

Seismic processing with continuous wavelet transform maxima

Kris Innanen

ABSTRACT

Sophisticated signal analysis methods have been in existence since the 1990s, formalizing the older idea of *edge detection and characterization*. Amongst the better known of these is one focusing on the portions of continuous wavelet transforms thought to contain most of the signal's information content—its maxima. While *analysis* of continuous wavelet transform maxima (CWTMs) has been applied in exploration seismology, *processing* of these data have not. Simple methods exist to do this, and we discuss and implement one—the model of Mallat and Zhong—in this paper. The approach requires a bit of conceptual explanation, as it has not appeared in the exploration geophysics literature before, and we devote most of this paper to that task; our primary goal is logically develop the implementation, and leave relatively open the many possible applications to the seismic trace it seems to have in potentia. However, we end the paper with a simple example of *thresholding* of CWTMs, and the resulting reduction in noise. In a companion paper we use such data to produce internal multiple prediction operators.

INTRODUCTION

In this paper we examine the ability of continuous wavelet transforms, and a certain subset of the information they contain, to characterize and process events in a measured signal such as a seismic trace.

Wavelet transforms (Mallat, 1989; Daubechies, 1992; Chui, 1992) can be either discrete (DWT) or continuous (CWT). A DWT is non-redundant: N samples in the input signal lead to N wavelet coefficients, and from these the original signal can be uniquely reconstructed. A CWT is redundant: an N -length input signal is convolved with the wavelet at each desired scale, so that if there are N_S scales to be analyzed, the CWT will contain $N_S \times N$ data. The continuity of the CWT allows for some important signal analysis tasks to be stably carried out, but in order to do so, we must find some way of meaningfully dealing with the redundant information it carries

One approach is to postulate that the *edges* of the signal, i.e., its regions of most rapid change, contain most of its information (Canny, 1986). If this is true it follows that we should focus only on the regions of the CWT which represent edges, which, as pointed out by Mallat and Hwang (1992) and Mallat and Zhong (1992), are the *local maxima of the modulus of the CWT*: its positive and negative peaks. Those authors, Mallat and Zhong in particular, developed a fully realized numerical and analytical framework in which the CWT maxima play the central role.

Broadly speaking we can do two things with maxima within the framework of Mallat and Zhong: analyze them and process them. The former is “easier”—we do not need a well behaved inverse transform for analysis, since we're not coming back. The decay or growth of the maxima with scale is related to the local *Lipschitz* (or *Hölder*) *regularity* of

the signal, which characterizes the edge type. Innanen (2003) discussed the evolution of Lipschitz regularity of a nonstationary seismic trace with time, and suggested via arguments involving the Strick approximation that it be used to estimate Q . This was pursued by Izadi and collaborators (Izadi et al., 2011a,b), and a strategy for Q estimation of VSP data based on the output of this signal model has been developed (Izadi et al., 2012, 2013).

Processing of continuous wavelet transform maxima, that is, altering the maxima associated with a signal and then reconstructing the signal, is less straightforward. Inversion of the CWT itself is linear and easy to formulate. But if we deal only with the less redundant maxima, we do more than transform: we also implicitly mute all *non*-maxima in the forward part of the transformation. Muting data, even highly redundant data, using a data-dependent criterion, is a highly nonlinear procedure. One of the key contributions of Mallat and Zhong (1992) is an approach to the inversion of the transform-then-mute forward problem. They formulate an iterative reconstruction scheme whereby CWT maxima are transformed to an approximate version of the signal which gave rise to them. Those authors report an inability to prove mathematically the correctness of the reconstruction, but they also report an inability to find numerical examples in which the reconstruction differed significantly from the input.

No seismic examples of processing of CWT maxima can be found in the geophysics literature, to the best of our knowledge. In a companion paper (Innanen, 2013) the opportunities the CWT maxima domain presents for more surgical prediction of internal multiples is examined. In this paper we carry out the groundwork needed to attempt such applications. We review the formulation of the problem: how to perform the CWT and manipulate the maxima, for example through what is termed *soft thresholding*, and most importantly how to reconstruct the signal from its maxima. We write a MATLAB implementation of this transform and reconstruction, and use it to ask some questions relevant to seismic application. In particular we consider the accuracy of signal reconstruction from its edge information, and comment on the impact of small amounts of reconstruction error to seismic data analysis.

MALLAT-ZHONG RECONSTRUCTION FROM CWT MAXIMA

In this section we will provide an intuitive discussion of the Mallat-Zhong approach. The input to the transform and the subject of the subsequent analysis is a single seismic trace, which we will discuss as a continuous function of time $d(t)$ or a discrete vector \mathbf{d} as needed. Reconstructions will be labelled with a tilde $\tilde{\cdot}$. Ultimately we will consider the error $E(t) = d(t) - \tilde{d}(t)$ associated with Mallat-Zhong reconstruction, and whether its magnitude is or is not such that seismic application is warranted.

The continuous wavelet transform and the operator W

The wavelets we consider are n th order derivatives of smoothing functions. A legitimate smoothing function must (1) have an integral of one, and (2) decay to zero at $t \rightarrow \pm\infty$. A Gaussian function $\theta(t)$ normalized such that

$$\int_{-\infty}^{\infty} \theta(t) dt = 1 \quad (1)$$

fits the bill, and is used by Mallat and Zhong. The mother wavelet $\psi(t)$ is chosen to be the first derivative of $\theta(t)$:

$$\psi(t) = \frac{d\theta(t)}{dt}, \quad (2)$$

which can then be compressed and dilated according to a scaling parameter s :

$$\psi_s(t) = \frac{1}{s} \psi\left(\frac{t}{s}\right). \quad (3)$$

In DWT theory, the next step is to endow the mother wavelet with a translation parameter τ , and the inner product of the input signal with the scaled and translated wavelet would give the wavelet transform coefficient for that (s, τ) pair. In a CWT the translation parameter is continuous, and is essentially indistinguishable from the original time axis. One row (i.e., one scale) of the CWT is obtained by convolving the input signal $d(t)$ by the wavelet at that scale:

$$\psi_s(t) * d(t) = \int_{-\infty}^{\infty} dt' d(t') \psi_s(t - t'). \quad (4)$$

Let us consider the operator \mathbf{W}_S when acting on $d(t)$ to produce a matrix with rows corresponding to the set of scales $S = \{s_1, s_2, \dots, s_N\}$; each row being the result of the convolution in equation (4), but with s replaced by the current scale s_1 , or s_2 , etc. This transform is illustrated in Figure 1.

