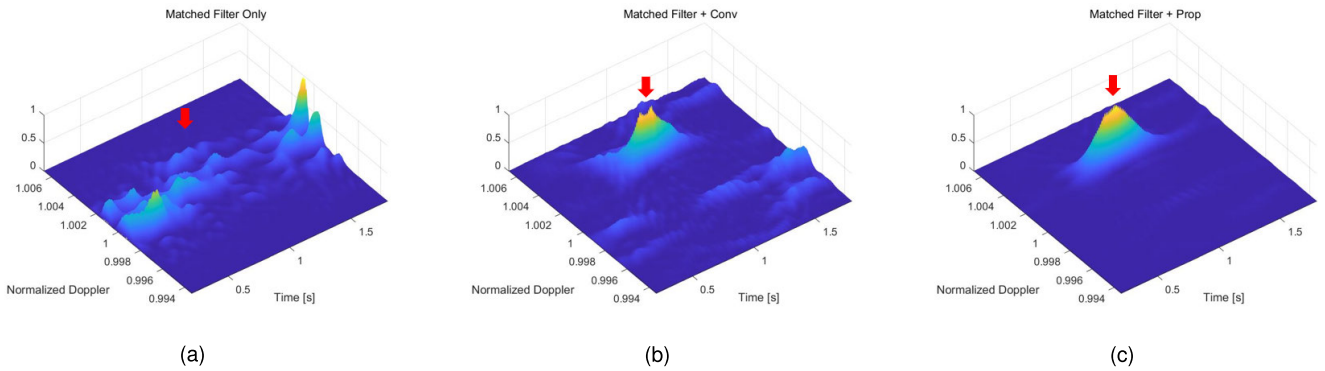
We used a normalized matched filter to compare the proposed method’s detection performance when the reverberation suppression algorithm was and was not applied. The normalized matched filter is defined as follows [24]

$$\chi(t, \eta) = \frac{\int x(t) s_p^*(t, \eta) dt}{\sqrt{\int x^2(t) dt \int s_p^2(t, \eta) dt}}, \quad (33)$$

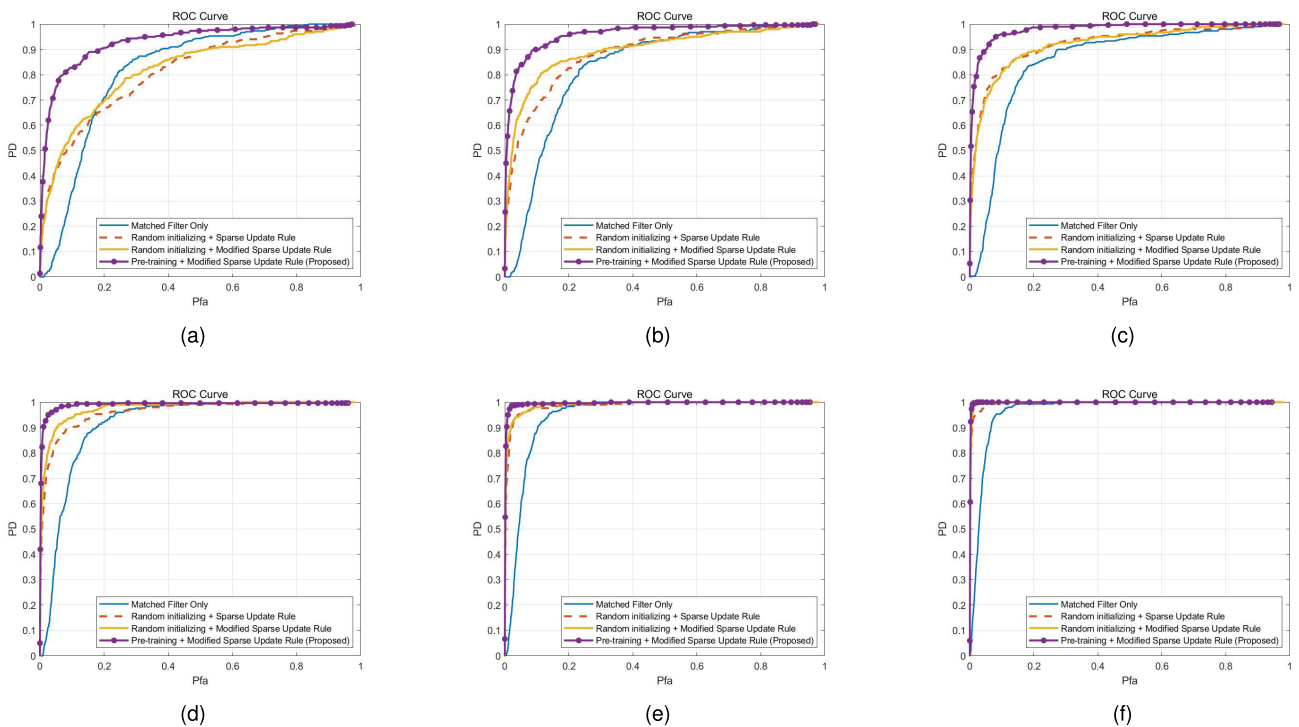
where  $s_p(t, \eta)$  is the Doppler replica of the transmitted CW pulse in (1), and  $x(t)$  is the received target echo signal in (2). In the simulation, the normalized Doppler range was set to  $-0.9947$  to  $1.0053$ .

Fig. 10 shows the normalized matched filter results at an SRR of  $-19$  dB. The normalized matched filter results





**FIGURE 10.** Comparison of matched filter results using (a) received signal, (b) output signal of the conventional method, and (c) output signal of the proposed method with simulated data.



**FIGURE 11.** Analysis of receiver operating characteristics (ROC) curve. This analysis was conducted at signal-to-reverberation ratio (SRR) of (a–f)  $-27$ ,  $-25$ ,  $-23$ ,  $-21$ , and  $-19$ ,  $-17$  dB, respectively.

using the original signal could not show a clear target echo peak. However, the normalized matched filter results using the output signal of the proposed method showed a clear target echo peak because reverberation was effectively suppressed. Although the normalized matched filter results using the output signal of the conventional method also showed a target echo peak, its performance was inferior to the proposed method.

