

Automatic noise-removal/signal-removal based on general cross-validation thresholding in synchrosqueezed domain and its application on earthquake data

S. Mostafa Mousavi¹ and Charles A. Langston²

ABSTRACT

Recorded seismic signals are often corrupted by noise. We have developed an automatic noise-attenuation method for single-channel seismic data, based upon high-resolution time-frequency analysis. Synchrosqueezing is a time-frequency reassignment method aimed at sharpening a time-frequency picture. Noise can be distinguished from the signal and attenuated more easily in this reassigned domain. The threshold level is estimated using a general cross-validation approach that does not rely on any prior knowledge about the noise level. The efficiency of the thresholding has been improved by adding a preprocessing step based on kurtosis measurement and a postprocessing step based on adaptive hard thresholding. The proposed algorithm can either attenuate the noise (either white or colored) and keep the signal or remove the signal and keep the noise. Hence, it can be used in either normal denoising applications or preprocessing in ambient noise studies. We tested the performance of the proposed method on synthetic, microseismic, and earthquake seismograms.

INTRODUCTION

During the acquisition process, seismic data are often corrupted by noise. Seismic denoising aims at increasing the signal-to-noise ratio (S/N) by eliminating this additive noise through some signal-processing steps, while preserving important features of the seismic signal. Spectral filtering, as a common approach for improving the S/N, is not effective for suppressing noise that has the same frequency content as the signal. Moreover, it can distort the signal

(Douglas, 1997) and/or generate artifacts prior to impulsive arrivals (Scherbaum, 2001).

A more effective noise suppression can be achieved through thresholding methods in the time-frequency domains, such as using the S-transform (Pinnegar and Eaton, 2003; Schimmel and Gallart, 2007; Parolai, 2009; Ditommaso et al., 2010, 2012; Tselentis et al., 2012), the Radon transform (Sabbione et al., 2013, 2015; Zhang et al., 2015), the wave-packet transform (WPT) (Galiana-Merino et al., 2003; Shuchong and Xun, 2014), f - x or f - k filtering (Bekara and van der Baan, 2009; Naghizadeh, 2011; Naghizadeh and Sacchi, 2012; Chen and Ma, 2014), singular spectrum analysis (Oropeza and Sacchi, 2011), sparse transform-based denoising (Chen et al., 2016), a mathematical morphology-based denoising approach (Li et al., 2016), reduced-rank filtering (Velis et al., 2015), damped multichannel singular spectrum analysis (Huang et al., 2016), the nonlocal means algorithm (Bonar and Sacchi, 2012), or the continuous wavelet transform (CWT) (Pazos et al., 2003; Sobolev and Lyubushin, 2006; Mousavi and Langston, 2016a, 2016b, 2016d).

Bekara and van der Baan (2009), Han and Van Der Baan (2015), and Gómez and Velis (2016) show that seismic noise can be removed effectively using empirical-mode decomposition (EMD). EMD (Huang et al., 1998) is a data driven time-frequency analysis technique that adaptively decomposes a signal into a set of localized, modulated oscillations called intrinsic mode functions.

Recently, a new reassignment technique called synchrosqueezing (SS) was introduced as a powerful alternative to EMD (Daubechies et al., 2011). SS produces a sharpened time-frequency representation (TFR) of the signal that highly localizes modulated oscillations. It has better mathematical support and adaptability properties compared with EMD (Thakur et al., 2013; Herrera et al., 2014, 2015).

Meignen et al. (2012) and Ahrabian and Mandic (2015) introduce denoising techniques based on SS for univariate and multivariate signals, respectively. These methods are based on identifying common modulated oscillations in elements of data. They outperform wavelet

Manuscript received by the Editor 12 August 2016; revised manuscript received 15 February 2017; published online 16 May 2017; corrected version published online 7 July 2017.

¹Stanford University, Geophysics Department, Stanford, California, USA. E-mail: mmousavi@stanford.edu.

²University of Memphis, Center for Earthquake Research and Information (CERI), Memphis, Tennessee, USA. E-mail: clangstn@memphis.edu.

© 2017 Society of Exploration Geophysicists. All rights reserved.

and EMD-based methods. Mousavi et al. (2016a) show that a simple normalization step in the synchrosqueezed domain can improve the S/N of microseismic events.

Here, we introduce an adaptive, fast algorithm for automatic noise or signal removal based on the synchrosqueezed-CWT (SS-CWT), incorporating higher order statistics (HOS), general cross validation (GCV), and wavelet hard-thresholding (WHT) for seismic data. The proposed method takes advantage of the mode decomposition property of the SS-CWT. Major components present in recorded data are thresholded separately based on data characteristics. Synthetic and real simulations show that the proposed method is effective for accurate denoising and increasing the S/N of microseismic and ocean-bottom seismic (OBS) data, as well as filtering out the seismic signal in the case of noise studies.

THEORETICAL BACKGROUND

Time-frequency representation

We assume that real signals can be modeled by time-varying oscillatory components defined as (Herrera et al., 2014)

$$s(t) = A(t) \cos(2\pi\phi(t)), \quad (1)$$

where $\phi(t)$ and $A(t)$ are the instantaneous phase and amplitude of a time series $s(t)$, respectively. The derivative of the instantaneous phase is referred to as instantaneous frequency $f = 1/2\pi(d\phi(t)/dt)$. To identify the associated $A(t)$ and $\phi(t)$ for a given signal $y(t)$, the Hilbert transform can be used to generate the analytic signal (Gabor, 1946):

$$y_+^a(t) = y(t) - iH[y(t)], \quad (2)$$

where $H[\cdot]$ is the Hilbert transform. The analytic signal is complex and can be used to find the $A(t)$ and the $\phi(t)$.

However, real-world signals, such as seismic traces, usually consist of many components and are generally contaminated with noise. Hence, the recorded signal can be represented as a combination of components plus some additive noise $\varepsilon(t)$:

$$y(t) = \sum_{k=1}^K S_k(t) + \sigma\varepsilon(t) = \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t)) + \sigma\varepsilon(t), \quad (3)$$

where K is the number of components in the recorded signal and σ is the noise level. Therefore, the analytic signal represents a mixture of the amplitudes and phases of individual components of the observed seismic traces. Time-frequency transforms (TFTs) aim to localize individual oscillatory components of the recorded signals. The nonstationary nature of seismic signals indicates that instead of considering a signal in the frequency or time domain (1D), it is often more informative to study their TFR. The 2D evolution of the spectral content of the seismic data can be tracked in a TFR.

The advantage of denoising in a time-frequency domain over traditional spectral filtering is that it allows for separating the noise from the signal even in the same passband as long as they are temporally separated.

Many TFTs exist for this purpose, such as the short-time Fourier transform (STFT) and the CWT. A comprehensive review and comparison of application of different TFTs on seismic data can be found in Tary et al. (2014). Here, we just give a short description of STFT

and CWT to briefly address resolution problems and the enhancement achieved by the SS step.

Short-time Fourier transform

The Fourier transform decomposes a signal into sine and cosine basis functions. The most common TFT used for time-frequency analysis is the STFT (also known as the windowed Fourier transform) (Gabor, 1946; Allen and Rabiner, 1977). Indeed, the STFT is the Fourier transform of successive windows of the signal:

$$F_y(\tau, \xi) = \int_{-\infty}^{\infty} y(t) G(t - \tau) \exp(-i\xi t) dt, \quad (4)$$

where t is the time, ξ is the angular frequency, τ is the time delay, and $G(t)$ is a chosen window function that is usually either a Gaussian or Hann function. The sliding window is held constant during the analysis, irrespective of the investigated frequencies. Thus, the time and frequency resolutions are kept constant and depend directly on the window size (Reine et al., 2009). Tary et al. (2014) define time and frequency resolutions as the ability to distinguish two wavefronts and two spectral peaks, respectively. However, these resolutions are always limited by the Heisenberg/Gabor uncertainty principle (Gabor, 1946; Mallat, 1999), i.e., $\Delta_t \Delta_f \geq 4\pi$, where Δ_t represents the time resolution and Δ_f stands for the frequency resolution.

The TFR resolution obtained from any transform depends on the intrinsic characteristics of the analyzed signal and on the specific properties of the chosen transform; hence, the TFT should be viewed as a measurement device (Auger et al., 2013).

This time-frequency trade-off and the fixed resolution of the STFT imply that one will lose time resolution (accurate timing of frequency changes) if one wants to accurately identify spectral peaks (high-frequency resolution) (Tary et al., 2014).

Continuous wavelet transform

Multiresolution transforms such as the CWT (Daubechies and Heil, 1992) can obtain a better TFR for signals with low- and high-frequency content because the signal is analyzed under different resolutions (or scales) at different frequencies. The CWT is accomplished through a prototype analyzing function known as the mother wavelet ψ , which can be interpreted as a bandpass. The CWT of y in equation 3 at scale a and time shift τ is given by (Daubechies and Heil, 1992; Mallat, 1999)

$$W_y(a, \tau) = \langle y, \psi_{a,\tau} \rangle = \int_{-\infty}^{+\infty} y(t) a^{-\frac{1}{2}} \psi^* \left(\frac{t - \tau}{a} \right) dt, \quad (5)$$

where $*$ denotes the complex conjugate, $\langle y, \psi \rangle$ is the inner product, and W_y is the coefficient representing the finite energy of the signal y in a concentrated time-frequency picture. The mother wavelet ψ should be a square integrable function, in which its Fourier transform $\hat{\psi}(\xi)$, should vanish at zero frequency:

$$\hat{\psi}(0) = \int \psi(t) dt = 0. \quad (6)$$

This is called the admissibility condition (Daubechies and Heil, 1992; Farge, 1992). According to Plancherel's theorem, equation 5 can be written in the frequency domain as (Daubechies et al., 2011; Herrera et al., 2014)

$$W_y(a, \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{y}(\xi) a^{-\frac{1}{2}} \hat{\psi}^*(a\xi) \exp(i\xi\tau) d\xi, \quad (7)$$

where $\hat{y}(\xi)$ is the Fourier transform of the signal. The inversion of the CWT can be expressed as

$$y(t) = \frac{1}{C_\psi} \iint W_y(a, \tau) d\tau \frac{da}{a^2}, \quad (8)$$

where the constant C_ψ is given by (Thakur et al., 2013)

$$C_\psi = \int_0^\infty \xi^{-1} \hat{\psi}^*(\xi) d\xi. \quad (9)$$

The CWT can be thought of as the crosscorrelation of y with stretched (or compressed) and shifted mother wavelets, to capture oscillatory features of the signal at different frequencies. The variable length of ψ leads to a flexible trade-off between frequency and time-localization compared with the STFT (Tary et al., 2014). However, it still displays spectral smearing due to the finite size of the operator (Hall, 2006).

Synchrosqueezing

The SS is a relatively new technique introduced by Daubechies and Maes (1996) and Daubechies et al. (2011) as a powerful tool for precisely decomposing and analyzing a signal. It can be classified as a time-frequency reassignment method aiming at a sharpened TFR by applying a postprocessing reallocation on the original TFR. However, unlike classic reassignment methods (e.g., Auger and Flandrin, 1995; Chassande-Mottin et al., 1997), SS is adaptive to different types of data, is visually informative, and enjoys a simple and efficient reconstruction formula (Yang, 2015).

At each time or space location, the SS process reassigns values of the TFR based on their local oscillation. The idea behind SS is that concentrating a spectrogram's energy around instantaneous frequencies will decrease spectral smearing, thus sharpening the TFR, while still allowing its reconstruction. SS can be used to enhance many classic TFRs, e.g., the SS-CWT as in Daubechies et al. (2011), Thakur et al. (2013), and Iatsenko et al. (2015); the SS-STFT as in Thakur and Wu (2011) and Iatsenko et al. (2015); the SS-WPT as in Yang (2015); the synchrosqueezed-curvelet transform (SS-CT) as in Yang and Ying (2014); and the synchrosqueezed S-transform (SS-ST) as in Huang et al. (2015).

Thakur et al. (2013) show that compared with the STFT and CWT, the SS-CWT has superior frequency resolution and can be used to distinguish oscillatory components of complicated signals. Rigorous analysis has proven the stability and robustness of SS for analyzing 1D signals corrupted by noise or perturbations in the signal (Hou et al., 2012; Thakur et al., 2013).

Following Daubechies et al. (2011), the SS-CWT is performed in three steps. First, wavelet coefficients $W_y(a, \tau)$, of the recorded signal y , are calculated, i.e., by using equation 5 or 7. In the next step, a candidate instantaneous frequency $\omega_y(a, \tau)$ can be computed for wavelet coefficients of y at any point (a, τ) as

$$\omega_y(a, \tau) = -\frac{i}{2\pi W_y(a, \tau)} \frac{\partial W_y(a, \tau)}{\partial \tau}, \text{ for } W_y(a, \tau) \neq 0. \quad (10)$$

The instantaneous frequencies are known as ridges in the TFR (Auger et al., 2013). In practice, the very small wavelet coefficients

$W_y(a, \tau)$ need to be removed to make the division operator numerically stable. SS squeezes the energy around these ridges (condensing the CWT coefficients at each time point along the scale axis) to decrease the smearing. To do this, in the last step, the information from the time-scale plane is transformed to the time-frequency plane, $(a, \tau) \rightarrow \langle \omega_y(a, \tau), \tau \rangle$. This operation is called SS and has been shown to improve the concentration of energy and, as a result, readability of the TFR (Daubechies et al., 2011). If the number of scales used in the CWT and the sampling frequency are N and s_f , respectively, frequencies on the SS-CWT would be $\omega_\ell = \ell s_f / N$, $\ell \in [1, N]$ because $W_y(a, \tau)$ is calculated at discrete values a_k . The CWT coefficients within the frequency range $\Delta\omega = \omega_\ell - \omega_{\ell-1} = s_f / N$, will be added up to the center frequency ω_ℓ to construct each instantaneous frequency. Hence, the synchrosqueezed transform is defined as

$$T_y(\omega_\ell, \tau) = \Delta\omega^{-1} \sum_{a_k | \omega(a_k, \tau) - \omega_\ell | \leq \Delta\omega / 2} W_y(a_k, \tau) a_k^{-3/2} \Delta a_k, \quad (11)$$

where ω_ℓ is the ℓ th discrete frequency, a_k is the k th scale, and $\Delta a = a_k - a_{k-1}$. We can recover individual component y_k from the T_y by integrating the coefficients over frequencies ω_ℓ that correspond to the k th component. Following Thakur et al. (2013), let the $l \in L_k(t)$, a small frequency band around the ridge of k th component in the SS-CWT. This band can be estimated using a standard least-squares ridge extraction method (e.g., Carmona et al., 1997) or defined manually. Because y_k is real, then we will have

$$y_k(t) = 2C_\psi^{-1} \operatorname{Re} \left(\sum_{l \in L_k(t)} T_y(\omega_\ell, t) \right). \quad (12)$$

Doing so, one can decompose a signal into its constituent components. It is clear that the highly structured TFR provided by SS (Figure 1) can be exploited for classic signal processing applications, such as denoising (Auger et al., 2013).

