JASA EXPRESS LETTERS

https://doi.org/10.1121/10.0001330

# Importance of phase guides from beamformed data for processing multi-channel data in highly scattering media

**Andrey Bakulin,[1,a] Ilya Silvestrov,[1,b] and Dmitry Neklyudov[2,c]**

[1]*The Exploration and Petroleum Engineering Center—Advanced Research Center, Saudi Aramco, Dhahran, 31311, Saudi Arabia*

[2]*Institute of Petroleum Geology and Geophysics of Siberian Branch Russian Academy of Sciences, Novosibirsk, 630090, Russia*

*andrey.bakulin@aramco.com, ilya.silvestrov@aramco.com, neklyudovda@ipgg.sbras.ru*

**Abstract:** Small- and medium-scale scattering can be extremely damaging for acoustic imaging. The mechanism for distortions is related to severe frequency-dependent phase perturbations making signals obscure on multi-channel records. Methods are proposed to compensate for such effects in seismic reflection imaging in two steps. First, the rough signal guide is constructed using massive beamforming. Second, trace-by-trace specialized time-frequency masking is employed to reconstruct the corrupted phase and amplitude spectra. Phase masks are of paramount importance to make target events coherent and visible. Amplitude masks remove scattered background noise from the amplitude spectra. Unlike beamforming itself, these methods avoid smearing of signals across channels and preserve full frequency bandwidth. *© 2020 Acoustical Society of America*

[Editor: Buye Xu] Pages: EL447–EL452

**Received:** 21 January 2020 **Accepted:** 13 May 2020 **Published Online:** 1 June 2020

## 1. Introduction

Reflection seismology experiment utilizes sources and receivers on the surface of the Earth. Since subsurface velocity is unknown, multiple receivers at different ranges record at each source [Fig. 1(a)] to invert for velocity as well as to perform imaging of subsurface layers. In the case of horizontally layered earth, delay times as a function of range are described by a hyperbola with parameters directly related to subsurface velocity. An example of such a hyperbola is seen in Fig. 1(b), which shows multi-channel gather (a seismogram with fixed source and receivers at various ranges) for a synthetic layered model. Seismic processing requires extensive manipulations of the recorded gathers to extract subsurface velocities and prepare them for final imaging. Usually, deviations from hyperbolic assumption caused by heterogeneity or noise can be compensated in terms of travel time as well as waveform variations.[1] However, conventional seismic processing approaches often fail when shallow near-surface layers become complex with heterogeneities comparable to or smaller than the dominant wavelength of the signals. For example, desert areas often possess near surfaces peppered with karsts. Multiple scattering distorts all coherent arrivals, making them hard to distinguish. The goal of this Letter is to present a new alternative to beamforming that removes damaging effects of near-surface scattering on multi-channel gathers but preserves original amplitudes and full frequency bandwidth for target reflection signals. This approach does not require any knowledge of a subsurface velocity and can be applied in a variety of other applications using acoustic and seismic waves.

## 2. Statement of the problem

Let us illustrate the problem using a simple five-layer model, where the top layer is represented by a random distribution of heterogeneities referred to as clutter[2] [Fig. 1(a)]. If a homogeneous layer replaces the clutter—we should observe a simple wavefield with four hyperbolic reflection arrivals and associated multiple bouncing events [Fig. 1(b)]. In the presence of the clutter, the wavefield becomes extremely complex [Fig. 1(c)]. While perhaps hyperbola associated with the reflector 3 can be recognized, the presence of other hyperbolae is not apparent. Knowledge of exact velocity structure in the top layer may enable finding and imaging reflectors; however, its estimation is too challenging in the presence of small- and medium-scale heterogeneity. There is one family of data-driven approaches that could offer a rough solution to find hidden reflectors.

[a]ORCID: 0000-0002-6638-7821.
[b]ORCID: 0000-0002-2394-3576.
[c]Author to whom correspondence should be addressed.

