

# Separating Signal from Noise and from Other Signal Using Nonlinear Thresholding and Scale-Time Windowing of Continuous Wavelet Transforms

by Charles A. Langston and Seyed Mostafa Mousavi

**Abstract** A procedure for removing noise or signal from seismic time series using the continuous wavelet transform (CWT) is developed through the common assumption of noise stationarity for pre-event or postevent estimates of the noise. Noise and signal are efficiently separated using nonlinear thresholding of the CWT avoiding computationally intensive block thresholding algorithms on the wavelet scale-time plane. Efficiency is gained by estimating the characteristic statistics of pre-event noise using empirical cumulative distribution functions and then using these characteristics to threshold the entire time series using hard or soft nonlinear thresholding. In addition, scale-time windowing of the CWT scalogram and inverse transforming into the time domain allows unprecedented control in partitioning a seismogram into component wave types that can subsequently be used to infer characteristics of Earth structure and source excitation. Noise can be separated from signal and signals decomposed into discrete wave groups. CWT techniques offer unique and intuitive alternatives to traditional Fourier methods for analyzing noise and signal useful for structure and source studies, event detection, and ambient-noise interferometry.

## Introduction

The separation of noise from signal remains a fundamental problem in seismic analysis. Recently, we have investigated a number of different ways to determine noise and signal using time–frequency representations (TFRs) of seismic signals (Mousavi and Langston, 2016a, 2016b, 2017; Mousavi *et al.*, 2016). The basic idea behind all techniques involves manipulation of 2D mapping in which the location of a “signal” is discriminated from “noise” on a plane consisting of complex amplitudes determined from the transform of seismic data. A common transform is the short time-window Fourier transform (STFT), but other transforms such as the continuous wavelet transform (CWT; Grossmann *et al.*, 1989; Starck *et al.*, 2010) and the synchrosqueezed CWT (SS-CWT; Daubechies and Maes, 1996; Daubechies *et al.*, 2011) may represent the time-series data in more compact form. The STFT is the basis for the common spectrogram in which short running time windows of the data are Fourier transformed and then the amplitude or power spectra are plotted as a function of time. The CWT may be implemented with a choice of different wavelet functions with wavelet coefficients at different scales being plotted against time. The SS-CWT represents another processing step in which the CWT is modified by assigning the energy of closely adjacent wavelet coefficients to ridges in the CWT time-scale map. Both the CWT and the SS-CWT can be manipulated

to represent the seismic signal (and the noise) by a smaller number of transform coefficients.

One theme that occurred within these studies was the issue of algorithm speed versus fidelity of the denoised signal. More operations on a data trace and its TFR increased the number of processor cycles. Higher fidelity noise reduction with extra processing steps, such as using the SS-CWT with general cross-validation thresholding (Mousavi and Langston, 2017), significantly increased processing time that made denoising large array datasets impractical. Another drawback of these studies was the relative difficulty in obtaining an intuitive view of the signal and noise on the TFR plane. Gaining insight into the signal and noise is routinely done by a seismic analyst assuming various band-pass filters and observing their effect on the data time series. Characteristics of the noise can be quickly seen and appropriate filters for the task at hand can be used to consistently process the data. Although adaptive processing of the noise using wavelet methods somewhat removes the analyst from the picture, there still remain a number of parameters to vary with judgments to make on the quality of the results. The analyst remains part of the process with additional needs for translating complex mathematical rules into acceptable data outcomes.

The purpose of this article is to present a relatively simple way to both visualize and manipulate seismic data using

nonlinear thresholding of CWTs of the data. We use a common assumption that both simplifies and speeds up the processing while at the same time makes denoising much more intuitive using the CWT. Many seismic signals have distinct onsets in time for particular frequency or wavelet scale bands. It is then a straightforward matter to choose a pre-event time window to estimate the noise, assuming the noise process is stationary, and then use the estimated noise in subsequent thresholding. There is nothing new in this assumption because this has been the basis of signal-to-noise ratio (SNR) estimates using Fourier methods for countless source and structure studies throughout the years. Using two different nonlinear thresholding criteria from the noise estimate allows efficient removal of the noise throughout the time series or, if wanted, removal of the signal.

Furthermore, it becomes a simple step to directly manipulate the CWT by choosing portions of the CWT scale-time plane to dissect phases from interesting seismic traces. The fact that the CWT has an inverse transform to reconstruct the time-domain signal allows for creative ways to analyze seismic data. We will demonstrate the utility of these techniques on local distance seismic data from the 2016 Incorporated Research Institutions for Seismology (IRIS) wavefields seismic experiment in northern Oklahoma (Sweet *et al.*, 2018). The waveform data come from a series of small explosions detonated by the Air Force Research Laboratory (AFRL) (C. Zeiler, personal comm., 2017) within 70 km of the instruments of the experiment. The IRIS experiment is described in detail by Sweet *et al.* (2018). We will be using example waveforms from only a small subset of the 247 three-component nodal seismometers that were deployed to demonstrate CWT techniques. A more thorough analysis of structure and array techniques will be the subject of a future article.

### The CWT and Nonlinear Thresholding

The Fourier transform is a staple in seismology because of its relation to theory of linear systems and use in solving theoretical problems in wave propagation. Fourier transformation of a time-domain signal into the frequency domain allows integral operations in time, such as convolution or correlation, to become simple algebraic operations in the frequency domain. But because the Fourier basis functions are infinitely long sinusoids for the Fourier transform, or periodic sinusoids for the discrete Fourier transform, signal information at particular times spreads over the entire frequency band.

The CWT, on the other hand, is a transform with two independent variables, scale  $a$ , and time lag  $\tau$ , which produces a map of amplitude at scale versus time lag for the signal. The Morlet–Grossmann definition (Grossmann *et al.*, 1989; Starck *et al.*, 2010) for the CWT is

$$W(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi^* \left( \frac{t - \tau}{a} \right) dt, \quad (1)$$

in which the asterisk represents the complex conjugate of the function. The CWT is simply a correlation of the signal,  $f(t)$ , with a scaled basis function  $\psi(t)$ . In general, the basis function is complex and is termed the “mother wavelet.” The wavelet coefficient  $W(a, \tau)$  is also complex and can be represented in the Fourier domain as

$$\hat{W}(a, \omega) = \sqrt{a} \hat{f}(\omega) \hat{\psi}^*(a\omega). \quad (2)$$