Time-frequency denoising

Recalling the model in equation 3 for the recorded data (observation) y , the goal of denoising is to remove as much of the additive noise e as possible, while preserving the main features of the signal of interest s where, for arguments sake, we have dropped the subscript k . Hence, the denoising problem can be viewed as a nonparametric regression problem and any denoising algorithm can be thought of as an operator D that maps the noisy data y onto an estimate of the signal of interest $\tilde{s} = D(y)$. The precision of the estimate is measured by the expected squared error:

$$R(\tilde{s} - s) = E\|\tilde{s} - s\|_2^2. \quad (13)$$

Donoho and Johnstone (1994, 1995) in their pioneering work show that nonlinear thresholding estimators operating in the wavelet domain achieve nearly minimax risk over a large class of functions that cannot be improved upon over an order of magnitude by any other estimator (Johnstone and Silverman, 1997). This is based on the energy compaction or the sparsity property of the wavelet transform, which can concentrate the signal's energy into a few large magnitude coefficients, whereas the small coefficients are more likely to

be associated with the noise. Hence, noise power can be suppressed by selecting a suitable threshold level λ , and thresholding rule η . The simplest but still most popular thresholding rules are hard and soft thresholding, respectively, given by

$$\eta_\lambda^h(W_y) = W_y \cdot I(|W_y| > \lambda) \quad (14)$$

and

$$\eta_\lambda^s(W_y) = \text{sgn}(W_y) \cdot (|W_y| - \lambda)_+, \quad (15)$$

where $\text{sgn}(\cdot)$ is the sign function and W_y are the wavelet coefficients of observation y .

THE PROPOSED METHOD

In our previous study (Mousavi and Langston, 2016a), we showed that the efficiency of denoising for seismic data can be significantly improved by hybrid approaches that incorporate preprocessing, thresholding, and postprocessing steps. Following this strategy, our denoising approach in the SS-CWT domain is proposed as follows.

Preprocessing

In this step, after transforming the observed data y , into the CWT domain, scales, which purely consist of coefficients associated to the Gaussian noise, are detected and removed from the TFR using the HOS and Kurtosis criteria, leaving the scales with a combination of noise and signal. The kurtosis kurt of N observed coefficients W_y is calculated by (Bickel and Doksum, 1977)

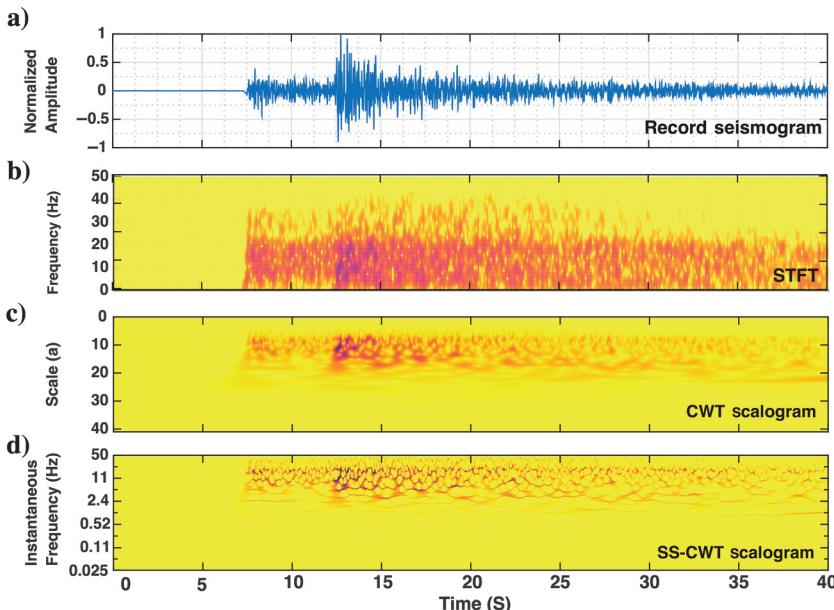


Figure 1. The top panel shows a high-S/N vertical component accelerogram for an M3.5 event recorded at station HALT, Tennessee, USA, in the New Madrid Cooperative Seismic Network. The next panels downward show the STFT, CWT, and SS-CWT, respectively, representing TFRs of the same signal. The variable length of ψ leads to a relatively higher time-frequency resolution of (c) CWT compared with the (b) STFT. However, it still displays spectral smearing due to the finite size of the operator. (d) Note the sharpened TFR obtained by the SS-CWT.

$$\text{kurt}_y = \frac{\sum_{n=1}^N (W_{y_n} - \mu_{W_y})^4}{N\sigma_{W_y}^4} - 3, \quad (16)$$

where σ_{W_y} and μ_{W_y} are, respectively, the estimated standard deviation and mean of wavelet coefficients W_y . The HOS criterion for distinguishing a Gaussian distribution from a non-Gaussian distribution then is defined by

$$|\text{kurt}_y| \leq \frac{\sqrt{24/N}}{\sqrt{1-\alpha}}, \quad (17)$$

where α is the level of confidence. Ravier and Amblard (2001) numerically estimate an optimum value for α , as 90%. This process acts as the equivalent of an automatic band-pass filter and removes noise with lower and higher frequency content compared with the frequency range of seismic signals. This step improves wavelet thresholding results by removing high-power coherent noise (usually associated with colored noise), outside of the frequency range of the seismic signal, from the TFR of data.

General cross-validation thresholding

Then, the preprocessed coefficients are processed to obtain SS-CWT coefficients T_y using equation 11. Major oscillatory components of the signal represented by coefficients in a narrow frequency band along the ridges are then thresholded using Donoho's hard-thresholding scheme (equation 14). This is based on the widely accepted idea that noise is best characterized across the instantaneous frequencies (Ahrabian and Mandic, 2015). The optimal threshold level λ_0 is automatically determined using the GCV approach, proposed and developed by Nason (1996) and Weyrich and Warhol (1995), for each component (ridge). The GCV is used in the statistics as an automatic procedure for selecting optimal smoothing parameters. In the GCV procedure, a data point is systematically excluded from the construction of an estimate, and then the value of the excluded data point is predicted and compared with the true value. Following Jansen et al. (1997), the GCV function is defined as

$$\text{GCV}(\lambda) = \frac{1}{N} \frac{\|T_y - \tilde{T}_\lambda\|^2}{\left\| \frac{N_0}{N} \right\|^2}, \quad (18)$$

where the \tilde{T}_λ are the thresholded coefficients using a threshold value of λ and N_0 is the number of coefficients that would be zeroed using the threshold value λ . This function mimics the errors between the estimation and the true signal; hence, its minimum can be used to select an optimal threshold value. Equation 18 is only a function of λ , and does not rely on any noise-level estimation, which is not a trivial task in the synchrosqueezed domain. Jansen et al. (1997) show that threshold values determined by finding the minimum of the GCV are asymptotically optimal and minimize the mean-square-error R . A grid search or minimization procedure, such as a Fibonacci

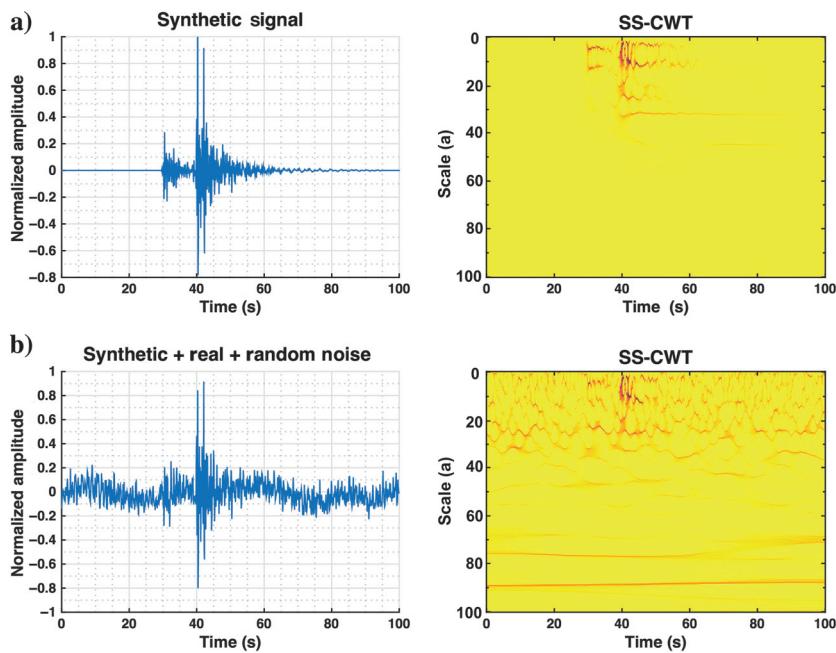


Figure 2. This figure illustrates the process of adding noise to the synthetic seismogram. Each panel shows the annotated time series with its associated SS-CWT. The resulting noisy seismogram contains the long-period and broadband stochastic noise signals.

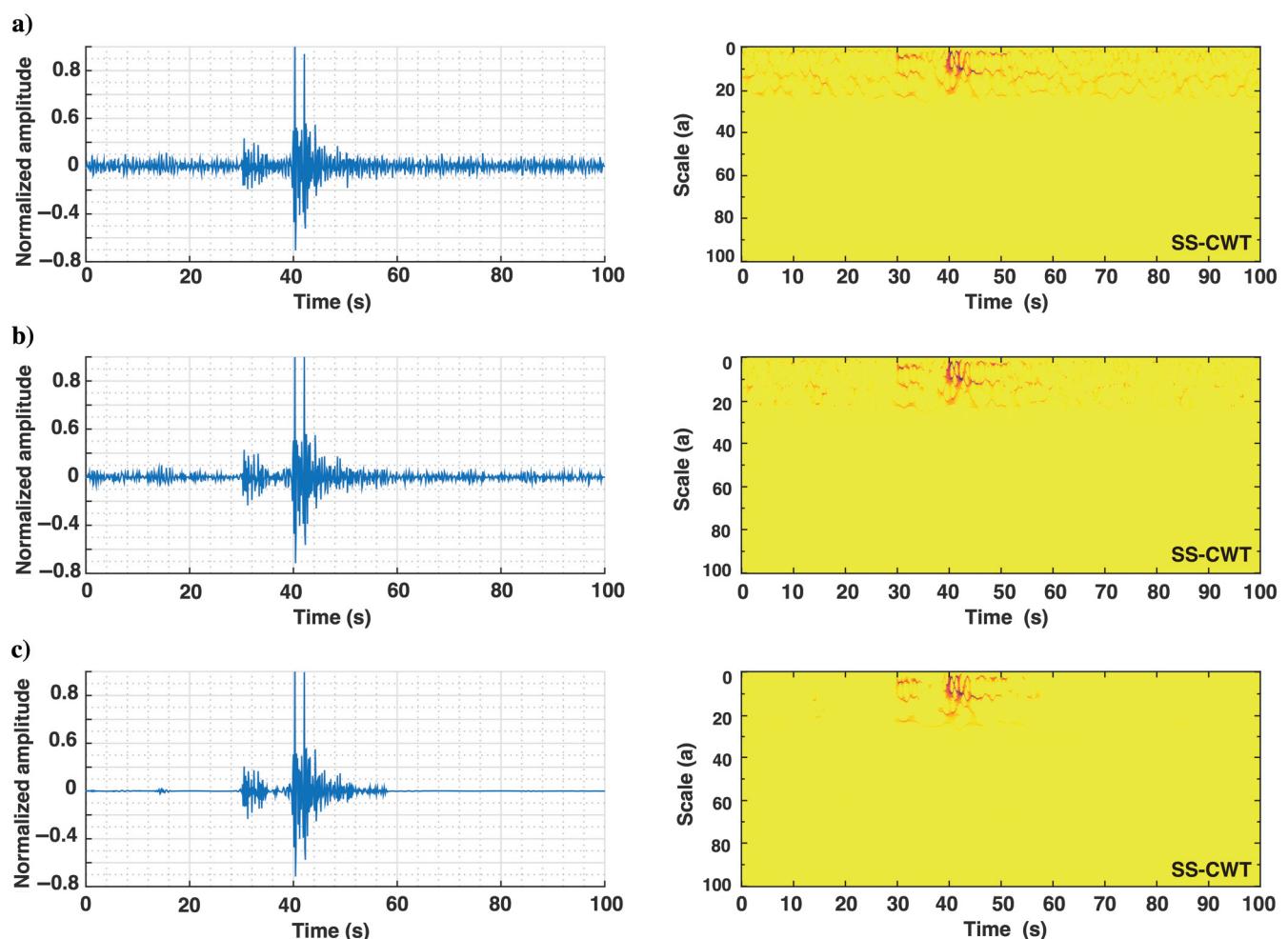


Figure 3. This figure illustrates the noisy synthetic after the (a) preprocessing, (b) GCV thresholding, and (c) postprocessing steps.

search, can then be used to find the optimal threshold λ_0 producing a minimum GCV. Thresholding major components of the signal, in this manner, provides a fast and effective method for increasing the localization of the TFR and obtains an initial estimate of the signal using the inverse transform in equation 12.

