RESEARCH   **Open Access**

Reverberation suppression using non-negative matrix factorization to detect low-Doppler target with continuous wave active sonar

Seokjin Lee[1] and Jun-seok Lim[2*]

Abstract

In active sonar systems, the detection of echo from targets can deteriorate due to reverberation. Detection becomes more difficult if targets have low-Doppler frequency and are located near the reverberation band, especially in an environment with low signal-to-reverberation ratio. In this paper, we propose an algorithm for the reverberation suppression of continuous wave signals using non-negative matrix factorization. To extract the target echo signal mixed with reverberations, the bases for the target echo and the reverberation are independently defined, and different constraints are applied for their corresponding estimation. We also derive constraints on temporal continuity and temporal length to estimate bases for the target echo. Experiments using simulated reverberations are performed to evaluate the proposed algorithm, and the results show an enhancement in the signal-to-noise ratio by 6–15 dB, as well as in the detection probability at several signal-to-reverberation ratios. Moreover, an experiment is conducted using reverberation measured from an ocean, and the results show that the proposed algorithm can effectively suppress reverberation and enhance detection performance in practical settings.

Keywords: Active sonar, Reverberation, Continuous wave signal, Non-negative matrix factorization

1 Introduction

Active sonar is used to detect a target under water using a reflected ping. However, in addition to reflection from the target, the received signal contains reverberations caused by reflection from scatterers. Conventional target detection algorithms, including the matched filter, are sensitive to reverberations, and extensive research has been conducted to improve detection in this condition [1–11].

Although the reverberation can be addressed by the beamformer, adequate suppression of the reverberation is not achieved. In case of the low-Doppler target, the target echo is located in the mainlobe and sidelobe of the beamformer, so additional algorithm is required to suppress the reverberation [1, 2]. Methods of reverberation mitigation are twofold: transmission pulse design at the

transmitter and signal processing for reverberation suppression at the receiver. Continuous wave (CW) pulses are commonly used active sonar signals used to detect targets because this type of pulse allows the easy location of targets in the range-Doppler resolution but is vulnerable to reverberations [3]. On the contrary, linear frequency-modulated pulses are robust against reverberation but deliver a low-Doppler resolution [4]. Some research has focused on the generation of specific types of pulses, such as the geometric comb [5], triplet-pair comb [4], and other Doppler-sensitive pulses [2, 3], to handle reverberations. Although such pulses can mitigate the effects of reverberations, they cannot suppress them. Moreover, CW pulses and some variants [2] are still widely used in active sonar systems. Consequently, signal processing for suppressing reverberation at the receiver is needed, and some algorithms have been proposed in this vein [6–11].

For instance, algorithms based on whitening with autoregressive (AR) models have been developed [6, 7] for reverberation suppression. However, the effectiveness

*Correspondence: jslim@sejong.ac.kr
[2]Department of Electrical Engineering, Sejong University, Neungdong-ro 209, 05006 Seoul, Republic of Korea
Full list of author information is available at the end of the article

of this approach deteriorates when the Doppler shifts of the reverberation and the target echo are similar. To overcome this limitation, algorithms based on principal component inversion (PCI) have been developed [8–10]. The signal subspace extraction (SSE) algorithm has recently been proposed as an improvement to the PCI-based algorithm by expressing the reverberation model as a sum of higher and lower reverberation echoes [11].

In this paper, we focus on reverberation suppression for the detection of low-Doppler targets. Such targets can be easily confused with clutter in reverberation. The low-Doppler target detection problem occurs more frequently when the receiver moves because the energy of the reverberation spreads over a frequency range proportional to the speed of the receiver [12]. Therefore, we also consider a moving receiver to develop a reverberation-suppression algorithm. Specifically, we use non-negative matrix factorization (NMF), a segmented signal representation [13], to suppress reverberations. While PCI-based algorithms detect the desired signal using only low ranks, we propose using signal characteristics in the time–frequency domain.

NMF decomposes a non-negative matrix into a multiplication of two non-negative matrices [13]. The spectrogram of frequency and the time basis matrices can then be analyzed, as has been shown for sound signals [14]. Iterative algorithms for the NMF based on the multiplicative update rule were first developed for the cost functions of the Euclidean distance and the Kullback–Leibler divergence [15] and were then expanded to other distance measures, such as Itakura–Saito divergence [16] and beta divergence [17]. Several update rules have also been researched, such as the alternating least squares [18] and expectation-maximization [19] algorithms. The application of NMF to music and speech signal processing have been researched actively to analyze music signals [20, 21], separate the desired source signals from the received signals [22, 23], and to denoise speech signals [24, 25]. NMF-based algorithms are not limited to the processing of mono-channel signals and can also deal with multichannel acoustic signals [19, 26]. Although NMF-based algorithms for music and speech signals have been researched, it is challenging to find NMF-based sonar signal processing algorithms, excluding those for sonar image recognition [27]. Both the PCA and NMF algorithms are matrix decomposition techniques, and the NMF algorithm has advantage for analyzing the sparse components of a matrix, which is called "the part-based representation," compared to the PCA algorithm [13]. It is expected that the NMF algorithm is also suitable for finding the target echo because the spectrogram of the target echo is a sparse component of the spectrogram of the entire received signal.

The main contribution of this research article is the design of NMF-based reverberation suppression in continuous wave active sonar, as no research has applied the NMF method to this problem, to the best of the authors' knowledge. To develop the reverberation suppression algorithm, we divide NMF bases into two parts—echo bases and reverberation bases—and design a constraint to discriminate between them. We apply two constraints, on the temporal length and temporal continuity. The former constraint and its estimation algorithm are novel whereas the latter constraint is developed by Virtanen [22].

The remainder of this paper is organized as follows: Section 2 presents the problem statement for this study and describes the time–frequency characteristics of the target echo and reverberation considering a moving receiver. Section 3 details the proposed NMF-based algorithm, and Section 4 describes the results of simulation and measurements using the proposed algorithm. Finally, we draw the conclusions of this article in Section 5.

Notation: We denote vectors and matrices by boldface lowercase and boldface uppercase letters, respectively. Table 1 specifies the symbols and their meanings.
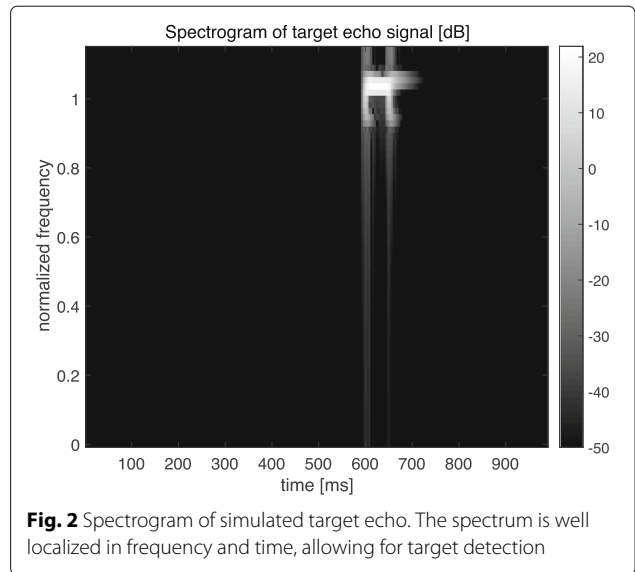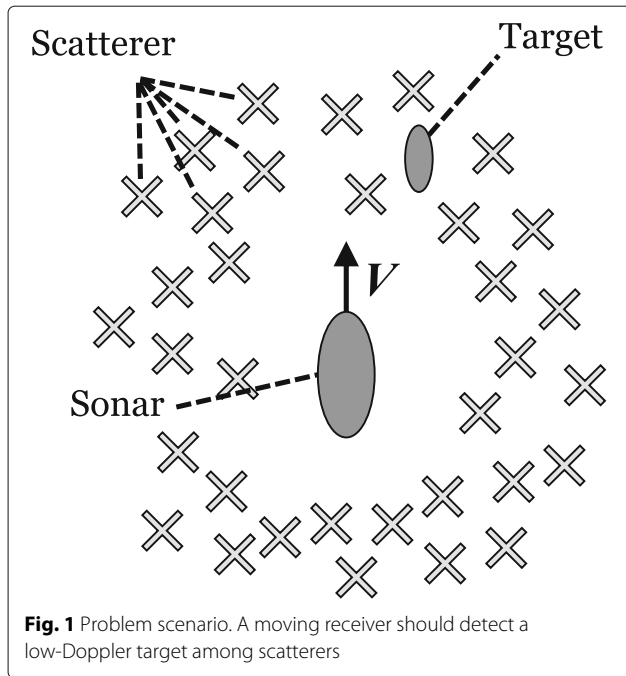
2 Problem statement

We assume that a sonar (receiver) is moving with constant velocity through a field with stationary scattering elements, as shown in Fig. 1. Echo signal $s_e(t)$ from the target is a replica of transmitted ping signal $s_t(t)$ with Doppler shift f_d:

$$s_e(t) = as_t(t - t_d) \exp\left(j2\pi f_d t\right), \qquad (1)$$

Table 1 Notation used in the paper

$\mathbf{A}$	Matrix $\mathbf{A}$
$\mathbf{a}$	Vector $\mathbf{a}$
$\mathbf{A} \in \mathbb{R}_{K \times N}^+$	$K \times N$ matrix whose elements are non-negative real numbers
$\mathbf{A} \in \mathbb{R}_{K \times N}$	$K \times N$ matrix whose elements are real numbers
$\mathbf{1}_{K \times N}$	A $K \times N$ matrix whose elements are all one.
$a_{(k,n)}$ or $[\mathbf{A}]_{(k,n)}$	An element of the matrix $\mathbf{A}$ in k-th row and n-th column
$a^{(t)}$	The updated value of a after t iterations
$\mathbf{A} \otimes \mathbf{B}$	Hadamard (elementwise) multiplication of $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{A}/\mathbf{B}$	Elementwise division of $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{A}^T$	Transpose of $\mathbf{A}$
$\mathbf{A}_{\uparrow d}$	Upward shift of matrix $\mathbf{A}$ by d rows
$\mathbf{A}_{\leftarrow d}$	Left shift of matrix $\mathbf{A}$ by d columns
$\mathbf{A}_{\rightarrow d}$	Right shift of matrix $\mathbf{A}$ by d columns
$\nabla_{\mathbf{A}} C$	The gradient of C with respect to $\mathbf{A}$
$\nabla_{\mathbf{A}}^+ C$	Positive parts of the gradient of C with respect to $\mathbf{A}$
$\nabla_{\mathbf{A}}^- C$	Negative parts of the gradient of C with respect to $\mathbf{A}$

Fig. 1 Problem scenario. A moving receiver should detect a low-Doppler target among scatterers

Fig. 2 Spectrogram of simulated target echo. The spectrum is well localized in frequency and time, allowing for target detection

where a is an attenuation factor, t_d is time delay, and j is the unit imaginary number. The frequency of $s_e(t)$ is given by

$$S_e(f) = aS_t(f - f_d) \exp\left[-j2\pi \left(f - f_d\right) t_d\right]. \quad (2)$$

where $S_t(f)$ is the Fourier transform of $s_t(t)$.

If the transmitted ping signal is a CW signal with a center frequency f_0, the received ping signal is a narrowband signal with frequency $(f_0 + f_d)$. The reverberation signal $s_r(t)$ consists of a large number of replicas from scatterers:

$$s_r(t) = \sum_i a_i s_t(t - t_{d_i}) \exp\left(j2\pi f_{d_i} t\right), \quad (3)$$

with spectrum

$$S_r(f) = \sum_i a_i S_t(f - f_{d_i}) \exp\left[-j2\pi \left(f - f_{d_i}\right) t_{d_i}\right], \quad (4)$$

where a_i and t_{d_i} are an attenuation factor and a time delay of ith scatterer, respectively, and the Doppler shift f_{d_i} is given by
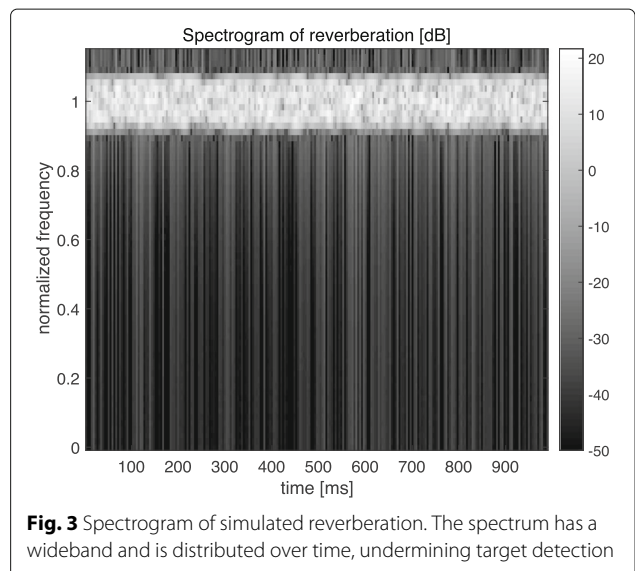
$$f_{d_i} = \frac{2V \cos \psi_i}{c} f_0, \quad (5)$$

where V is the speed of the receiver, c is the speed of sound, and ψ_i is the angle between the direction of motion of the sonar and each scatterer. Equation (5) shows that the spectrum of reverberation has a range of frequency $-\left(\frac{2V}{c}\right)f_0 < f < \left(\frac{2V}{c}\right)f_0$ and, hence, has a wider band than the echo signal when the sonar receiver moves.