We quantitatively analyzed the detection performance by calculating the receiver operating characteristics (ROC) curve performing 300 Monte–Carlo simulations per SRR in Fig. 11. In this figure, four cases are compared corresponding to without pre-processing, random initializing with the conventional

sparse update rule, random initializing with the proposed sparse update rule, and pre-trained initializing with the proposed sparse update rule. From these figures, the proposed sparse update rule is superior to the conventional sparse update rule, and when combined with the pre-trained initializing method proposed in this paper, the proposed NMF-based reverberation suppression method performs better than the conventional method in all conditions.

### B. SEA EXPERIMENTS ANALYSIS

To further evaluate the proposed NMF-based reverberation suppression method, data of sea experiments conducted at the Eastern Sea of Pohang, Republic of Korea, were analyzed.

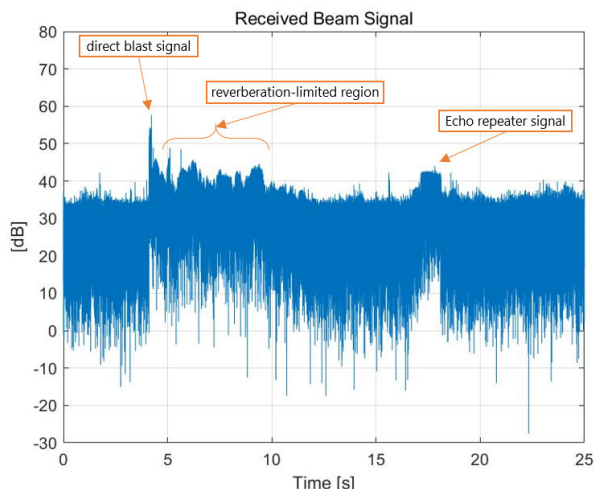


FIGURE 12. Time signal of sea experiment data.

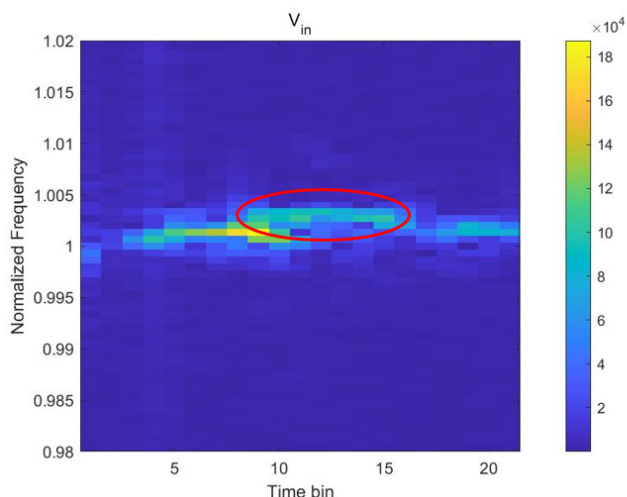


FIGURE 14. Spectrogram of received signal of the sea experiment data (NMF input).

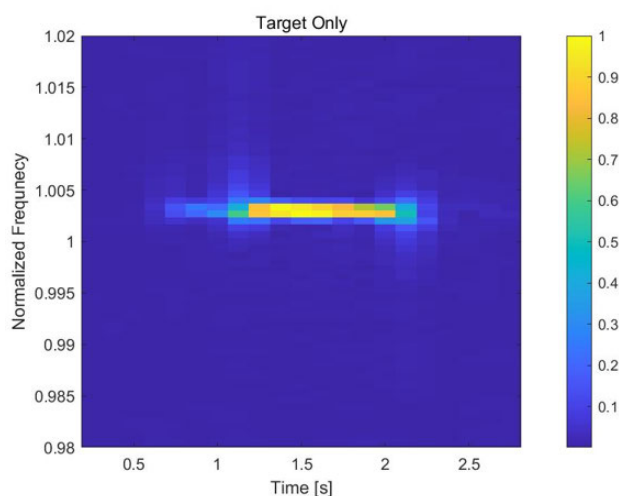


FIGURE 13. Spectrogram of target echo signal of sea experiment data.

A 1.0 s CW pulse and an echo repeater were used to imitate the target echo signal.

Fig. 12 shows the received beam signal of sea experiment data. A direct blast signal arrived at 4 s, and reverberation was subsequently received. Unfortunately, the received echo repeater signal arrived far from the direct blast arrival, approximately 17 s. Consequently, it was in the noise-limited region. Therefore, we must cut out the echo repeater signal and relocate it to the reverberation-limited region to reasonably evaluate the reverberation suppression methods.

Fig. 13 shows the spectrogram of the echo repeater signal, and Fig. 14 shows the signal synthesized by cutting the target echo repeater signal. The target echo signal was set to arrive at the 10<sup>th</sup> time bin, and Doppler was estimated to be 1.003.

Figs. 15 (a) and (b) show the NMF output of the conventional and proposed methods, respectively. When the conventional method was used, reverberation components remained,

whereas they were effectively removed when the proposed method was used and target echo was emphasized. From Fig. 15 (b), the component around normalized frequency 1.002 was significantly attenuated because the component corresponding to the pre-trained reverberation basis component was effectively removed. This phenomenon was predicted in the simulation results.

Fig. 16 shows the normalized matched filter results of sea experiment data. Fig. 16 (a) shows the normalized matched filter results using original sea experiment data, and no target echo peak appeared because the reverberation remained strong. It masked the target echo signal. Figs. 16 (b) and (c) show the normalized matched filter results using the output signal of the conventional and proposed methods, respectively. Although the conventional method suppressed the reverberation components, superior reverberation suppression performance was seen when using the proposed method. Therefore, the proposed method outperforms the conventional method, even in the actual sea environment.

## V. FURTHER DISCUSSIONS

### A. DISCUSSION FOR HYPERPARAMETER SETTINGS

In the proposed NMF method, three major hyperparameters  $\alpha$ ,  $\beta$  and  $\gamma$  affect the NMF performance results. To evaluate the performance according to hyperparameters, we conducted 100 Monte–Carlo simulations and calculated ROC curves at SRR  $-25$  dB.