Postprocessing

Similar to Ghael et al. (1997), signal estimation is improved with a postprocessing step by applying a simple level-dependent wavelet threshold on the signal obtained from the previous step. For this, the initial estimate of the seismic signal is CWT transformed again and coefficients at all scales are thresholded using the hard-thresholding rule, scale-by-scale. In this step, the threshold value is estimated using the universal threshold of Donoho and Johnstone (1994)

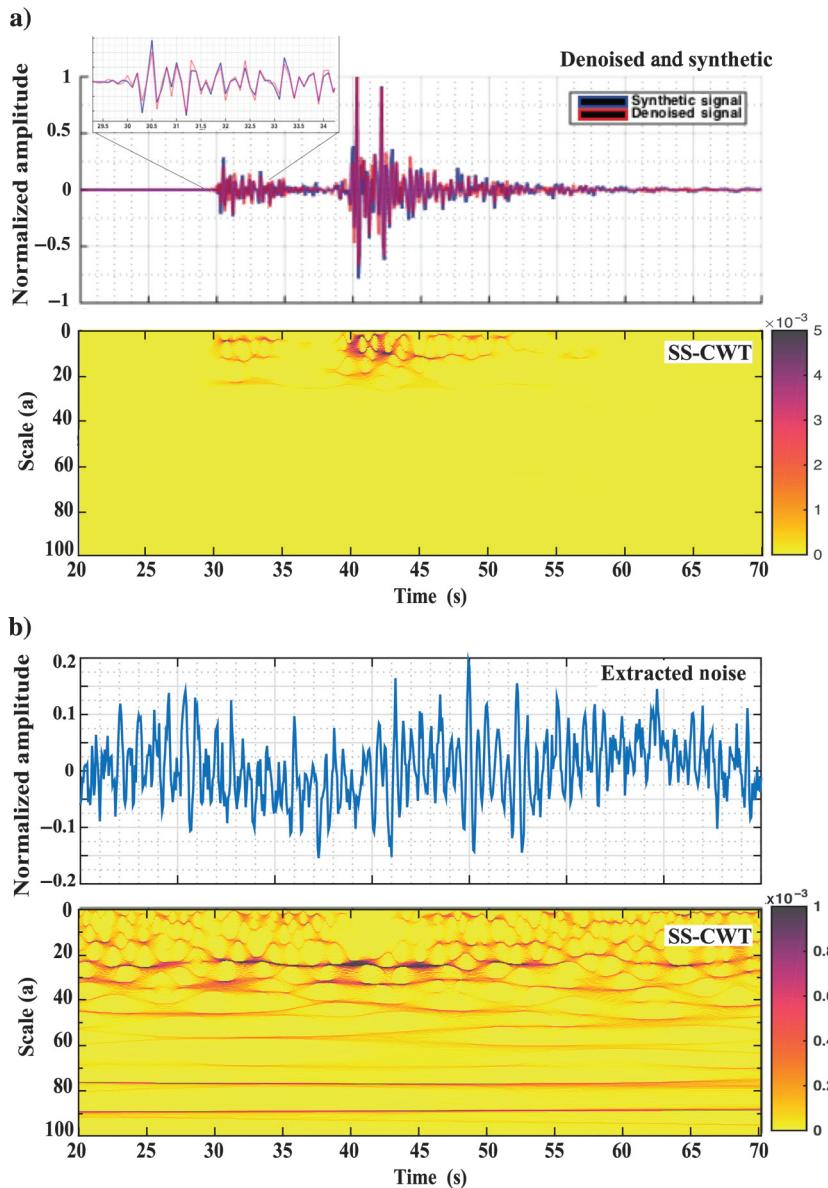


Figure 4. (a) Denoised seismogram and its associated SS-CWT. (b) Extracted noise and its SS-CWT found by removing the signal using a reverse approach.

$$\lambda = \sigma_n \sqrt{2 \ln N}, \quad (19)$$

where the variance of the noise σ_n^2 is estimated in each scale from the median of wavelet coefficients prior to the signal's arrival $\sigma_n = \text{median}(|W_y|)/0.6745$. The final estimate of the denoised seismic signal \tilde{s} is obtained by applying the inverse CWT transform, i.e., equation 8, over the thresholded coefficients.

Using the hard-thresholding scheme makes it easy to implement the method in a reverse manner to remove the signal's energy and keep the noise:

$$\eta_\lambda^{HR}(W_y) = \lambda \cdot I(|W_y| \leq \lambda). \quad (20)$$

The latter will have application in the ambient noise studies.

RESULTS

The algorithm is applied to one synthetic and three field seismograms. In our implementation of the CWT and SS-CWT, we use a Morlet wavelet as the mother wavelet with 100 scales. We have tested different numbers of scales for the wavelet transform to check the sensitivity of the method to the number of decomposition levels. Performance of the denoising algorithm was very similar in all cases, but fewer decomposition levels speed up the denoising process. Another important factor in time-frequency thresholding methods is the selection of an optimal mother wavelet to emphasize the energy of the signal of interest. The proper number of vanishing moments, size of support, regularity, and wavelet families are important parameters in selecting an optimal wavelet type (Luo and Zhang, 2012). We have tested several mother wavelets, such as Morlet, Shannon, Mexican hat, and bump wavelets for the CWT and selected Morlet as the optimal wavelet based on preserving the phase-arrival shapes, S/N, and crosscorrelation coefficients (CCs) between the denoised and original signal. However, based on the characteristics of the signal and the goal of denoising (maximizing the S/N or minimizing the errors), a mother wavelet can be selected.

Approximate arrival times are determined manually. However, this can be automated using any automatic onset picker in the wavelet domain, such as those proposed by Karamzadeh et al. (2013), Bogiatzis and Ishii (2015), or Mousavi et al. (2016a). The S/N is measured as the root-mean-square amplitude in a time window around the signal to a same length window of preceding noise.

Synthetic data

A local synthetic seismogram and its contaminated versions with random and real seismic noise (Figure 2) with a S/N of 2.5 are used for

the synthetic test. The synthetic seismogram is calculated using the frequency-wavenumber method (Zhu and Rivera, 2002). A point source was located at a depth of 12 km, and 3C seismograms were computed for a receiver located on the surface at an epicentral distance of 80 km. Real seismic noise recorded by the New Madrid Cooperative Seismic Network was added to the synthetic seismogram in a way to yield to S/N of 2.5 for the resulting seismogram (Figure 2).

Effects of each step on the noisy trace are presented in Figure 3. The preprocessing step removes those decomposition levels that consisted purely of noise. Hence, it acts like an automatic band-pass filtering (Figure 3a). In the GCV thresholding step, noisy coefficients are attenuated and a more distinct representation of noise and signal is provided (Figure 3b). In the post-processing step, the isolated noisy coefficients remaining after the previous steps are cleaned up (Figure 3c). The denoised and original signals are presented in Figure 4.

The method was successful in removing the random noise and improving the S/N. The denoised and synthetic signals match very well over the entire waveform (Figure 4a) except at the very beginning of the P-wave arrival and end of the P coda. The polarity and amplitude of the first two cycles of the P-wave arrival are preserved very well; however, a very small time shift (less than one sample interval) exists between the denoised and the synthetic signals (Figure 4a). The P-wave arrival that was buried under the noise became clear after the denoising. This can improve arrival time picking and as a result the source location estimation. However, the P coda is smoothed at the very end. The algorithm is also very successful in removing the seismic energy from the waveform (Figure 4b). Comparing scalograms in Figure 4a and 4b, the seismic energy between 30 and 50 s has been removed from the data without changing the time-frequency structure of the background noise in the surrounding areas.

Wavelet power spectra for the denoised (CWT_d) and original data (CWT_o) are shown in Figure 5. To compare the time-frequency structure of these two spectra, we construct a cross-wavelet spectrum (XWT) that highlights regions in time-frequency space in which the two spectra have high common power and represents their local relative phase (Figure 6a). To find regions where the two spectra covary (but do not necessarily have high power), we used wavelet squared coherency (WSC) (Figure 6b). WSC is equivalent to localized correlation in time-frequency space. The equations for XWT and WSC are given in Appendices A and B, respectively.

High-power areas in Figure 6a indicate the correlation of high-magnitude coefficients in the denoised and original CWT spectra. High-power regions within CWT_o and CWT_d coincide for arrival times between 30 and 40 s, indicating

the preservation of P and S energy after denoising. The two waveforms are in-phase for all sectors with significant common power (Figure 6a), but the phase relationship becomes mostly antiphase outside the common power areas. The cross-wavelet power indicates a strong link between the two spectra.

The WSC constructed from spectra of denoised and original data is presented in Figure 6b. Compared with the XWT in addition to

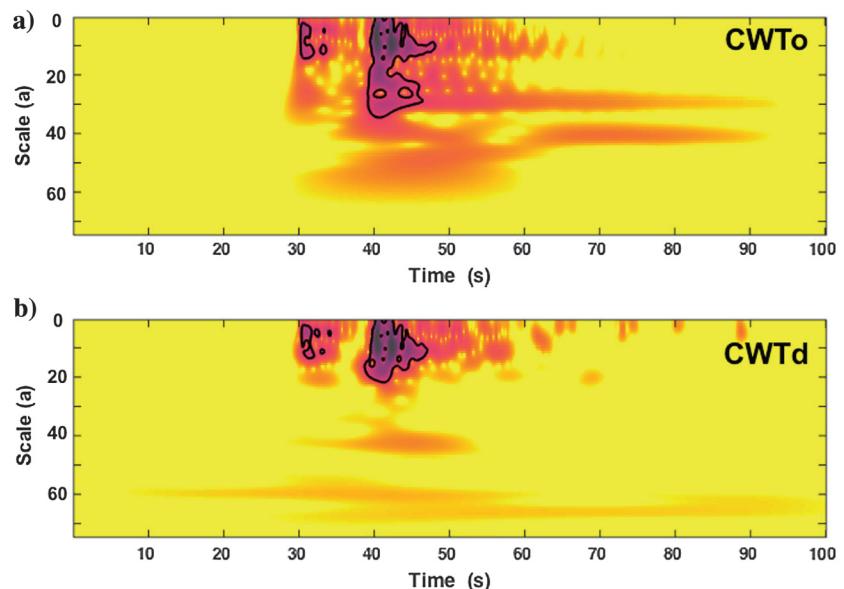


Figure 5. (a) Wavelet spectrogram of synthetic data before adding noise (CWT_o). (b) The same for denoised data (CWT_d).

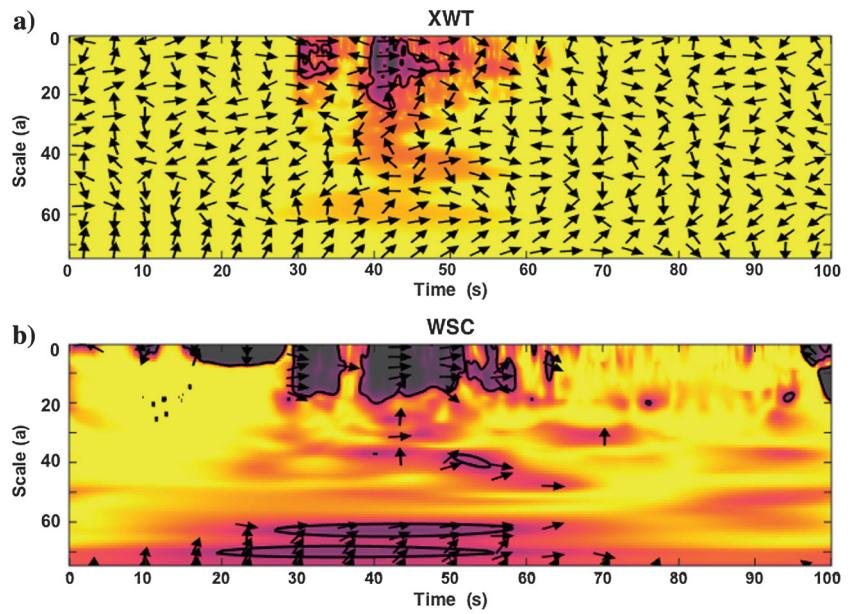


Figure 6. (a) The XWT. The XWT finds regions in time-frequency space, in which two wavelet representations (CWT_o and CWT_d) show high common power. (b) The WSC. The WSC finds regions in time-frequency space, in which the two representations covary (but do not necessarily have high power). The arrows represent local relative phase information. Arrows pointing right, in-phase; left, antiphase, down, original phase leading the denoised phase by 90° ; and up, denoised phase leading original phase by 90° .

the high-power region in lower periods, a relatively large high-power section is present at higher period. An in-phase relationship exists in high-power areas. Low-power regions coincide with low-wavelet powers in the original and denoised scalograms.

To investigate denoising effects on wave polarization, we perform hodogram analyses of the P- and S-wave windows using 3C data before and after the denoising (Figure 7). The particle motions cannot be exactly the same because the amplitude of the denoised and synthetic signals does not remain the same after modification of CWT coefficients during the thresholding. However, the overall direction of motions shown by dashed lines, which represent the average of relative directions of motion for P and S, are quite similar. The relative motions (angles between dashed lines) remained approximately the same after the denoising.

The proposed algorithm has a better performance compared with band-pass filtering, hard thresholding, soft thresholding, and hybrid block thresholding (Figure 8). Band-pass filtering removes noise with frequencies higher and lower than the signal's frequency range, but noise

within the same frequency range of the signal remains untouched (Figure 8b). The poor performance of soft and hard thresholding (Figure 8c and 8d) is due to the presence of high-power features at high scales. These high-power features can be due to either ground-roll (e.g., Chen et al., 2015), tilt (e.g., Crawford and Webb,

Table 1. Root-mean-square (rms) error, signal-to-noise ratio (S/N), and maximum correlation coefficients between the denoised and original signal (CC) from the synthetic test using band-pass filtering between 5 and 20 Hz, hard and soft thresholding (Donoho and Johnstone, 1994), neighboring thresholding (Mousavi and Langston, 2016b), and hybrid block thresholding (Mousavi and Langston, 2016a).