Figures 2 and 3 show the spectrograms of the simulated target echo and the reverberation, respectively, for a transmitted CW ping signal. Clearly, the target echo

is a narrowband signal with frequency $f_0 + f_d$, whereas the reverberation is a broadband signal in the range $-\left(\frac{2V}{c}\right)f_0 < f < \left(\frac{2V}{c}\right)f_0$, as verified in (2) and (4), respectively. Further, unlike reverberations, the spectrum of the target echo signal does not randomly fluctuate over time in response to a received signal. Moreover, the target echo signal last for a short period, whereas reverberation is more persistent.

In this study, we extract the target echo signal from a received signal that contains reverberation. To this end, we analyze the spectrogram of the mixed signal by using NMF with distinctive characteristics of the target echo signal, as detailed below.

Fig. 3 Spectrogram of simulated reverberation. The spectrum has a wideband and is distributed over time, undermining target detection

3 Proposed method

3.1 Non-negative matrix factorization

NMF allows for the decomposition of a non-negative matrix into a multiplication of two non-negative matrices by using the following model:

$$\mathbf{V} = \mathbf{WH} + \mathbf{E}, \tag{6}$$

where $\mathbf{V} \in \mathbb{R}^+_{K \times N}$, $\mathbf{W} \in \mathbb{R}^+_{K \times R}$, $\mathbf{H} \in \mathbb{R}^+_{R \times N}$, and $\mathbf{E} \in \mathbb{R}_{K \times N}$. When NMF is applied to acoustic signal processing, $\mathbf{V}$ is the spectrogram of the input signal with K frequency bins and N time frames, $\mathbf{W}$ and $\mathbf{H}$ are the frequency and the time basis matrices, respectively, and R is the number of basis vectors. Then, NMF determines matrices $\mathbf{W}$ and $\mathbf{H}$ from $\mathbf{V}$ by alternating estimations

$$\mathbf{W} \leftarrow \arg\min_{\mathbf{W}} C(\mathbf{V}|\mathbf{WH}), \tag{7}$$

$$\mathbf{H} \leftarrow \arg\min_{\mathbf{H}} C(\mathbf{V}|\mathbf{WH}), \tag{8}$$

where $C(\mathbf{A}|\mathbf{B})$ is a divergence function between $\mathbf{A}$ and $\mathbf{B}$. Lee and Seung [15] introduced the multiplicative update rule for two cost functions—namely, the Euclidean distance and the Kullback–Leibler divergence—and Virtanen [22] subsequently generalized the update rule as

$$\mathbf{W} \leftarrow \mathbf{W} \otimes \frac{\nabla^-_{\mathbf{W}} C(\mathbf{W}, \mathbf{H})}{\nabla^+_{\mathbf{W}} C(\mathbf{W}, \mathbf{H})}, \tag{9}$$

$$\mathbf{H} \leftarrow \mathbf{H} \otimes \frac{\nabla^-_{\mathbf{H}} C(\mathbf{W}, \mathbf{H})}{\nabla^+_{\mathbf{H}} C(\mathbf{W}, \mathbf{H})}, \tag{10}$$

where $\otimes$ and the fractions represent Hadamard multiplication and elementwise divisions, respectively, $\nabla^+_{\mathbf{W}} C(\mathbf{W}, \mathbf{H})$, $\nabla^-_{\mathbf{W}} C(\mathbf{W}, \mathbf{H})$, $\nabla^+_{\mathbf{H}} C(\mathbf{W}, \mathbf{H})$, and $\nabla^-_{\mathbf{H}} C(\mathbf{W}, \mathbf{H})$ are elementwise non-negative terms that are part of gradients

$$\nabla_{\mathbf{W}} C(\mathbf{W}, \mathbf{H}) = \nabla^+_{\mathbf{W}} C(\mathbf{W}, \mathbf{H}) - \nabla^-_{\mathbf{W}} C(\mathbf{W}, \mathbf{H}), \tag{11}$$

$$\nabla_{\mathbf{H}} C(\mathbf{W}, \mathbf{H}) = \nabla^+_{\mathbf{H}} C(\mathbf{W}, \mathbf{H}) - \nabla^-_{\mathbf{H}} C(\mathbf{W}, \mathbf{H}), \tag{12}$$

and $C(\mathbf{W}, \mathbf{H})$ is an arbitrary cost function containing a divergence function and additional constraints.

NMF decomposes the spectrogram of the input signal into several basis components in acoustic signal processing, with the number of components usually larger than that of the sources. Recently proposed algorithms apply different constraints to each basis group to handle the basis components, depending on the source signal, as illustrated in Fig. 4.

3.2 Basis estimation

As shown in Figs. 2 and 3, the target echo signal has distinctive characteristics due to reverberation. The target echo signal is as follows:

- Is a frequency-shifted replica of the transmitted ping
- Has continuous values along the time axis (unlike those of reverberations, which randomly fluctuate)
- Is short and limited in the time domain (unlike reverberations, which are persistent)

To estimate each basis group separately, consider that the frequency and time basis matrices consist of the echo and the reverberation basis groups, respectively, as

$$\mathbf{W} = \begin{bmatrix} \mathbf{W_P} & \vdots & \mathbf{W_R} \end{bmatrix}, \tag{13}$$

$$\mathbf{H} = \begin{bmatrix} \mathbf{H_P}^T & \vdots & \mathbf{H_R}^T \end{bmatrix}^T, \tag{14}$$

where $\mathbf{W_P} \in \mathbb{R}^+_{K \times R_P}$ and $\mathbf{H_P} \in \mathbb{R}^+_{R_P \times N}$ are the frequency and time bases of the echo basis group, $\mathbf{W_R} \in \mathbb{R}^+_{K \times R_R}$ and $\mathbf{H_R} \in \mathbb{R}^+_{R_R \times N}$ are those of the reverberation basis group, respectively, and R_P and R_R are numbers of echo and reverberation bases, respectively. Then, each basis group should be estimated separately as described below.

Applying Eqs. (13) and (14) into Eq. (6), matrix $\mathbf{V}$ can be expressed as

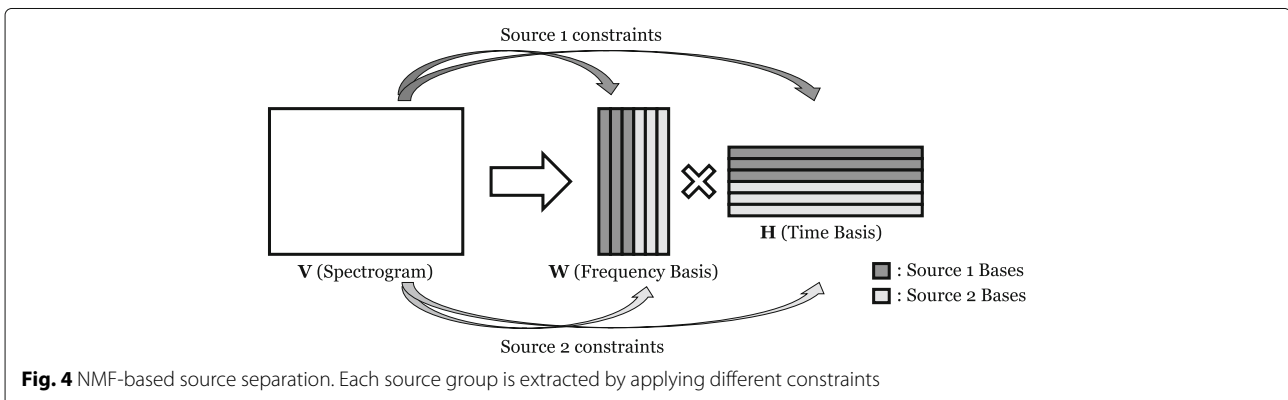$$\mathbf{V} = \mathbf{W_P H_P} + \mathbf{W_R H_R} + \mathbf{E}. \tag{15}$$

Fig. 4 NMF-based source separation. Each source group is extracted by applying different constraints

If we define the target echo portion as $\mathbf{V_P} = \mathbf{W_P H_P}$ and the reverberation portion as $\mathbf{V_R} = \mathbf{W_R H_R}$ in the magnitude spectrogram, Eq. (15) becomes

$$\mathbf{V} = \mathbf{V_P} + \mathbf{V_R} + \mathbf{E}. \tag{16}$$

Therefore, the strategy of dividing into basis groups can help identify the energy contribution of each group to the spectrogram, if we apply appropriate constraints. This strategy enables easy application of prior knowledge of frequency and temporal structures, so it is widely used in current research for source separation and speech signal denoising problems [28].

3.2.1 Echo bases

As the target echo signal has a frequency similar in structure to the Doppler-shifted transmitted ping, the frequency basis matrix consists of Doppler-shifted replicas of frequency basis $\mathbf{w_T}$ of the transmitted ping.

Frequency basis $\mathbf{w_T} \in \mathbb{R}_{K \times 1}^+$ and time basis $\mathbf{h_T} \in \mathbb{R}_{1 \times N}^+$ are modeled by rank–one NMF as $\mathbf{V_T} \approx \mathbf{w_T h_T}$, where $\mathbf{V_T} \in \mathbb{R}_{K \times N}^+$ is the spectrogram of transmitted ping $s_t(t)$. The cost function is defined by the Kullback–Leibler divergence between $\mathbf{V_T}$ and $\mathbf{w_T h_T}$, and $\mathbf{w_T}$ and $\mathbf{h_T}$ can be iteratively estimated as [15]

$$\mathbf{w_T} \leftarrow \mathbf{w_T} \otimes \frac{[\mathbf{V_T}/(\mathbf{w_T h_T})]\,\mathbf{h_T}^T}{\mathbf{1}_{K \times N}\mathbf{h_T}^T}, \tag{17}$$

$$\mathbf{h_T} \leftarrow \mathbf{h_T} \otimes \frac{\mathbf{w_T}^T\,[\mathbf{V_T}/(\mathbf{w_T h_T})]}{\mathbf{w_T}^T \mathbf{1}_{K \times N}}. \tag{18}$$

where $\mathbf{1}_{K \times N}$ denotes a $K \times N$ matrix whose elements are all one. Following convergence, the echo frequency-basis matrix $\mathbf{W_P}$ consists of frequency-shifted replicas of frequency basis $\mathbf{w_T}$ as follows:

$$\mathbf{W_P} = \left[\mathbf{w}_{\mathbf{T},\uparrow -D} \cdots \mathbf{w}_{\mathbf{T},\uparrow 0} \cdots \mathbf{w}_{\mathbf{T},\uparrow D}\right], \tag{19}$$

where $\mathbf{w}_{\mathbf{T},\uparrow d}$ is a d-bin-shifted version of $\mathbf{w_T}$, and D is the maximum number of bins of the Doppler shift to be observed. As shown in Eq. (19), the number of echo bases R_P is $2D + 1$. Equations (17)–(19) depend only on the transmitted ping and not on the received signal. Hence, the iterative process can be carried out initially and $\mathbf{W_P}$ is fixed during runtime.