The CWT is a linear operation and has an exact inverse transform given by the double integral

$$f(t) = \frac{1}{C} \int_0^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{a}} W(a, \tau) \psi \left( \frac{t - \tau}{a} \right) \frac{dad\tau}{a^2}, \quad (3)$$

in which  $C$  is found from a Parseval-like integral

$$C = \int_0^{+\infty} \frac{\hat{\psi}^*(\omega) \hat{\psi}(\omega)}{\omega} d\omega \quad (4)$$

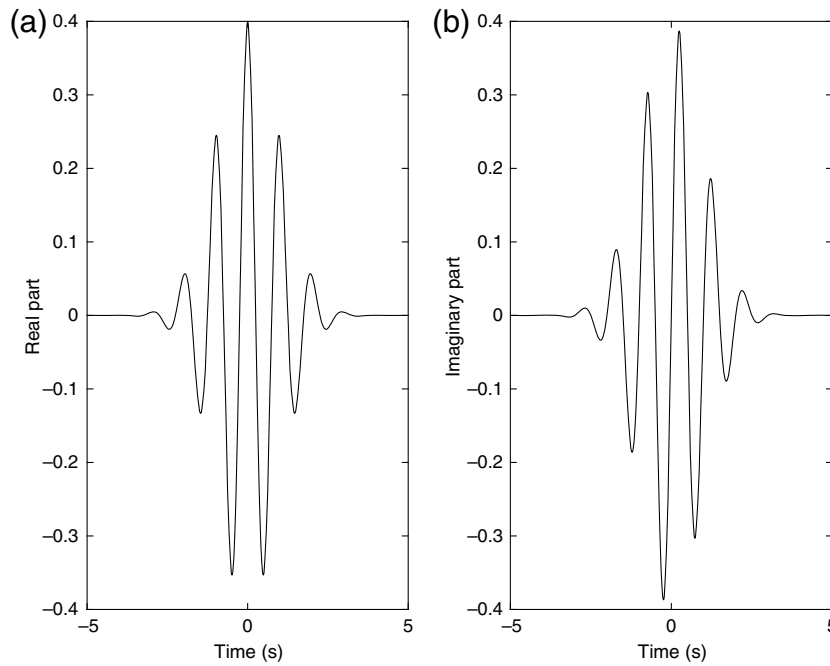
that requires the basis function to have zero mean for this integral to be bounded at  $\omega = 0$ . We use Morlet’s wavelet (Goupillaud *et al.*, 1985) in which

$$\begin{aligned} \text{Re}(\psi(t)) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \cos(2\pi\vartheta_0 t) \\ \text{Im}(\psi(t)) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \sin(2\pi\vartheta_0 t). \end{aligned} \quad (5)$$

As can be seen (Fig. 1), Morlet’s wavelet is a heavily damped sinusoid around  $t = 0$ , allowing the scale parameter to be interpreted as Fourier period (inverse frequency) in the CWT domain because we use  $\vartheta_0 = 1$ . Morlet’s wavelet is not strictly admissible because it has a finite mean, but the mean is negligible for  $\vartheta_0 > 0.8$  and gives satisfactory results. Other wavelets may be used, and the techniques outlined here will work in a similar way. However, the strength of the Morlet wavelet is that it is harmonic-like so that the CWT can be interpreted by relating scale to Fourier period.

Figure 2 illustrates the use of the CWT. The vertical-component data come from a 2000 lbs explosion recorded at about 35 km distance from station 5014 within the IRIS experiment. The raw time-series data are transformed into the CWT domain to form a map of wavelet amplitude as a function of scale and time lag. The signal is seen both in the seismogram and in the wavelet map (scalogram) between 65 and 110 s arrival time and 0.1–1 s scale. This TFR is an intuitive way of viewing the characteristics of the signal that is seemingly buried by the noise. The signal is evident at scales shorter than 1 s and is seen above the horizontal bands of lower frequency noise. Essentially, the map is a tool to view the signal in regions where the noise has much less amplitude than the signal. These areas of high SNR are of interest in partitioning the signal from the noise.

Partitioning can be accomplished several ways. An estimate of the noise power is made in a time window before the



**Figure 1.** (a) Real and (b) imaginary parts of the Morlet wavelet.

first arrival of the event (Fig. 2a) by summing the absolute value of the amplitude of the coefficients at each scale over time lag. This can then be used with several different criteria to remove wavelet coefficients from the map.

“Hard thresholding” is the nonlinear process of keeping wavelet coefficients if they are greater than a threshold criterion  $\beta(a)$ ; otherwise, they are set to zero (Donoho *et al.*, 1995). Mathematically, this is represented by

$$\tilde{W}(a, \tau) = \begin{cases} W(a, \tau) & \text{if } |W(a, \tau)| \geq \beta(a) \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Alternatively, thresholding can be done in a somewhat less severe manner by shrinking all coefficients that survive by the inferred noise level. This is called soft thresholding (Weaver *et al.*, 1991) and is given by

$$\tilde{W}(a, \tau) = \begin{cases} \text{sign}[W(a, \tau)](|W(a, \tau)| - \beta(a)) & \text{if } |W(a, \tau)| \geq \beta(a) \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

in which

$$\text{sign}[W(a, \tau)] = \frac{W(a, \tau)}{|W(a, \tau)|} \quad (8)$$

Soft thresholding minimizes outliers in the noise better than hard thresholding because all surviving coefficients are reduced. Noise outliers just above the threshold amplitude will have an effective SNR close to unity and will experience greater reduction compared to high SNR coefficients.

The threshold function  $\beta(a)$ , is determined based on the statistics of the absolute value of the noise estimate and is determined for each wavelet scale  $a$ . Mousavi and Langston (2016a) describe a denoising method that breaks up a wavelet plane into separate blocks and attempts to determine which blocks are noise-like and which contain signal. A threshold level is determined for each block based on the ensemble and applied to remove noise using a separate method from the hard or soft thresholding methods shown earlier. Determining the threshold requires solving an inverse problem for each scale, which tends to take significant time.