Fig. 17 (a), (b), and (c) shows the ROC curves with various  $\alpha$ ,  $\beta$  and  $\gamma$ , respectively. While analyzing certain parameters, other parameters are the same Sec. IV. From these figures, the value of  $\alpha$  and  $\beta$  does not significantly affect the performance. Therefore, the proposed method converges stably, regardless of hyperparameter tuning, unlike the conventional method. This robustness is because the frequency basis matrix is trained in advance using a reference signal. For the hyperparameter  $\gamma$  case, ranges of 0.7 to 0.9 show good

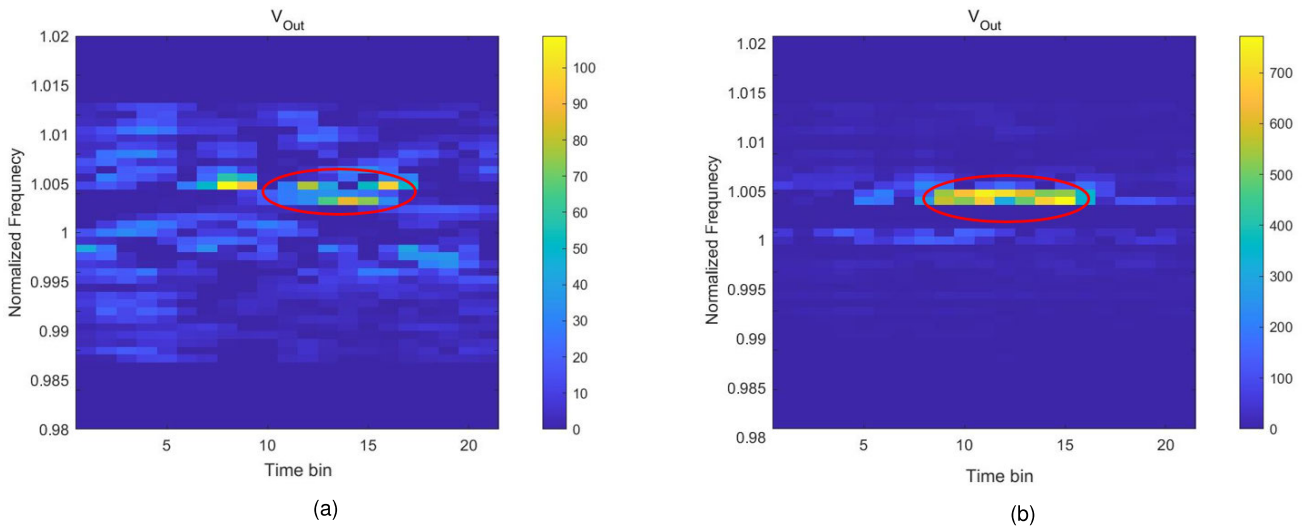


FIGURE 15. NMF output of (a) the conventional method and (b) the proposed method with sea experiment data.

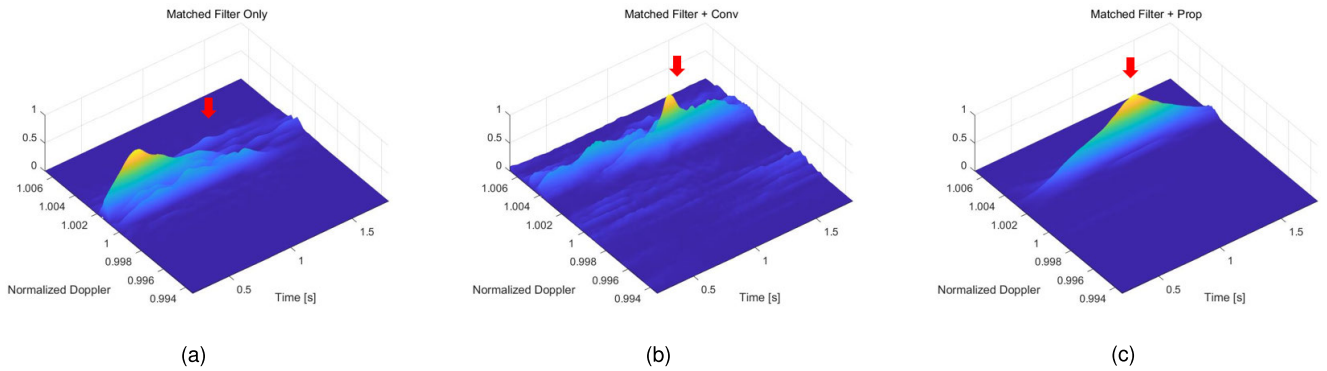


FIGURE 16. Comparison of matched filter results using (a) the received signal, (b) the output signal of the conventional method, and (c) the output signal of the proposed method with sea experiment data.