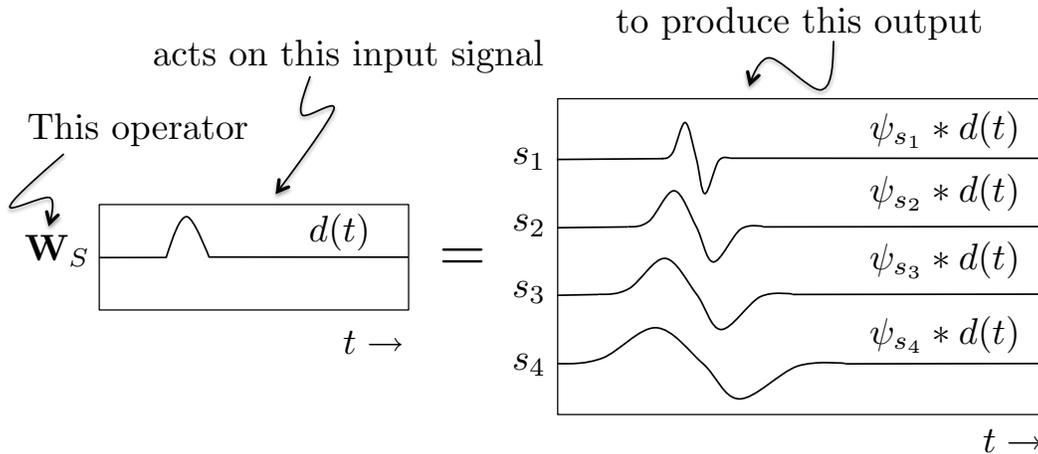


FIG. 1. A schematic illustration of the action of a general wavelet transform operator \mathbf{W}_S on an input signal $d(t)$. S represents N_S scales, where in this case $N_S = 4$, i.e., $S = \{s_1, s_2, s_3, s_4\}$. The operator \mathbf{W}_S acts on a length- N_T input signal $d(t)$, creating an $N_S \times N_T$ matrix, whose rows are the convolutions of $d(t)$ with the wavelet at the associated scale.

Fast forward and inverse transforms: \mathbf{W} and \mathbf{W}^{-1} on dyadic scales

Much of the redundancy of the continuous wavelet transform $\mathbf{W}_S d(t)$ derives from the unspecified set of scales S . If row s and row s' correspond to scales which are very close to being the same, those two rows will not contain much independent information.

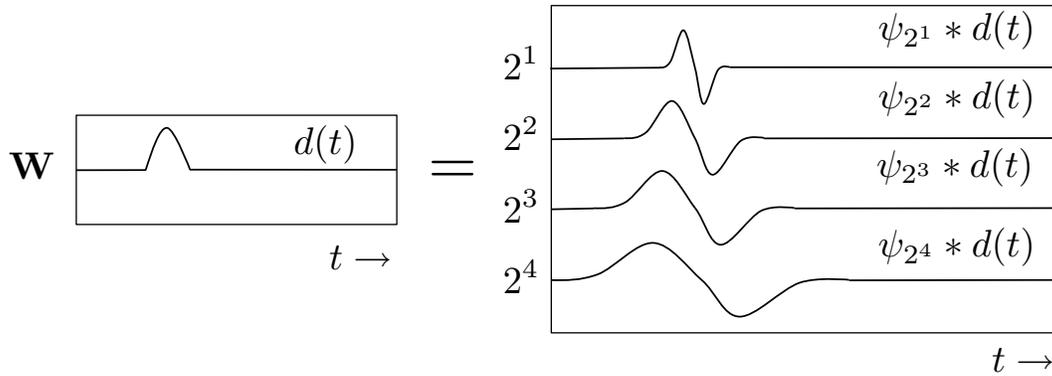


FIG. 2. A schematic illustration of the action of a dyadic wavelet transform operator \mathbf{W} on an input signal $d(t)$. The CWT on the right hand side can be expressed as $\mathbf{W}d(t)$.

In order to obtain efficient CWTs, the Mallat-Zhong approach calls for a discrete “dyadic” scaling system $s = 2^j$, $j = \{\dots, -2, -1, 0, 1, 2, \dots\}$. Let us refer to the CWT operator which involves this set of dyadic scales as the un-indexed \mathbf{W} , as illustrated in Figure 2.

On the scales $s = 2^j$, $j = \{1, 2, 3, \dots\}$, it is known that an *algorithme à trous*—“algorithm with holes” (Mallat, 1998)—can be applied to numerically perform the wavelet transform of a discrete input signal. At a scale j , a set of short filters (see Appendix B, Mallat and Zhong, 1992) can be applied to an input signal, with the outputs being (1) the signal component at that scale, and (2) the remaining larger scale signal components. At the next scale $j + 1$, this is repeated, producing the signal at scale $j + 1$ and the remaining signal components. This continues until the desired number of signal components have been determined. These components are the rows of $\mathbf{W}d(t)$. There will always remain a “coarse” vector containing all signal components at scales larger than the largest chosen CWT scale. These are kept for the inverse transform.

The inverse transform is carried out sequentially also. The largest scale (i.e., bottom) row of $\mathbf{W}d(t)$ is combined with the coarse vector, again using the short filters and an *algorithme à trous*; this combination is next itself combined with the second largest scale row of $\mathbf{W}d(t)$. This sequential inclusion of scale information continues, moving from large to small scales. When the smallest scale $j = 1$ has been incorporated, the original signal has been recovered. The procedure is illustrated schematically in Figure 3. The filter coefficients are derived by making spline approximations to the actual mathematical form for the wavelet, and hence a non-zero reconstruction error is introduced. This error, as we shall illustrate below, is small compared to the reconstruction error we shall be considering presently.

To summarize, we now have access to fast, accurate operators \mathbf{W} (and \mathbf{W}^{-1}) which can be applied to a given signal (or CWT) as needed. In other words we now may contemplate the matrix quantity