In this article, we suggest a simpler method in which the statistics of the noise are estimated from a time window before or after the signal and then used to estimate the threshold function. Once the threshold function is determined, nonlinear thresholding of the entire signal becomes straight-

forward. Much of the signal processing literature starts from the assumption of Gaussian noise. In this case, the threshold function can be computed using the mean and standard deviation of the absolute value of the wavelet transform at each scale:

$$\beta(a) = \text{mean}(|W(a, \tau)|) + c \text{ st.dev.}(|W(a, \tau)|), \quad (9)$$

in which

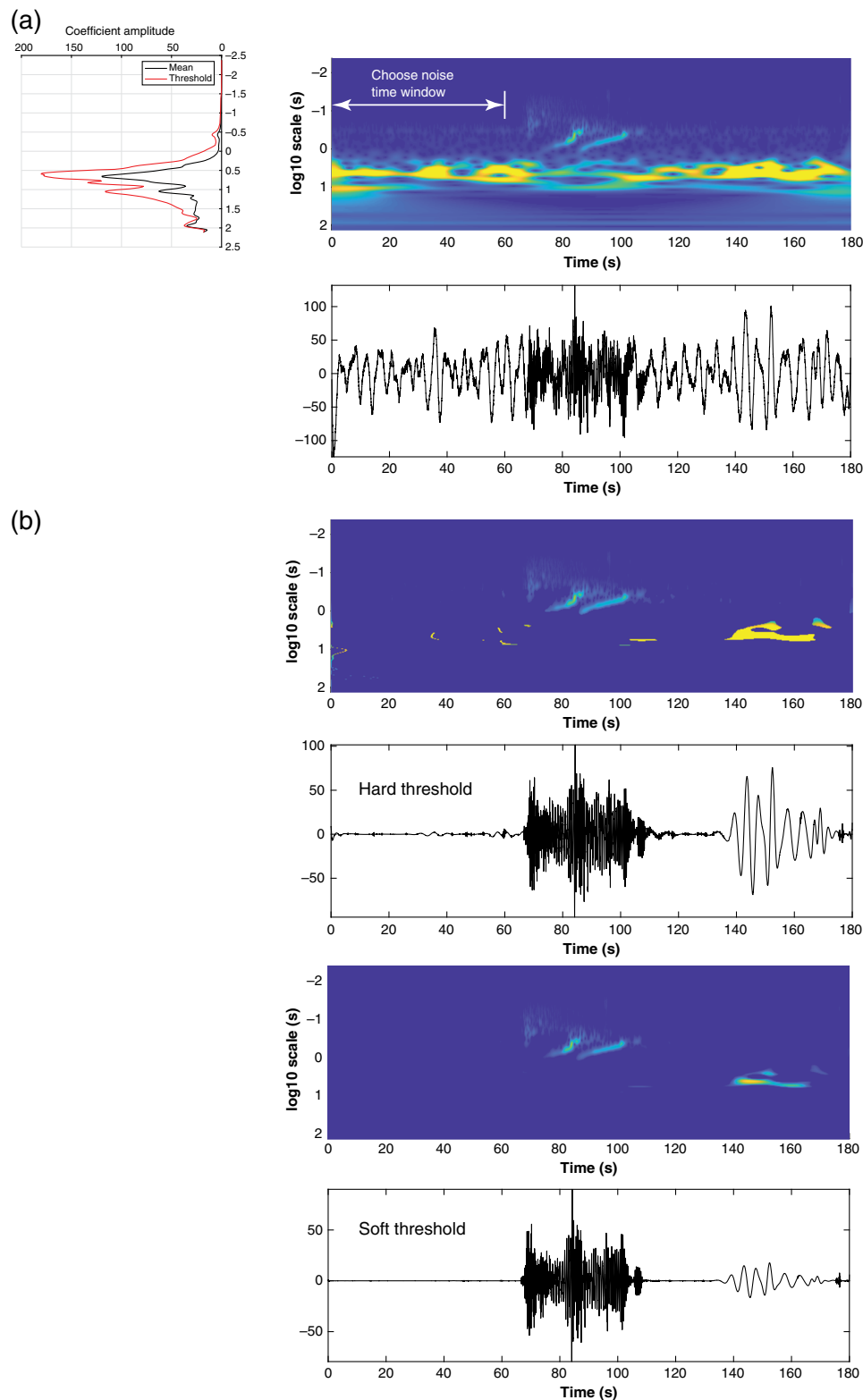
$$\text{mean}(|W(a, \tau)|) = \frac{1}{(t_2 - t_1)} \int_{t_1}^{t_2} |W(a, \tau)| d\tau \quad (10)$$

$$\begin{aligned} \text{st.dev.}(|W(a, \tau)|) &= \left[ \frac{1}{(t_2 - t_1)} \int_{t_1}^{t_2} (|W(a, \tau)| - \text{mean}(|W(a, \tau)|))^2 d\tau \right]^{\frac{1}{2}}, \end{aligned} \quad (11)$$

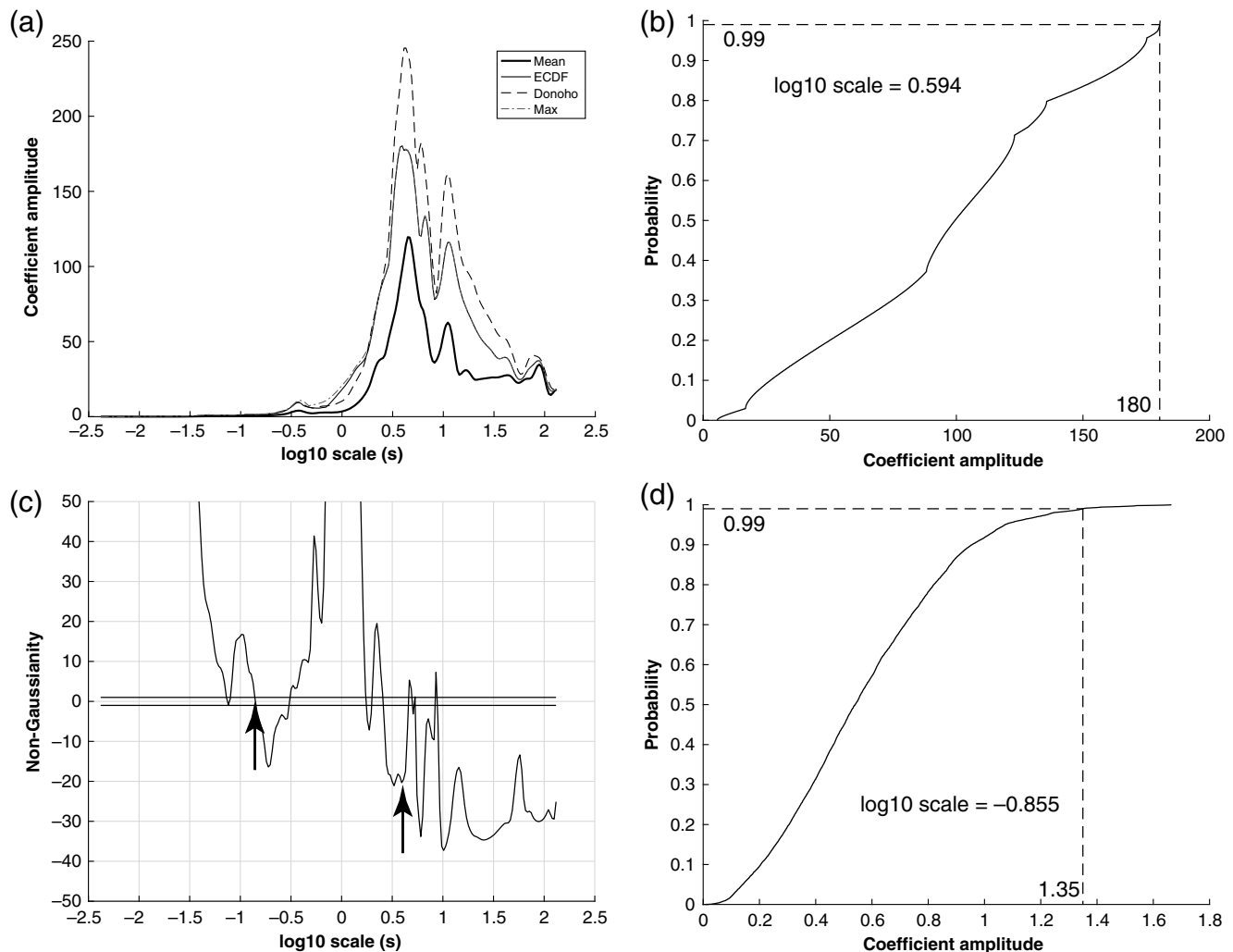
and  $c$  is a constant that controls the threshold. The limits  $t_2$  and  $t_1$  represent the limits of the noise time lag window.