Method	Rms error	S/N	CC	Time (s)	TFT
Band-pass filtering	0.063	5.441	0.683	0.19	—
Hard thresholding	0.061	2.864	0.796	0.50	CWT
Soft thresholding	0.048	3.457	0.833	0.51	CWT
Neighboring thresholding	0.060	750.5	0.721	0.87	STFT
Hybrid block thresholding	0.027	42.831	0.935	9.23	CWT
CGV thresholding	0.025	136.174	0.945	3.38	SS_CWT

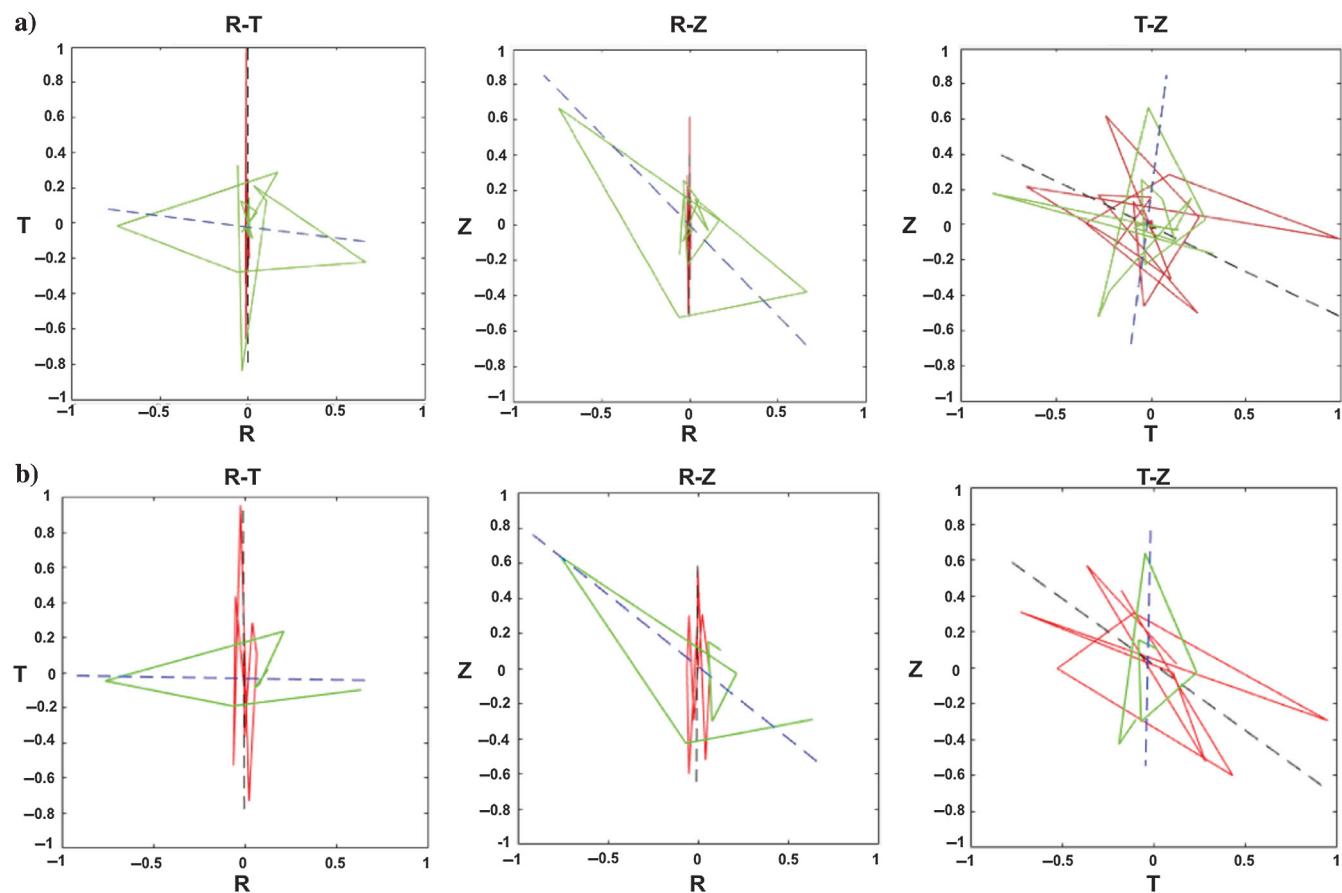


Figure 7. Hodogram analysis of the (a) original signal and (b) denoised one. The P-wave (red) and S-wave (green) particle motions are displayed (radial component R versus transverse T versus vertical Z). The dashed lines show the average particle motions as computed using principal component analysis. The P-wave, S-wave polarization azimuths, and the angle between the P-wave and S-wave particle motions for the seismogram before the denoising are measured 0°, -79°, and 94°, respectively, and changes to 359°, -90°, and 88° for the denoised seismograms.

2000), noises in marine experiments, strong electrical noises (e.g., Castellanos and van der Baan, 2013), long-period-long-duration signals (e.g., Zoback et al., 2012; Caffagni et al., 2015; Zecevic et al., 2016) in microseismic monitoring, very-long-period signals in mining-induced microearthquakes (e.g., Mousavi et al., 2015), or

volcano seismology. These types of features cannot be affected by either a global or level-dependent thresholding. However, their existence can affect severely the performance of denoising in removing high-frequency components of the noise. This is because an assumption of sparsity is the key point in wavelet thresholding

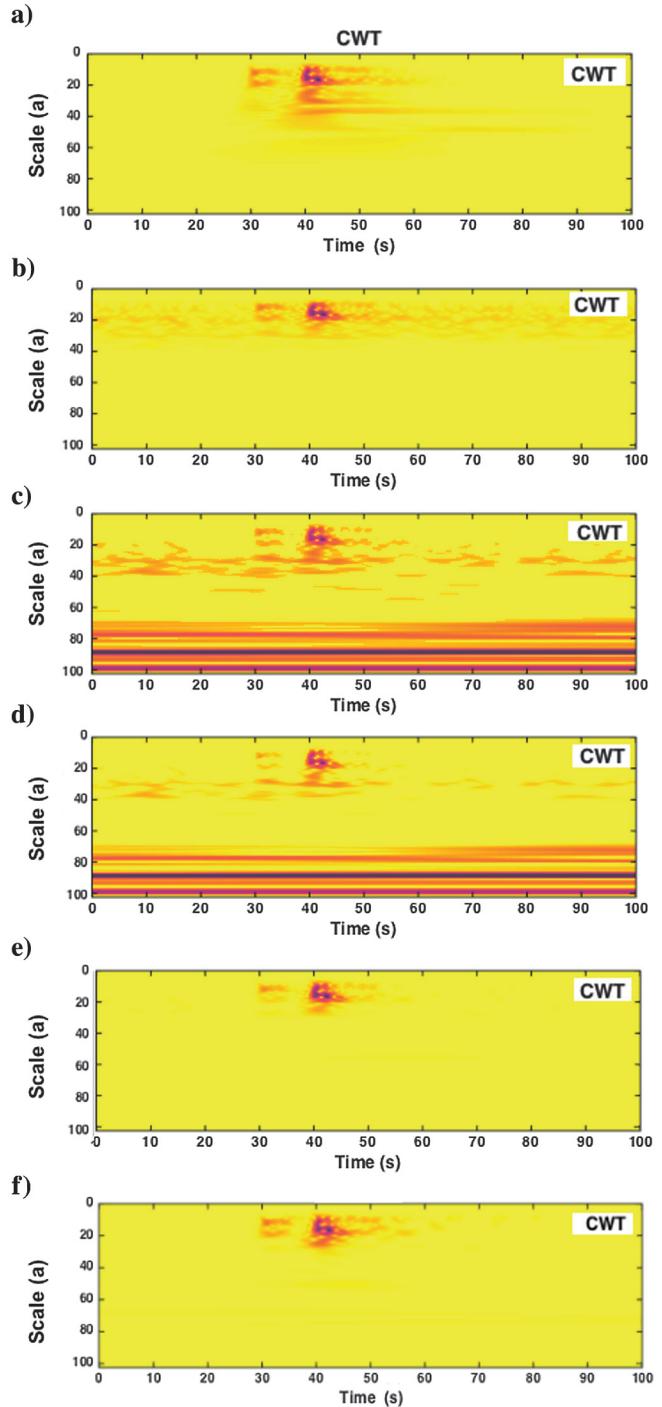
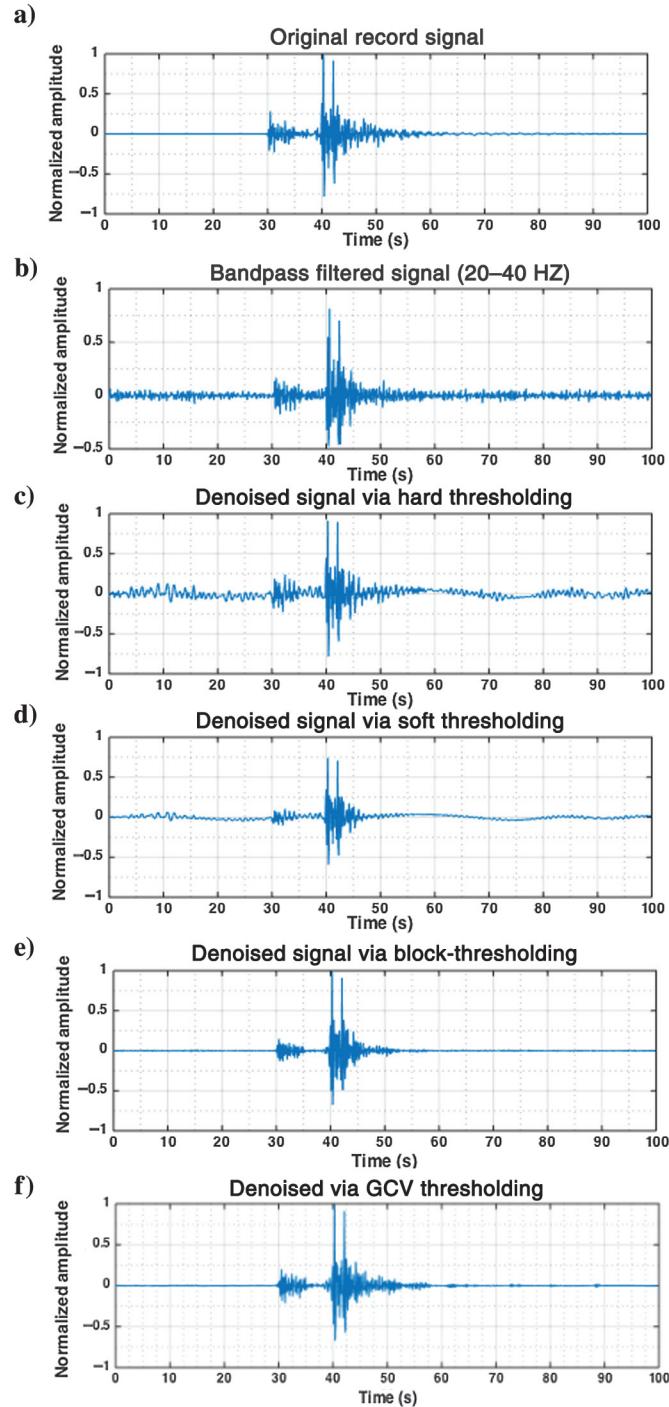


Figure 8. (a) Waveforms and CWT spectrogram of synthetic data. (b) Waveforms and CWT spectrogram after band-pass filtering between 20 and 40 Hz. (c) Waveforms and CWT spectrogram of data after hard thresholding. (d) Waveforms and CWT spectrogram of data after soft thresholding. (e) Waveforms and CWT spectrogram of data after hybrid block thresholding. (f) Waveforms and CWT spectrogram of data after denoising using the GCV thresholding of this study.

because the threshold level is set to separate small-magnitude coefficients, which are assumed to be due to the noise, from high-power coefficients, which are assumed to be due to the signal. Thresholding uses the data itself to decide which coefficients are significant and which are not. The best results were obtained by hybrid block-thresholding (Figure 8e) and the GCV-thresholding

method of this study (Figure 8f). In both methods, high-power long-period features have been removed from TFR by implementing the preprocessing step. However, GCV thresholding obtains much higher S/N (136.17) compared with hybrid block thresholding (42.83) in addition to higher crosscorrelation with the original data (0.945), lower root-mean-squared error (0.025), and reduced