$$\mathbf{W}d(t). \tag{5}$$

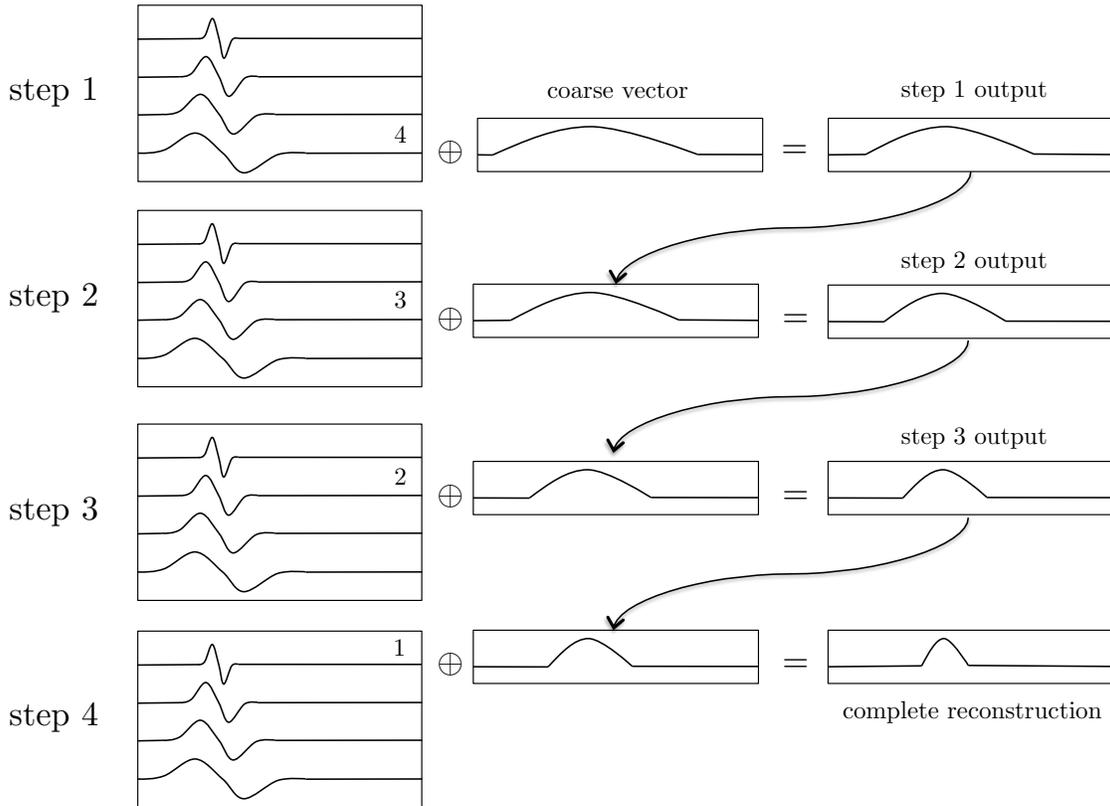


FIG. 3. Scale by scale reconstruction of a signal from its CWT. The input (top left) is the CWT, in this case a set of four scales, and the coarse vector, which contains information from all scales > 4 . In step 1, the bottom row of the CWT matrix $\mathbf{W}d(t)$, which contains signal information at scale index 4, is combined with the coarse vector to create a partial reconstruction of $d(t)$. This partial reconstruction now involves information from all scales > 3 . The symbol \oplus here denotes the combination of two rows using the short reconstruction filters and the à trous algorithm (Mallat, 1998). In step 2, the output of step 1 takes the place of the coarse vector. This is combined with the next row up (scale 3) of the CWT, and the result is a signal reconstruction containing information from all scales > 2 . This continues, until, in step 4, the output of step 3 is combined with the smallest scale information in the CWT (i.e., the top row), and the result is a complete reconstruction of the input signal.

Wavelet modulus maxima and the operator M

The processing domain we will consider involves the wavelet transform *maxima* only, which we can effect by applying a second operator, \mathbf{M} , which seeks values along each row of $\mathbf{W}d(t)$ whose neighbours on either side in time are smaller in magnitude. The matrix $\mathbf{M}\mathbf{W}d(t)$ is therefore of the same dimensions as $\mathbf{W}d(t)$, but whose entries, in contrast to the CWT, are almost all zero. $\mathbf{M}\mathbf{W}d(t)$ is “spiky”, as illustrated in Figure 4.

Analysis of the wavelet transform modulus maxima of a signal event could now take place if desired. For instance Izadi et al. (2013) collects these values (the scale 2 value m_2 is illustrated in Figure 4) by following one of the diverging paths in the scale direction, and uses them to estimate the local regularity of the event corresponding to the path.

Our current interest is to move beyond analyzing these spiky entries, and consider instead processing them. This introduces two questions: (1) how do we understand the meaning of the spikes, such that we could intelligently alter them, and should we alter them, (2) how do we reconstruct the new signal in the original time domain from the altered spikes?

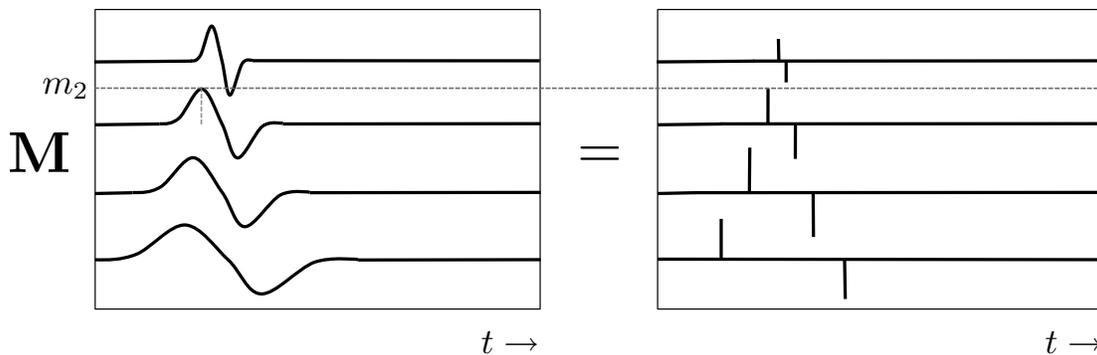


FIG. 4. The operator \mathbf{M} acting on a CWT matrix such as $\mathbf{W}d(t)$ produces a matrix of the same dimensions, whose only non-zero values occur at the maxima of $\mathbf{W}d(t)$. At these points the entries of $\mathbf{M}\mathbf{W}d(t)$ are equal to the entries of $\mathbf{W}d(t)$.

To summarize, we now contemplate the CWT maxima only, i.e., the quantity

$$\mathbf{M}\mathbf{W}d(t). \quad (6)$$

Thresholding and the operator T

Much of the answer to the first question above, namely how exactly should we process the spikes in the CWT maxima domain, will have to wait, as it is surely application dependent. However, we will for the moment restrict the type of alterations we will consider, namely *thresholding*. The delay of any of the spikes in time will not be altered, but the magnitude of the spikes will be. This will occur by defining a thresholding operator \mathbf{T} , which will set to zero spikes whose magnitudes fall below a certain defined size, and which will multiply those that remain by a certain defined factor (which may be 1). After applying this operator, we have produced the quantity

$$\mathbf{T}\mathbf{M}\mathbf{W}d(t). \quad (7)$$

Reconstruction and the operator $(\mathbf{MW})^{-1}$

Having altered the wavelet transform maxima (i.e., modifying the heights of the “spikes” on the right hand side of Figure 4) through the operator \mathbf{T} , we now need a way to return to the original space of signals $d(t)$. This will take place through repeated projections of signals onto two spaces, Γ and V .