There are a number of ways to estimate the threshold coefficient  $c$  contained in equation (9). If the wavelet coefficients at a particular scale follow a normal distribution, then setting  $c = 3$  will yield a signal at the 99.7% confidence level (Starck *et al.*, 2010). Donoho and Johnstone (1994) suggest a somewhat less stringent criterion called the “universal” threshold that is related to the number of noise samples  $N$  at each scale. The universal threshold is given as

$$c = \sqrt{2 \log_{10} N}. \quad (12)$$



**Figure 2.** (a) Vertical-component data from station 5014 for a 2000 lbs Air Force Research Laboratory (AFRL) explosion located about 35 km to the northwest of the Incorporated Research Institutions for Seismology (IRIS) wavefields experiment (14 July 2016 at 14:26:00 UT; lower panel). The explosion is evident as the higher frequency signal riding on the ambient background noise. The upper panel shows the continuous wavelet transform (CWT) scalogram displaying the absolute value of wavelet coefficient amplitude as a function of scale and time. The small plot in the upper left displays the average amplitude of noise within the 60 s time window before the explosion event. The threshold amplitude computed using the empirical cumulative distribution function (ECDF) technique described in equation (14) is also shown. (b) Denoising results using the hard threshold (top two panels) and soft threshold (bottom two panels) methods. The color version of this figure is available only in the electronic edition.



**Figure 3.** (a) Comparison of the mean amplitude of the noise used in Figure 2 with three different thresholding functions. The threshold computed using the ECDF method is compared to Donoho's universal threshold assuming Gaussian statistics and the max of  $ECDF^{-1}$ . (c) Display of the non-Gaussianity versus scale for the noise using relation (13). The two horizontal lines delineate the narrow region for which the ECDF is likely Gaussian. Most of the noise is distinctly non-Gaussian, justifying use of the ECDF method. Arrows point to portions of the noise that are Gaussian-like (left arrow) and strongly non-Gaussian (right arrow). (b) The ECDF of the strongly non-Gaussian noise at log10 scale of 0.594. (d) The ECDF of the Gaussian-like noise for log10 scale of -0.855. The coefficient amplitude is shown on both (b) and (d) for the  $P = 0.99$  level.

Normally,  $c$  from this relation is close to the value 3 for practical problems.

Unfortunately, the assumption of Gaussian statistics for the noise is rarely seen in seismic data. Using the noise time window shown in Figure 2a, the excess kurtosis was calculated at each scale and compared to the Gaussian excess kurtosis estimator presented in Mousavi and Langston (2016a, their equation 12). Figure 3 displays the “non-Gaussianity” versus scale determined by

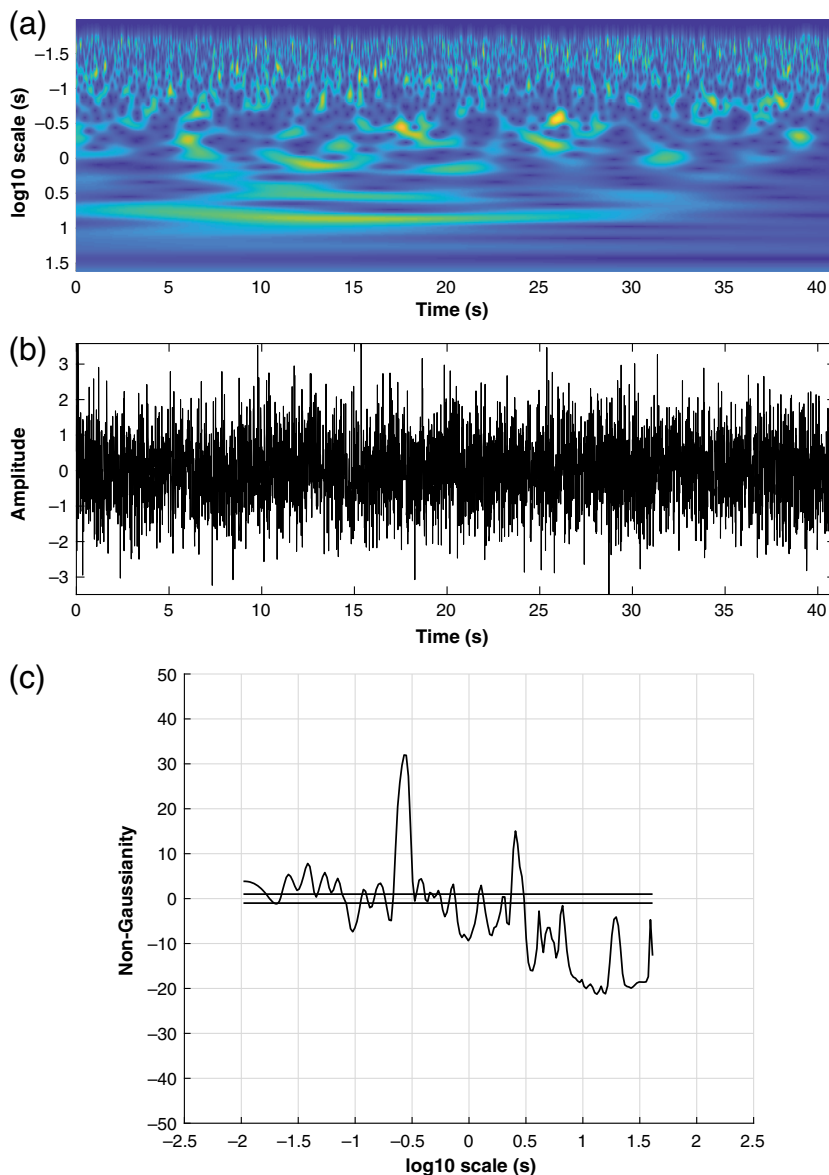
$$\frac{\text{kurt}\sqrt{0.9}}{\sqrt{24/N}}, \quad (13)$$