  

JASA EXPRESS LETTERS

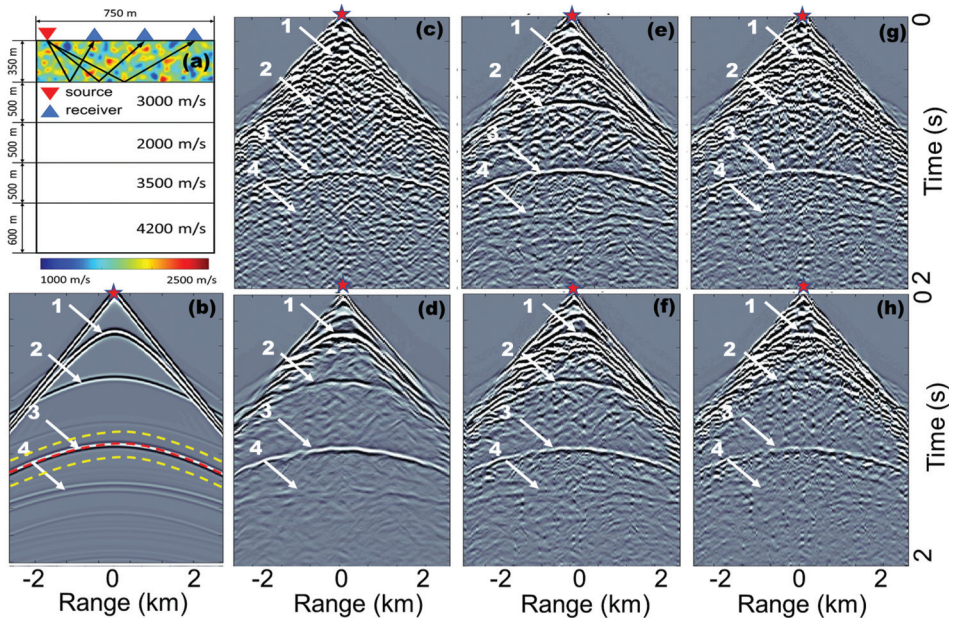


Fig. 1. (a) A fragment of the five-layer synthetic acoustic model. The near-surface layer is modeled as a random, cluttered media using a homogeneous Gaussian isotropic random distribution with mean value 1700 m/s, standard deviation 200 m/s, and correlation length 30 m. (b) Seismogram with fixed source and line of receivers when homogeneous layer replaces the clutter. The red line shows a typical hyperbolic relationship between delay time and range. Yellow lines denote the time gate used for quantitative metrics in Table 1. (c) Seismogram in the model (a) with the clutter layer present. Seismogram from (c) after applying various methods: (d) massive nonlinear beamforming, (e) phase substitution, (f) phase substitution + IRM mask, (g) phase sign-correction method, (h) phase sign-correction method + IRM mask. Time frames of 160 ms with Hann window tapering and overlap of 144 ms were used in STFT.

In geophysics, they are referred to as multi-dimensional local stacking.[3–5] They are closely related to well-known time-delay-and-sum beamforming with local delay times estimated directly from the data assuming locally hyperbolic behavior. In this study, we use so-called nonlinear beamforming[5] that evaluates the coherence of the data along locally hyperbolic surfaces on 3D subvolumes of the data with two spatial and one temporal coordinate. When surface with the largest coherence is found, then the algorithm performs local stacking along these surfaces and assigns the resulting beamformed output into a reference trace in the middle of the ensemble. Then the process is repeated for every single time and a spatial sample of the multi-channel data. As a result, we obtain beamformed or enhanced gather [Fig. 1(d)] with reliable detection of all four reflectors. While such massive multi-dimensional stacking methods are efficient at finding events in the jumble of original data, they are rarely used in actual seismic processing because of highly undesirable side effects: (1) original amplitudes at each receiver point might be heavily averaged, (2) data are overly smoothed whereas local travel times and amplitudes characterizing near surface or subsurface are distorted, and (3) higher frequencies are lost during beamforming. While beamforming preconditions the data for conventional seismic processing methods, distortions introduced by local stacking are often considered as too severe for reliable quantitative analysis.

## 3. Methods

Here, we propose an alternative approach that uses massively beamformed data as a signal guide and heals original data corrupted by scattering but only to the extent required for conventional methods to work. We further aim to preserve original amplitude information as well as full frequency bandwidth so that subsurface targets can be characterized at full vertical resolution and with minimum distortions. Consider two multi-channel seismic datasets: (1)"noisy" data $x(t)$ and (2) "enhanced" data $s(t)$ obtained from the noisy data by local stacking or massive beamforming. The enhanced dataset represents our best available estimate of the signal suffering from the limitations above. Our goal is to extract an improved estimate of the signal contained in trace $x(t)$ using corresponding enhanced trace $s(t)$ as a guide. The propagation of seismic signals is a nonstationary process, so our procedure is designed in the time-frequency (TF) domain. Applying the short-time Fourier transform (STFT) to $x(t)$ and $s(t)$, we obtain complex-valued TF spectra of the traces $X(k, l)$, $S(k, l)$ with $k$, $l$ representing the discrete frequency bin and time frame indexes, respectively. TF spectrum of the input signal, $X(k, l)$, is represented as a superposition of desired signal and noise: $X(k,l) = \tilde{S}(k,l) + \tilde{N}(k,l)$. It is useful to formulate the healing