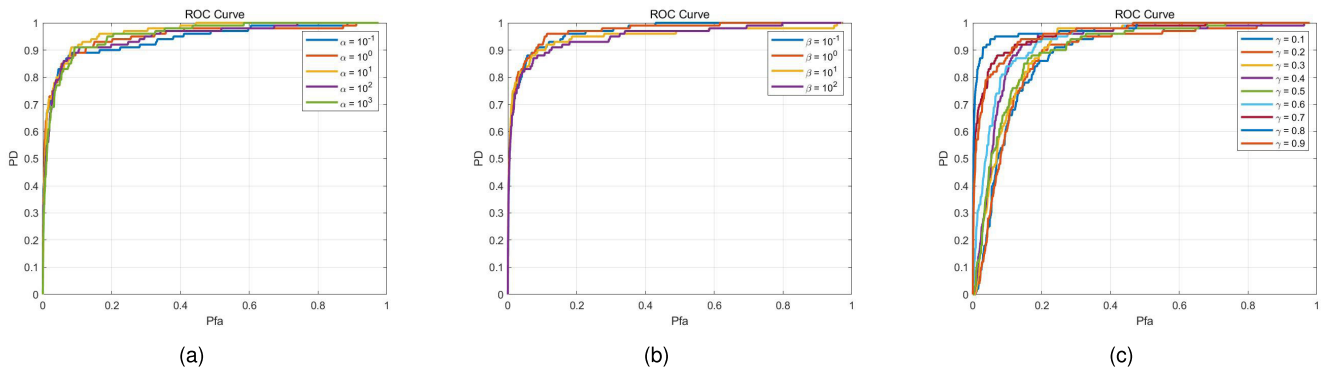
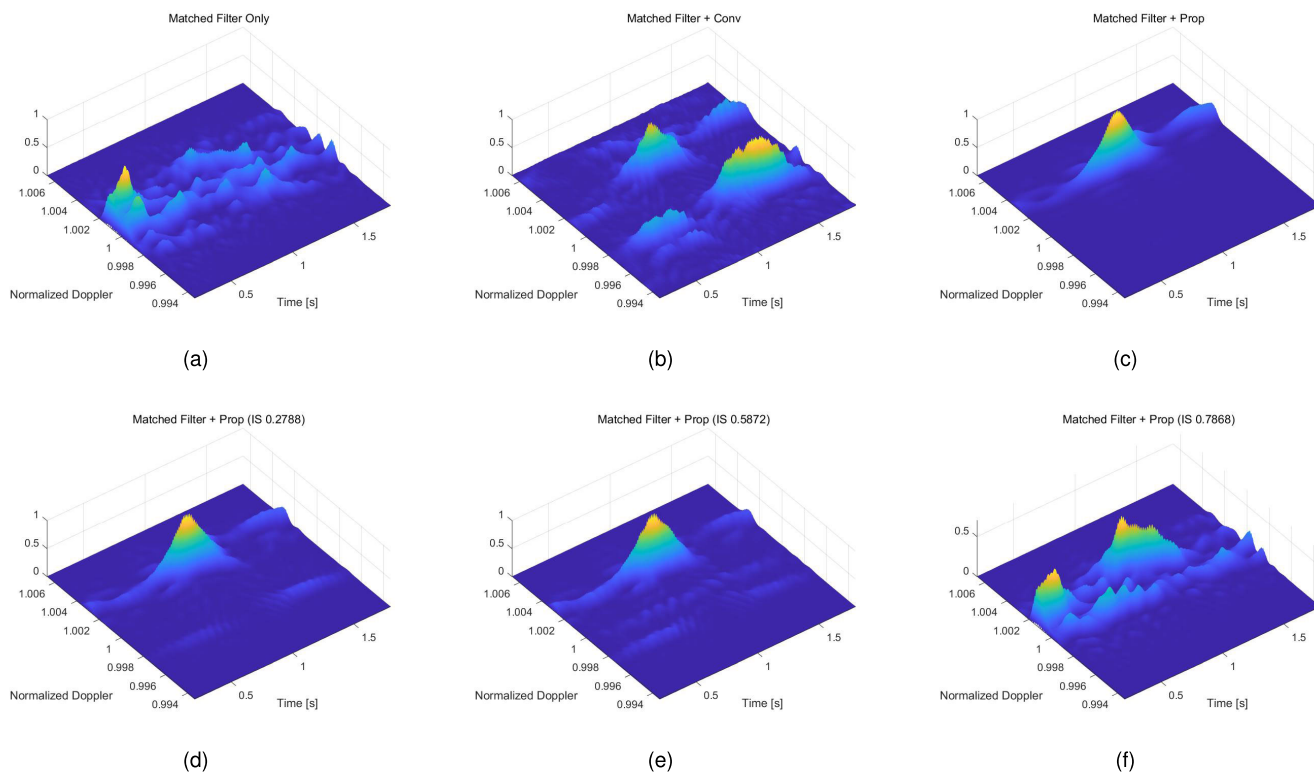


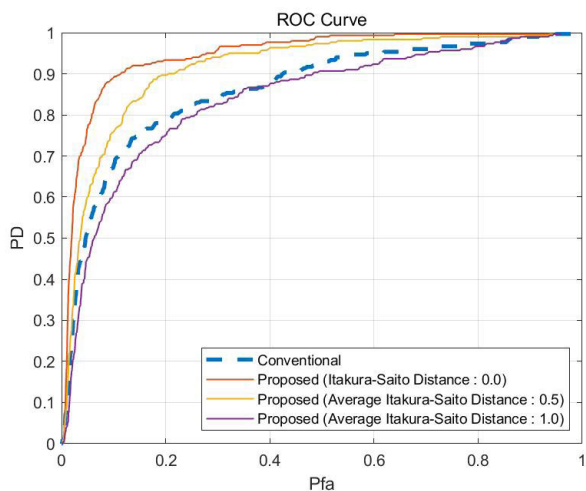
FIGURE 17. Analysis of the ROC curve according to parameter (a)  $\alpha$ , (b)  $\beta$  and (c)  $\gamma$ .

performance. With our further investigations, the optimal value for  $\gamma$  depends on the reverberation level. Therefore, it can be fine-tuned if we know the reverberation level in

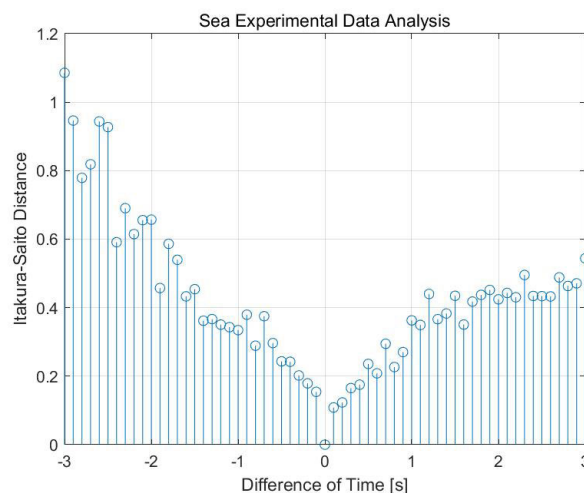
advance. Such hyperparameter tuning typically occurs in the other reverberation suppression method, like the PCI algorithm [9], [10].



**FIGURE 18.** Comparison of matched filter results. (a) Without pre-processing. (b) With conventional NMF. (c) With proposed NMF according to the Itakura-Saito distance of (c) 0.00, (d) 0.27, (e) 0.58, and (f) 0.78.



**FIGURE 19.** ROC curve according to various Itakura-Saito distances.



**FIGURE 20.** Measured Itakura-Saito distance in sea experimental data according to the time difference.

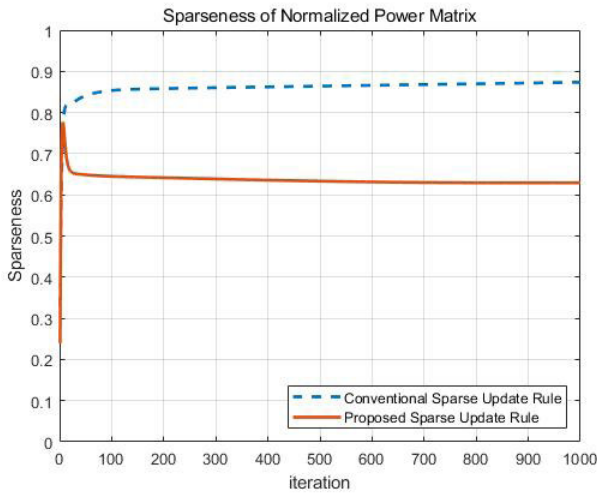
**B. DISCUSSION FOR MISMATCH ISSUE OF PRE-TRAINED REVERBERATION BASIS MATRIX**

In this paper, we proposed a pre-trained reverberation basis matrix using a reference signal. Therefore, the spectrum similarity between the reverberation signals in the target beam and reference signals directly affects the performance. To analyze the reference mismatch issue, we simulated some reference

reverberations, which have various spectrum similarity conditions. Note that spectrum similarity is measured using the Itakura–Saito distance [25].