The time basis matrix of the echo signal is estimated using constraints on temporal continuity and temporal length limitation (TLL) to utilize the second and third characteristics of the echo signal, i.e., time continuity and limited duration, respectively. The cost function consists of error terms

$$C(\mathbf{W_P}, \mathbf{H_P}) = C_E(\mathbf{W_P}, \mathbf{H_P}) + \alpha C_T(\mathbf{H_P}) + \beta C_L(\mathbf{H_P}), \tag{20}$$

where $C_E(\mathbf{W_P}, \mathbf{H_P})$, $C_T(\mathbf{H_P})$, and $C_L(\mathbf{H_P})$ are the costs of the reconstruction error, temporal continuity, and TLL,

respectively. α and β are the weights of the costs of temporal continuity and TLL, respectively. The gradient of total cost is the weighted sum of the gradients of each cost:

$$\nabla_{\mathbf{H_P}} C = \nabla_{\mathbf{H_P}} C_E + \alpha \nabla_{\mathbf{H_P}} C_T + \beta \nabla_{\mathbf{H_P}} C_L, \tag{21}$$

where we omit the parameters of the cost functions for convenience. Assuming that the derivative of each cost function can be expressed as a sum of positive and negative terms, the total cost gradient is given by

$$\nabla_{\mathbf{H_P}} C = \left(\nabla_{\mathbf{H_P}}^+ C_E - \nabla_{\mathbf{H_P}}^- C_E\right) + \alpha \left(\nabla_{\mathbf{H_P}}^+ C_T - \nabla_{\mathbf{H_P}}^- C_T\right)$$
$$+ \beta \left(\nabla_{\mathbf{H_P}}^+ C_L - \nabla_{\mathbf{H_P}}^- C_L\right)$$
$$= \nabla_{\mathbf{H_P}}^+ C - \nabla_{\mathbf{H_P}}^- C, \tag{22}$$

where

$$\nabla_{\mathbf{H_P}}^+ C = \left(\nabla_{\mathbf{H_P}}^+ C_E + \alpha \nabla_{\mathbf{H_P}}^+ C_T + \beta \nabla_{\mathbf{H_P}}^+ C_L\right), \tag{23}$$

$$\nabla_{\mathbf{H_P}}^- C = \left(\nabla_{\mathbf{H_P}}^- C_E + \alpha \nabla_{\mathbf{H_P}}^- C_T + \beta \nabla_{\mathbf{H_P}}^- C_L\right), \tag{24}$$

and $\mathbf{H_P}$ is iteratively estimated from (10) as

$$\mathbf{H_P} \leftarrow \mathbf{H_P} \otimes \frac{\nabla_{\mathbf{H_P}}^- C}{\nabla_{\mathbf{H_P}}^+ C}. \tag{25}$$

Reconstruction error C_E is defined by the Kullback–Leibler divergence, which is often used in NMF for source separation [15]:

$$C_E(\mathbf{V}|\mathbf{WH}) = \sum_{k,n} \left(v_{(k,n)} \log \frac{v_{(k,n)}}{[\mathbf{WH}]_{(k,n)}} - v_{(k,n)} + [\mathbf{WH}]_{(k,n)}\right), \tag{26}$$

where $v_{(k,n)}$ and $[\mathbf{WH}]_{(k,n)}$ are elements of $\mathbf{V}$ and $\mathbf{WH}$, respectively, in the kth row ($1 \le k \le K$) and the nth column ($1 \le n \le N$). In the proposed algorithm, $\mathbf{V}$ is the spectrogram of received signal $s(t) = s_e(t) + s_r(t)$. Therefore, the positive and negative terms of the cost function derivative with respect to $\mathbf{H_P}$ are given by [22]

$$\nabla_{\mathbf{H_P}}^+ C_E = \mathbf{W_P}^T \mathbf{1}_{K \times N}, \tag{27}$$

$$\nabla_{\mathbf{H_P}}^- C_E = \mathbf{W_P}^T \frac{\mathbf{V}}{\mathbf{WH}}. \tag{28}$$

The temporal continuity constraints are defined as [22]

$$C_T = \sum_{r=1}^R \frac{\sum_{n=2}^N \left(h_{p,(r,n)} - h_{p,(r,n-1)}\right)^2}{\frac{1}{N}\sum_{n=1}^N h_{p,(r,n)}^2}, \tag{29}$$

where $h_{p,(r,n)}$ is the element of $\mathbf{H_P}$ in the rth row and nth column. The gradient of C_T with respect to $\mathbf{H_P}$ is derived as [22]

$$\nabla_{\mathbf{H_P}} C_T = 2N \frac{2\mathbf{H_P} - \mathbf{H}_{\mathbf{P} \rightarrow 1} - \mathbf{H}_{\mathbf{P} \leftarrow 1}}{\mathbf{H_P}^2 \mathbf{1}_{N \times N}}$$
$$- N \frac{2\mathbf{H_P} \otimes \left[\{(\mathbf{H_P} - \mathbf{H}_{\mathbf{P} \rightarrow 1})^2\}\mathbf{1}_{N \times N}\right]}{\left\{\mathbf{H_P}^2 \mathbf{1}_{N \times N}\right\}^2}, \tag{30}$$

where $\mathbf{H_{P\leftarrow 1}}$ and $\mathbf{H_{P\rightarrow 1}}$ are left and right shifts of $\mathbf{H_P}$ by one, respectively, $\mathbf{1}_{N\times N}$ denotes a $N \times N$ matrix whose elements are all one, and $\mathbf{H_P}^2$ means an elementwise square of $\mathbf{H_P}$. By grouping the positive and negative terms, $\nabla_{\mathbf{H_P}}^+ C_T$ and $\nabla_{\mathbf{H_P}}^- C_T$ are expressed as [22]

$$\nabla_{\mathbf{H_P}}^+ C_T = \frac{4N\mathbf{H_P}}{\mathbf{H_P}^2 \mathbf{1}_{N\times N}}, \tag{31}$$

$$\nabla_{\mathbf{H_P}}^- C_T = 2N\frac{\mathbf{H_{P\rightarrow 1}} - \mathbf{H_{P\leftarrow 1}}}{\mathbf{H_P}^2 \mathbf{1}_{N\times N}}$$
$$+ \frac{2N\mathbf{H_P}\otimes\left[\{(\mathbf{H_P}-\mathbf{H_{P\rightarrow 1}})^2\}\mathbf{1}_{N\times N}\right]}{\left\{\mathbf{H_P}^2\mathbf{1}_{N\times N}\right\}^2}. \tag{32}$$

The TLL constraint is defined as

$$C_L = \sum_{r=1}^{R}\left[1 - \sum_{m=n}^{n+l_n-1}\hat{h}_{p,(r,m)}\right], \tag{33}$$

where l_n is the expected length of a ping. The maximum value indicator function $\hat{h}_{p,(r,n)}$ is given as

$$\hat{h}_{p,(r,n)} = \frac{e^{\bar{h}_{p,(r,n)}}}{\sum_{m=1}^{N}e^{\bar{h}_{p,(r,m)}}}, \tag{34}$$

where $\bar{h}_{p,(r,n)}$ is the moving sum of each time basis signal calculated as

$$\bar{h}_{p,(r,n)} = \sum_{m=n-l_n+1}^{n}h_{p,(r,m)}. \tag{35}$$

The individual elements from the positive part $\nabla_{\mathbf{H_P}}^+ C_L$ and negative part $\nabla_{\mathbf{H_P}}^- C_L$ are given by (see Appendix 1)

$$\left[\nabla_{\mathbf{H_P}}^+ C_L\right]_{(r,n)} = \sum_{m=n}^{n+l_n-1}\left(\frac{e^{\bar{h}_{p,(r,m)}}}{\sum_{i=1}^{N}e^{\bar{h}_{p,(r,i)}}}\right)^2, \tag{36}$$

$$\left[\nabla_{\mathbf{H_P}}^- C_L\right]_{(r,n)} = \sum_{m=n}^{n+l_n-1}\frac{e^{\bar{h}_{p,(r,m)}}}{\sum_{i=1}^{N}e^{\bar{h}_{p,(r,i)}}}, \tag{37}$$

where $[\mathbf{A}]_{(r,n)}$ is the (r,n) element of matrix $\mathbf{A}$.

3.2.2   Reverberation bases
We consider the basis matrices of reverberation as components of the signal that are longer and more fluctuating than the target echo signal. Therefore, the frequency and time basis matrices can be estimated by the multiplicative update rule without additional constraints as

$$\mathbf{H_R} \leftarrow \mathbf{H_R} \otimes \frac{\mathbf{W_R}^T\left[\mathbf{V}/\left(\mathbf{WH}\right)\right]}{\mathbf{W_R}^T\mathbf{1}_{K\times N}} \tag{38}$$

and

$$\mathbf{W_R} \leftarrow \mathbf{W_R} \otimes \frac{\left[\mathbf{V}/\left(\mathbf{WH}\right)\right]\mathbf{H_R}^T}{\mathbf{1}_{K\times N}\mathbf{H_R}^T}. \tag{39}$$

3.2.3   Reconstruction of the target echo signal
Following the iterative estimation of the basis matrices for the echo and reverberation, (25), (38), and (39) converge, and the estimated signal of the target echo can be reconstructed from its basis matrices. The spectrogram of the reconstructed echo signal is calculated from these matrices as

$$\mathbf{V}_{\text{out}} = \mathbf{W_P}\mathbf{H_P}. \tag{40}$$

To construct a complex spectrogram, phase information is needed. In several NMF-based algorithms, either the phase from the original spectrogram is used [22] or an approach based on the Wiener filter is applied to the original spectrogram [19]. We observed a similar performance by both methods, but prefer the phase of the original spectrogram because it allows the easy construction of the complex spectrogram. The estimated target echo signal $\hat{s}_e(t)$ is then obtained by the inverse short-time Fourier transform.

The proposed algorithm is summarized in Table 2, where $\mathbf{H_P}$, $\mathbf{W_R}$, and $\mathbf{H_R}$ are initialized using non-negative random values and iteratively estimated, whereas $\mathbf{W_P}$ is initialized and fixed by the frequency-shifted replicas of the transmitted ping spectrum $\mathbf{w_T}$.

3.2.4   Considerations on the convergence of the algorithm
Lee and Seung [15] proved that the cost function is non-increasing when the multiplicative update processes (Eqs. (9) and (10)) minimize the Kullback–Leibler divergence. Convergence is proved by designing an auxiliary function for the objective functions. The detailed proof can be found in [15, 17] and is briefly reviewed in Appendix 2.

It is difficult to show that convergence is guaranteed when the NMF contains additional constraints because finding an appropriate auxiliary function is difficult. However, we expect that the iterative algorithm converges well if α and β are small, owing to the non-increasing property of the multiplicative update rules for the Kullback–Leibler divergence. Instead of a theoretical analysis of the

Table 2 Proposed algorithm

$\mathbf{H_P}$, $\mathbf{W_R}$, and $\mathbf{H_R}$ are randomly initialized with non-negative values, and $\mathbf{W_P}$ is initialized using (17), (18), and (19). Following initialization, the following process is iteratively performed until convergence:

1. $\nabla_{\mathbf{H_P}}^+ C_E$, $\nabla_{\mathbf{H_P}}^- C_E$, $\nabla_{\mathbf{H_P}}^+ C_T$, $\nabla_{\mathbf{H_P}}^- C_T$, $\nabla_{\mathbf{H_P}}^+ C_L$, and $\nabla_{\mathbf{H_P}}^- C_L$ are calculated using (27), (28), (31), (32), (36), and (37), respectively;
2. $\nabla_{\mathbf{H_P}}^+ C$ and $\nabla_{\mathbf{H_P}}^- C$ are calculated using (23) and (24), respectively, and $\mathbf{H_P}$ is updated using (25);
3. $\mathbf{H_R}$ is updated using (38);
4. $\mathbf{H}$ is updated with $\mathbf{H_P}$ and $\mathbf{H_R}$ using (14);
5. $\mathbf{W_R}$ is updated using (39);
6. $\mathbf{W}$ is updated with $\mathbf{W_R}$ using (13).

Following convergence, the resulting spectrogram $\mathbf{V}_{\text{out}}$ is calculated using (40).

convergence condition, we check the convergence of the objective function with varying α and β in the simulation experiments presented in Section 4. We observe that the objective function converges well in most cases except for those with very large values of α and β, such as 10^4.

4 Results and discussion

4.1 Simulation with synthesized reverberation

To evaluate the proposed algorithm, we perform simulations with synthesized reverberations. Specifically, a sonar system transmits CW ping signals of 50 ms and moves forward with speed V such that $\frac{2V}{c} = 0.07$. The reverberation is synthesized using the non-Rayleigh underwater reverberation model [29]. Moreover, the target echo signal is received at 600 ms and the normalized Doppler (f_d/f_0) of the target echo is 0.036 for this simulation.