https://doi.org/10.1121/10.0001330

procedure using the concept of time-frequency masking (TFM) widely used in speech processing for extracting clean speech from noisy data. TFM provides an estimate of signal TF spectrum $\hat{S}(k,l)$ as a multiplication,

$$\hat{S}(k,l) = M(k,l)X(k,l), \tag{1}$$

where a time-frequency mask $M(k, l)$ is typically a real-valued function, $0 \leq M(k,l) \leq 1$. In essence, amplitude TFM attempts to preserve signal while suppressing noise by applying a filter close to unity in a "signal dominance" region and close to zero in a "noise dominance" area.[6,7]

Estimates of signal and noise power spectra are required for TFM design. In speech processing, a number of approaches have been developed to estimate the noise power spectrum directly from noisy data using some unique features of human speech. When the noise power spectrum is found, the power spectrum of clean speech (signal) may be estimated from a noisy speech by spectral subtraction. Such an approach is not easily transferable to seismic applications. Instead, we propose to utilize an enhanced or massively beamformed dataset as a crude analog of clean speech. Having an ancillary dataset from beamforming identical in the number of channels, we can perform single-channel filtering where each noisy trace is subject to specialized TFM derived solely based on corresponding enhanced trace from beamforming, thus simplifying the processing sequence to trace-by-trace transformations.

### *3.1 Phase substitution method*

Accurate phase information has been vividly demonstrated to have an outsized role for multichannel images.[8] In the presence of accurate phase information, 2D photo images preserve full visual content even with nearly arbitrary amplitude spectrum. In the absence of phase information, images become completely uninterpretable even when the exact amplitude spectrum is preserved. This clue leads us to our first method of "phase substitution," where we take phase of the beamformed trace as our best estimate of the signal phase, whereas the amplitude spectrum of the original noisy data remains untouched. TF spectrum of the desired signal trace is given as

$$\hat{S}(k,l) = |X(k,l)| \exp\left[i\phi_S(k,l)\right], \tag{2}$$

where $|X(k,l)|$ is the amplitude TF spectrum of original trace and $\phi_S$ is the phase spectrum of the enhanced trace after beamforming. Corrected time-domain signal estimation $\hat{s}(t)$ is obtained by inverse STFT of $\hat{S}(k,l)$. Phase substitution method can be considered as a special case of complex-valued phase-only TFM with

$$M(k,l) = \exp\left[i(\phi_S(k,l) - \phi_X(k,l))\right], \tag{3}$$

as can be easily observed by substituting Eq. (3) into Eq. (1). Here, $\phi_X$ and $\phi_S$ are phase spectra of noisy and enhanced traces, respectively. Complex-valued TFMs that modify both phase and amplitude recently appear in speech processing.[9]

Figure 1(e) shows the resulting seismogram after applying phase substitution. Reflectors become visible and trackable. In contrast to beamforming, there is no oversmoothing and no loss of higher frequencies since the amplitude spectrum of the original data is fully preserved.

### *3.2 Phase correction method*

Full phase substitution is equivalent to propagating *a priori* phase information from enhanced data directly to signal estimate. As a more delicate alternative, we can apply phase masks that correct the phase of noisy traces using a guide but do not force them to be equal to the enhanced phase. Specifically, we find that near-surface scattering often makes data look incomprehensible when scattering alters the sign of the phase at specific frequency bins along the multi-channel gathers. In this case, the promising approach is "phase correction," which corrects the phase of original data using a phase sign-correction mask (PSM)

$$\hat{S}(k,l) = X(k,l)PSM(k,l), \quad PSM(k,l) = \operatorname{sgn}\left[\cos\left(\phi_S(k,l) - \phi_X(k,l)\right)\right]. \tag{4}$$