in which  $N$  is the number of samples in the noise estimate (here, 6001), 0.9 is for the 90% confidence level, and kurt

is the kurtosis. Values of non-Gaussianity between -1 and +1 represent Gaussian noise distributions. The kurtosis of a probability density function represents both the width of a possible central peak and the behavior of outliers. A negative kurtosis often means very short tails to the distribution, whereas positive kurtosis represents flat tails. The main idea in noise removal is to include as much of the noise distribution as possible to avoid outliers that would give rise to false arrivals in the denoised signal. Thus, for example, high positive kurtosis could mean long flat tails to the distribution of the noise so that assuming Gaussian statistics would let more noise through the thresholding operation.

At this point, it is important to emphasize that the statistical nature of the noise in the time domain can become quite different in the wavelet domain. Figure 4 shows a synthetic





**Figure 4.** (b) A synthetic random time series constructed using Gaussian statistics for the random excursions assuming a mean of 0 and standard deviation of 1. The CWT of this time series is shown in (a). Note that for high scale or long period, the scalogram shows very smooth variation with time. Variation increases with decreasing scale. (c) The statistics of the absolute value of the wavelet coefficients for this realization of time-domain Gaussian noise show distinctly non-Gaussian behavior over much of the range of CWT scale. The color version of this figure is available only in the electronic edition.

time series and resulting CWT. This random time series was constructed using a Gaussian distribution for the excursions with zero mean and a standard deviation of unity, creating what is known as white noise in the Fourier domain. After transformation into the scale-time domain, the distribution of the absolute amplitude of the wavelet coefficients can become distinctly non-Gaussian, as revealed by the kurtosis test (Fig. 4). In effect, what is usually considered as a Gaussian white noise process in the time domain becomes distinctly non-Gaussian in the scale-time domain, necessitating a

more careful look at the criteria used for thresholding the CWT.

Because the distribution for real data is often unpredictable, we take the approach of empirically estimating the cumulative distribution of noise in the time window at each scale then calculating the 99% confidence value for the distribution. The empirical cumulative distribution function ( $ECDF_a$ ) is determined by ordering the  $N$  noise values and then assigning a probability jump of  $1/N$  when a value is attained, starting with the smallest value. The threshold function becomes

$$\beta(a) = ECDF_a^{-1}(P = 0.99), \quad (14)$$

in which  $ECDF_a^{-1}$  is the inverse cumulative distribution function or quantile function. Figure 3 shows the comparison between threshold functions assuming Gaussian statistics (equation 9) and non-Gaussian statistics (equation 14). The result in which only the maximum of the inverse ECDF is chosen is also shown. Thresholds from distributions with negative kurtosis closely track the maximum of each distribution, as expected. Thresholds from distributions with positive kurtosis are underestimated by as much as half if Gaussian statistics are assumed. Overall, there are significant differences between the inverse ECDF and inverse Gaussian threshold functions suggesting that the ECDF method would lead to better estimates of the threshold. Just using the maximum of the inverse ECDF would also be acceptable and certainly simpler to implement.

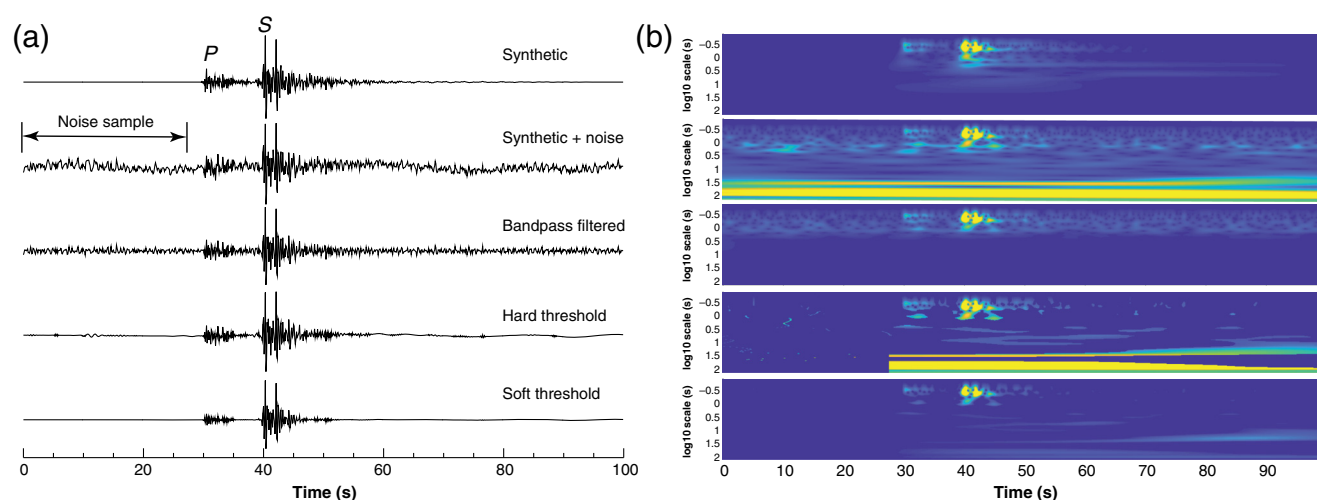
An interesting aspect of these denoising techniques is that they can be reversed: signal can be extracted from the time series leaving only noise. This could be useful in removing earthquake signals from ground-motion time series for ambient-noise processing (e.g., Benson *et al.*, 2007). For

the hard thresholding case, signal is removed using

$$\tilde{W}(a, \tau) = \begin{cases} 0 & \text{if } |W(a, \tau)| \geq \beta(a) \\ W(a, \tau) & \text{otherwise} \end{cases}, \quad (15)$$

and for soft thresholding

$$\tilde{W}(a, \tau) = \begin{cases} \text{sign}[W(a, \tau)]\beta(a) & \text{if } |W(a, \tau)| \geq \beta(a) \\ W(a, \tau) & \text{otherwise} \end{cases}. \quad (16)$$