The space Γ and getting there with the operator P_Γ

Consider that we have in hand a matrix $\mathbf{TMW}d(t)$ of wavelet transform maxima. Let the rectangle in Figure 5 represent the space Γ of all matrices *with the same maxima as $\mathbf{TMW}d(t)$* . We require an operator P_Γ which takes some matrix which is *not* an element of Γ (i.e., the black dot in Figure 5), and projects it orthogonally into Γ . P_Γ 's job, in other words, would be to find a matrix maximally similar to the nonmember, but whose maxima now match those of $\mathbf{TMW}d(t)$.

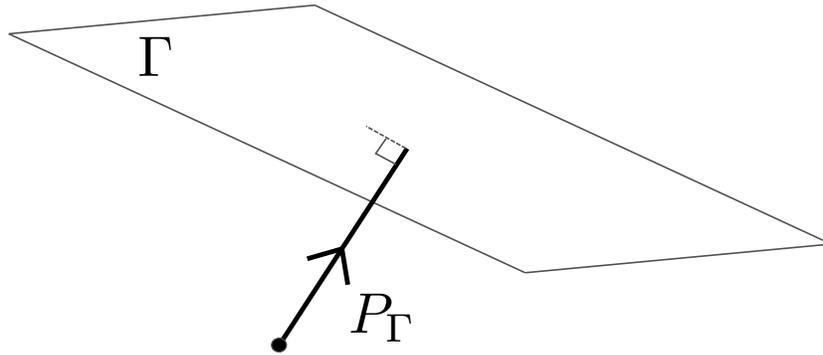


FIG. 5. Schematic representation of the space Γ and the orthogonal projection operator P_Γ . The space Γ contains all matrices whose maxima match those of the maxima matrix whose signal we would like to reconstruct. If we are presented with a matrix whose maxima do *not* match those of our starting point (e.g., the black dot sitting outside the rectangle Γ), we may apply the projection operator P_Γ to it. In doing so we find the element of Γ which is maximally similar to this nonmember of Γ , but whose maxima now do in fact match those of our starting point.

The way this is done in the Mallat-Zhong approach is by adding, to the non-element matrix, exponential functions designed to generate the correct maxima (i.e., the maxima of $\mathbf{TMW}d(t)$), exponential functions whose decay to the values of the original matrix is as rapid as possible (Figure 6).

The space V and getting there with the operator P_V

In the space of matrices with the dimensions of $\mathbf{W}d(t)$ must be many elements which are not legitimate wavelet transforms. In other words, which cannot have been created by applying the operator \mathbf{W} to any signal $d(t)$. Let V be the space of all matrices which *are* legitimate continuous wavelet transforms of the type we have been examining. A second rectangle can represent this space (Figure 7).

We seek the orthogonal projection operator P_V to be enacted upon a non-element of

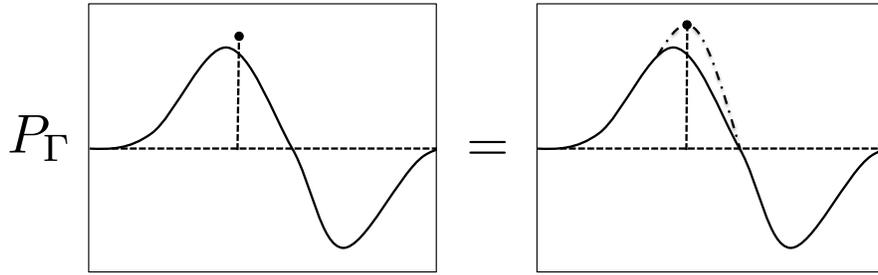


FIG. 6. A matrix whose maxima do not match those of $\text{TMW}d(t)$ is projected orthogonally into Γ by adding to the rows of the matrix one exponential function for each of the required maxima. The exponential functions are designed to match the maxima as required, and to decay away from the maxima to the original values of the matrix as rapidly as possible.

V (the black dot in Figure 7). The output of this operator is to be maximally close to the input, but, now, a legitimate element of V , a legitimate wavelet transform. This projection operator is simply $P_V = \mathbf{W}\mathbf{W}^{-1}$ (Mallat and Zhong, 1992). In other words, to find the desired continuous wavelet transform, one applies the inverse wavelet transform (in spite of the fact that the input to it is not a real wavelet transform) to the input, and then immediately applies the forward wavelet transform.

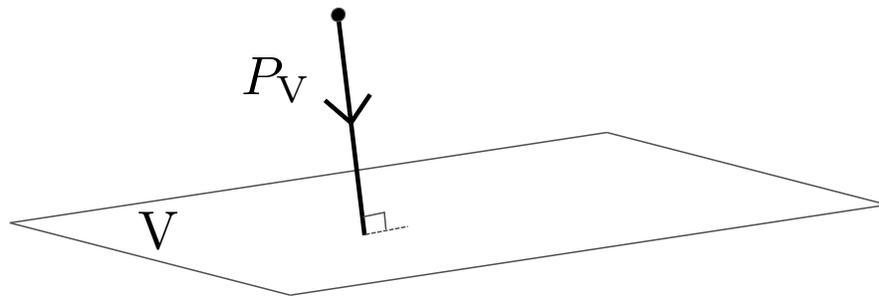


FIG. 7. Representation of the space V of all legitimate continuous wavelet transforms, and the operator P_V which maps a non-element of V (namely, a matrix which cannot have been produced with the operator \mathbf{W}) into V with as few alterations as possible.

Iterative reconstruction

The reconstruction of a signal from its wavelet transform maxima involves iterative application of the two projection operators P_Γ and P_V . The two spaces Γ and V are not the same, but they must intersect (Figure 8a). Beginning with the zero element of space V , that is a “wavelet transform” completely populated with zeros, we apply the operator P_Γ , projecting the zeros into Γ (first leg of the path in Figure 8b). The element of Γ produced by this is a matrix whose rows are exponential functions “draped” over the maxima, like a tablecloth. Of course, this is not the wavelet transform we are looking for—in fact, there is no guarantee it is a legitimate wavelet transform at all. We next ensure that the signal is a legitimate wavelet transform by applying the operator P_V (the second leg of the path in Figure 8b). The application of $\mathbf{W}\mathbf{W}^{-1}$ produces a *bona fide* wavelet transform, but this is still not guaranteed to be the one we seek. In fact, there is no guarantee that after the application of $\mathbf{W}\mathbf{W}^{-1}$ the resulting wavelet transform has its maxima in the same place

or with the same amplitude as the maxima matrix we began with. We next ensure that the maxima are correct by projecting it into Γ with P_Γ (the third leg of the path in Figure 8b). But this is not certain to be a real wavelet transform, so we apply P_V , etc. This is repeated until we attain an acceptable convergence.