Figure 1(g) shows a seismogram after phase corrections. Correcting only signs frequency by frequency and frame by frame using the "phase sign-correction mask" from Eq. (4) can greatly improve event tracking. If original and enhanced data are in phase (the phase difference is less than $\pm\pi/2$) at a specific frequency, then no correction is made (PSM = 1). If they are out of phase (a difference of more than $\pm\pi/2$), then the phase at this frequency is flipped by $\pm\pi$ (PSM = −1). We assume that the phase difference is wrapped within the interval of $[-\pi, \pi]$. Unlike the phase-sensitive mask of the form $\cos\left[\phi_S - \phi_X\right]$,[10] the proposed phase sign-correction mask is amplitude-preserving since no values of amplitude spectra are modified. We observe a big improvement in reflector identification while doing less invasive corrections to the phase

compared to the phase substitution method. After reflectors become trackable in Fig. 1(e)–1(g), data become processable with conventional methods.

### 3.3 Assessment of phase methods

To demonstrate the benefit of the proposed approaches, we compute three quantitative metrics:

- Coherence computed as

$$\frac{1}{M}\frac{\sum_{j=1}^{N}\left(\sum_{i=1}^{M}u_{ij}\right)^2}{\sum_{j=1}^{N}\sum_{i=1}^{M}u_{ij}^2},$$

  where $u_{ij}$ is a $j$th time sample in $i$th channel, N is a number of time samples within a window, and M is a number of channels.
- Normalized difference in amplitudes between processed data $u_{ij}$ and original clutter $d_{ij}$ computed as

$$\frac{1}{M}\sum_{i=1}^{M}\frac{\sum_{j=1}^{N}(u_{ij}-d_{ij})^2}{\sum_{j=1}^{N}d_{ij}^2}.$$

- Dominant frequency of the spectrum $f_{dom}$.

Table 1 presents metrics computed for 150 ms window surrounding the third reflector and shown in Fig. 1(b). Clutter destroys data coherence. While beamforming achieves the highest coherence, it drops the dominant frequency and does not preserve amplitudes. The new phase methods maintain the dominant frequency and better preserve original amplitudes, but have somewhat lower coherence that is still acceptable for successful processing. In other words, the new methods correct for most corrupting details interpreted as small-to-medium scale near-surface scattering, but preserve medium-to-large scale details demanded by conventional methods to characterize near surface as well as deeper subsurface. We further maintained original amplitude information as well as higher frequencies without smearing across multiple channels. Such improvements are also critical for seismic prospecting that rely on tracking subtle amplitude anomalies to map properties of subsurface targets. As a result, the proposed methods strike a balance not achievable with beamforming: obtaining sufficient coherence required for processing while preserving amplitudes and higher frequencies.

### 3.4 Amplitude TFM using beamformed data as a guide

Phase masks helped us to make reflections visible and coherent, while thoroughly preserving original amplitude spectra containing both target reflections as well as undesired arrivals from near-surface scattering. If removing scattered noise for further imaging is required, then TFM offers additional opportunities in terms of conventional amplitude masks.[7] Specifically, we can attack amplitude remnants of small-scale near-surface scattering now on power spectra. One of the popular amplitude-only TFMs is the ideal rationale mask (IRM),[11]

$$IRM(k,l)=\sqrt{\frac{|\tilde{S}(k,l)|^2}{|\tilde{S}(k,l)|^2+|\tilde{N}(k,l)|^2}}, \tag{5}$$

where $|\tilde{S}(k,l)|^2$ and $|\tilde{N}(k,l)|^2$ are estimates of signal and noise power spectrum. To derive the noise power spectrum, we apply the minimal statistic (MS) approach[12] that is well-known in

Table 1. Quantitative metrics for the performance of the new methods compared with beamforming. For reference, data without clutter has the coherence of 0.95 and dominant frequency 22 Hz.

| Metrics | Clutter | Beamforming | Phase Substitution | Phase corrections |
|---|---|---|---|---|
| Coherence | 0.1 | 0.64 | 0.51 | 0.31 |
| Ampl. difference | 0 | 0.8 | 0.64 | 0.35 |
| $f_{dom}$ | 19 Hz | 16 Hz | 19 Hz | 19 Hz |

https://doi.org/10.1121/10.0001330

speech processing directly to the original noisy traces. In our case, the most straightforward estimate of the signal TF power spectrum can be obtained by taking it from enhanced data after beamforming, e.g., $|\tilde{S}(k,l)|^2 = |S(k,l)|^2$. Amplitude IRM is additionally applied together with phase correction masks, Eq. (3) or (4), providing an improved signal estimate

$$\hat{S}(k,l) = X(k,l)PSM(k,l)IRM(k,l). \tag{6}$$

Figures 1(f)–1(h) demonstrate IRM capability to suppress further amplitudes associated with small-scale near-surface scattering. Amplitude noise in between the main reflection events is significantly reduced after the application of the IRM mask [cf. Figs. 1(e)–1(g) and Figs. 1(f)–1(h)]. As expected, amplitudes and local features of key reflection events remain largely untouched since IRM is expected to be close to unity.