**Figure 5.** (a) The time series for the synthetic seismogram, synthetic with noise, band-pass filtered (1 Hz zero-phase Butterworth high-pass filter with two poles), hard and soft threshold results using the ECDF method. (b) The corresponding CWT scalograms. These results are comparable to results shown in figure 3 of Mousavi and Langston (2016a), who used a more complex denoising method. The color version of this figure is available only in the electronic edition.

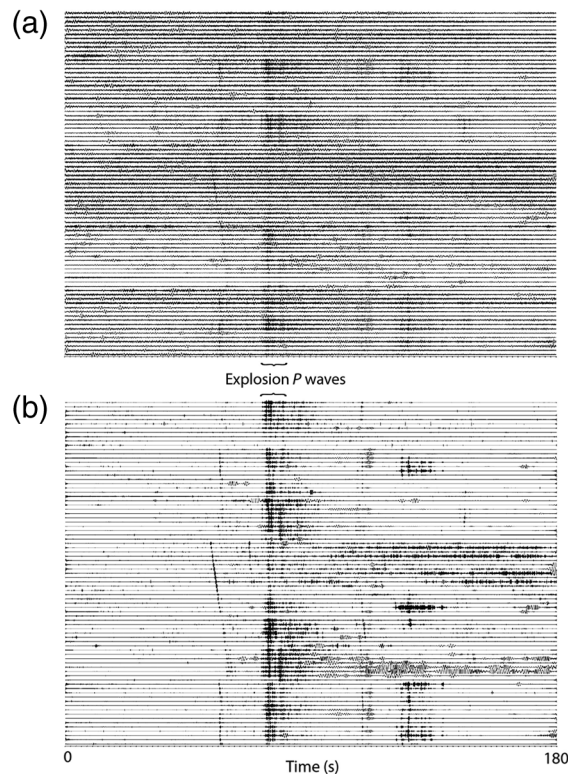
Hard thresholding leaves a “hole” in the CWT by zeroing out all coefficients above the noise level, whereas soft thresholding fills in the hole with an oscillatory version of the background noise.

### Nonlinear Thresholding Examples

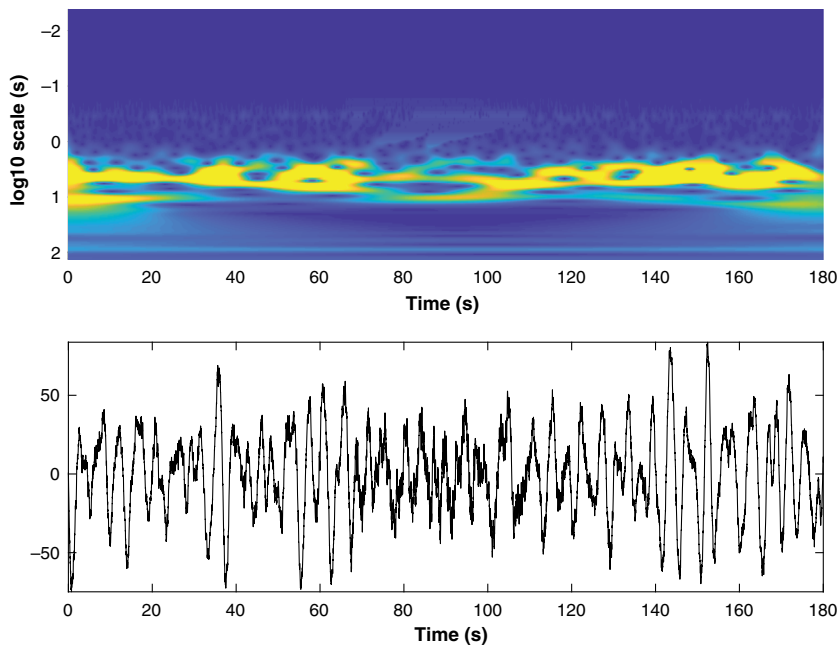
We demonstrate each thresholding method first using the synthetic seismogram example in Mousavi and Langston (2016a, their fig. 3). Implementation of equation (14) for both hard and soft thresholding is equally efficient and show processed seismograms that are comparable to the six-step block thresholding procedure of that article (Fig. 5). The synthetic seismogram comes from a wavenumber integration solution (Zhu and Rivera, 2002) for a double-couple point source in plane-layered structure appropriate for the central United States. Observed seismic noise from a station in the New Madrid Cooperative Seismic Network was added to the synthetic. A noise time window 28 s in length was taken to estimate the scale-dependent  $ECDF_a$  and resulting threshold. Although both hard and soft thresholding methods are successful in removing most of the pre-event noise, soft thresholding outperforms hard thresholding for removing postevent noise above 10 s in scale. Even so, some of the *P*- and *S*-wave signals are reduced by soft thresholding.

Figure 2 shows results using the ECDF method on the local explosion waveform recorded by the IRIS wavefields experiment. The time-domain signal is dominated by ambient background noise. A 60 s time window is used to estimate the noise threshold function. Again, soft thresholding outperforms hard thresholding, leaving fewer artifacts in the pre-event and postevent time series.

The utility of soft thresholding and added scale-time band rejection for improving event detection can be seen in Figure 6. Here, an array composed of 81 vertical-component elements



**Figure 6.** An array of 81 stations were subset for a 2000 lbs AFRL explosion that was detonated at 67 km from the IRIS wavefields experiment (14 July 2016, 15:12:00 UT). (a) A shot gather for the 81 array elements after high-pass filtering with a 1 Hz zero-phase Butterworth filter with four poles. Explosion *P* waves are discernable through the fog of high-frequency noise. (b) The result after denoising the original data using noise estimates from the first 60 s of record at each array element separately with the  $P = 0.99$  quantile value. In addition, wavelet coefficients for scales greater than 1 s were zeroed to remove obvious noise. The resulting shot gather is much clearer, showing the explosion *P* waves and many other local microseisms that occurred near the explosion in time.