Fig. 18 shows a comparison of matched filter results with various conditions at SRR  $-19$  dB. Figs. 18 (a) and (b) show matching filter results without pre-processing and with conventional NMF, respectively. These results are depicted





**FIGURE 21.** Comparison of sparseness of diagonal components of normalized power vector  $\mathbf{d}$  in (50).

for comparison. Figs. 18 (c), (d), (e), and (f) show the matched filter results when the proposed NMF uses reference signals with Itakura–Saito distances of 0.00, 0.27, 0.58, and 0.78. From these figures, no significant difference occurs until the Itakura–Saito distance is 0.58, but when it is increased to about 0.78, the performance deteriorates noticeably. This phenomenon is interpreted that the reference signal should have similar spectral characteristics as the reverberation signal of the target beam signal.

Fig. 19 shows the ROC curve calculated using 100 Monte–Carlo simulations at SRR  $-25$  dB using Itakura–Saito distances of 0, 0.5, and 1. In this figure, the performance decreases as the Itakura–Saito distance increases, but a value of 0.5 guarantees better performance than the conventional method.

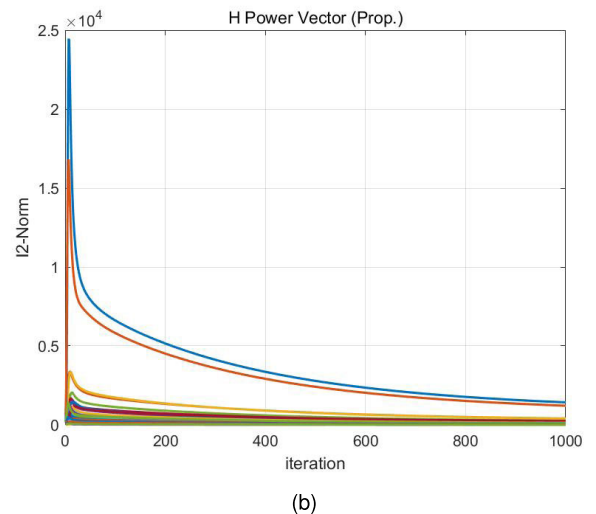
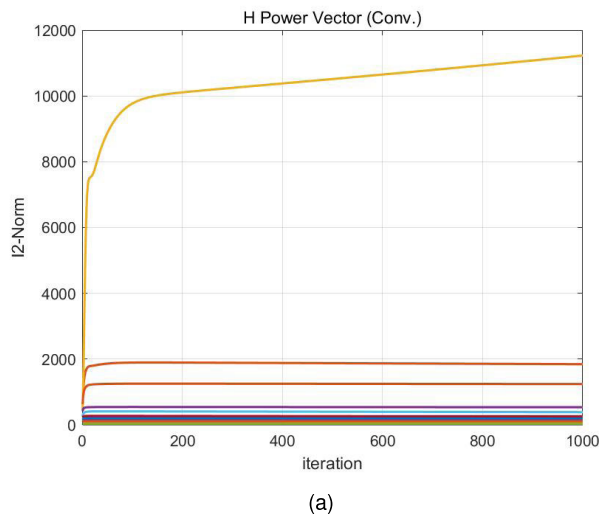
Fig. 20 shows the Itakura–Saito distance according to the time difference measured from sea experimental data. In the figure, as the time difference increases, the Itakura–Saito distance gradually increases. However, the signal within 1 s has an Itakura–Saito distance as low as 0.2 to 0.4, and most signals within 3 s have an Itakura–Saito distance lower than 0.6. From the ROC analysis in Figs. 18 and 19, it is desirable to use a reference signal with an Itakura–Saito distance of 0.6 or less. Therefore, it is desirable to use a signal close to the target signal within roughly 3 s as a reference signal.

**C. DISCUSSION FOR THE NEW SPARSE UPDATE RULE**

In this subsection, the update process and results of the algorithm were analyzed to verify that the proposed sparse update rule mechanism is valid. Fig. 21 compares the sparseness of the normalized power vector  $\mathbf{d}$ , whose elements are the power of each time basis vector. Therefore,  $\mathbf{d}$  being sparse means that only a certain time basis vector is being strongly activated. In this figure, the sparseness continuously increases in the conventional case and decreases at some point in the proposed case. This phenomenon can be interpreted as appearing because the step-size is adjusted to a lower level when the sparseness becomes too large in the proposed case.

Figs. 22 (a) and (b) show the  $l_2$ -norm of each row vector of  $\mathbf{H}_P$  when using the conventional and proposed cases, respectively. In this figure, only one specific basis power continues to increase in the conventional case because, if it initially converges erroneously, excessive sparsity continuously emphasizes the corresponding component. However, in the proposed case, the power of multiple bases converges to a lower level as the update is repeated, reducing the problem of excessive sparsity.

Figs. 23 (a) and (b) show NMF results using the conventional and proposed cases, respectively. While the conventional method emphasized only one non-target component,



**FIGURE 22.**  $l_2$ -norm of each row of  $\mathbf{H}_P$  when using (a) conventional and (b) proposed methods, respectively.