To apply the proposed algorithm, we use the short-time Fourier transform for the received signal with a 16-ms Hamming window, a 75% overlap, and 128 Fourier transform points. The number of basis vectors for the echo R_P and reverberation R_R of the proposed algorithm are set to 19 and 60, respectively, and the weight factors for

temporal continuity and TLL constraints, α and β, are set to 1.0 and 10.0, respectively. The expected ping length l_n is set to 50 ms. The proposed algorithm is compared with the PCI [9], autoregressive (AR) pre-whitening [10], and SSE [11] algorithms. The length of the time frame and the order of the AR model are set to 64 ms and 20, respectively, and these values are selected as the optimal ones through trial and error. The thresholds for eigenvalues of the PCI and SSE algorithms are set to optimal values, assuming that the power of the reverberation is known.

Figures 5a–f and 6a–f show examples of the basis matrices of the echo estimated by the proposed algorithm with an input signal-to-reverberation ratio (SRR) of −12 dB, from $r = 7$ to $r = 12$, respectively. As shown in Fig. 5, the ninth frequency basis vector correspond to the frequency structure of the transmitted CW ping signal, whereas the other basis vectors are shifted from the CW ping in d bins. Figure 6 shows that the 11th time basis vector has large values between 600 and 650 ms, whereas the other basis vectors have relatively small values. As the difference in frequency between adjacent bins is $0.018f_0$ in this experiment, the 11th frequency and time basis vectors
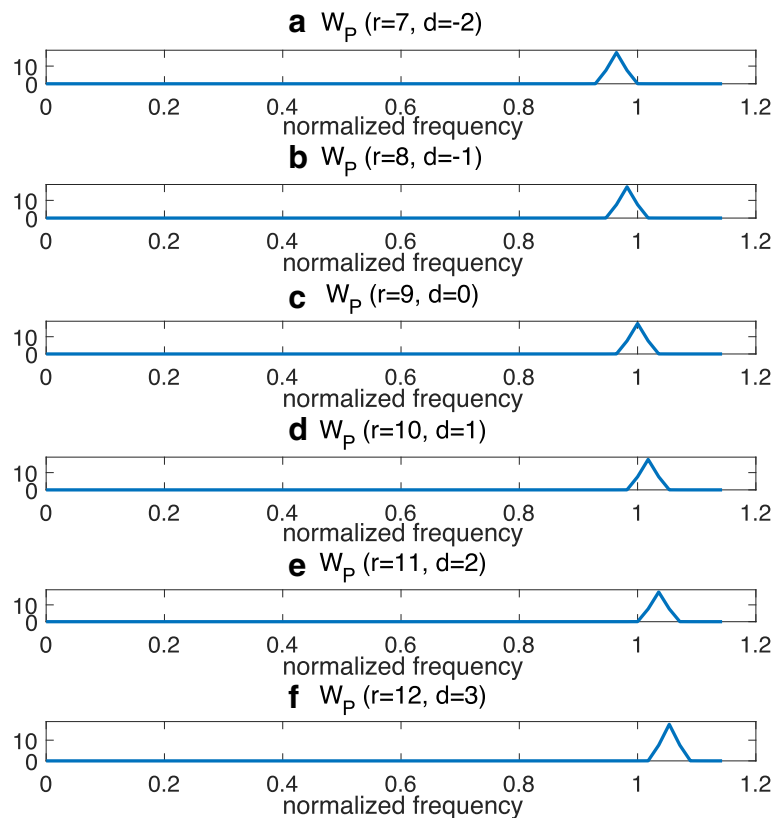

Fig. 5 Echo frequency basis vectors at **a** $r = 7$, **b** $r = 8$, **c** $r = 9$, **d** $r = 10$, **e** $r = 11$, and **f** $r = 12$ (input SRR = −12 dB). The *x*-axis of each graph denotes frequency normalized by the CW ping frequency
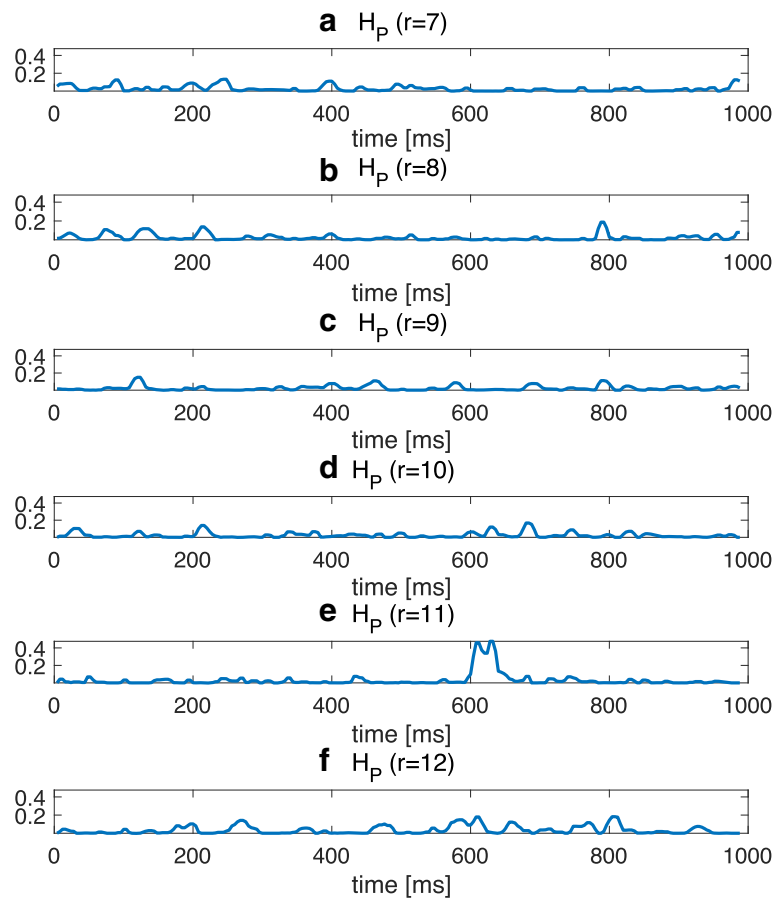
Fig. 6 Echo time basis vectors at **a** $r = 7$, **b** $r = 8$, **c** $r = 9$, **d** $r = 10$, **e** $r = 11$, and **f** $r = 12$ (input SRR $= -12$ dB)

corresponded to the target echo. Thus, the graphs show that the proposed constraints are suitable to find the target echo signal.

Figure 7 shows the waveforms of the ideal target echo (ground truth, Fig. 7a), the received signal comprising the ideal target echo and reverberation (Fig. 7b), the output signal obtained using the proposed algorithm (Fig. 7c), the SSE algorithm (Fig. 7d), the AR pre-whitening algorithm (Fig. 7e), and the PCI algorithm (Fig. 7f). For the algorithms, we consider the input signal shown in Fig. 7b.
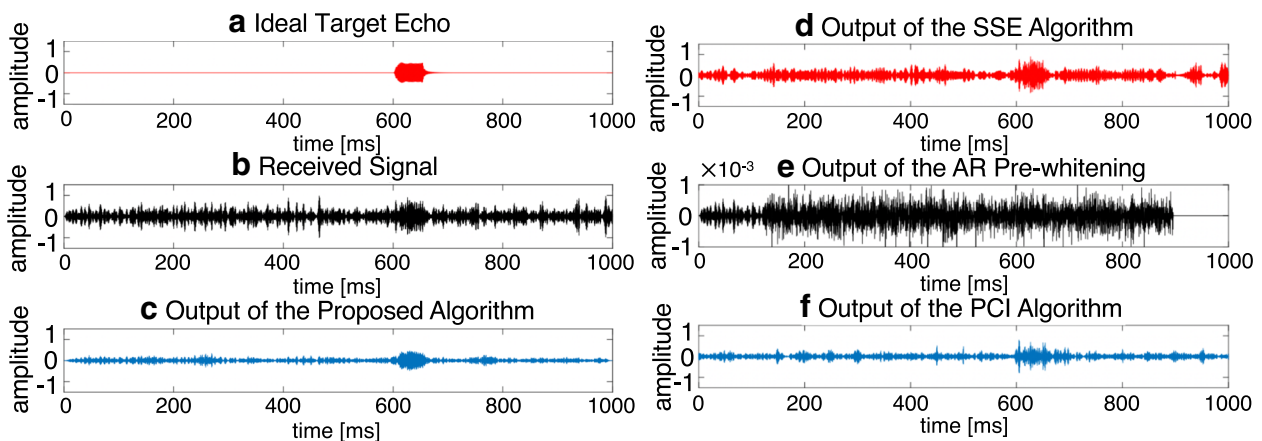

Fig. 7 Waveforms of **a** ideal target echo signal, **b** received signal, **c** output signal using the proposed algorithm, **d** output by the SSE algorithm, **e** output by the AR pre-whitening, and **f** output by the PCI algorithm (input SRR $= -12$ dB)

Reverberations with amplitudes similar to that of the target echo signal can be seen before applying the algorithm, but the reverberations are removed following the application of the proposed algorithm. Although the SSE and PCI algorithms can also suppress reverberation, parts of the reverberation signals persist upon application of these algorithms.

To evaluate the proposed algorithm for various input SRRs, we measure the output signal-to-noise ratio (SNR) using Monte Carlo simulations with 100 iterations per SRR ranging from -20 to -6 dB, with increments of 2 dB. The SNR is calculated as

$$\text{SNR} = \frac{\sum_n |s_e(n)|^2}{\sum_n |s_e(n) - \hat{s_e}(n)|^2}, \tag{41}$$

where $s_e(n)$ is the echo signal of the ideal target and $\hat{s_e}(n)$ is the output signal obtained after applying the algorithm. Figure 8 shows the SNR of the proposed, the SSE, and the PCI algorithms. The proposed algorithm achieves an SNR gain of 6 to 15 dB compared to the input SRR and approximately 5 to 7 dB compared to the SSE and PCI algorithms. The SNR results of the AR pre-whitening algorithm could not be obtained because the output amplitude of the algorithm is too small. Thus, it is not appropriate to evaluate AR pre-whitening algorithm using the SNR values.

Given that reverberation suppression aims to improve target detection, we calculate the range-Doppler by applying a matched filter to the outputs of the algorithms. To calculate the map, we use the block-normalized matched filter [9, 11], with the Doppler-shifted transmitted ping signal being

Fig. 8 SNR retrieved from the proposed, SSE, and PCI algorithms for various input SRRs. The vertical error bars indicate the standard deviation

$$L(i,f_d) = \frac{\sum_{n=0}^{N-1} s_{f_d}^*(n)x_i(n)}{(1/2N)\sum_{n=0}^{N-1} |s_{f_d}(n)|^2 \sum_{n=0}^{N-1} |x_i(n)|^2}, \tag{42}$$

where N is the length of a block, $s_{f_d}(n)$ is a Doppler-shifted transmitted ping signal with Doppler frequency f_d, and $x_i(n)$ is the ith block of the received signal. Figure 9a–e show the results of applying the matched filter only and applying the matched filter with the proposed algorithm, the SSE algorithm, the AR pre-whitening algorithm, and the PCI algorithm, respectively. Figure 9a shows that the range-Doppler map has a large value at the location of the target echo, but has many false peaks. Figure 9b shows that the SSE algorithm enhances the target echo, but several false peaks still remain, and Fig. 9c shows that the PCI algorithm retrieves some false peaks over several Doppler frequencies. Figure 9d shows that the AR pre-whitening algorithm yields relatively good performance for enhancing the target echo, but several large false peaks persist near the zero-meter range, for certain Doppler frequencies. Figure 9e shows that the proposed algorithm determines the true peak corresponding to the ideal target echo with better intensity than for the other peaks. Therefore, it can reduce false positives when applying a threshold to the matched filter during target detection.

To evaluate detection performance, we calculate the probabilities of detection and false alarm after 1000 Monte Carlo simulations per SRR input. A *positive* is defined as the range-Doppler bin containing the target, whereas a *negative* is defined as a bin with no target. When the value of any bin in the range-Doppler map is greater than a predefined threshold, we define those bins as true positives (TPs) if they are positives, and as false positives (FPs) otherwise. The probabilities of detection and false alarm are respectively defined as

$$\text{Detection prob.} = \frac{\text{Number of TPs}}{\text{Number of positives}}, \tag{43}$$

$$\text{False alarm prob.} = \frac{\text{Number of FPs}}{\text{Number of negatives}}. \tag{44}$$

These probabilities are dependent on the threshold, and we thus calculate them for various thresholds to determine the receiver operating characteristic (ROC) curves as shown in Fig. 10.