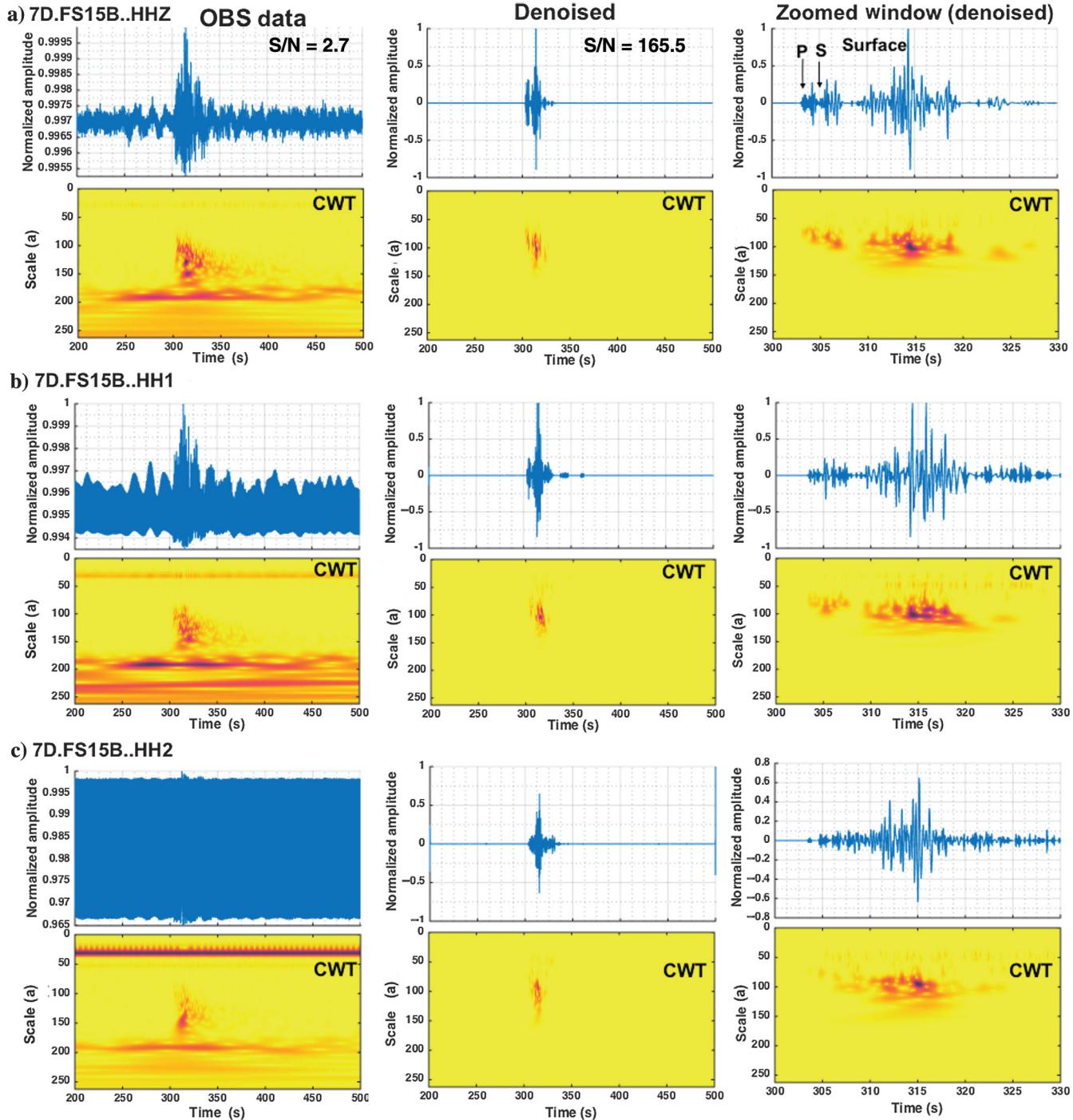
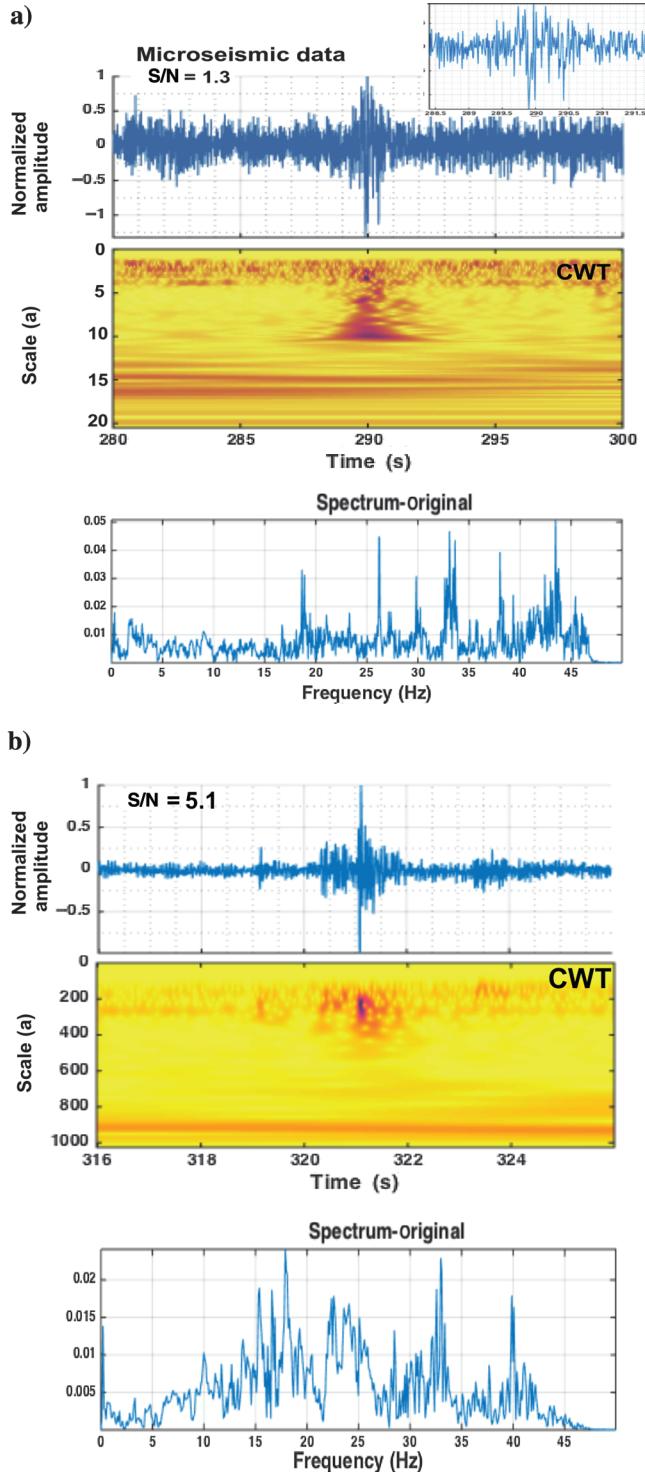


Figure 9. Denoising OBS data of an M4.3 earthquake occurring offshore Petrolia, California, in May 2013 recorded by an OBS (7D.FS15B) from the Cascadia initiative experiment (Toomey et al., 2014). From left to right, each column shows the original time series data and associated CWT of raw data, denoised data using proposed method, and magnified windows around the event on denoised data, respectively. (a) Is the vertical and (b and c) are the horizontal components of motion.

computational time (3.38 s). Moreover, fewer coda were attenuated in the GCV thresholding compared with the block-thresholding method. Quantified comparisons of the proposed algorithm with other methods are presented in Table 1.



Field seismic data

Background noise is a challenging problem encountered in surface monitoring of microseismic events. We have applied the

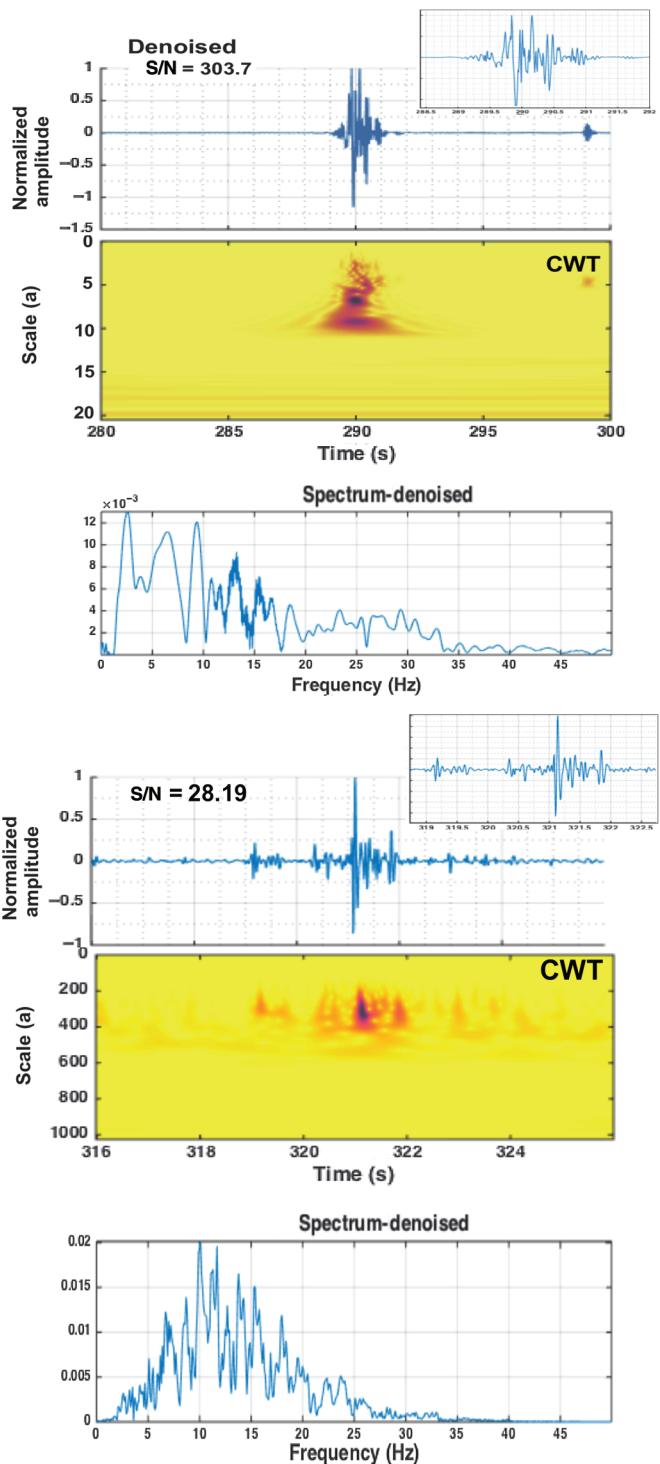


Figure 10. Denoising of two microseismic events induced by wastewater injection in central Arkansas (in 2010). For each event (a and b), the time series and associated CWT representation and a single side amplitude spectra are presented on the left and denoised results are shown on the right. There are high-power bursts at 19, 27, 30, 34, 38, and 44 Hz in the spectra of raw microseismic data that could be associated with electronic noise. The proposed method removed these noises.

algorithm to real seismic data including two microseismic data sets induced during wastewater injection in central Arkansas (Horton, 2012), and an underground collapse of a cavern in Bayou Corne, Louisiana (Mousavi et al., 2016b). Another group of seismic experiments typically known to have high background seismic-noise levels concerns seismic measurements made at the seafloor. Seismic noise at the seafloor is usually long period with frequencies less than 1 Hz. Hence, we have also tested the algorithm on one M4.3 earthquake on the west coast recorded on an OBS during the Cascadia initiative experiment (Toomey et al., 2014). The selected station for OBS data (7D.FS15B) is the shallowest OBS deployed during the Cascadia initiative experiment, hence, resembling the worst-case scenario (A. Barclay and S. Webb, personal communication, 2016).

In addition to high-magnitude noise at very low frequency, some high-frequency noise components are present around scale 40 in the OBS data (Figure 9). This high-frequency noise is stronger on the horizontal components. The proposed hybrid scheme was successful in attenuating coherent and high-power noise at the higher and lower scales and increasing the S/N from 2.7 to 165.5. Moreover, our method was able to automatically remove noise within the same frequency band as the signal and significantly improve the S/N for all 3C. We note that the P, S, and surface waves are preserved. However, some coda are removed from S and surface waves.

In the case of the microearthquakes induced by wastewater injection (Figure 10), noise with higher and lower frequency content compared with the signal frequency was attenuated from the TFR of

data with the S/N increasing from 1.3 to 303.7 (Figure 10a) and from 5.1 to 28.19 (Figure 10b). For the first event, high-energy noise components at 19, 27, 30, 33, 38, and 43 Hz were attenuated from the spectra in addition to some low-frequency components. In the second example (Figure 10b), the dominant energy of the noise is concentrated between 34 and 40 Hz. The denoising algorithm improves the S/N by removing the dominant noise energy and modifying the spectral content of the signal. The P and S arrivals are much clearer after denoising. This can help the phase-arrival time picking and improve source-location estimates.

Our fourth example concerns a case of mining-induced microseismic events with a 7.5 min long vertical component data trace, recorded on 1 November 2013 at Bayou Corne, Louisiana. Seven events associated with an underground collapse of a cavern can be observed on the top seismograms recorded by a 3C, 2 Hz geophone (LA17.01) at the bottom of a borehole (approximately 287 m deep) located at 30.0134°N, 91.1439°W (Figure 11). However, these events are not clear on other raw data (left column) recorded by two other sensors: One is a 3C broadband sensor in the same borehole near the surface (LA17.2 at approximately 190 m deep), and another is a 3C broadband sensor at the surface (LA14) located at 30.0087°N, 91.1398°W (1 km southeast of LA17). From the left panel in Figure 10, we can see that the noise level increases as the sensors become closer to the surface. In addition to coherent and strong long-period noise, some high-frequency noise exists within the surface data. The S/N for the near-surface data was improved after band-pass filtering between 2 and 15 Hz (the middle panel in