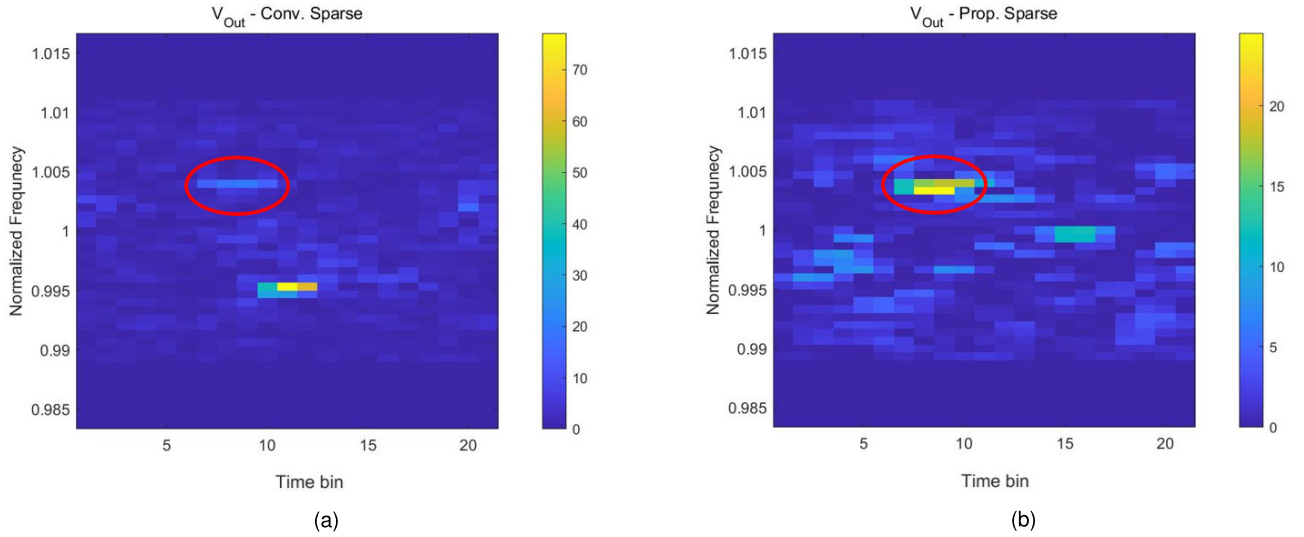


FIGURE 23. NMF output when using (a) conventional and (b) proposed methods, respectively.

the proposed method successfully emphasized the target component because, as analyzed in Figs. 21 and 22, the proposed method adjusts the step-size according to sparseness.

## VI. CONCLUSION

An NMF-based reverberation suppression method with a pre-trained frequency basis matrix was proposed. The proposed method improved the initialization problem of the conventional method and the sparse update rule. The proposed method was verified by evaluating it using simulation and sea experiment data. In the simulation, the comparative analyses of spectrogram, normalized matched filter, and ROC curve are conducted and the proposed method showed superior detection performance than the conventional method when it converges in valid SRR conditions. In the spectrogram and normalized matching filter analysis using sea experimental data, the proposed method showed superior performance as expected in the simulation. Hyperparameter setting, pre-trained reverberation basis mismatch issue, and new sparse update rule were discussed to analyze the proposed method from various viewpoints.

## APPENDIX A

### REVIEW OF COST FUNCTION AND GRADIENT TERMS IN THE NMF METHOD

The cost function comprises three parts. The reconstruction error (RE)  $C_{RE}$  is defined by the KL divergence, which is generally used in the NMF method as [17]

$$C_{RE}(\mathbf{W}_P, \mathbf{H}_P) = \sum_{k=1}^K \sum_{n=1}^N \left( [\mathbf{V}]_{(k,n)} \log \frac{\mathbf{V}_{(k,n)}}{[\mathbf{WH}]_{(k,n)}} - [\mathbf{V}]_{(k,n)} + [\mathbf{WH}]_{(k,n)} \right), \quad (34)$$

and its gradient terms are given as

$$\nabla_{\mathbf{H}_P}^+ C_{RE} = \mathbf{W}_P^T \mathbf{1}_{K \times N}, \quad (35)$$

$$\nabla_{\mathbf{H}_P}^- C_{RE} = \mathbf{W}_P^T \frac{\mathbf{V}}{\mathbf{WH}}. \quad (36)$$

The temporal continuity (TC) is measured by assigning the cost of large changes between adjacent frames for each row component in  $\mathbf{H}_P$ . The TC cost function is defined as [23]

$$C_{TC}(\mathbf{H}_P) = \sum_{k=1}^K \left( \frac{\sum_{n=2}^N (h_{(k,n)} - h_{(k,n-1)})^2}{\frac{1}{N} \sum_{n=1}^N h_{(k,n)}^2} \right), \quad (37)$$

and its gradient terms are given as

$$\nabla_{\mathbf{H}_P}^+ C_{TC} = \frac{4N\mathbf{H}_P}{\mathbf{H}_P^2 \mathbf{1}_{N \times N}}, \quad (38)$$

$$\nabla_{\mathbf{H}_P}^- C_{TC} = 2N \frac{\mathbf{H}_{P \rightarrow 1} - \mathbf{H}_{P \leftarrow 1}}{\mathbf{H}_P^2 \mathbf{1}_{N \times N}} + \frac{2N\mathbf{H}_P \circ [(\mathbf{H}_P - \mathbf{H}_{P \rightarrow 1})^2 \mathbf{1}_{N \times N}]}{(\mathbf{H}_P^2 \mathbf{1}_{N \times N})^2}, \quad (39)$$

where  $\mathbf{H}_{P \leftarrow 1}$ ,  $\mathbf{H}_{P \rightarrow 1}$  are variants of  $\mathbf{H}_P$  shifted one column to the left and right, respectively, and  $\mathbf{1}_{N \times N}$  is an  $N \times N$  matrix whose elements are all ones. Furthermore,  $\mathbf{A}^2$  indicates an element-wise square of matrix  $\mathbf{A}$ .

The temporal length limiting (TLL) is used to penalize activated components longer than the expected length of the target echo signal. The TLL cost function is defined as [12]

$$C_{TLL}(\mathbf{H}_P) = \sum_{k=1}^K \left( 1 - \sum_{m=n}^{n+l_T-1} \hat{h}_{(k,n)} \right), \quad (40)$$

and its gradient terms are given as

$$\left[ \nabla_{\mathbf{H}_P}^+ C_{TLL} \right]_{(k,n)} = \sum_{m=n}^{n+l_T-1} \left( \frac{e^{[\hat{\mathbf{H}}_P]_{(k,m)}}}{\sum_{i=1}^N e^{[\hat{\mathbf{H}}_P]_{(k,i)}}} \right)^2, \quad (41)$$

$$\left[ \nabla_{\mathbf{H}_P}^- C_{TLL} \right]_{(k,n)} = \sum_{m=n}^{n+l_T-1} \left( \frac{e^{[\tilde{\mathbf{H}}_P]_{(k,m)}}}{\sum_{i=1}^N e^{[\tilde{\mathbf{H}}_P]_{(k,i)}}} \right), \quad (42)$$

where  $[\mathbf{A}]_{(k,n)}$  is the  $(k, n)$  element of matrix  $\mathbf{A}$ , and  $l_T$  is the expected target echo length.  $[\tilde{\mathbf{H}}_P]_{(k,n)}$  is calculated using the moving sum of each target echo time basis, i.e., the rows of  $\mathbf{H}_P$ . It is expressed as

$$[\tilde{\mathbf{H}}_P]_{(k,n)} = \sum_{m=n-l_T+1}^n [\mathbf{H}_P]_{(k,m)}. \quad (43)$$