Figure 10 shows the ROC results at -12 dB, -14 dB, -16 dB, and -18 dB because the detection performance of the algorithms rapidly changed in this region. The graphs show that the proposed algorithm with matched filter enhances detection performance compared with the AR pre-whitening, PCI, and SSE algorithms. Under certain false alarm conditions (greater than a probability of 30%), the AR pre-whitening algorithm shows slightly higher detection probability than the proposed algorithm, but the difference in the high false alarm conditions are much
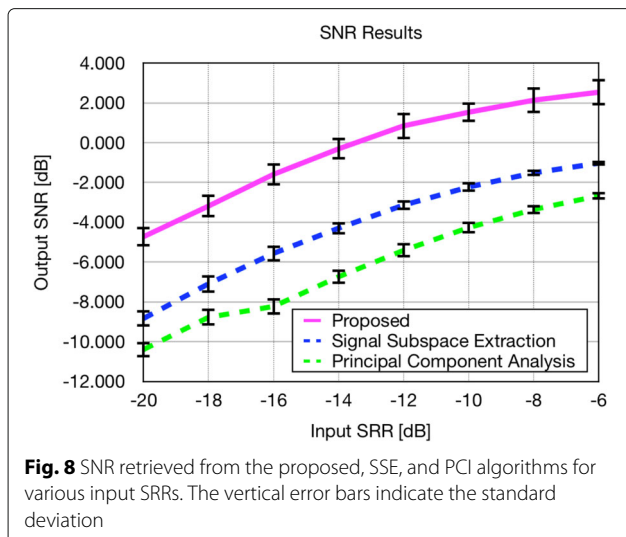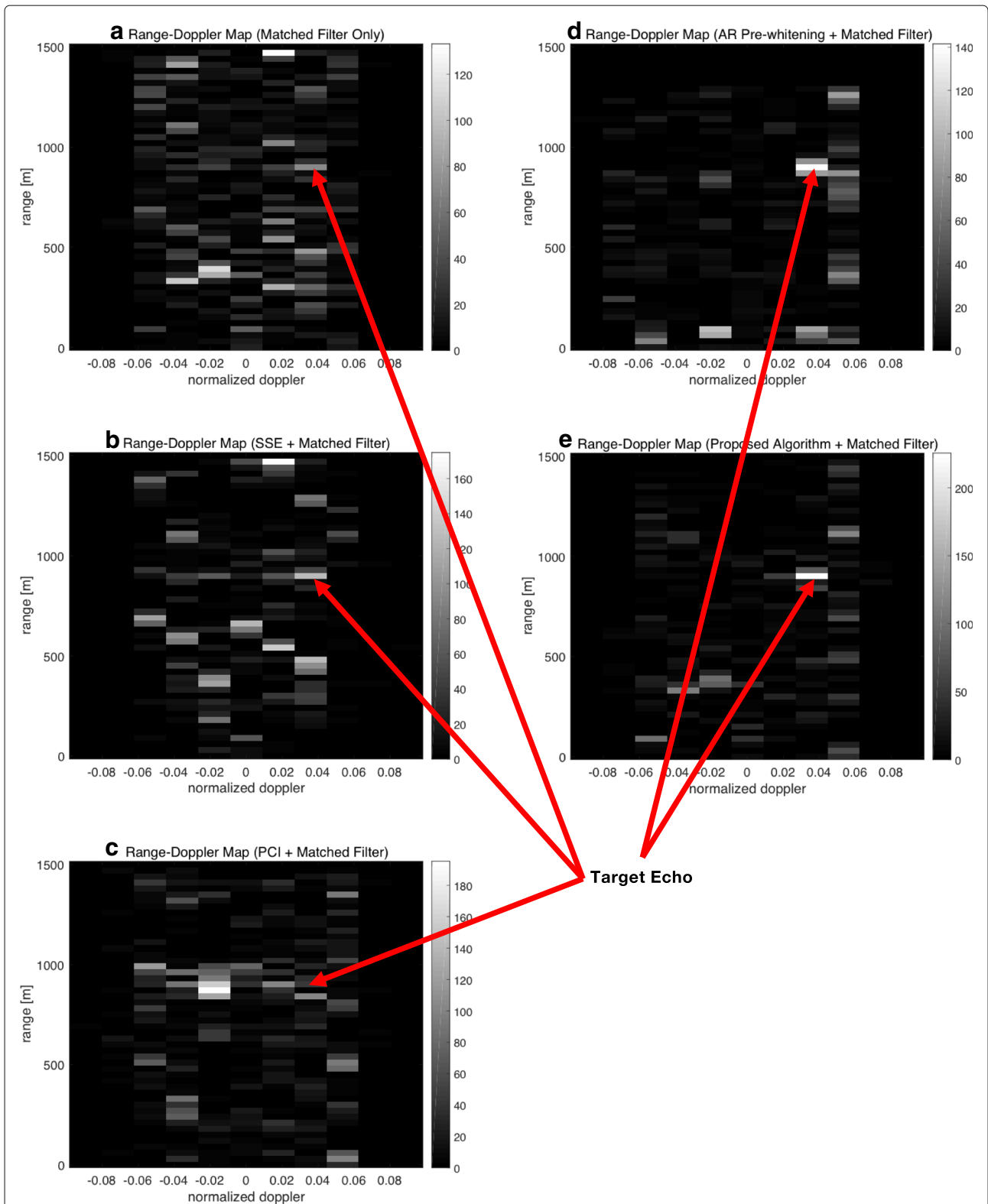
Fig. 9 Range-Doppler maps from simulations (input SRR = −12 dB). Maps of the **a** received signal, **b** output signal from the SSE algorithm, **c** output signal from the PCI algorithm, **d** output signal from the AR pre-whitening, and **e** output signal from the proposed algorithm. Each range-Doppler map is acquired using the block-normalized matched filter [9]
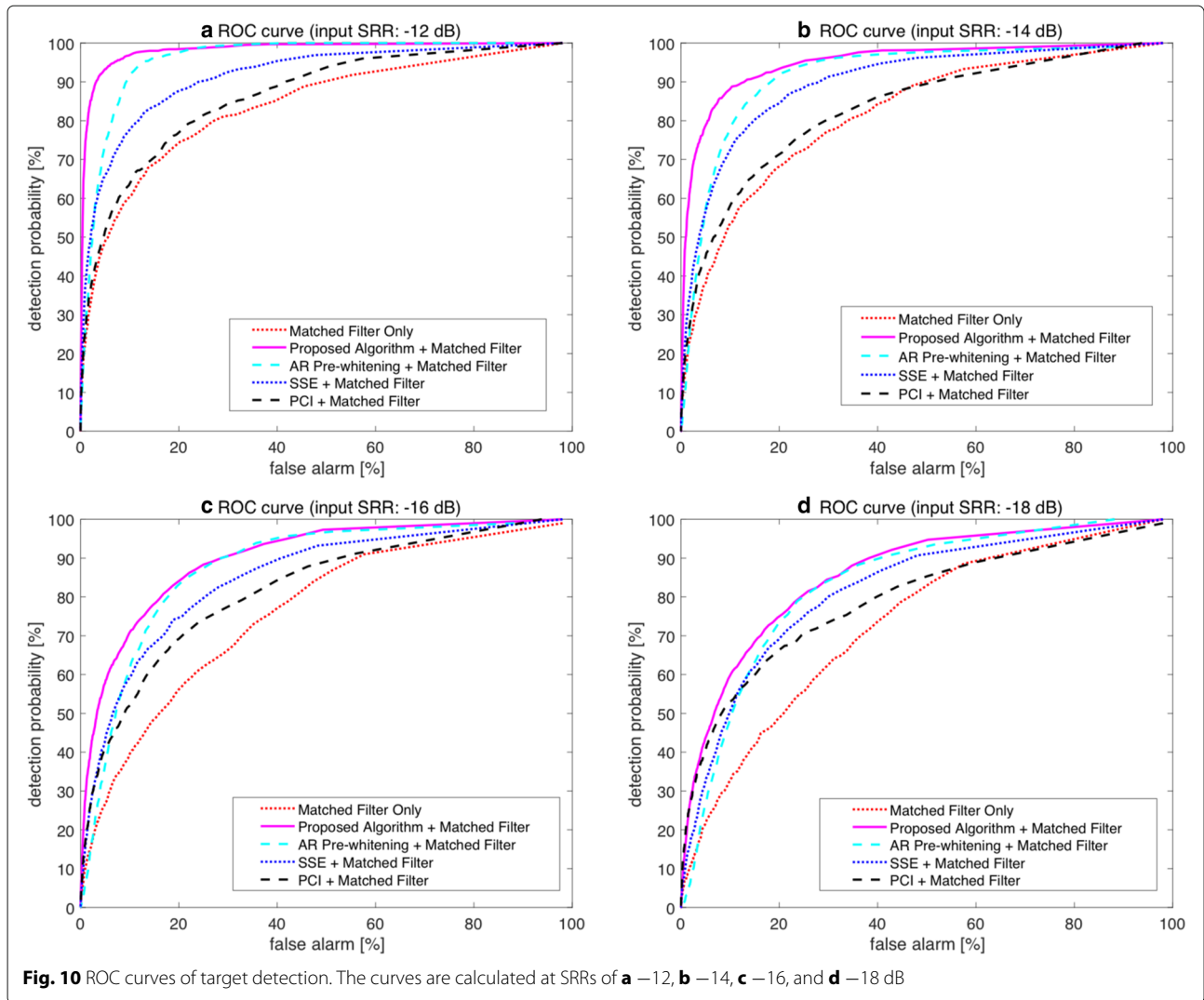
Fig. 10 ROC curves of target detection. The curves are calculated at SRRs of **a** −12, **b** −14, **c** −16, and **d** −18 dB

smaller compared to the difference in the low false alarm conditions. Therefore, the proposed algorithm is considered to be more advantageous than the AR pre-whitening algorithm. Figure 11 shows the detection probability with respect to input SRR conditions at a false alarm probability of 1%, and the results show that the proposed algorithm can significantly enhance detection performance at SRR inputs between −18 dB and −8 dB with a low probability of false alarms.

To verify the convergence of the proposed algorithm, we calculate the value of the objective functions during iterative estimation. Figure 12a and b show the Kullback–Leibler divergence between $\mathbf{V}$ and $\mathbf{WH}$ used to estimate $\mathbf{W_R}$ and $\mathbf{H_R}$ (Eqs. (38) and (39)), respectively, and Fig. 12c and d show the objective function, including the costs of temporal continuity and TLL (Eq. (20)), used to estimate $\mathbf{H_P}$. The curves are calculated by an ensemble averaged

over 1000 Monte Carlo simulations. We check the value of the objective function by varying the SRR input from −20 dB to 0 dB, but the results are barely affected by the input SRR. We thus show the graph corresponding to −10 dB as input SRR only. As shown in Fig. 12a and c, the objective functions do not diverge when the value of α is less than 10^4. Figure 12b and d show that the objective functions converge with β less than 10^4. We test the convergence of the objective functions under several combinations of α and β, and the results show similar behaviors regardless of the value of each parameter.