## 4. Real data example from land seismic in the desert environment

Recorded land seismic data from the desert environment typically are very challenging due to strong near-surface scattering and variable source and receiver coupling that create additional frequency-dependent phase distortions. A seismogram from such a dataset after standard processing shows only clues of a few of the strongest reflection events [Fig. 2(a)]. Massive nonlinear beamforming[5] reveals coherent reflections, but the events appear overly smoothed and loss of higher frequencies above 40 Hz is observed [Fig. 2(e)]. In contrast, phase substitution and phase correction approaches also make the reflections coherent and visible [Fig. 2(c)–2(d)] but preserve higher frequencies, avoid oversmoothing, and keep local details of the reflections. The phase sign corrections achieved by multiplication of TF spectra by 1 or −1, according to Eq. (4), maintain the absolute amplitudes of the spectrum. Computed spectra validate that new approaches preserve higher frequencies [Fig. 2(e), green and blue lines]. Application of amplitude IRM mask from Eq. (6) enables additional cleanup of the reflection amplitudes from scattered noise. For example, this effect is seen in the amplitude spectrum after IRM [Fig. 2(e), green and black lines]

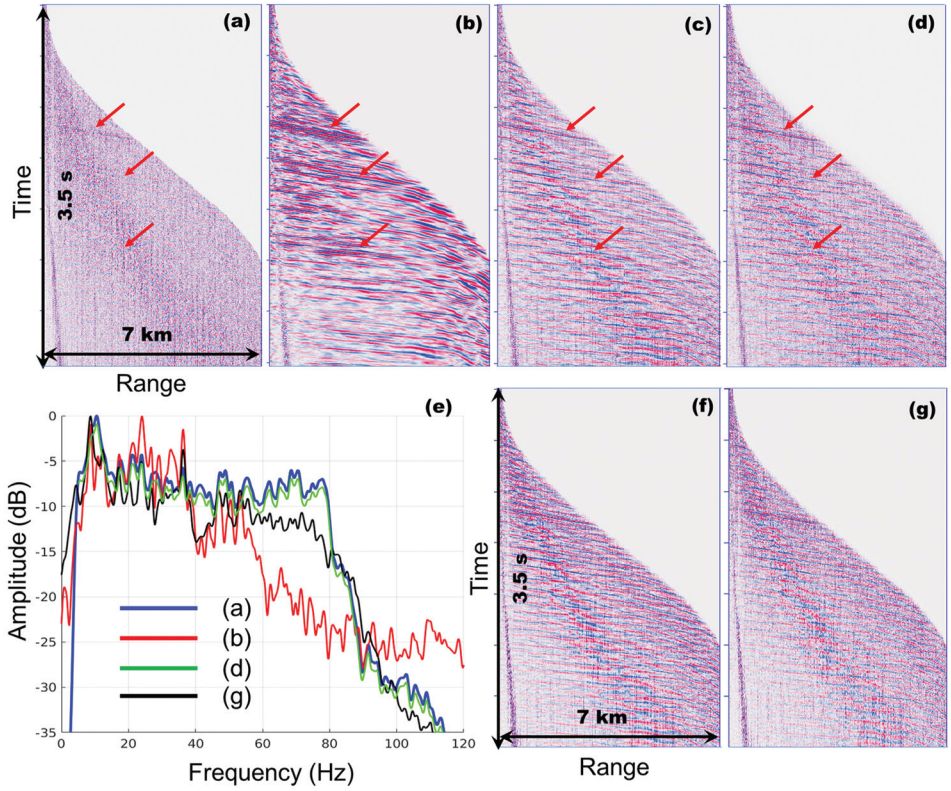


Fig. 2. Example from real land seismic data: (a) input seismogram after conventional processing, (b) seismogram after massive beamforming, (c) seismogram after phase substitution method, (d) seismogram after phase sign-correction method, (e) comparison of amplitude spectra, (f) same as (c) but with IRM added, (g) same as (d) but with added IRM. Massive nonlinear beamforming (b) unveils reflections hidden beneath scattered noise but leads to loss of higher frequencies and oversmoothed character of the signals. In contrast, phase masking approaches (c) and (d) also uncover coherent reflections avoiding oversmoothing, preserving higher frequencies, and maintaining the original character of the gathers. Amplitude spectra (e) shows how higher frequencies are suppressed by massive beamforming and restored by phase masking methods [curves for (c) and (f) are not shown but close to (d) and (g)].