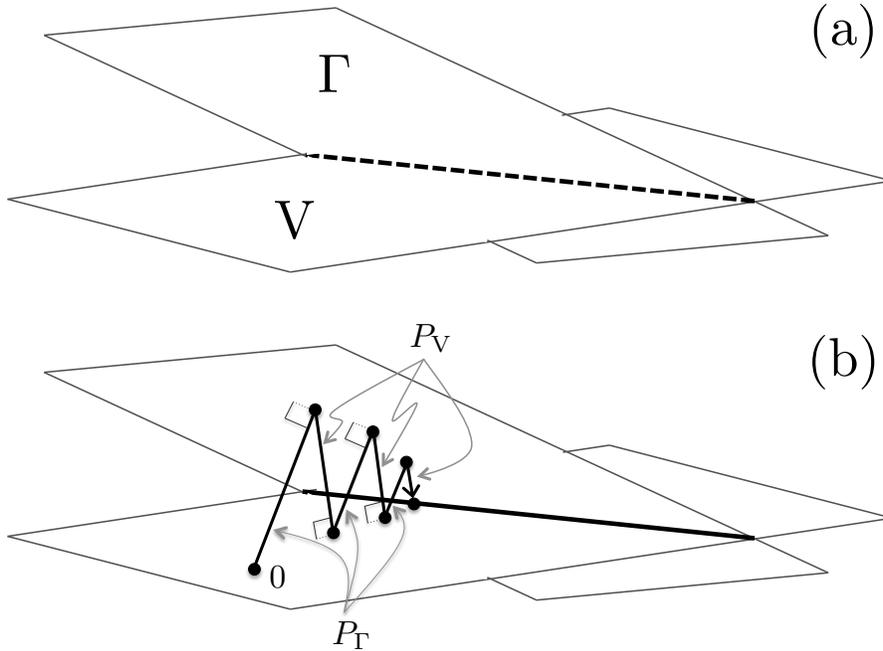


FIG. 8. The two spaces Γ and V intersect, and at this point of intersection is the wavelet transform we seek. Beginning with the zero element of V , i.e., a matrix of the correct dimensions but with zero elements, we apply the projection operators P_Γ and P_V sequentially until an acceptable convergence is obtained.

The process is understood (Mallat and Zhong, 1992) to converge to $W\tilde{d}(t)$ at the intersection of Γ and V , and there is reason to expect that the wavelet transform which (1) has the correct maxima, and (2) is a legitimate wavelet transform is the one we seek. With a last application of W^{-1} the desired output $\tilde{d}(t)$ is produced.

It is sensible to wonder if the reconstruction has necessarily converged to the right answer—is there one element in the intersection $\Gamma \cap V$, or are there several, or many? We will leave this as a question to be answered with our numerical analysis in the coming sections.

MALLAT-ZHONG RECONSTRUCTION OF SEISMIC DATA

In this section we present a numerical analysis of Mallat-Zhong reconstruction, involving synthetic data and post-stack land data. The land data set is an extraction from a 2D 3C land survey carried out in North Eastern British Columbia, Canada (Zuleta, 2012). The purpose is to illustrate the processing flow discussed schematically in the previous section, and investigate the accuracy of various reconstruction and denoising efforts.

Synthetic calculations of the wavelet transform and its maxima

In Figure 9a a synthetic seismic trace is plotted, containing a more or less random distribution of spike events convolved with a 40Hz zero-phase Ricker wavelet. The continuous wavelet transform is applied to this input, and the smallest 6 scales are plotted in Figure 9b. In the terminology of the previous section, Figure 9b is produced by applying the operator \mathbf{W} to Figure 9a.

In Figure 10a the continuous wavelet transform (i.e., Figure 9b) is reproduced and plotted alongside its maxima, in Figure 10b. The postulate of the Mallat-Zhong reconstruction approach is that these spikes contain most of the information in the signal. The maxima “domain” is where we will consider carrying out processing. In the terminology of the previous section, Figure 9b is produced by applying the operator sequence \mathbf{MW} to Figure 9a.

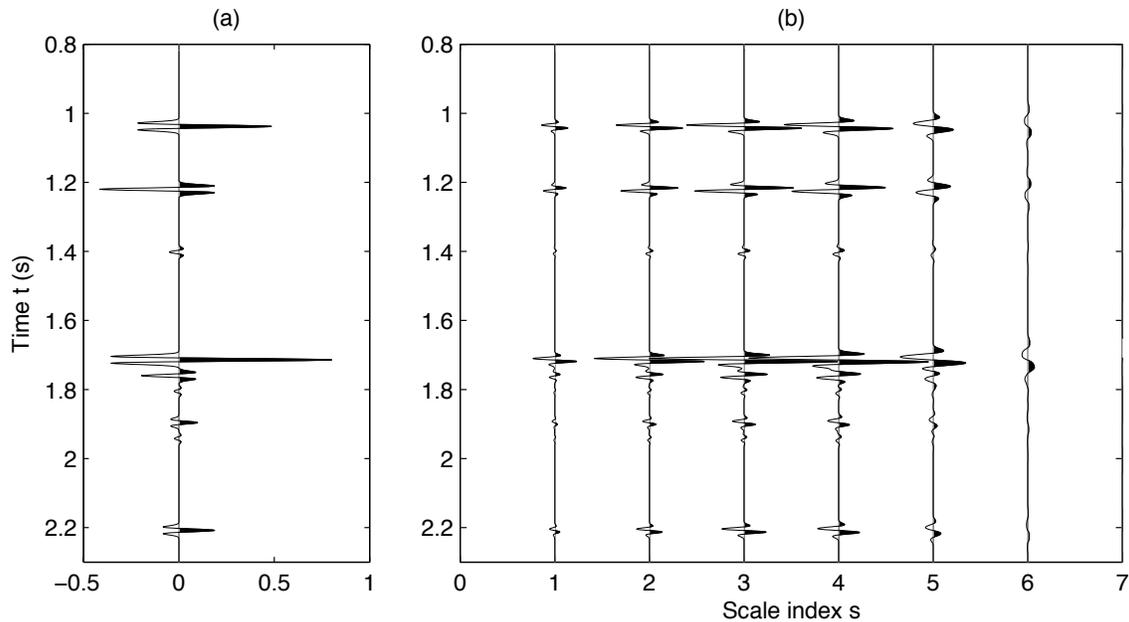


FIG. 9. Synthetic example of the continuous wavelet transform. (a) Synthetic trace, dominant frequency 40Hz; (b) continuous wavelet transform of the synthetic trace, with a plot range containing the 6 smallest scales. In operator terms, panel (b) is generated by applying \mathbf{W} to panel (a).

The culling of non-maxima from the wavelet transform (i.e., the operator \mathbf{M}) is carried out scale by scale, by seeking values of the modulus of the wavelet transform whose nearest neighbours are both smaller in modulus. These means we retain negative peaks as well as positive peaks (Figure 11). The reconstruction problem is to recover the blue curve from the black spikes.