Because the computational complexities of the SSE and PCI algorithms depend on the singular-value decomposition used, it is challenging to compare the complexities of the algorithms. We thus measure the computation times on a laptop with a 3.5-GHz CPU and 16 GB of memory. All the tested algorithms are implemented in MATLAB,
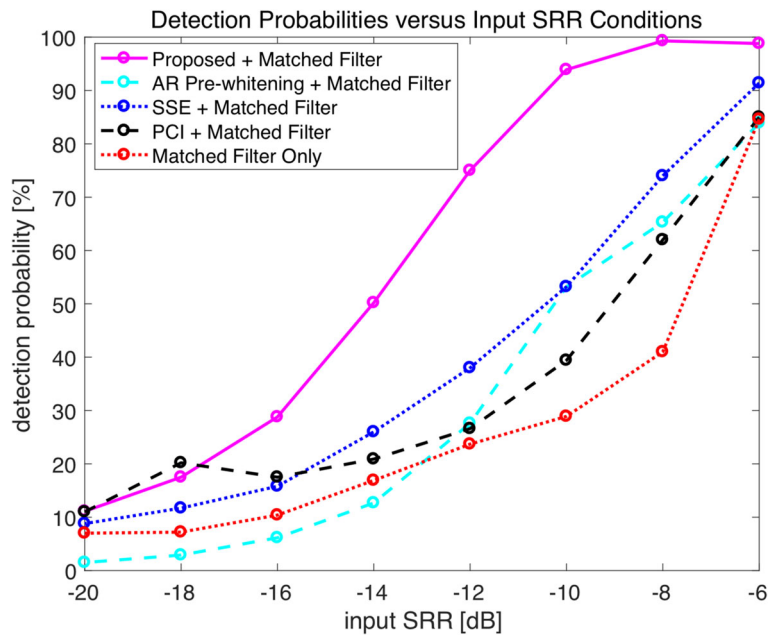
Fig. 11 Detection probabilities versus input SRR conditions at a false alarm probability of 1%
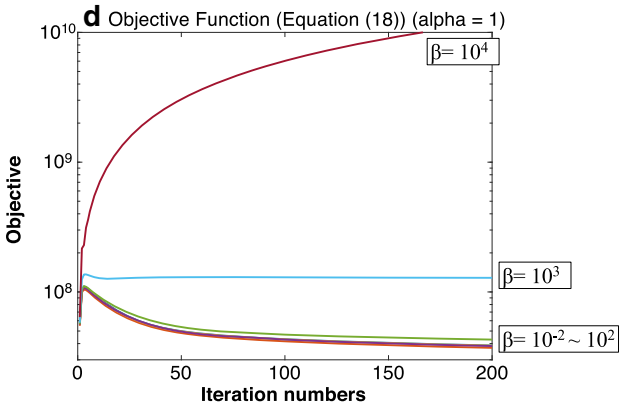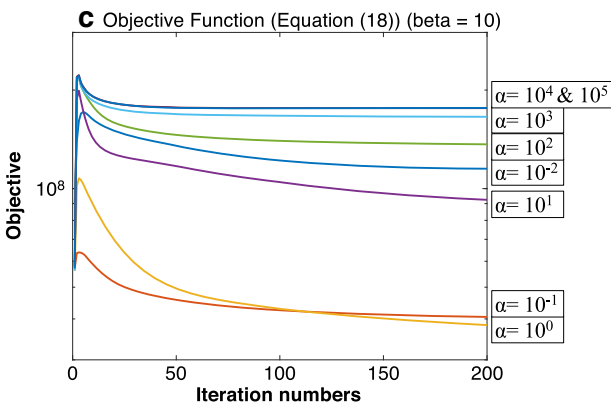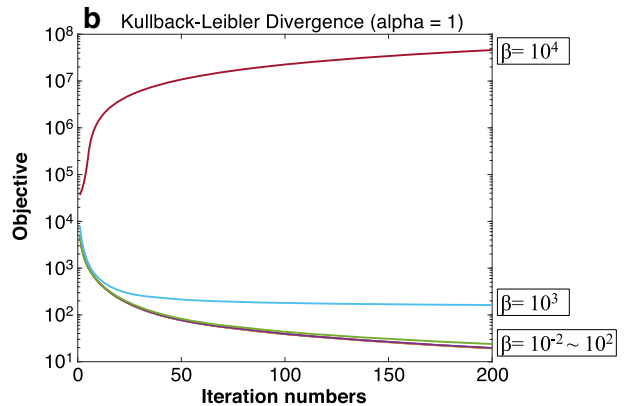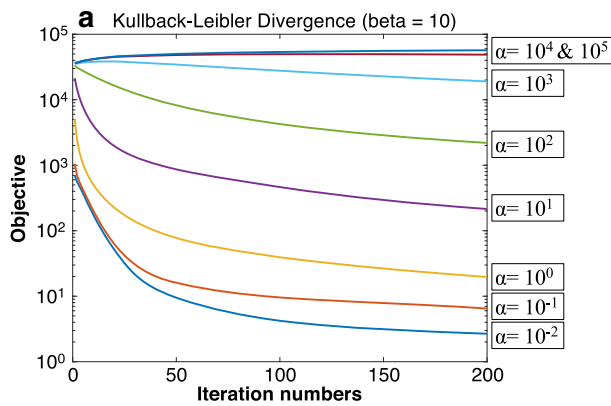

Fig. 12 Objective functions with varying α and β. The values of the objective functions are calculated for **a** the Kullback–Leibler divergence with $\beta = 10$ and varying α, **b** the Kullback–Leibler divergence with $\alpha = 1$ and varying β, **c** Eq. (20) with $\beta = 10$ and varying α, and **d** Eq. (20) with $\alpha = 1$ and varying β

and the SSE and PCI algorithms use the built-in functions for singular-value decomposition. For a 1-s long signal, the proposed algorithm with 200 iterations requires 0.7 s, whereas the AR and SSE (PCI) algorithms require 0.025 and 0.6 s, respectively. The calculation times of the SSE and PCI algorithms are nearly identical. The proposed algorithm thus requires more calculation time than the conventional algorithms.
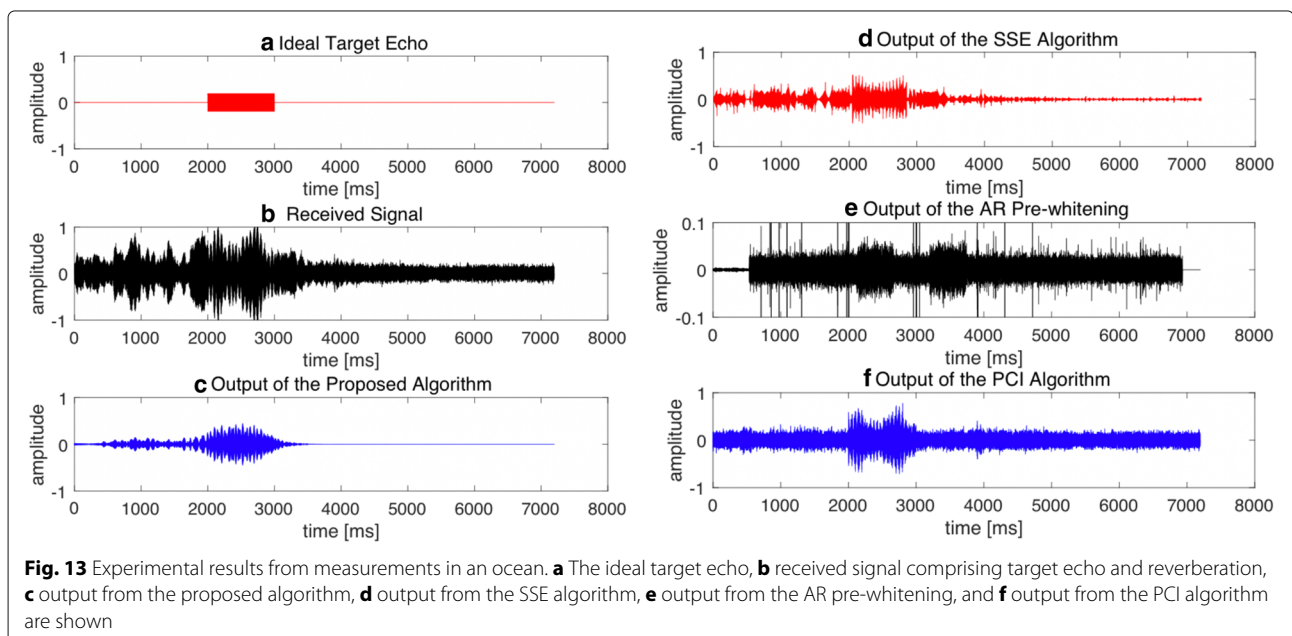
4.2　Simulation with measured reverberation

Having verified the performance of the proposed algorithm through a simulation, we test it on reverberation measurements acquired at the Eastern Sea of Pohang in the Republic of Korea from a 1-s transmitted CW ping. The reverberation of the CW signal is measured by a nested towed array consisting of 48 sensors moving at four knots.
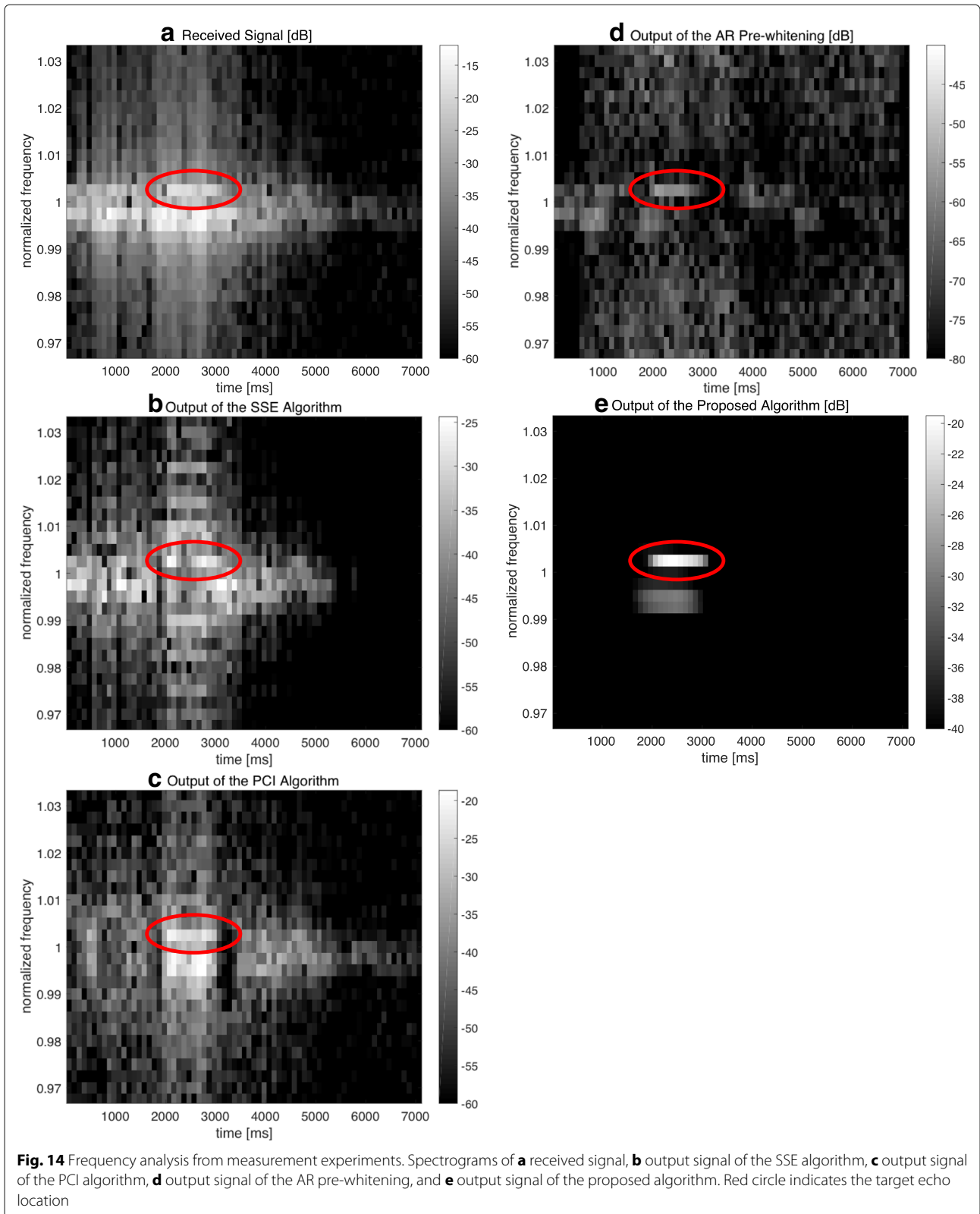
During this experiment, we assume an input SRR of −12 dB, and the target echo is received after 2 s with a normalized Doppler of 1.0025. Its signal is thus interfering with the reverberation. We apply the short-time Fourier transform to the received signal with a 133-ms Hamming window, 75% overlap, and 2048 transform points. As for the simulations with synthesized reverberation, temporal continuity, and TLL weights, α and β are set to 1.0 and 10.0, respectively, and the expected ping length l_n is set to 1 s. The length of the time frame and the order of the AR model are set to 530 ms and 50, respectively, and these values are selected as the optimal ones through trial and error. The thresholds for eigenvalues of the PCI and SSE algorithms are set to optimal values, assuming that the power of the reverberation is known.