indicating suppression of dominant high-frequency environmental noise above 60 Hz typical for land seismic surveys. While massive beamforming at higher frequencies suppressed both signal and noise due to mis-stacking, amplitude masks combined with phase corrections were able to reduce the noise while leaving some expected signal exhibiting more natural roll-off due to intrinsic attenuation [Fig. 2(e), black line].

## 5. Conclusion

We presented a new approach for reconstructing an acoustic or seismic wavefield with scattering distortions caused by small- and medium-scale heterogeneities. First, we utilized massive beamforming with large apertures to roughly unveil hidden events. The beamformed data form an ancillary multi-channel dataset serving as an approximate signal model to guide the reconstruction process. We exploited this guide to correct the original distorted data using specially designed time-frequency masking. We demonstrated that the repair of the phase spectrum is critical for making signal events coherent and trackable. The phase substitution method uses the guide phase as a direct estimate of the uncorrupted phase, whereas the phase sign-correction method only fixes apparent polarity reversals based on this guide. Both methods are formulated as special cases of phase-only time-frequency masking, augmenting the TFM arsenal already used in speech processing. The proposed phase masking methods maintain original amplitude spectra of each channel, thus avoiding smearing of information along traces and preserving the full frequency bandwidth. Amplitude masks are further designed to tackle the remnants of multiple scattering on amplitude spectra, which could be beneficial for imaging. Likewise, traces from the beamformed dataset are used as guides to design amplitude masks to additionally correct noise-dominated areas while preserving the signals. Examples from challenging synthetic and real seismic data validate the ability of new methods to mitigate the effects of small- and medium-scale scattering caused by complex near-surface layers. Corrected multi-channel data become acceptable for conventional seismic processing using existing methods.

## References and links

[1]O. Yilmaz, *Seismic Data Analysis* (Society of Exploration Geophysics, Tulsa, 2001).
[2]J. G. Berryman, L. Borcea, G. C. Papanicolaou, and C. Tsogka,"Statistically stable ultrasonic imaging in random media," J. Acoust. Soc. Am. **112**, 1509–1522 (2002).
[3]A. Berkovitch, K. Deev, and E. Landa, "How non-hyperbolic MultiFocusing improves depth imaging," First Break **29**, 103–111 (2011).
[4]V. Buzlukov and E. Landa, "Imaging improvement by prestack signal enhancement," Geophys. Prospect. **61**, 1150–1158 (2013).
[5]A. Bakulin, I. Silvestrov, M. Dmitriev, D. Neklyudov, M. Protasov, and K. Gadylshin, "Nonlinear beamforming for enhancing prestack seismic data with a challenging near surface or overburden," First Break **36**(12), 121–126 (2018).
[6]O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," IEEE Trans. Sign. Process. **52**, 1830–1846 (2004).
[7]D. Wang, "Time-frequency masking for speech separation and its potential for hearing aid design," Trends Amplif. **12**, 332–353 (2008).
[8]A. V. Oppenheim and J. S. Lim," The importance of phase in signals," Proc. IEEE **69**, 529–541 (1981).
[9]D. S. Williamson and D. Wang, "Time-frequency masking in the complex domain for speech dereverberation and denoising," IEEE/ACM Trans. Audio Speech Lang. Process. **25**(7), 1492–1501 (2017).
[10]H. Erdogan, J. R. Hershey, S. Watanabe, and J. Le Roux, "Phase-sensitive and recognition-boosted speech separation using deep recurrent neural networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2015), pp. 708–712.
[11]S. Liang, W. Liu, W. Jiang, and W. Xue, "The optimal ratio time-frequency mask for speech separation in terms of the signal-to-noise ratio," J. Acoust. Soc. Am. **134**(5), EL452–EL458 (2013).
[12]R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," IEEE Trans. Speech Audio Process. **9**(5), 504–512 (2001).