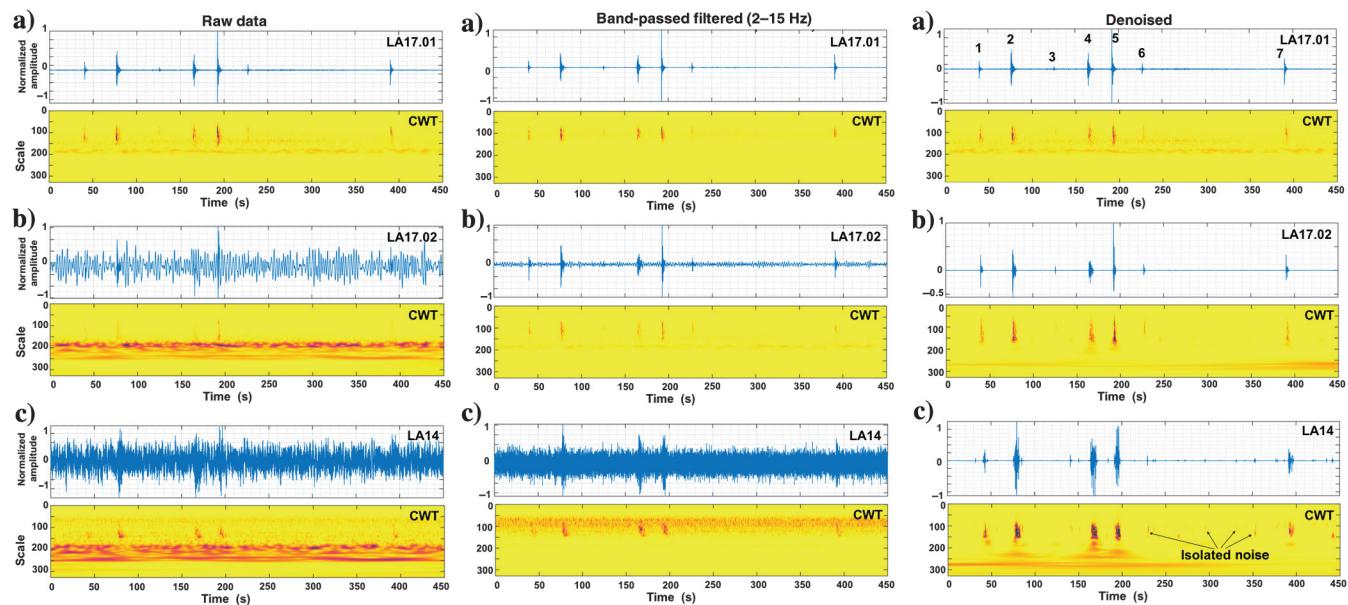


Figure 11. The 7.5 min long vertical seismograms passively recorded in 1 November 2013 at Bayou Corne, Louisiana. (a) A 3C 2 Hz geophone (LA17.01) at the bottom of a borehole (approximately 287 m deep) located at 30.0134°N, 91.1439°W. (b) A 3C broadband sensor (Trillium-compact) (LA17.02) at the top of the same borehole (approximately 190 m deep). (c) A 3C broadband sensor (Trillium-compact) (LA14) at the surface located at 30.0087°N, 91.1398°W (1 km southeast of LA17). The left panels are raw data recorded in these stations and their CWTs. The seven microearthquakes induced by underground collapse of the cavern in the area are observable on the borehole data (LA17.01), whereas near-surface (LA17.02) and surface (LA14) data are much noisier. Middle panels are the same traces after band-pass filtering between 2 and 15 Hz. Because most of the noise in LA17.02 (b) has lower frequencies compared with microseismic events, spectral filtering helps reveal most of the events that were covered under background noise. However, filtering does not improve the S/N at the LA14 because of the presence of some high-frequency noises within the frequency range of seismic events. In the right panel, data are presented after denoising by proposed method of this study. Denoising is successful in removing the noise and significant improvement of the S/N in (b and c). However, some isolated noises were left in (b). Magnified windows around each event are presented in Figure 12.

Figure 11). This is because most of the noise recorded by LA17.02 has a lower frequency compared with the microseismic events. Hence, simple spectral filtering helps to reveal most of the events that were covered by background noise. However, filtering does not improve the S/N of LA14 data because of the presence of some high-frequency noise within the same frequency band as the seismic events. In the right panel (Figure 11), the data are presented after denoising using our method. Significant improvement of the S/N occurs in all three cases indicating that our proposed method is not just limited to random noise, but it is effective for removing colored noise. By applying the method on single-channel data, noise that is not coherent across an array can be detected and attenuated. This S/N enhancement has special importance for microseis-

mic detection, which is a challenging problem in surface monitoring of microseismic events. However, some isolated noise was left even after processing. Such noise can be removed by block thresholding. Magnified windows around each event are presented in Figure 12. Revealing the P-wave arrival that was buried under background noise can facilitate the picking of first-arrival times that will improve the source location estimates and fracture imaging.

In Figure 13, we present results of applying the algorithm in the reverse manner using the OBS and Arkansas microearthquake data to remove the signal's energy and preserve the noise. As one can see from Figure 13, the algorithm is successful in removing the signal even when it is completely buried under the background noise (Figure 13a-3).

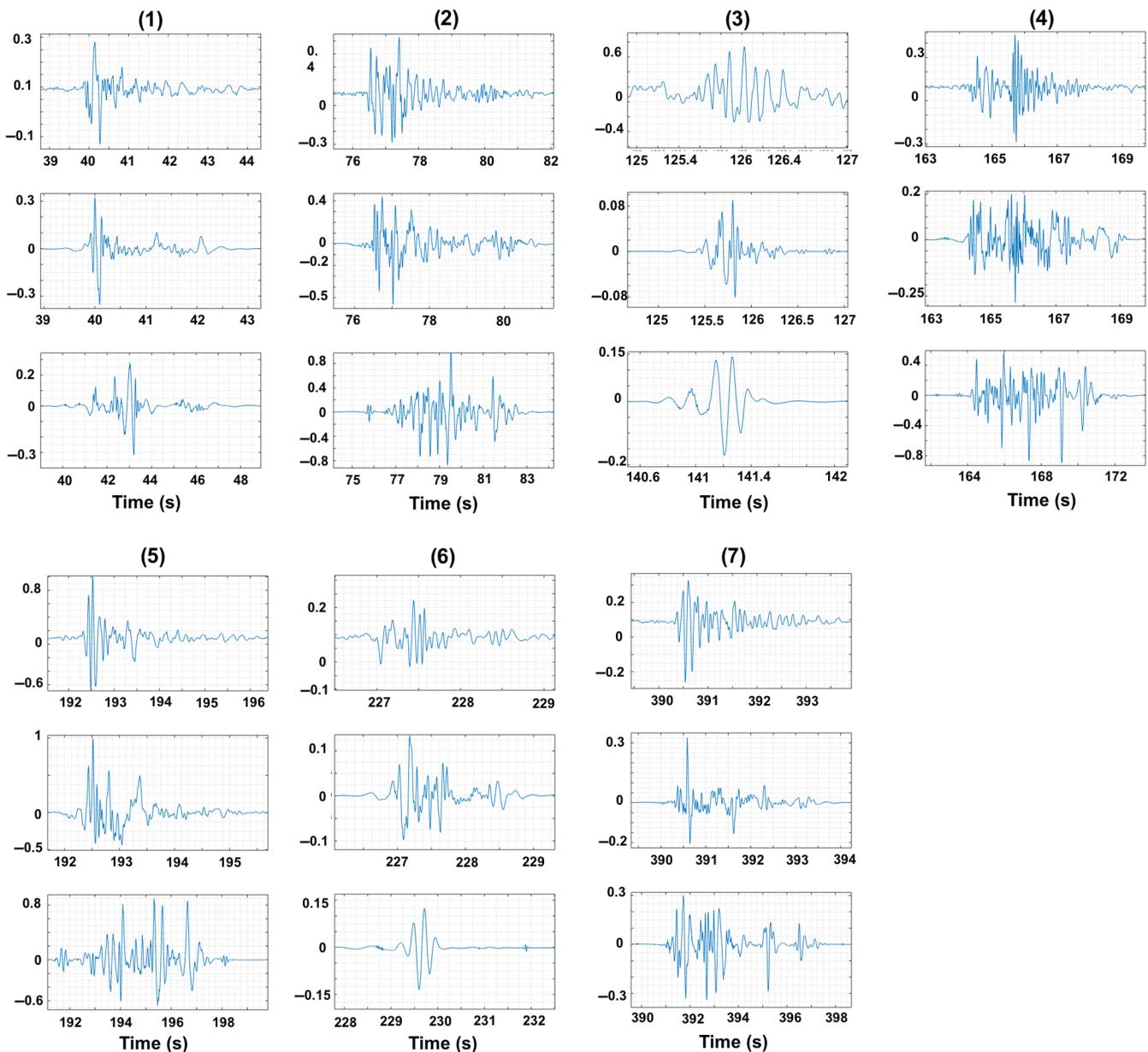


Figure 12. Magnified windows around each microseismic events presented in continues record of Figure 11. In each panel, from top to bottom are raw data recorded on LA17.01 (bottom of the borehole), denoised data recorded on LA17.02 (same location, near the surface), and denoised data recorded on LA14 (at the surface 1 km southeast of the LA17).

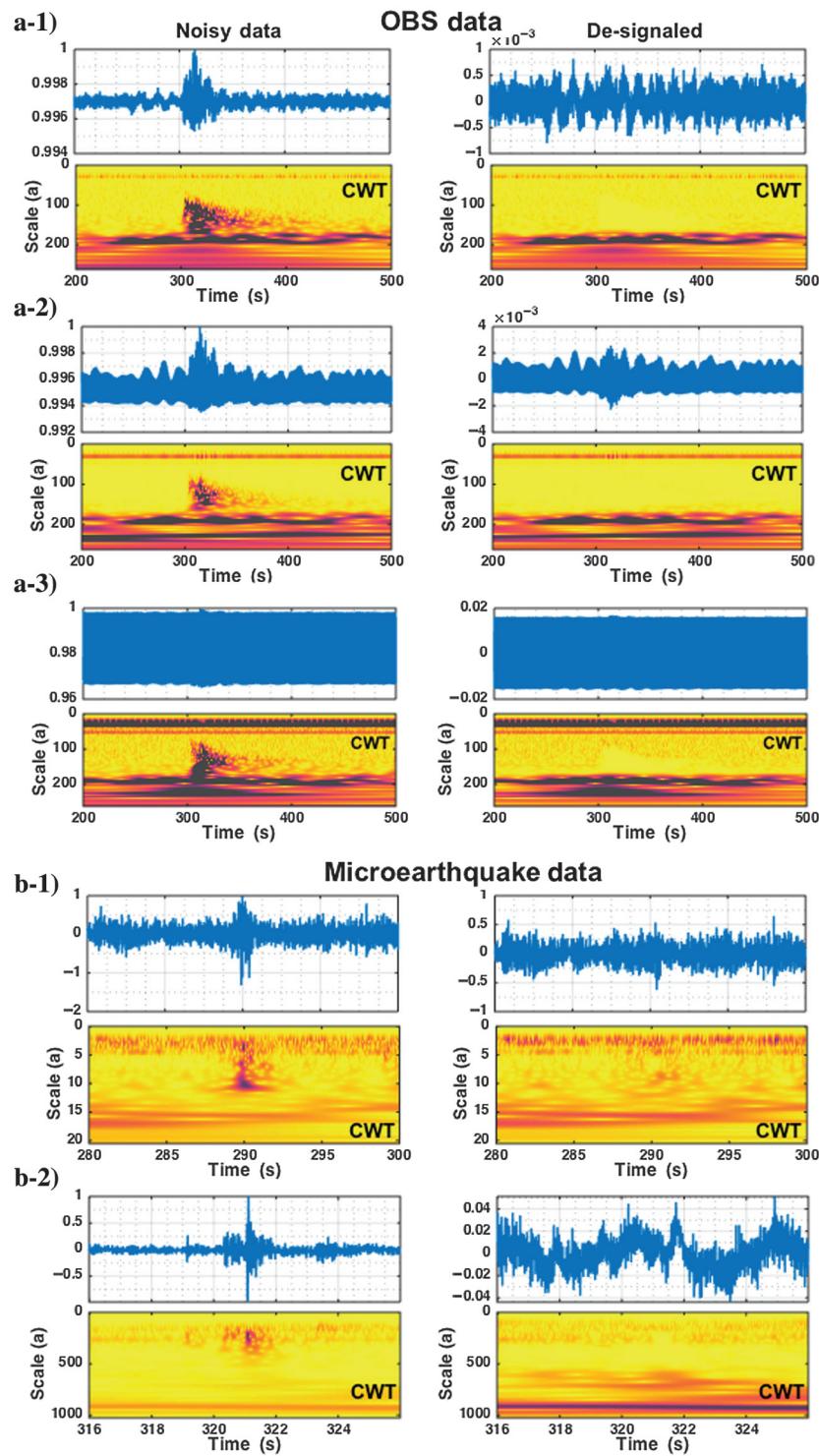


Figure 13. (a) The OBS and (b) microearthquake data as presented in Figures 9 and 10. The left column shows raw data, and the right column shows data after removing the signal's energy (de-signaling) from the traces. As you can see, the algorithm is successful in removing the signal from the waveform event in the case that the signal is completely buried under the background noise (a-3) and it is hard to identify the presence of the signal and removing it using the commonly used time normalizations in ambient noise studies.

DISCUSSION

It is well-known that the typical background seismic noise level is much higher at the seafloor than on land, especially for frequencies less than 1 Hz (e.g., Sutton and Barstow, 1990; Webb et al., 1994; Romanowicz et al., 1998). Efforts have concentrated on attenuating some specific types of noise that have strong effects on the S/N of data in seafloor seismic measurements. Some examples include the proposed method of Crawford and Webb (2000) for identifying and removing tilt noises with frequencies less than 0.1 Hz from vertical data, the proposed method of Webb (1998) for removing compliance noise from OBS data, or the proposed method of Chen et al. (2015) for attenuating ground-roll noise. The effectiveness of our hybrid algorithm for denoising OBS data has been shown in the previous section. Our proposed method is not limited to any specific type of noise, but it can effectively identify and then remove noise from data.

Comparing our denoising method with common methods used for microseismic denoising, the proposed method does not need coherent arrivals in an array, a master event with high S/N, or parameters that need to be tuned manually. It is adaptive to data type and can effectively attenuate the background noise through an automatic procedure. This method can be applied to single-channel data more appropriately for noise that is not coherent within an array and consequently improve source detection and location, which are crucial goals in microseismic monitoring. Moreover, also the analyses performed in this study were limited to the microearthquake and earthquake seismograms, but this method can be applied on the typical reflection data. The single-channel approach of this method makes it possible to combine it with other array-based denoising techniques.

In ambient noise studies, broadband ambient noise needs to be accentuated by attenuating or removing earthquake signals that tend to dominate. This is usually done by band-pass filtering the raw data followed by a temporal normalization to reduce the effect of crosscorrelating earthquakes. Time normalization is one of the most important steps in data preparation and can be done by one-bit normalization, clipping, automatic event detection and removing, running-absolute-mean normalization, or water-level normalization (Bensen et al., 2007). The first three methods are based on the simplified assumptions and remove a large amount of information from the waveform. The last two methods are known to be more effective, but they rely on more parameters that need to be tuned for optimal performance (Bensen et al., 2007). However, the signal removal approach proposed in this paper

can be an effective procedure. The data can be processed automatically with high flexibility and adaptability. It can find buried signals and remove them from the waveform without affecting the time-frequency structure of the original data. Moreover, after crosscorrelation, the denoising scheme can also be used to improve the S/N of the recovered Green's functions. In ambient crosscorrelation, the input time series are often very long. This makes it hard to recover high-fidelity signals. Baig et al. (2009) show that time-frequency denoising of correlograms can alleviate this problem. Hence, the proposed method can be used again in a straightforward way to improve the S/N of Green's functions and make it possible to constructed high-fidelity Green's functions from shorter time series. This will be the subject of a future study.