The accuracy of the inverse wavelet transform

The implementation of the continuous wavelet transform is made fast through the use of filter banks (in addition to previous references, see Chapter 7 of Mallat, 1998). The filter coefficients are derived using spline approximations of the wavelet and scaling functions,

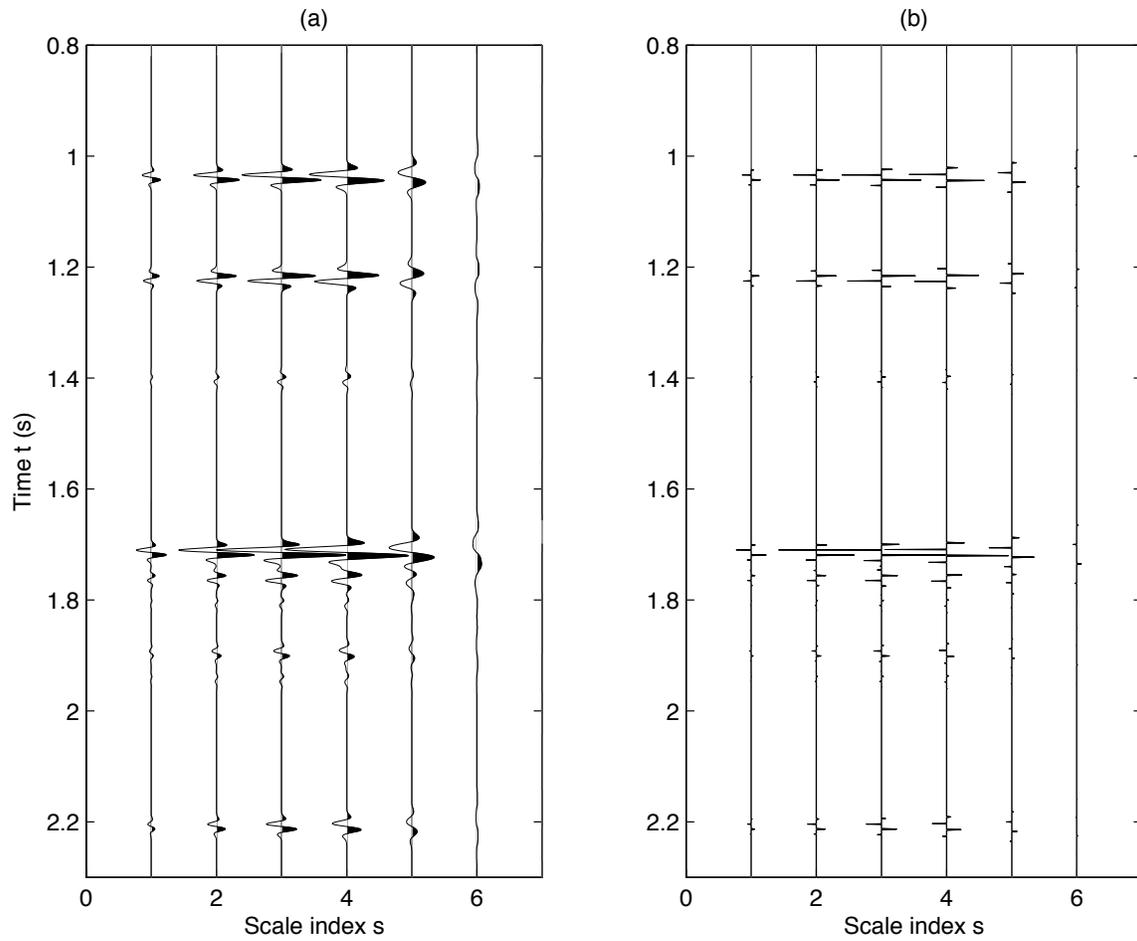


FIG. 10. Culling all but the maxima of the continuous wavelet transform. (a) The continuous wavelet transform from Figure 9b is re-plotted; (b) the maxima of (a) are retained, with all other points set to zero. In operator terms, panel (b) is generated by applying M to panel (a). Alternatively, panel (b) can be thought of as applying the sequence MW to panel (a) of Figure 9.