**APPENDIX B  
DERIVATION OF SPARSE UPDATE RULE WITH ADJUSTING UPDATE RATE USING SPARSENESS**

The original MU rule in (27) is a modified element-wise gradient descent algorithm, expressed as follows:

$$[\mathbf{H}_P]_{(k,n)} \leftarrow [\mathbf{H}_P]_{(k,n)} - \eta(k,n) \left( \nabla_{[\mathbf{H}_P]_{(k,n)}}^+ C - \nabla_{[\mathbf{H}_P]_{(k,n)}}^- C \right), \quad (44)$$

where  $\eta(k,n)$  is the step-size parameter for updates.  $\nabla_{[\mathbf{H}_P]_{(k,n)}}^+ C$  and  $\nabla_{[\mathbf{H}_P]_{(k,n)}}^- C$  are the element-wise positive and negative parts of the gradient, which have non-negative values. Typically,  $\eta(k,n)$  is set to a small positive scalar value; however, if we replace it as  $\eta(k,n) = [\mathbf{H}_P]_{(k,n)} / \nabla_{[\mathbf{H}_P]_{(k,n)}}^+ C$ , we can obtain (27). Using this process, if we appropriately control the step-size, we can reflect our intention.

Recently, Kim *et al.* [13] proposed a sparse constraint update rule. They implemented sparse constraints by setting the step-size parameter proportional to the power of each basis vector as

$$\eta(k,n) = d_r \frac{[\mathbf{H}_P]_{(k,n)}}{\nabla_{[\mathbf{H}_P]_{(k,n)}}^+ C}, \quad (45)$$

where  $d_r$  is the normalized power of each basis vector. It is expressed as

$$d_r = \frac{\|[\mathbf{H}_P]_r\|_2}{\max(\|[\mathbf{H}_P]_1\|_2, \dots, \|[\mathbf{H}_P]_{L_P}\|_2)}, \quad (46)$$

where  $[\mathbf{H}_P]_r$  is the  $r$ th row vector of  $\mathbf{H}_P$  and  $\|\cdot\|_2$  indicates the  $l_2$ -norm. Notably, we constrain each power of the echo time basis, not each element of  $\mathbf{H}_P$ . If we substitute (44) into (45), it yields the new element-wise update rule as

$$[\mathbf{H}_P]_{(k,n)} \leftarrow (1 - d_r) [\mathbf{H}_P]_{(k,n)} + d_r \frac{\nabla_{[\mathbf{H}_P]_{(k,n)}}^- C}{\nabla_{[\mathbf{H}_P]_{(k,n)}}^+ C} [\mathbf{H}_P]_{(k,n)}. \quad (47)$$

We can formulate these element-wise update rules in a matrix form as follows:

$$\mathbf{H}_P \leftarrow (\mathbf{I} - \mathbf{D}) \mathbf{H}_P + \mathbf{D} \frac{\nabla_{\mathbf{H}_P}^- C}{\nabla_{\mathbf{H}_P}^+ C} \mathbf{H}_P, \quad (48)$$

where  $\mathbf{I}$  is an identity matrix,  $\nabla_{\mathbf{H}_P}^+ C$  and  $\nabla_{\mathbf{H}_P}^- C$  are positive and negative total gradient terms in (25) and (26), respectively.  $\mathbf{D}$  is a diagonal matrix, comprising the normalized

power of each basis vector. It can be expressed as

$$\mathbf{D} = \text{diag}(\mathbf{d}), \quad (49)$$

$$\mathbf{d} = [d_1, \dots, d_{L_P}]^T, \quad (50)$$

where  $\mathbf{d}$  is the normalized power vector and  $\text{diag}(\cdot)$  indicates a diagonal matrix.

The problem of the conventional sparse update rule in (48) is that it can highlight local non-target components that appeared during the NMF update because of too much emphasis on sparseness. This phenomenon might cause an inability to find the target information, even if the NMF algorithm converges. Therefore, we proposed a new step-size parameter, which is the weighted sum of the original and sparse step-size parameters. The weighting ratio was determined by measuring the sparseness that appeared according to the NMF update.

The proposed step-size parameter is expressed as

$$\eta(k,n) = (1 - S(\mathbf{d})) d_r \frac{[\mathbf{H}_P]_{(k,n)}}{\nabla_{[\mathbf{H}_P]_{(k,n)}}^+ C} + S(\mathbf{d}) \frac{[\mathbf{H}_P]_{(k,n)}}{\nabla_{[\mathbf{H}_P]_{(k,n)}}^+ C} \quad (51)$$

where  $S(\mathbf{d})$  is the sparseness of the updated matrix. From this equation, the original and sparse step-size parameters are combined via a sparsity term. As sparsity increases, the preceding weights become smaller; consequently, the sparse step-size has a smaller impact and vice versa.

The sparseness of basis power vector  $\mathbf{d}$  can be measured as follows [26]:

$$S(\mathbf{d}) = \frac{\sqrt{L_P} - (\sum_r |d_r|) / (\sum_r \sqrt{d_r^2})}{\sqrt{L_P} - 1}, \quad (52)$$

$$\mathbf{d} = [d_1, \dots, d_{L_P}], \quad (53)$$

where  $L_P$  is the number of elements in vector  $\mathbf{d}$ , and  $|\cdot|$  is the absolute value.