This method has many different applications. For instance, in the attenuation estimation, it can assemble the spectral content of the phase of interest more precisely and decrease the uncertainties (McNamara et al., 2012; Mousavi et al., 2014; Tary et al., 2016).

CONCLUSION

We have proposed a new and fast algorithm for accurate noise removal/signal removal based on HOS, GCV, and WHT in the synchrosqueezed domains. Performance of the proposed algorithm was tested using synthetic and real seismic data and showed improvements over our previous method. The denoising procedure proposed here is a powerful data-driven method that can significantly improve S/N and lower the detection threshold for small seismic events. This automatic algorithm can remove the high- and low-frequency seismic noise and retrieve the seismic signal with full features, such as dominant phases, polarity, and spectral content. The method can be used for single-component data and is applicable to land and ocean-bottom data processing.

ACKNOWLEDGMENTS

We thank P. Ogwari for providing the information about microseismic events. C. A. Langston was partially supported by the AFRL Contract #FA9453-16-C-0015 during this research. We thank D. Velis and three anonymous reviewers for the careful reviews and constructive comments.

APPENDIX A

CROSS-WAVELET TRANSFORM

The XWT is constructed from the wavelet coefficients of the original signal (CWT_o) and the denoised signal (CWT_d), and it will expose their common power and relative phase in the time-frequency space. The XWT measures the similarity of the wavelet representations of two signals and provides the ability to account for temporal (or spatial) variability in spectral character. Following Torrence and Compo (1998), the cross-wavelet transform is defined as

$$\text{XWT}_{o,d} = \text{CWT}_o \times \text{CWT}_d^*, \quad (\text{A-1})$$

where $*$ denotes the complex conjugation. The cross-wavelet power is defined as $|\text{XWT}_{o,d}|$, and the complex argument $\arg(\text{XWT}_{o,d})$ can be interpreted as the local relative phase between the denoised and original seismograms (Grinsted and Jevrejeva, 2004).

APPENDIX B

WAVELET-SQUARED COHERENCY

The coherency between two CWTs can be measured by wavelet coherency. The WSC is defined as the absolute value squared of the smoothed XWT, normalized by the smoothed wavelet power spectra (Torrence and Webster, 1999):

$$R^2(a) = \frac{|\langle a^{-1}\text{XWT}(a) \rangle|^2}{\langle a^{-1}|\text{CWT}_o(a)|^2 \rangle \langle a^{-1}|\text{CWT}_d(a)|^2 \rangle}, \quad (\text{B-1})$$

where $\langle \cdot \rangle$ indicates the smoothing in time and scale. The factor a^{-1} is used to convert to an energy density. The wavelet-coherency phase difference is given by

$$\phi = \tan^{-1} \left(\frac{\Im\{\langle a^{-1}\text{XWT} \rangle\}}{\Re\{\langle a^{-1}\text{XWT} \rangle\}} \right), \quad (\text{B-2})$$

where \Im is the smoothed imaginary part and \Re is the smoothed real part of cross-wavelet coefficients. The smoothing is done using a weighted running average (or convolution) in the scale and time. The time smoothing uses a filter given by the absolute value of wavelet coefficient at each scale, normalized to have a total weight of unity. For the Morlet mother wavelet, this is just a Gaussian $\exp(-t^2/(2a^2))$. For scale smoothing of the Morlet wavelet, a boxcar filter of width $\delta j_0 = 0.60$ (from Torrence and Webster, 1999) was used.

REFERENCES

- Ahrabian, A., and D. P. Mandic, 2015, A class of multivariate denoising algorithms based on synchrosqueezing: IEEE Transactions on Signal Processing, **63**, 2196–2208, doi: [10.1109/TSP.2015.2404307](https://doi.org/10.1109/TSP.2015.2404307).
- Allen, J. B., and L. R. Rabiner, 1977, A unified approach to short-time Fourier analysis and synthesis: Proceedings of the IEEE, **65**, 1558–1564, doi: [10.1109/PROC.1977.10770](https://doi.org/10.1109/PROC.1977.10770).
- Auger, F., and P. Flandrin, 1995, Improving the readability of time-frequency and time-scale representations by the reassignment method: IEEE Transactions on Signal Processing, **43**, 1068–1089, doi: [10.1109/78.382394](https://doi.org/10.1109/78.382394).
- Auger, F., P. Flandrin, Y. T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H. T. Wu, 2013, Time-frequency reassignment and synchrosqueezing: An overview: IEEE Signal Processing Magazine, **30**, 32–41, doi: [10.1109/MSP.2013.2265316](https://doi.org/10.1109/MSP.2013.2265316).
- Baig, A., M. Campillo, and F. Brenguier, 2009, Denoising seismic noise cross correlations: Journal of Geophysical Research, **114**, B08310, doi: [10.1029/2008JB006085](https://doi.org/10.1029/2008JB006085).
- Bekara, M., and M. van der Baan, 2009, Random and coherent noise attenuation by empirical mode decomposition: Geophysics, **74**, no. 5, V89–V98, doi: [10.1190/1.3157244](https://doi.org/10.1190/1.3157244).
- Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang, 2007, Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements: Geophysical Journal of International, **169**, 1239–1260.
- Bickel, P. J., and K. A. Doksum, 1977, Mathematical statistics: Bickel.
- Boagtzis, P., and M. Ishii, 2015, Continuous wavelet decomposition algorithms for automatic detection of compressional- and shear-wave arrival times: Bulletin of Seismological Society of America, **105**, 1628–1641, doi: [10.1785/0120140267](https://doi.org/10.1785/0120140267).
- Bonar, D., and M. Sacchi, 2012, Denoising seismic data using the nonlocal means algorithm: Geophysics, **77**, no. 1, A5–A8, doi: [10.1190/geo2011-0235.1](https://doi.org/10.1190/geo2011-0235.1).
- Caffagni, E., D. Eaton, M. van der Baan, and J. P. Jones, 2015, Regional seismicity: A potential pitfall for identification of long-period long-duration events: Geophysics, **80**, no. 1, A1–A5, doi: [10.1190/geo2014-0382.1](https://doi.org/10.1190/geo2014-0382.1).
- Carmona, R. A., W. L. Hwang, and B. Torresani, 1997, Characterization of signals by the ridges of their wavelet transforms: IEEE Transactions on Signal Processing, **45**, 2586–2590.
- Castellanos, F., and M. van der Baan, 2013, Microseismic event locations using the double-difference algorithm: CSEG Recorder, **38**, 26–37.
- Chassande-Mottin, E., I. Daubechies, F. Auger, and P. Flandrin, 1997, Differential reassignment: IEEE Signal Processing Letters, **4**, 293–294, doi: [10.1109/97.633772](https://doi.org/10.1109/97.633772).
- Chen, Y., S. Jiao, J. Ma, H. Chen, Y. Zhou, and S. Gan, 2015, Ground-roll noise attenuation using a simple and effective approach based on local

- band-limited orthogonalization: IEEE Geoscience and Remote Letters, **12**, 2316–2320, doi: [10.1109/LGRS.2015.2475280](https://doi.org/10.1109/LGRS.2015.2475280).
- Chen, Y., and J. Ma, 2014, Random noise attenuation by F-X empirical mode decomposition predictive filtering: *Geophysics*, **79**, no. 3, V81–V91, doi: [10.1190/geo2013-0080.1](https://doi.org/10.1190/geo2013-0080.1).
- Chen, Y., J. Ma, and S. Fomel, 2016, Double-sparsity dictionary for seismic noise attenuation: *Geophysics*, **81**, no. 2, V17–V30, doi: [10.1190/geo2014-0525.1](https://doi.org/10.1190/geo2014-0525.1).
- Chen, Y. C., M. Y. Cheng, and H. T. Wu, 2014, Non-parametric and adaptive modelling of dynamic periodicity and trend with heteroscedastic and dependent errors: *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, **76**, 651–682, doi: [10.1111/rssb.12039](https://doi.org/10.1111/rssb.12039).
- Crawford, W. C., and S. C. Webb, 2000, Identifying and removing tilt noise from low-frequency (<0.1 Hz) seafloor vertical seismic data: *Bulletin of Seismological Society of America*, **90**, 952–963.
- Daubechies, I., and C. Heil, 1992, Ten lectures on wavelets: *Computers in physics*: AIP Publishing, 697.
- Daubechies, I., J. Lu, and H. Tieng Wu, 2011, Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool: *Applied and Computational Harmonic Analysis*, **30**, 243–261, doi: [10.1016/j.acha.2010.08.002](https://doi.org/10.1016/j.acha.2010.08.002).
- Daubechies, I., and S. Maes, 1996, A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models: *Wavelets in Medicine and Biology*, 527–546.
- Ditommaso, R., M. Mucciarelli, M. R. Gallipoli, and F. C. Ponzo, 2010, Effect of a single vibrating building on free-field ground motion: Numerical and experimental evidences: *Bulletin of Earthquake Engineering*, **8**, 693–703, doi: [10.1007/s10518-009-9134-5](https://doi.org/10.1007/s10518-009-9134-5).
- Ditommaso, R., M. Mucciarelli, and F. C. Ponzo, 2012, Analysis of non-stationary structural systems by using a band-variable filter: *Bulletin of Earthquake Engineering*, **10**, 895–911, doi: [10.1007/s10518-012-9338-y](https://doi.org/10.1007/s10518-012-9338-y).
- Donoho, D. L., and I. M. Johnstone, 1994, Ideal spatial adaption by wavelet shrinkage: *Biometrika*, **81**, 425–455, doi: [10.1093/biomet/81.3.425](https://doi.org/10.1093/biomet/81.3.425).
- Donoho, D. L., and I. M. Johnstone, 1995, Adapting to unknown smoothness via wavelet shrinkage: *Journal of the American Statistical Association*, **90**, 1200–1224.
- Douglas, A., 1997, Bandpass filtering to reduce noise on seismograms: Is there a better way?: *Bulletin of Seismological Society of America*, **87**, 770–777.
- Farge, M., 2003, Wavelet transforms and their applications to turbulence, <http://www.annualreviews.org/doi/abs/10.1146/annurev.fl.24.010192.002143>, accessed January 2016.
- Gabor, D., 1946, Theory of communication. Part 1: The analysis of information: *Journal of the Institution of Electrical Engineers — Part III: Radio and Communication Engineering*, **93**, 429–441, doi: [10.1049/ji-3-2.1946.0074](https://doi.org/10.1049/ji-3-2.1946.0074).
- Galiana-Merino, J. J., J. Rosa-Herranz, J. Giner, S. Molina, and F. Botella, 2003, De-noising of short-period seismograms by wavelet packet transform: *Bulletin of Seismological Society of America*, **93**, 2554–2562, doi: [10.1785/0120010133](https://doi.org/10.1785/0120010133).
- Ghoshal, S. P., A. M. Sayeed, and R. G. Baraniuk, 1997, Improved wavelet denoising via empirical wiener filtering: *Proceedings of SPIE, Mathematical Imaging*, 389–399, doi: [10.1117/12.29799](https://doi.org/10.1117/12.29799).
- Gómez, J. L., and D. R. Velis, 2016, A simple method inspired by empirical mode decomposition for denoising seismic data: *Geophysics*, **81**, no. 6, V403–V413, doi: [10.1190/GEO2015-0566.1](https://doi.org/10.1190/GEO2015-0566.1).
- Grinsted, A., J. C. Moore, and S. Jevrejeva, 2004, Application of the cross wavelet transform and wavelet coherence to geophysical time series: *Nonlinear Processes in Geophysics*, **11**, 561–566, doi: [10.5194/npg-11-561-2004](https://doi.org/10.5194/npg-11-561-2004).
- Hall, M., 2006, Resolution and uncertainty in spectral decomposition: First Break, **24**, 43–47, doi: [10.3997/1365-2397.2006027](https://doi.org/10.3997/1365-2397.2006027).
- Han, J., and M. van der Baan, 2015, Microseismic and seismic denoising via ensemble empirical mode decomposition and adaptive thresholding: *Geophysics*, **80**, no. 6, KS69–KS80, doi: [10.1190/geo2014-0423.1](https://doi.org/10.1190/geo2014-0423.1).
- Herrera, R. H., J. Han, and M. van der Baan, 2014, Applications of the synchrosqueezing transform in seismic time-frequency analysis: *Geophysics*, **79**, no. 3, V55–V64, doi: [10.1190/geo2013-0204.1](https://doi.org/10.1190/geo2013-0204.1).
- Herrera, R. H., J. B. Tary, M. van der Baan, and D. W. Eaton, 2015, Body wave separation in the time-frequency domain: IEEE Geoscience and Remote Sensing Letters, **12**, 364–368, doi: [10.1109/LGRS.2014.2342033](https://doi.org/10.1109/LGRS.2014.2342033).
- Horton, S., 2012, Disposal of hydrofracking waste fluid by injection into subsurface aquifers triggers earthquake swarm in central Arkansas with potential for damaging earthquake: *Seismological Research Letters*, **83**, 250–260.
- Hou, T., Z. Shi, and P. Tavallali, 2012, Nonparametric and adaptive modeling of dynamic seasonality and trend with heteroscedastic and dependent errors: arXiv:1210.4672 [math.ST].
- Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, 1998, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **454**, 903–995, doi: [10.1098/rspa.1998.0193](https://doi.org/10.1098/rspa.1998.0193).
- Huang, W., R. Wang, Y. Chen, H. Li, and S. Gan, 2016, Damped multichannel singular spectrum analysis for 3D random noise attenuation: *Geophysics*, **81**, no. 4, V261–V270, doi: [10.1190/geo2015-0264.1](https://doi.org/10.1190/geo2015-0264.1).
- Huang, Z. L., J. Zhang, T. H. Zhao, and Y. Sun, 2015, Synchrosqueezing s-transform and its application in seismic spectral decomposition: *IEEE Transactions on Geoscience and Remote Sensing*, **54**, 817–825, doi: [10.1109/TGRS.2015.2466660](https://doi.org/10.1109/TGRS.2015.2466660).
- Iatsenko, D., P. V. E. McClintock, and A. Stefanovska, 2015, Linear and synchrosqueezed time-frequency representations revisited: Overview, standards of use, resolution, reconstruction, concentration, and algorithms: *Digital Signal Processing: A Review Journal*, **42**, 1–26, doi: [10.1016/j.dsp.2015.03.004](https://doi.org/10.1016/j.dsp.2015.03.004).
- Jansen, M., M. Malfait, and A. Bultheel, 1997, Generalized cross validation for wavelet thresholding: *Signal Processing*, **56**, 33–44, doi: [10.1016/S0165-1684\(97\)83621-3](https://doi.org/10.1016/S0165-1684(97)83621-3).
- Johnstone, I. M., and B. W. Silverman, 1997, Wavelet threshold estimators for data with correlated noise: *Journal of the Royal Statistical Society*, **59**, 319–351.
- Karamzadeh, N., G. J. Doloei, and A. Reza, 2013, Automatic earthquake signal onset picking based on the continuous wavelet transform: *IEEE Transactions on Geoscience and Remote Sensing*, **51**, 2666–2674, doi: [10.1109/TGRS.2012.2213824](https://doi.org/10.1109/TGRS.2012.2213824).
- Li, H., R. Wang, S. Cao, Y. Chen, and W. Huang, 2016, A method for low-frequency noise suppression based on mathematical morphology in microseismic monitoring: *Geophysics*, **81**, no. 3, V159–V167, doi: [10.1190/geo2015-0222.1](https://doi.org/10.1190/geo2015-0222.1).
- Luo, G., and D. Zhang, 2012, Wavelet denoising, in *Advances in wavelet theory and their applications in engineering*, in D. Baleanu, ed., *Physics and technology*: InTech, doi: [10.5772/37424](https://doi.org/10.5772/37424).
- Mallat, S., 1999, A wavelet tour of signal processing, in *Wavelet analysis and its applications* (2nd ed.): Academic Press, 620.
- Meignen, S., T. Oberlin, and S. McLaughlin, 2012, A new algorithm for synchrosqueezing: With an application to multicomponent signals sampling and denoising: *IEEE Transactions on Signal Processing*, **60**, 5787–5798.
- McNamara, D. E., M. Meremonte, J. Z. Maharry, S.-L. Mildore, J. R. Altidore, D. Anglade, S. E. Hough, D. Given, H. Benz, and L. Gee, 2012, Frequency-dependent seismic attenuation within the Hispaniola Island region of the Caribbean sea: *Bulletin of the Seismological Society of America*, **102**, 773–782, doi: [10.1785/0120110137](https://doi.org/10.1785/0120110137).
- Mousavi, S. M., C. H. Cramer, and C. A. Langston, 2014, Average QLg, QSn, and observation of Lg blockage in the continental Margin of Nova Scotia: *Journal of Geophysical Research*, **119**, 7722–7744, doi: [10.1002/2014JB011237](https://doi.org/10.1002/2014JB011237).
- Mousavi, S. M., S. P. Horton, and C. L. Langston, 2015, Differences between locations from surface seismometers and downhole geophones at the Napoleonville salt dome Louisiana: Presented at the Annual Meeting, SSA.
- Mousavi, S. M., S. P. Horton, C. A. Langston, and B. Samei, 2016b, Seismic features and automatic discrimination of deep and shallow induced-micro-earthquakes using neural network and logistic regression: *Geophysical Journal International*, **207**, 29–46, doi: [10.1093/gji/ggw258](https://doi.org/10.1093/gji/ggw258).
- Mousavi, S. M., and C. A. Langston, 2016a, Hybrid seismic denoising using higher-order statistics and improved wavelet block thresholding: *Bulletin of the Seismological Society of America*, **106**, 1380–1393, doi: [10.1785/0120150345](https://doi.org/10.1785/0120150345).
- Mousavi, S. M., and C. A. Langston, 2016b, Adaptive noise estimation and suppression for improving microseismic event detection: *Journal of Applied Geophysics*, **132**, 116–124, doi: [10.1016/j.jappgeo.2016.06.008](https://doi.org/10.1016/j.jappgeo.2016.06.008).
- Mousavi, S. M., and C. A. Langston, 2016c, Fast and novel microseismic detection using time-frequency analysis: 86th Annual International Meeting, SEG, Expanded Abstracts, 2632–2636, doi: [10.1190/segam2016-13262278](https://doi.org/10.1190/segam2016-13262278).
- Mousavi, S. M., and C. A. Langston, 2016d, Adaptive microseismic noise estimation and denoising: 86th Annual International Meeting, SEG, Expanded Abstracts, 4772–4776, doi: [10.1190/segam2016-13262233.1](https://doi.org/10.1190/segam2016-13262233.1).
- Mousavi, S. M., C. A. Langston, and S. P. Horton, 2016a, Automatic microseismic denoising and onset detection using the synchrosqueezed-continuous wavelet transform: *Geophysics*, **81**, no. 4, V341–V355, doi: [10.1190/GEO2015-0598.1](https://doi.org/10.1190/GEO2015-0598.1).
- Naghizadeh, M., 2011, Seismic data interpolation and denoising in the frequency-wavenumber domain: *Geophysics*, **77**, no. 2, V71–V80, doi: [10.1190/GEO2011-0172.1](https://doi.org/10.1190/GEO2011-0172.1).
- Naghizadeh, M., and M. Sacchi, 2012, Multicomponent seismic random noise attenuation via vector autoregressive operators: *Geophysics*, **77**, no. 2, V91–V99, doi: [10.1190/geo2011-0198.1](https://doi.org/10.1190/geo2011-0198.1).
- Nason, G. P., 1996, Wavelet shrinkage using cross-validation: *Journal of the Royal Statistical Society. Series B (Methodological)*, **58**, 463–479, doi: [10.1111/j.144.5690](https://doi.org/10.1111/j.144.5690).