**Figure 7.** The explosion signal shown in Figure 2 is removed using the ECDF soft thresholding method. Same scheme as Figure 2. The color version of this figure is available only in the electronic edition.

of the IRIS experiment is constructed to act as a phased array in determining phase velocity and azimuth for local explosion seismic waves. An analysis of one data channel shows that ambient noise is large for scales greater than 1 s or frequencies less than 1 Hz. Application of a four-pole phaseless Butterworth filter shows considerable noise, but the explosion signal is evident. However, applying the ECDF denoising with an additional scale-time block rejection for wavelet coefficients at scales greater than 1 s clarifies both the explosion *P* waves and additional waves from microseismic events under the array.

Finally, we take the same explosion seismogram as shown in Figure 2 and remove the signal using soft thresholding (equation 16; Fig. 7). Clearly, this would have use in conditioning ambient-noise data for obtaining Green's functions through correlation (e.g., Benson *et al.*, 2007).

### Scale-Time Windowing and Phase Decomposition

An interesting attribute of the scale-time plane of the CWT of a signal is the ability to separate seismic phases. As an example, Figure 8 shows an expanded version of the soft thresholded denoised signal of Figure 2. The CWT scalogram clearly delineates the three major seismic phases within the seismogram. The *P* wavetrain and its coda are separate from the first higher Rayleigh mode, which is also separate from the fundamental Rayleigh mode. It becomes a simple matter to separate these three wavetrains on the scale-time plane by capturing the area of their phase energy through picking or by some algorithm. Here, we draw a polyhedron around the fundamental-mode energy to window it from the rest of the

seismogram. The complex wavelet coefficients within the polyhedron are kept and everything else discarded. Applying the inverse CWT produces the isolated fundamental-mode Rayleigh wave. The fundamental mode can then be analyzed for group and phase velocity to determine structure information. For example, a phased array of seismometers can be used to analyze the fundamental mode after it has been isolated through scale-time windowing at each element of the array. Clearly, this will be useful in increasing the resolution of array methods for analysis of local and regional seismic phases.

### Discussion and Conclusions

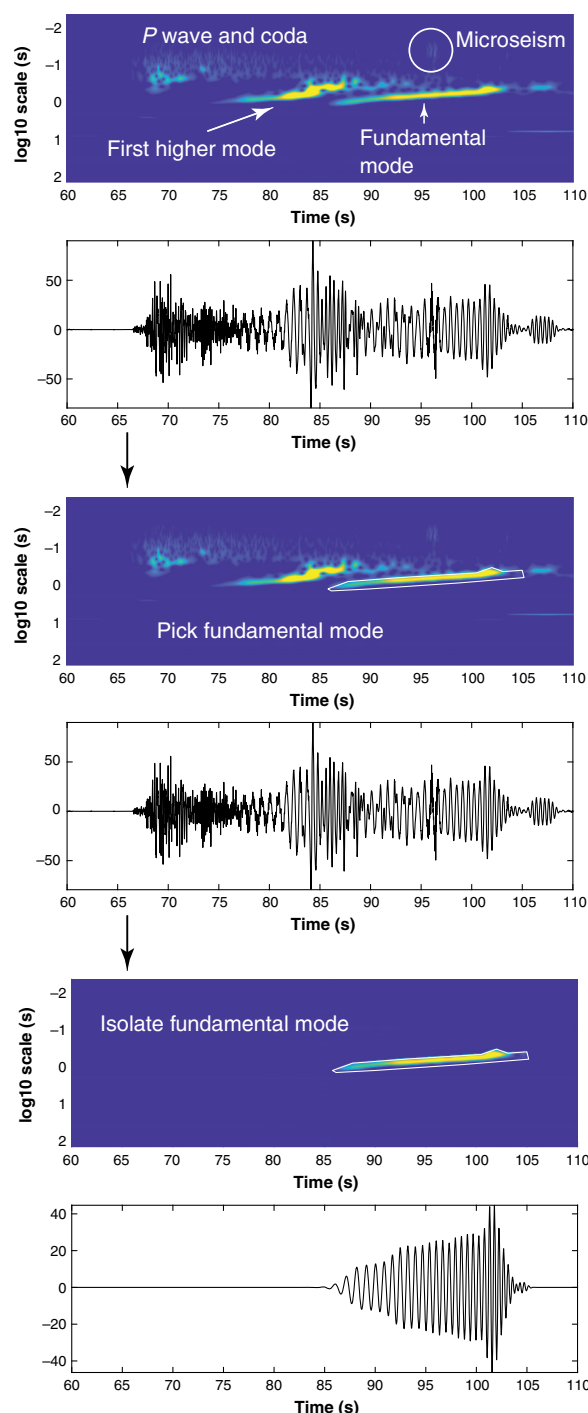
The performance of wavelet denoising of seismic data can look “too good to be true,” in a sense, as shown in Figures 2 and 6. Pre-event noise may be almost perfectly removed, particularly if the pre-event noise is used in the noise estimate.

Seismologists are often interested in the fidelity of seismic phase arrival times for travel-time studies or preserving wave shape for waveform or correlation studies. The accuracy of travel times is controlled largely by the energy of signal at fine scales or high frequency. If the SNR is large at fine scale, as it is in the explosion example shown here, then nonlinear denoising is an ideal method for filtering the record to obtain a travel-time estimate. However, if the event signal is less than the background noise at fine scales, then it is very likely that nonlinear thresholding will remove the signal along with the noise. For cases in between in which the noise overlaps with the signal on the time-scale plane (Fig. 5), nonlinear thresholding will degrade the waveforms of small individual phases.

A number of these issues were investigated by Mousavi and Langston (2016a). The preservation of waveform depended somewhat on the choice of the mother wavelet and the level of the noise. However, removing moderate amounts of noise from a synthetic three-component earthquake signal showed that *P*- and *S*-wave particle motions were preserved after noise removal. If there are regions on the time-scale plane of the CWT that have high SNRs, then nonlinear thresholding is very effective in removing the noise in adjacent areas and exposing the signal. However, characteristics of the signal in those low SNR regions will be lost in the processing.