Figure 13 shows the waveforms of the ideal target echo (ground truth, Fig. 13a), the received signal comprising the ideal target echo and reverberation (Fig. 13b), the output signal obtained using the proposed algorithm (Fig. 13c), the SSE algorithm (Fig. 13d), the AR prewhitening algorithm (Fig. 13e), and the PCI algorithm (Fig. 13f). It is challenging to identify the target echo from the received signal, whereas it is clearly distinguishable when the proposed algorithm is applied. Although the SSE and PCI algorithms can also suppress reverberation, parts of the reverberation signals persist following its application.

To analyze the proposed algorithm in the frequency domain, we calculate the spectrograms of the received signal and the output signals as shown in Fig. 14. Clearly, the received signal has many undistinguishable components as shown in Fig. 14a, and parts of the reverberation signals persist following the application of SSE and PCI algorithms, as shown in Fig. 14b and c, respectively. The AR pre-whitening algorithm suppresses the target echo as well as the reverberation as shown in Fig. 14d, whereas the target echo signal is clearly visible in the spectrogram of the algorithm's output as shown in Fig. 14e. Therefore, the proposed algorithm can estimate both the Doppler frequency and the temporal location of the target echo.

Figure 15 shows the range-Doppler maps of the received signal without reverberation suppression and the processed signal using the SSE algorithm, PCI algorithm, AR pre-whitening, and the proposed algorithm, respectively. Each range-Doppler map is determined from the block-normalized matched filter, as for the simulation of synthesized reverberation. Figure 15a shows significant

Fig. 13 Experimental results from measurements in an ocean. **a** The ideal target echo, **b** received signal comprising target echo and reverberation, **c** output from the proposed algorithm, **d** output from the SSE algorithm, **e** output from the AR pre-whitening, and **f** output from the PCI algorithm are shown

Fig. 14 Frequency analysis from measurement experiments. Spectrograms of **a** received signal, **b** output signal of the SSE algorithm, **c** output signal of the PCI algorithm, **d** output signal of the AR pre-whitening, and **e** output signal of the proposed algorithm. Red circle indicates the target echo location
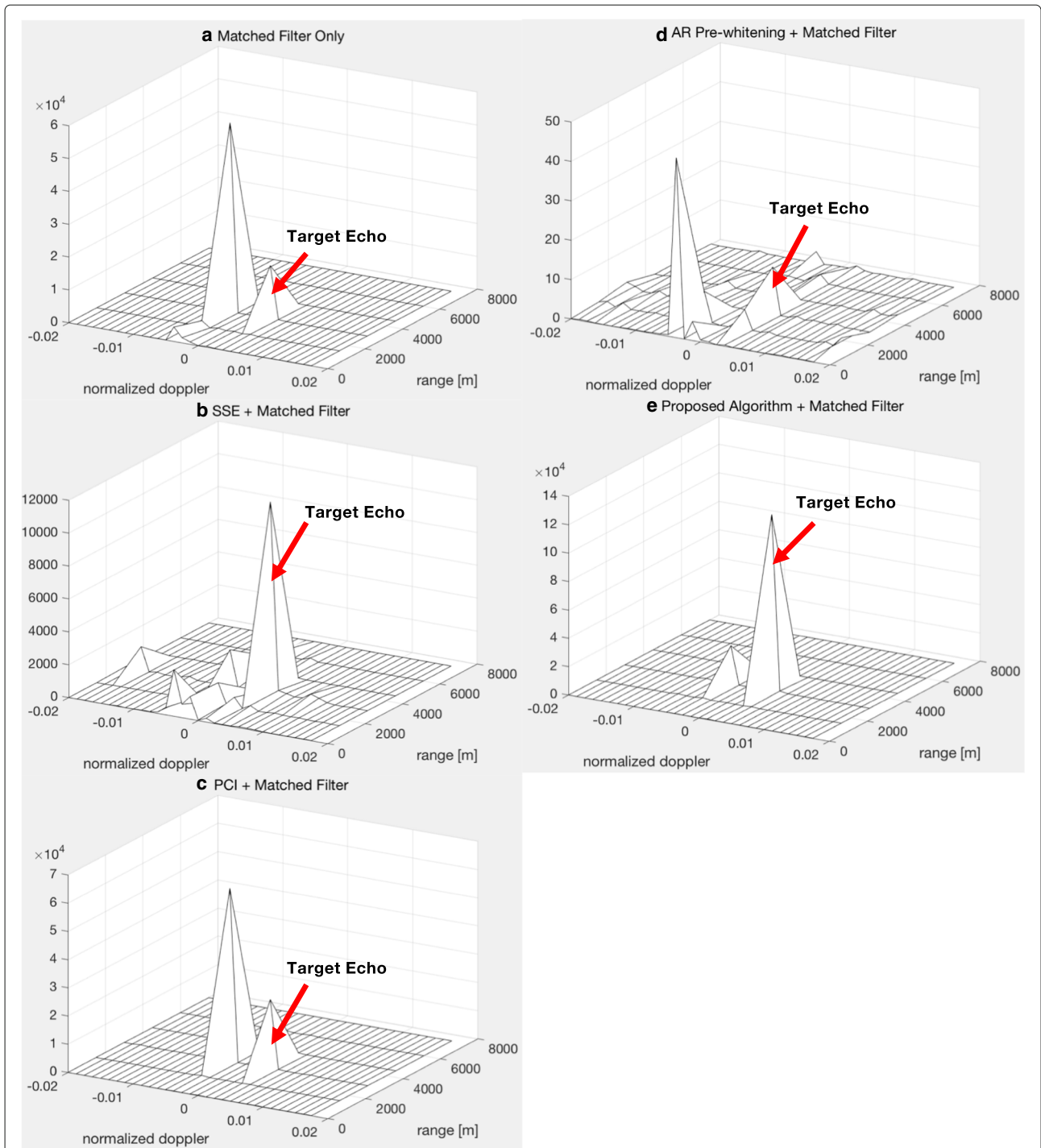
Fig. 15 Range-Doppler maps from measurement experiments. Maps of **a** received signal, **b** output signal of the SSE algorithm, **c** output signal of the PCI algorithm, **d** output signal of the AR pre-whitening, and **e** output signal of the proposed algorithm that is determined using a block-normalized matched filter

values in the two Doppler frequency bins, where values at the negative Doppler bin correspond to reverberations and those at the positive Doppler bin to the target echo. The values related to reverberations are large and can lead to target misidentification. Figure 15b shows that the SSE algorithm reduces the peak value corresponding to the reverberation, but several false peaks persist. Figure 15c shows that the PCI algorithm cannot effectively

suppress the reverberation peak, and Fig. 15d shows that the result of AR pre-whitening has a large value near zero meters. Figure 15e shows that the proposed algorithm suppresses the peak of the reverberation and enhances that of the target echo, thus confirming that it performs well in practice.

5 Conclusion

In this work, the detection performance of the target echo from CW ping signals is enhanced in the presence of reverberations. To this end, a reverberation suppression algorithm is proposed based on NMF and distinguishing characteristics between the target echo and reverberations. Specifically, three characteristics of the target are considered: the frequency structure of the target echo is similar to that of the Doppler-shifted transmitted signal, the target echo has time-continuous values, and the target echo has relatively short and finite duration. To use these characteristics, the frequency bases of the target echo are fixed by frequency-shifted transmitted pings, and constraints on the temporal continuity and TLL are developed.

Experiments are conducted with both simulated and measured reverberations to evaluate the proposed algorithm. The results of the simulations show the ability of the proposed algorithm to find the target echo signal by enhancing SNR from 6 to 15 dB compared to that of the received signal. ROC curves calculated by a Monte Carlo simulation confirm that the proposed algorithm enhances detection probability with a low false alarm rate. The performance of the proposed algorithm is also verified in an ocean by measuring the reverberations. Overall, the results show that the proposed algorithm suppresses reverberations and enhances detection performance in simulations and practice. Expanding our algorithm to multichannel sonar signals and analyzing the convergence of the constrained NMF are expected to be considered in our future work.

Appendix 1: Gradients of the TLL constraints

For the TLL constraint, the backward moving sum of each time basis signal is calculated as

$$\bar{h}_{p,(r,n)} = \sum_{m=n-l_n+1}^{n} h_{p,(r,m)}, \tag{45}$$

where l_n is the expected length of a ping, and $\bar{h}_{p,(r,n)}$ represents the sum of the nth moving block that has a frame index ranging from $(n - l_n + 1)$ to n. The maximum value indicator function $\hat{h}_{p,(r,n)}$ is defined as

$$\hat{h}_{p,(r,n)} = \begin{cases} 1, & n = \arg\max_m \bar{h}_{p,(r,m)}, \\ 0, & \text{otherwise}. \end{cases} \tag{46}$$

To penalize all data from each time basis except for the moving block that has maximum energy, the TLL constraint is defined as

$$C_L = \sum_{r=1}^{R} \left[1 - \sum_{m=n}^{n+l_n-1} \hat{h}_{p,(r,m)} \right]. \tag{47}$$

The TLL penalty takes values between zero and one per basis r. As shown in (46), the maximum value of $\sum_{m=n}^{n+l_n-1} \hat{h}_{p,(r,m)}$ is one, thus limiting the value of the constraint.

Figure 16 shows an example of a simulated TLL constraint on the basis of limited length corrupted by estimation errors. The time basis has a length-limited signal from the 21st to the 30th index, as shown in Fig. 16a. The TLL constraint should ideally have large values outside this region because the TLL is designed to penalize the longer signal. The graphs in Fig. 16b–d describe (45)–(47), respectively. The TLL constraint (Fig. 16e) has large values for every time index except for the region of the signal between the 21st and 30th indices and is thus consistent with the purpose of the design.

Gradient $\nabla_{\mathbf{H_P}} C_L$ can be determined by the chain rule, but the derivative of (46) is challenging to find. Therefore, we use function *softmax* instead of *max*:

$$\hat{h}_{p,(r,n)} = \frac{e^{\bar{h}_{p,(r,n)}}}{\sum_{m=1}^{N} e^{\bar{h}_{p,(r,m)}}}. \tag{48}$$

The TLL penalty with *softmax* is also between zero and one for each basis r because

$$1 - \sum_{m=n}^{n+l_n-1} \hat{h}_{p,(r,m)} \geq 1 - \sum_{m=1}^{N} \hat{h}_{p,(r,m)} = 0, \tag{49}$$

and $\forall \hat{h}_{p,(r,m)} \geq 0$.

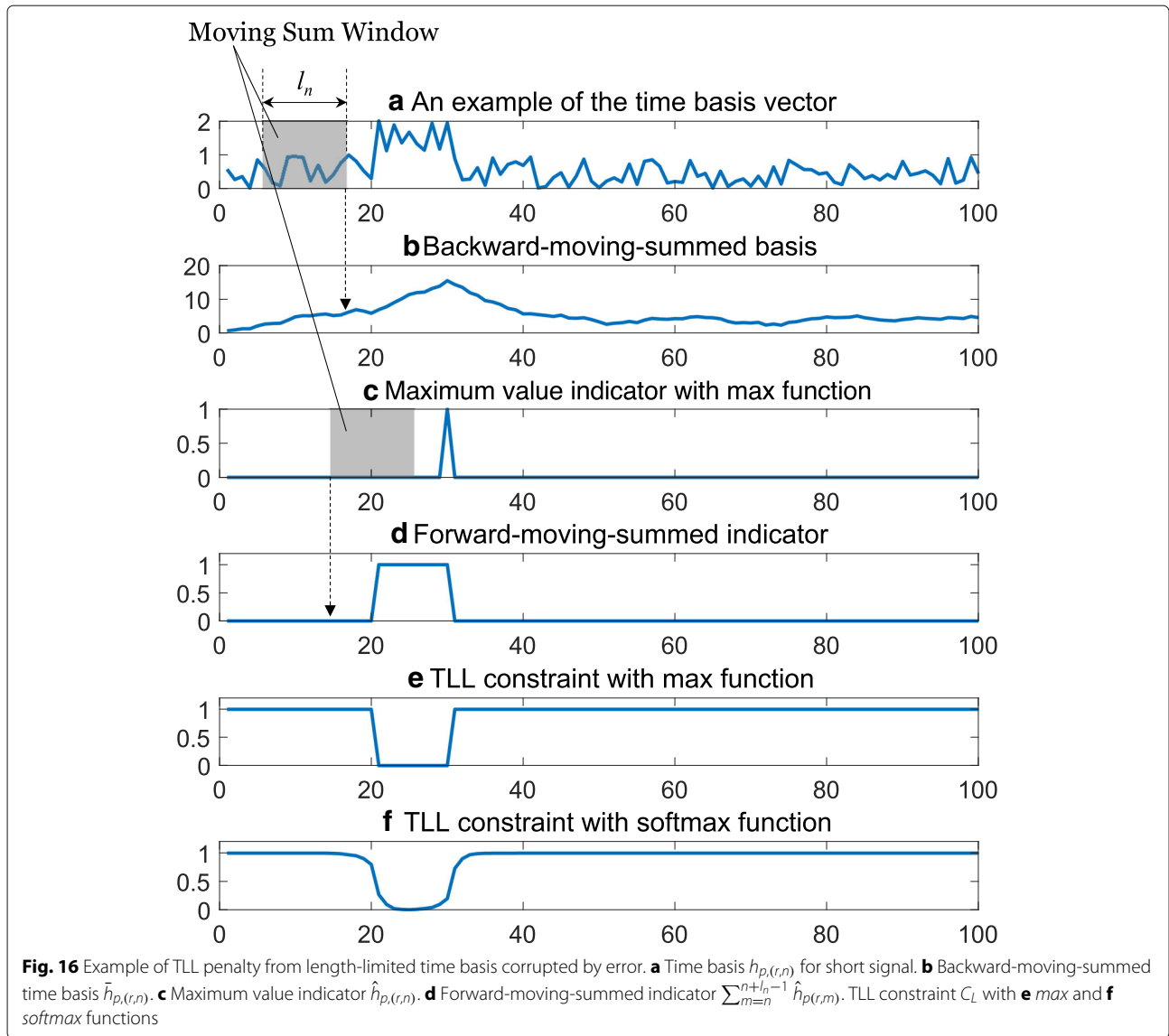Then, the derivative of C_L with respect to $h_{p,(r,n)}$ can be easily determined by the chain rule as