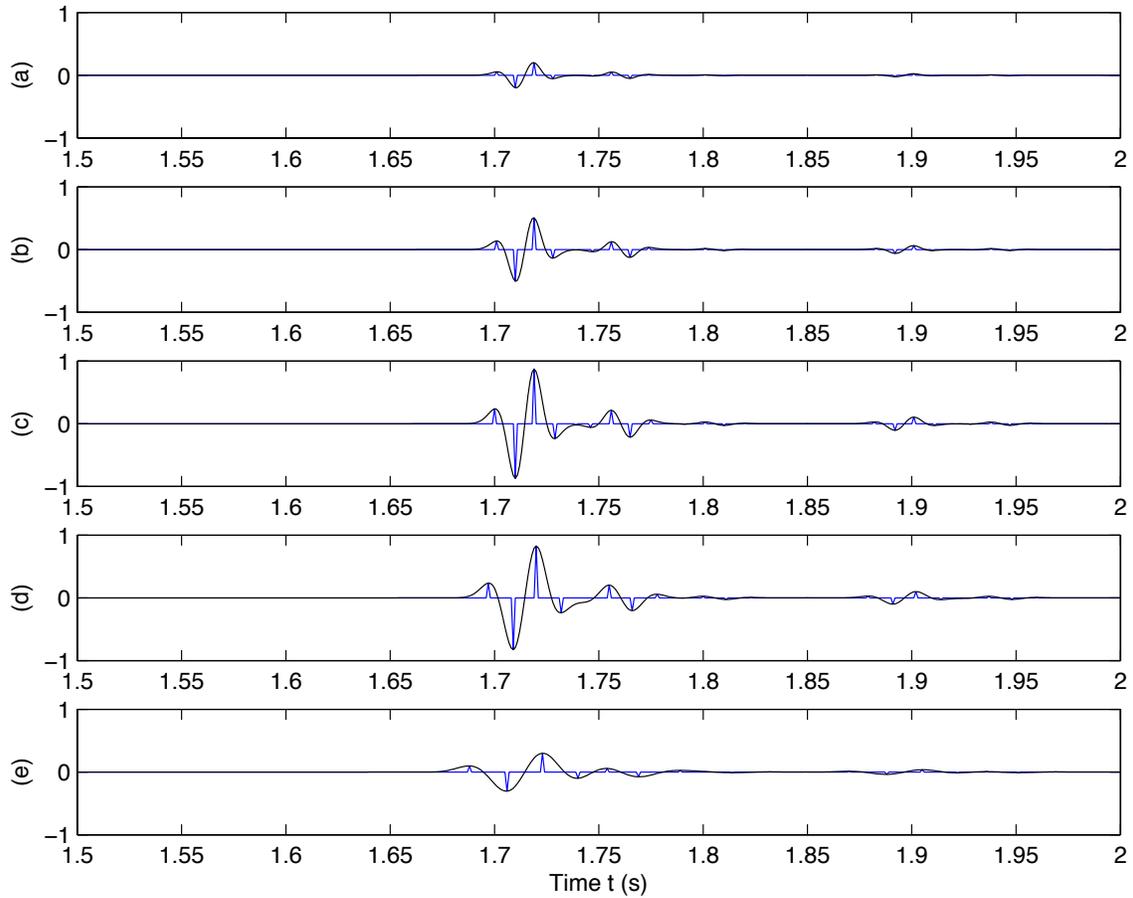


FIG. 11. Detail of the extraction of maxima (black) from the full continuous wavelet transform (blue), which are plotted a-e with scale index increasing from 1 to 5.

and so in principle there will be reconstruction error associated with the wavelet transform itself. In the operator terminology, $d(t) - \mathbf{W}^{-1}\mathbf{W}d(t) \neq 0$.

We consequently begin by comparing the input synthetic trace to its reconstruction to get a sense of the size of the errors. The result is summarized in Figure 12. In Figure 12a, the input trace is plotted in solid black, and the result of a forward followed by an inverse wavelet transform is plotted, upshifted by 0.1 amplitude units so it can be seen. In Figure 12b the percent error

$$\% \text{ Error} = 100\% \times \left(\frac{\text{Input} - \text{Inverse CWT}}{\text{Input}} \right), \quad (8)$$

is plotted as a function of time. The maximum error is on the order of 1/1000th of a percent. We conclude that for our purposes the fast forward and inverse continuous wavelet transform is essentially exact.

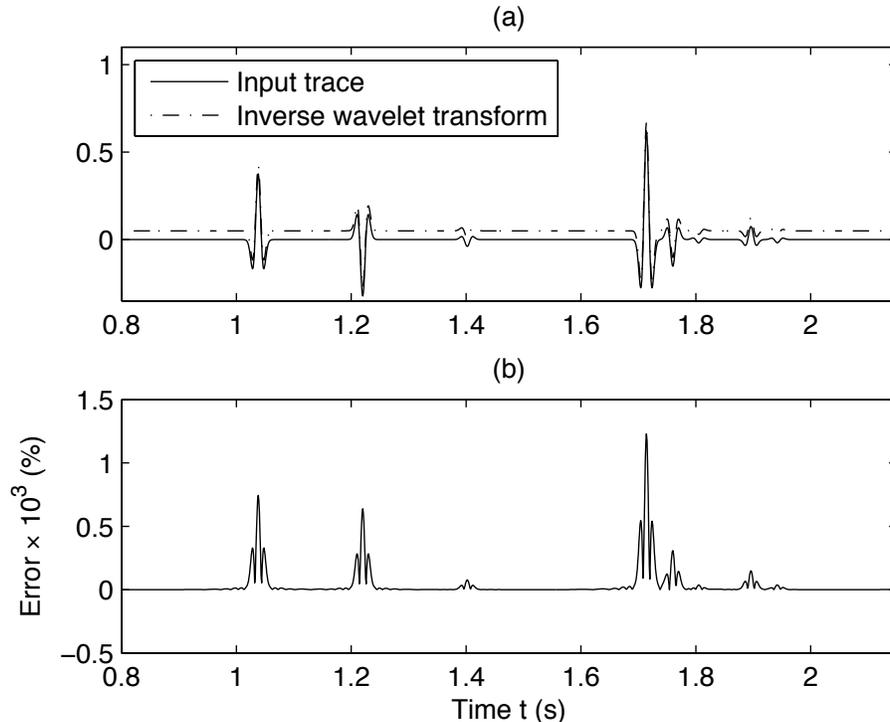


FIG. 12. The accuracy of the fast forward and inverse continuous wavelet transform. (a) The input trace (solid) vs. the result of $\mathbf{W}^{-1}\mathbf{W}$ acting on the input trace (dashed); (b) the percent error as calculated by equation (8).

The accuracy of reconstruction from maxima

We also face an a priori unknown level of error in the Mallat-Zhong reconstruction from maxima, since proofs regarding what member of the space $\Gamma \cap \mathbf{V}$ appear to be absent. Performing the iterative reconstruction as outlined in the previous section, and using the maxima in Figure 10b, namely $\mathbf{M}\mathbf{W}d(t)$, as input, we reconstruct $d(t)$, generating $\tilde{d}(t) = [\mathbf{M}\mathbf{W}]^{-1}[\mathbf{M}\mathbf{W}d(t)]$. The maximum iteration was chosen to be 50, though numerical convergence is attained at around 10; certainly, the reconstruction error remaining is not due to insufficient iterations.

In Figure 13a, the original signal is plotted as a dashed curve, with the reconstruction a solid curve. Visually the reconstruction is very good. Zooming in by a factor of roughly $10\times$, in Figure 13b, the differences emerge more clearly. Retaining the same scale as (b) and we plot their difference in Figure 13c.