- Oropeza, V., and M. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: *Geophysics*, **76**, no. 3, V25–V32, doi: [10.1190/1.3552706](https://doi.org/10.1190/1.3552706).
- Parolai, S., 2009, Denoising of seismograms using the s transform: *Bulletin of Seismological Society of America*, **99**, 226–234, doi: [10.1785/0120080001](https://doi.org/10.1785/0120080001).
- Pazos, A., M. J. Gonz, and G. Alguacil, 2003, Non-linear filter, using the wavelet transform, applied to seismological records: *Journal of Seismology*, **7**, 413–429.
- Pinnegar, C. R., and D. W. Eaton, 2003, Application of the S transform to prestack noise attenuation filtering: *Journal of Geophysical Research*, **108**, 2422, doi: [10.1029/2002JB002258](https://doi.org/10.1029/2002JB002258).
- Ravier, P., and P. O. Amblard, 2001, Wavelet packets and denoising based on higher-order-statistics for transient detection: *Signal Processing*, **81**, 1909–1926.
- Reine, C., M. van der Baan, and R. Clark, 2009, The robustness of seismic attenuation measurements using fixed- and variable-window time-frequency transforms: *Geophysics*, **74**, no. 2, WA123–WA135, doi: [10.1190/1.3043726](https://doi.org/10.1190/1.3043726).
- Romanowicz, B., D. Stakes, J. P. Montagner, P. Tarits, R. Urhammer, M. Begnaud, and E. Stutzmann, 1998, MOISE: A pilot experiment towards long term sea-floor geophysical observatories: *Earth Planets Space*, **50**, 927–937.
- Sabbione, J. I., M. D. Sacchi, and D. R. Velis, 2013, Microseismic data denoising via an apex-shifted hyperbolic Radon transform: 83rd Annual International Meeting, SEG, Expanded Abstracts, 2155–2161.
- Sabbione, J. I., M. D. Sacchi, and D. R. Velis, 2015, Radon transform-based microseismic event detection and signal-to-noise ratio enhancement: *Journal of Applied Geophysics*, **113**, 51–63.
- Scherbaum, F., 2001, Of poles and zeros: Fundamentals of digital seismology: Kluwer Academic Publishers.
- Schimmel, M., and J. Gallart, 2007, Frequency-dependent phase coherence for noise suppression in seismic array data: *Journal of Geophysical Research: Solid Earth*, **112**, 1–14, doi: [10.1029/2006JB004680](https://doi.org/10.1029/2006JB004680).
- Shuchong, L., and C. N. Xun, 2014, Seismic signals wavelet packet denoising method based on improved threshold function and adaptive threshold: *Computer Modeling and New Technologies*, **18**, 1291–1296.
- Sobolev, G., and A. Lyubushin, 2006, Microseismic impulses as earthquake precursors: *Izvestiya: Physics of the Solid Earth*, **42**, 721–733, doi: [10.1134/S1069351306090023](https://doi.org/10.1134/S1069351306090023).
- Sutton, G. H., and N. Barstow, 1990, Ocean bottom ultra-low frequency (ULF) seismo-acoustic ambient noise: 0.002–0.4 Hz: *Journal of the Acoustical Society of America*, **87**, 2005–2012.
- Tary, J. B., R. H. Herrera, J. Han, and M. van der Baan, 2014, Spectral estimation: What is new? What is next?: *Reviews of Geophysics*, **52**, 723–749, doi: [10.1002/2014RG000461](https://doi.org/10.1002/2014RG000461).
- Tary, J. B., M. van der Baan, and R. H. Herrera, 2016, Attenuation estimation using high resolution time-frequency transforms: *Digital Signal Processing*, **60**, 46–55, doi: [10.1016/j.dsp.2016.08.007](https://doi.org/10.1016/j.dsp.2016.08.007).
- Thakur, G., E. Breydo, N. S. Fućkar, and H. T. Wu, 2013, The synchrosqueezing algorithm for time-varying spectral analysis: Robustness properties and new paleoclimate applications: *Signal Processing*, **93**, 1079–1094, doi: [10.1016/j.sigpro.2012.11.029](https://doi.org/10.1016/j.sigpro.2012.11.029).
- Thakur, G., and H. T. Wu, 2011, Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples: *SIAM Journal on Mathematical Analysis*, **43**, 2078–2095, doi: [10.1137/100798818](https://doi.org/10.1137/100798818).
- Toomey, D. R., R. M. Allen, A. H. Barclay, S. W. Bell, P. D. Bromirski, R. L. Carlson, X. Chen, J. A. Collins, R. P. Dziak, B. Evers, D. W. Forsyth, P. Gerstoft, E. E. E. Hooft, D. Livelybrooks, J. A. Lodewyk, D. S. Luther, J. J. McGuire, S. Y. Schwartz, M. Tolstoy, A. M. Tréhu, M. Weirathmueller, and W. S. D. Wilcock, 2014, The Cascadia Initiative: A sea change in seismological studies of subduction zones: *Oceanography*, **27**, 138–150.
- Torrence, C., and G. F. Compo, 1998, A practical guide to wavelet analysis: *Bulletin of the American Meteorological Society*, **79**, 61–78.
- Torrence, C., and P. Webster, 1999, Interdecadal changes in the ESNOMonsoon system: *Journal of Climate*, **12**, 2679–2690.
- Tselentis, G. A., N. Martakis, P. Paraskevopoulos, A. Lois, and E. Sokos, 2012, Strategy for automated analysis of passive microseismic data based on s-transform, Otsu's thresholding, and higher order statistics: *Geophysics*, **77**, no. 6, KS43–KS54, doi: [10.1190/geo2011-0301.1](https://doi.org/10.1190/geo2011-0301.1).
- Velis, D., J. I. Sabbione, and M. D. Sacchi, 2015, Fast and automatic microseismic phase-arrival detection and denoising by pattern recognition and reduced-rank filtering: *Geophysics*, **80**, no. 6, WC25–WC38, doi: [10.1190/GEO2014-0561.1](https://doi.org/10.1190/GEO2014-0561.1).
- Webb, S. C., 1998, Broadband seismology and noise under the ocean: *Reviews of Geophysics*, **36**, 105–142, doi: [10.1029/97RG02287](https://doi.org/10.1029/97RG02287).
- Webb, S. C., W. C. Crawford, and J. A. Hildebrand, 1994, Long period seismometer deployed at OSN-1, OSN Newsletter: *Seismic Waves*, **3**, 4–6.
- Weyrich, N., and G. Warhola, 1995, De-noising using wavelets and cross validation: Springer.
- Yang, H., 2015, Synchrosqueezed wave packet transforms and diffeomorphism based spectral analysis for 1D general mode decompositions: *Applied and Computational Harmonic Analysis*, **39**, 33–66, doi: [10.1016/j.acha.2014.08.004](https://doi.org/10.1016/j.acha.2014.08.004).
- Yang, H., and L. Ying, 2014, Synchrosqueezed curvelet transform for two-dimensional mode decomposition: *The SIAM Journal on Mathematical Analysis*, **46**, 2052–2083.
- Zecevic, M., G. Daniel, and D. Jurick, 2016, On the nature of long-period long-duration seismic events detected during hydraulic fracturing: *Geophysics*, **81**, no. 3, KS113–KS121, doi: [10.1190/geo2015-0524.1](https://doi.org/10.1190/geo2015-0524.1).
- Zhang, L., Y. Wang, Y. Zheng, and X. Chang, 2015, Deblending using a high-resolution radon transform in a common midpoint domain: *Journal of Geophysics and Engineering*, **12**, 167–174.
- Zhu, L., and L. A. Rivera, 2002, A note on the dynamic and static displacements from a point source in multi-layered media: *Geophysical Journal International*, **148**, 619–627.
- Zoback, M. D., A. Kohli, I. Das, and M. McClure, 2012, The importance of slow slip on faults during hydraulic fracturing stimulation of shale gas reservoirs: *American Unconventional Resources Conference, SPE*, doi: [10.2118/155476-MS](https://doi.org/10.2118/155476-MS).