$$\nabla_{h_{p,(r,n)}} C_L = \sum_{m=n}^{n+l_n-1} \left\{ \left(\frac{e^{\bar{h}_{p,(r,m)}}}{\sum_{m=1}^{N} e^{\bar{h}_{p,(r,m)}}} \right)^2 - \frac{e^{\bar{h}_{p,(r,m)}}}{\sum_{m=1}^{N} e^{\bar{h}_{p,(r,m)}}} \right\}. \tag{50}$$

Consequently, each element from $\nabla_{\mathbf{H_P}}^{+} C_L$ and $\nabla_{\mathbf{H_P}}^{-} C_L$ is given by

$$\left[\nabla_{\mathbf{H_P}}^{+} C_L \right]_{(r,n)} = \sum_{m=n}^{n+l_n-1} \left(\frac{e^{\bar{h}_{p,(r,m)}}}{\sum_{i=1}^{N} e^{\bar{h}_{p,(r,i)}}} \right)^2, \tag{51}$$

$$\left[\nabla_{\mathbf{H_P}}^{-} C_L \right]_{(r,n)} = \sum_{m=n}^{n+l_n-1} \frac{e^{\bar{h}_{p,(r,m)}}}{\sum_{i=1}^{N} e^{\bar{h}_{p,(r,i)}}}. \tag{52}$$

Figure 16f shows the TLL constraint using *softmax*, which has a similar structure to that obtained using *max*, as

Fig. 16 Example of TLL penalty from length-limited time basis corrupted by error. **a** Time basis $h_{p,(r,n)}$ for short signal. **b** Backward-moving-summed time basis $\bar{h}_{p,(r,n)}$. **c** Maximum value indicator $\hat{h}_{p,(r,n)}$. **d** Forward-moving-summed indicator $\sum_{m=n}^{n+l_n-1} \hat{h}_{p(r,m)}$. TLL constraint C_L with **e** *max* and **f** *softmax* functions

shown in Fig. 16e. Thus, *softmax* can be used instead of *max* to design the constraint for the length-limited basis.

Appendix 2: Review of convergence analysis of the NMF algorithm

We briefly review the auxiliary function approach to optimizing **H** with the Kullback–Leibler divergence only ($\alpha = \beta = 0$). Let $C_E(h_{(r,n)})$ denote the Kullback–Leibler divergence that is a function of $h_{(r,n)}$ with a fixed $w_{(k,r)}$. An auxiliary function $C^+(h_{(r,n)}, h_{(r,n)}^{(t)})$ for $C_E(h_{(r,n)})$ is defined by a function that satisfies

$$C^+\left(h_{(r,n)}, h_{(r,n)}^{(t)}\right) \geq C_E\left(h_{(r,n)}\right), \tag{53}$$

$$C^+\left(h_{(r,n)}, h_{(r,n)}\right) = C_E\left(h_{(r,n)}\right), \tag{54}$$

where $h_{(r,n)}^{(t)}$ is the updated value of $h_{(r,n)}$ after t iterations (definition 1 in [15]). If we update $h_{(r,n)}^{(t+1)}$ as

$$h_{(r,n)}^{(t+1)} = \arg\min_{h_{(r,n)}} C^+\left(h_{(r,n)}, h_{(r,n)}^{(t)}\right) \tag{55}$$

$C_E(h_{(r,n)})$ is non-increasing during the update because

$$C_E\left(h_{(r,n)}^{(t+1)}\right) \leq C^+\left(h_{(r,n)}^{(t+1)}, h_{(r,n)}^{(t)}\right) \leq C^+\left(h_{(r,n)}^{(t)}, h_{(r,n)}^{(t)}\right) = C_E\left(h_{(r,n)}^{(t)}\right). \tag{56}$$

Lee and Seung [15] and Nakano et al. [17] showed the non-increasing property of multiplicative update algorithms by designing an auxiliary function that satisfies Eqs. (53) and (54) based on Jensen's inequality [17] and by deriving the multiplicative update algorithm from Eq. (55).

Abbreviations

AR: Autoregressive; CW: Continuous wave; FP: False positive; NMF: Non-negative matrix factorization; PCI: Principal component inverse; ROC: Receiver operating characteristic; SNR: Signal-to-noise ratio; SRR: Signal-to-reverberation ratio; SSE: Signal subspace extraction; TLL: Temporal length limitation; TP: True positive

Authors' contributions

SL and JSL conceived and designed the study. JSL defined the problem and directed the study. SL developed the detailed algorithm and performed the experiments. Both authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

[1] School of Electronics Engineering, Kyungpook National University, Daehak-ro 80, 41566 Daegu, Republic of Korea. [2] Department of Electrical Engineering, Sejong University, Neungdong-ro 209, 05006 Seoul, Republic of Korea.

References

1. K. Mio, Y. Chocheyras, Y. Doisy, in *Oceans 2000 MTS/IEEE Conference and Exhibition, vol. 2*. Space time adaptive processing for low frequency sonar (IEEE, Piscataway, 2000), pp. 1315–1319
2. Y. Doisy, L. Deruaz, S. van IJsselmuide, S. Beerens, R. Been, Reverberation suppression using wideband Doppler-sensitive pulses. IEEE J. Oceanic Eng. **33**(4), 419–433 (2008)
3. T. Collins, P. Atkins, Doppler-sensitive active sonar pulse designs for reverberation processing. IEE Proc.-Radar, Sonar Navig. **145**(6), 347–353 (1998)
4. J. Alsup, in *33rd Asilomar Conference on Signals, Systems, and Computers, vol. 2*. Comb Waveforms for Sonar (IEEE, Piscataway, 1999), pp. 864–869
5. H. Cox, H. Lai, in *Proceedings of 28th Asilomar Conference on Signals, Systems and Computers, vol. 2*. Geometric comb waveforms for reverberation suppression (IEEE, Piscataway, 1994), pp. 1185–1189
6. S. Kay, J. Salisbury, Improved active sonar detection using autoregressive prewhiteners. J. Acoust. Soc. Am. **87**(4), 1603–1611 (1990)
7. V. Carmillet, G. Jourdain, in *OCEANS'96 MTS/IEEE Prospects for the 21st Century, vol. 3*. Wideband sonar detection in reverberation using autoregressive models (IEEE, Piscataway, 1996), pp. 1435–1440
8. G. Ginolhac, G. Jourdain, in *OCEANS 2000 MTS/IEEE Conference and Exhibition, vol. 2*. Detection in presence of reverberation (IEEE, Piscataway, 2000), pp. 1043–1046
9. G. Ginolhac, G. Jourdain, "Principal component inverse" algorithm for detection in the presence of reverberation. IEEE J. Oceanic Eng. **27**(2), 310–321 (2002)
10. W. Li, X. Ma, Y. Zhu, J. Yang, C. Hou, Detection in reverberation using space time adaptive prewhiteners. J. Acoust. Soc. Am. **124**(4), 236–242 (2008)
11. W. Li, Q. Zhang, X. Ma, C. Hou, Active sonar detection in reverberation via signal subspace extraction algorithm. EURASIP J. Wirel. Commun. Netw. **2010**(1), 11 (2010)
12. W. Burdic, *Underwater Acoustic System Analysis*. (Prentice Hall, NJ, 1991)
13. D. Lee, H. Seung, Learning the parts of objects by non-negative matrix factorization. Nature. **401**(6755), 788–791 (1999)
14. P. Smaragdis, J. C. Brown, in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, vol. 3*. Non-negative matrix factorization for polyphonic music transcription (IEEE, Piscataway, 2003), pp. 177–180
15. D. D. Lee, H. S. Seung, in *Advances in Neural Information Processing Systems*. Algorithms for non-negative matrix factorization (MIT Press, Cambridge, 2001), pp. 556–562
16. C. Févotte, N. Bertin, J.-L. Durrieu, Nonnegative matrix factorization with the Itakura-Saito divergence: with application to music analysis. Neural Comput. **21**(3), 793–830 (2009)
17. M. Nakano, H. Kameoka, J. Le Roux, Y. Kitano, N. Ono, S. Sagayama, in *Machine Learning for Signal Processing (MLSP), 2010 IEEE International Workshop On*. Convergence-guaranteed multiplicative algorithms for nonnegative matrix factorization with β-divergence (IEEE, Piscataway, 2010), pp. 283–288
18. H. Kim, H. Park, Nonnegative matrix factorization based on alternating nonnegativity constrained least squares and active set method. SIAM J. Matrix Anal. Appl. **30**(2), 713–730 (2008)
19. A. Ozerov, C. Févotte, Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation. IEEE Trans. Audio Speech Lang. Process. **18**(3), 550–563 (2010)
20. E. Vincent, N. Bertin, R. Badeau, in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference On*. Harmonic and inharmonic nonnegative matrix factorization for polyphonic pitch transcription (IEEE, Piscataway, 2008), pp. 109–112
21. E. Benetos, S. Cherla, T. Weyde, in *6th International Workshop on Machine Learning and Music*. An efficient shift-invariant model for polyphonic music transcription (ACM, New York, 2013)
22. T. Virtanen, Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria. IEEE Trans. Audio Speech Lang. Process. **15**(3), 1066–1074 (2007)
23. S. Ewert, M. Müller, in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference On*. Using score-informed constraints for NMF-based source separation (IEEE, Piscataway, 2012), pp. 129–132
24. K. W. Wilson, B. Raj, P. Smaragdis, A. Divakaran, in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference On*. Speech denoising using nonnegative matrix factorization with priors (IEEE, Piscataway, 2008), pp. 4029–4032
25. N. Mohammadiha, P. Smaragdis, A. Leijon, Supervised and unsupervised speech enhancement using nonnegative matrix factorization. IEEE Trans. Audio Speech Lang. Process. **21**(10), 2140–2151 (2013)
26. S. Lee, S. H. Park, K.-M. Sung, Beamspace-domain multichannel nonnegative matrix factorization for audio source separation. IEEE Sign. Process. Lett. **19**(1), 43–46 (2012)
27. Y. L. Hao, L. Wang, in *Advanced Materials Research, vol. 383*. Compressed sensing recognition algorithm for sonar image based on non-negative matrix factorization and adjacency spectra feature (Trans Tech Publ, Singapore, 2012), pp. 5818–5823
28. C. Joder, F. Weninger, F. Eyben, D. Virette, B. Schuller, in *International Conference on Latent Variable Analysis and Signal Separation*. Real-time speech separation by semi-supervised nonnegative matrix factorization (Springer, Berlin, 2012), pp. 322–329
29. D. Abraham, A. Lyons, Simulation of non-rayleigh reverberation and clutter. IEEE J. Oceanic Eng. **29**(2), 347–362 (2004)