For convenience, we reformulate (51) as

$$\eta(k,n) = (d_r + S(\mathbf{d})(1 - d_r)) \frac{[\mathbf{H}_P]_{(k,n)}}{\nabla_{[\mathbf{H}_P]_{(k,n)}}^+ C}. \quad (54)$$

The modified step-size parameter shows that  $d_r$  in (45) appears in a modified form  $(d_r + S(\mathbf{d})(1 - d_r))$ . Therefore, the proposed sparse update rule can be written as follows, with the new normalized power matrix  $\tilde{\mathbf{D}}$ :

$$\mathbf{H}_P \leftarrow (\mathbf{I} - \tilde{\mathbf{D}}) \mathbf{H}_P + \tilde{\mathbf{D}} \frac{\nabla_{\mathbf{H}_P}^- C}{\nabla_{\mathbf{H}_P}^+ C} \mathbf{H}_P, \quad (55)$$

where

$$\tilde{\mathbf{D}} = \text{diag}(\tilde{d}_1, \dots, \tilde{d}_{L_P}), \quad (56)$$

and

$$\tilde{d}_r = d_r + S(\mathbf{d})(1 - d_r). \quad (57)$$

## REFERENCES

- [1] R. J. Urick, *Principles of Underwater Sound for Engineers*. New York, NY, USA: McGraw-Hill, 1967.
- [2] R. O. Nielsen, *Sonar Signal Processing*. Norwood, MA, USA: Artech House, 1991.
- [3] T. Collins, "Active sonar pulse design," Ph.D. dissertation, School Electron. Elect. Eng., Dept. Eng., Univ. Birmingham, Birmingham, U.K., Nov. 1996.
- [4] A. D. Waite and A.D. Waite, *Sonar for Practising Engineers*, vol. 3. New York, NY, USA: Wiley, 2002.
- [5] D. A. Hague and J. R. Buck, "The generalized sinusoidal frequency-modulated waveform for active sonar," *IEEE J. Ocean. Eng.*, vol. 42, no. 1, pp. 109–123, Jan. 2017.
- [6] D. A. Abraham, *Underwater Acoustic Signal Processing: Modeling, Detection, and Estimation*. Cham, Switzerland: Springer, 2019.
- [7] S. Kay and J. Salisbury, "Improved active sonar detection using autoregressive prewhiteners," *J. Acoust. Soc. Amer.*, vol. 87, no. 4, pp. 1603–1611, Apr. 1990.
- [8] B. W. Choi, E. H. Bae, J. S. Kim, and K. K. Lee, "Improved prewhitening method for linear frequency modulation reverberation using dechirping transformation," *J. Acoust. Soc. Amer.*, vol. 123, no. 3, pp. EL21–EL25, Mar. 2008.
- [9] G. Ginolhac and G. Jourdain, "'Principal component inverse' algorithm for detection in the presence of reverberation," *IEEE J. Ocean. Eng.*, vol. 27, no. 2, pp. 310–321, Apr. 2002.
- [10] W. Li, Q. Zhang, X. Ma, and C. Hou, "Active sonar detection in reverberation via signal subspace extraction algorithm," *EURASIP J. Wireless Commun. Netw.*, vol. 2010, no. 1, pp. 1–10, Dec. 2010.
- [11] M. Wang, S. Wu, S. Guo, and D. Peng, "Study on an anti-reverberation method based on PCI-SVM," *Appl. Acoust.*, vol. 182, Nov. 2021, Art. no. 108189.
- [12] S. Lee and J.-S. Lim, "Reverberation suppression using non-negative matrix factorization to detect low-Doppler target with continuous wave active sonar," *EURASIP J. Adv. Signal Process.*, vol. 2019, no. 1, p. 11, Dec. 2019.
- [13] G. Kim, K. Lee, and S. Lee, "Linear frequency modulated reverberation suppression using non-negative matrix factorization methods, dechirping transformation and modulo operation," *IEEE Access*, vol. 8, pp. 110720–110737, 2020.
- [14] W. S. Burdick, *Underwater Acoustic System Analysis*. Upper Saddle River, NJ, USA: Prentice-Hall, 1991.
- [15] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, p. 788, 1999.
- [16] S. K. Mitra and Y. Kuo, *Digital Signal Processing: A Computer-Based Approach*, vol. 2. New York, NY, USA: McGraw-Hill, 2006.
- [17] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, pp. 556–562.
- [18] S. Wild, J. Curry, and A. Dougherty, "Improving non-negative matrix factorizations through structured initialization," *Pattern Recognit.*, vol. 37, no. 11, pp. 2217–2232, Nov. 2004.
- [19] D. Kitamura and N. Ono, "Efficient initialization for nonnegative matrix factorization based on nonnegative independent component analysis," in *Proc. IEEE Int. Workshop Acoustic Signal Enhancement (IWAENC)*, Sep. 2016, pp. 1–5.
- [20] C. Joder, F. Wening, F. Eyben, D. Virette, and B. Schuller, "Real-time speech separation by semi-supervised nonnegative matrix factorization," in *Proc. Int. Conf. Latent Variable Anal. Signal Separat.* Berlin, Germany: Springer-Verlag, 2012, pp. 322–329.
- [21] C.-W. Wu and A. Lerch, "Drum transcription using partially fixed non-negative matrix factorization," in *Proc. 23rd Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2015, pp. 1281–1285.
- [22] M. Fakhry, P. Svaizer, and M. Omologo, "Reverberant audio source separation using partially pre-trained nonnegative matrix factorization," in *Proc. 14th Int. Workshop Acoustic Signal Enhancement (IWAENC)*, Sep. 2014, pp. 273–277.
- [23] T. Virtanen, "Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria," *IEEE Trans. Audio, Speech, Language Process.*, vol. 15, no. 3, pp. 1066–1074, Mar. 2007.
- [24] R. Diamant, "Closed form analysis of the normalized matched filter with a test case for detection of underwater acoustic signals," *IEEE Access*, vol. 4, pp. 8225–8235, 2016.
- [25] C. Févotte, N. Bertin, and J.-L. Durrieu, "Nonnegative matrix factorization with the Itakura–Saito divergence: With application to music analysis," *Neural Comput.*, vol. 21, no. 3, pp. 793–830, 2009.
- [26] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *J. Mach. Learn. Res.*, vol. 5, no. 9, pp. 1457–1469, 2004.



**GEUNHWAN KIM** was born in Daegu, South Korea, in 1992. He received the B.S. and M.S. degrees in electronics engineering from Kyungpook National University, Daegu, in 2015 and 2017, respectively, where he is currently pursuing the Ph.D. degree in electrical and electronics engineering. His research interests include array signal processing, underwater acoustic signal processing, and continuous active sonar.



**SEOKJIN LEE** (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical and computer engineering from Seoul National University, in 2006, 2008, and 2012, respectively. From 2012 to 2014, he was a Senior Research Engineer with LG Electronics. From 2014 to 2018, he was an Assistant Professor with the Department of Electronics Engineering, Kyonggi University, Suwon, Republic of Korea. Since 2018, he has been an Assistant Professor (promoted to an Associate Professor, in 2020) with Electrical and Electronics Engineering, Kyungpook National University, Daegu, Republic of Korea. His research interests include acoustic, sound and music signal processing, array signal processing, and blind source separation.

• • •