As with any seismic data processing method, performance depends on how well assumptions of the method agree with the reality of the data. The key assumption is that the noise sample used to estimate the thresholding parameter is stationary throughout the time series. Changing noise sources or other data problems, such as data dropouts or high-amplitude transients, are not stationary by definition and will likely





**Figure 8.** Example of using scale-time windowing to obtain the fundamental-mode Rayleigh wave from the explosion data of Figure 2. A close look at the CWT scalogram in the top panel shows that seismic wavetrains in the seismogram are separated in scale and time. The seismogram consists of a *P* wavetrain and its coda, the first higher-mode Rayleigh wave, and the fundamental-mode Rayleigh wave. In addition, a high-frequency (low-scale) microseism is seen to mix with the fundamental-mode Rayleigh wave in time. The fundamental mode is chosen by placing a polygon around its high-amplitude signature (middle panels). The wavelet coefficients for all other scales and times are removed and the CWT inverse transformed to obtain the fundamental-mode waveform. The color version of this figure is available only in the electronic edition.

misrepresent the noise estimate. Although examining the CWT scalogram may help in assessing these nonstationary noise sources, effective denoising depends on an appropriate choice of the noise time window and quality-controlled seismic data.

The principal contribution of this article is proposing a simplified method for determining the threshold function for denoising or designating seismic time series. Constructing the ECDF from observed values of noise as a function of wavelet scale, and assuming that the noise is stationary through the time series, accounts for the non-Gaussian distribution of noise amplitude. Using the 99% confidence level of the ECDF gives a significantly better estimate of the threshold function than assuming Gaussian statistics. In the case that the noise is actually Gaussian, the method yields the correct result. The net effect is to speed up computations for noise removal by two orders of magnitude while yielding similar results compared to those in Mousavi and Langston (2016a). Simplifying the analysis also makes examining data for noise and signal more intuitive.

Use of the CWT in seismic signal processing shows great potential. Denoising allows for increasing event detection and designating may be used for conditioning ambient noise in correlation studies. Manipulation of the signal on the scale-time plane potentially allows for high-resolution analysis of discrete seismic phases and surface-wave modes within the seismogram.

## Data and Resources

Data used in this article can be obtained from the Incorporated Research Institutions for Seismology (IRIS) data center at <https://ds.iris.edu/ds/nodes/dmc/> (last accessed June 2019) and is described in Sweet *et al.* (2018). MATLAB software used in processing the data in this article is available from the author's website at <http://www.ceri.memphis.edu/people/clangstn/> (last accessed July 2019). Seismic Analysis Code (SAC) was used in this article (Goldstein *et al.*, 2003) and is gratefully acknowledged. The authors thank Associate Editor Eric Chael and two anonymous reviewers for their suggestions and comments that improved the article.

## Acknowledgments

This research was supported, in part, by the Center for Earthquake Research and Information, University of Memphis.

## References

- Benson, 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, *Geophys. J. Int.* **169**, 1239–1260, doi: [10.1111/j.1365-246X.2007.03374.x](https://doi.org/10.1111/j.1365-246X.2007.03374.x).
- Daubechies, I., and S. Maes (1996). A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models, in *Wavelets in Medicine and Biology*, A. Aldroubi and M. Unser (Editors), CRC Press, Boca Raton, Florida, 527–546.

- Daubechies, I., J. Lu, and H.-T. Wu (2011). Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool, *Appl. Comput. Harmon. Anal.* **30**, 243–261, doi: [10.1016/j.acha.2010.08.002](https://doi.org/10.1016/j.acha.2010.08.002).
- Donoho, D., and I. Johnstone (1994). Ideal spatial adaptation via wavelet shrinkage, *Biometrika* **81**, 425–455.
- Donoho, D., I. Johnstone, G. Kerkycharian, and D. Picard (1995). Wavelet shrinkage: Asymptopia? *J. Roy. Stat. Soc. B* **57**, 301–369.
- Goldstein, P., D. Dodge, and M. Firpo (2003). SAC2000: Signal processing and analysis tools for seismologists and engineers, in *IASPEI International Handbook of Earthquake and Engineering Seismology*, W. H. K. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger (Editors), Academic Press, London, United Kingdom.
- Goupillaud, P., A. Grossmann, and J. Morlet (1985). Cycle-octave and related transforms in seismic signal analysis, *Geoexploration* **23**, 85–102.
- Grossmann, A., R. Kronland-Martinet, and J. Morlet (1989). Reading and understanding the continuous wavelet transform, in *Wavelets: Time-Frequency Methods and Phase-Space*, J. Combes, A. Grossmann, and P. Tchamitchian (Editors), Springer, New York, New York, 2–10.
- Mousavi, S. M., and C. A. Langston (2016a). Hybrid seismic denoising using higher-order statistics and improved wavelet block thresholding, *Bull. Seismol. Soc. Am.* **106**, no. 4, 1380–1393, doi: [10.1785/B0120150345](https://doi.org/10.1785/B0120150345).
- Mousavi, S. M., and C. A. Langston (2016b). Adaptive noise estimation and suppression for improving microseismic event detection, *J. Appl. Geophys.* **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 (2017). Automatic noise-removal/signal-removal based on general cross-validation thresholding in synchrosqueezed domain and its application on earthquake data, *Geophysics* **82**, no. 4, V211–V227.
- Mousavi, S. M., S. P. Horton, and C. A. Langston (2016). 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).
- Starck, J.-L., F. Murtagh, and J. M. Fadili (2010). *Sparse Image and Signal Processing*, Cambridge University Press, New York, New York.
- Sweet, J. R., K. R. Anderson, S. Bilek, M. Brudzinski, X. Chen, H. DeShon, C. Hayward, M. Karplus, K. Keranen, C. Langston, *et al.* (2018). A community experiment to record the full seismic wavefield in Oklahoma, *Seismol. Res. Lett.* **89**, 1923–1930, doi: [10.1785/0220180079](https://doi.org/10.1785/0220180079).
- Weaver, J. B., X. Yansun, D. M. Healy, and L. D. Cromwell (1991). Filtering noise from images with wavelet transforms, *Magn. Reson. Med.* **21**, 288–295.
- Zhu, L., and L. A. Rivera (2002). A note on the dynamic and static displacements from a point source in mult-layered media, *Geophys. J. Int.* **148**, 619–627.

#### Charles A. Langston

Center for Earthquake Research and Information  
University of Memphis  
3876 Central Avenue, Suite 1  
Memphis, Tennessee 38152 U.S.A.  
clangstn@memphis.edu

#### Seyed Mostafa Mousavi

Department of Geophysics  
School of Earth, Energy and Environmental Sciences  
Stanford University  
Mitchell Building, 3rd Floor  
397 Panama Mall  
Stanford, California 94305 U.S.A.

Manuscript received 21 March 2019;

Published Online 6 August 2019