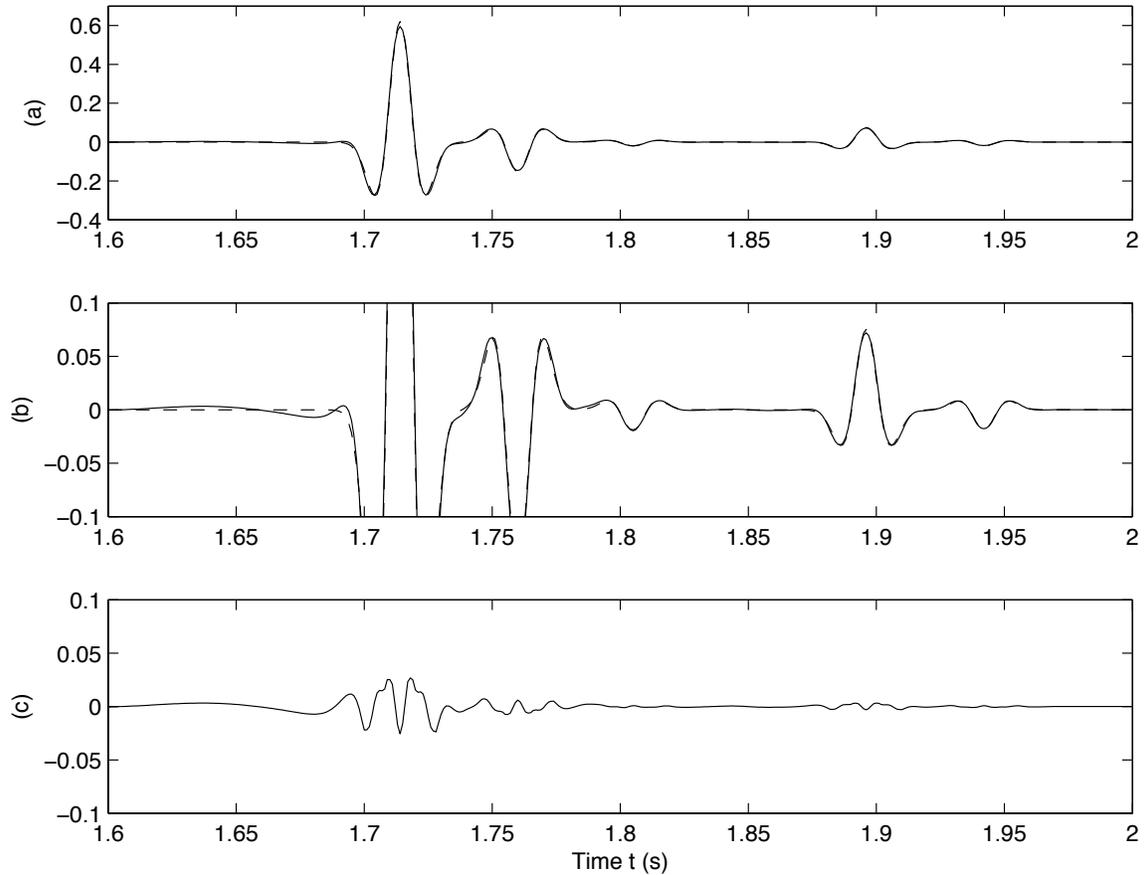


FIG. 13. Input vs. reconstruction from maxima. (a) The input trace (dashed) is plotted on top of the reconstructed trace (solid); (b) the comparison is repeated on a smaller scale; (c) the difference between the input and reconstructed traces is plotted.

Previously we considered the error associated with the inverse wavelet transform, calculated via equation (8). Let us repeat this for the reconstruction in Figure 13. In Figure 14a the inverse wavelet transform error is re-plotted for reference, and in Figure 14b the same calculation is carried out on the reconstruction from maxima. The error levels have increased by a factor of 10^3 , and now climb up to 1-5%. These errors are much more significant, and we must keep them in mind when considering application of reconstructions to our seismic traces.

Random noise removal on land data

We next consider some simple de-noising applications of processing in the wavelet transform maxima domain, carried out on poststack traces from land data. The data are pre-processed in preparation for 1D internal multiple prediction (Hernandez and Innanen,

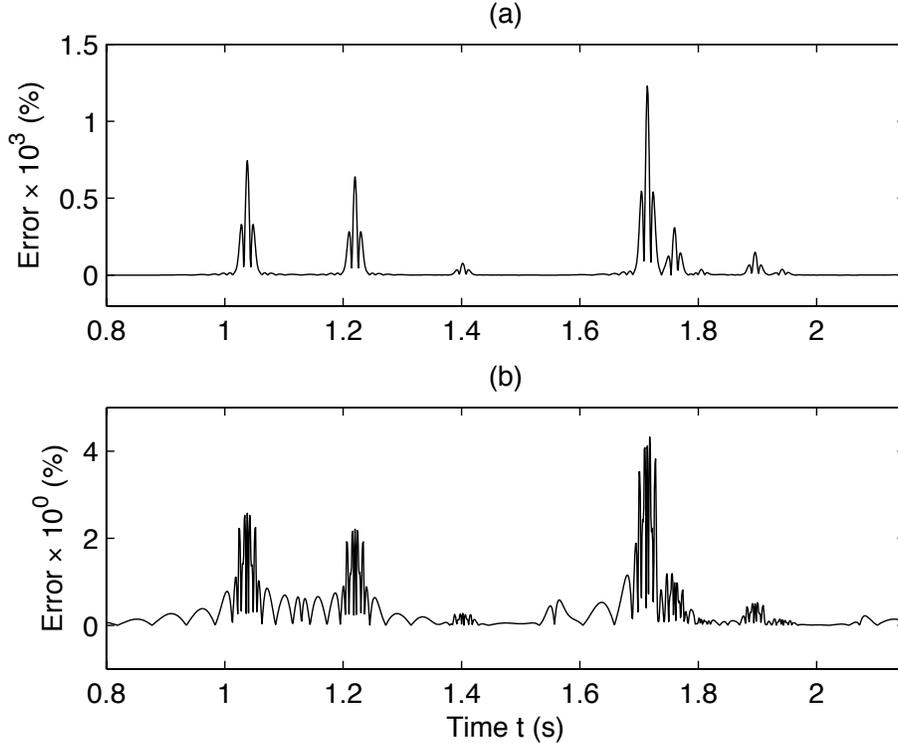


FIG. 14. Reconstruction errors. (a) The error due to the operator $\mathbf{W}^{-1}\mathbf{W}$, repeated from Figure 14a; (b) comparable error due to the iterative determination of $\hat{d}(t) = [\mathbf{MW}]^{-1}[\mathbf{MW}d(t)]$.

2013), and muted above 0.3s (a selection of traces from the post-stack section is displayed in Figure 15).

We extract from this selection the leftmost trace (Figure 16a). In spite of the muting, and in spite of the clear presence of events rising above the noise, this is not a particularly clean trace. We treat this trace as the input $d(t)$, and subject it to the continuous wavelet transform. The result, $\mathbf{W}d(t)$, is plotted in Figure 16b. The maxima of the signal in the CWT domain are then further extracted. The result, $\mathbf{MW}d(t)$, is plotted in Figure 16c. Our attempts at de-noising will now take place through some form of alteration of the “spikes” in this domain.

We will attempt to denoise this trace through simple *thresholding*, wherein all maxima below a certain amplitude are set to zero, and all maxima above that threshold are modified by a scalar multiple (we will use 1). The threshold can be a function of scale. To illustrate the process, we examine in detail a portion of the continuous wavelet transform and its maxima in Figure 17. The five smallest scales are included in panels (a)–(e), with the CWT, $\mathbf{W}d(t)$ in blue, and the maxima, $\mathbf{MW}d(t)$, in black. The output is $\mathbf{TMW}d(t)$.

We design a thresholding operator \mathbf{T} such that all maxima on scales $s = 2^j$ where $j = \{1, 2, 3, 4, 5, 6\}$ are set to zero which are below the moduli $\{1, 0.5, 0.5, 0.5, 0.1, 0.1\}$. In Figure 18a–e the results of the thresholding are displayed. The blue curves remain the same, the original CWT, but now not all of the maxima are represented, only those exceeding the threshold at each scale.

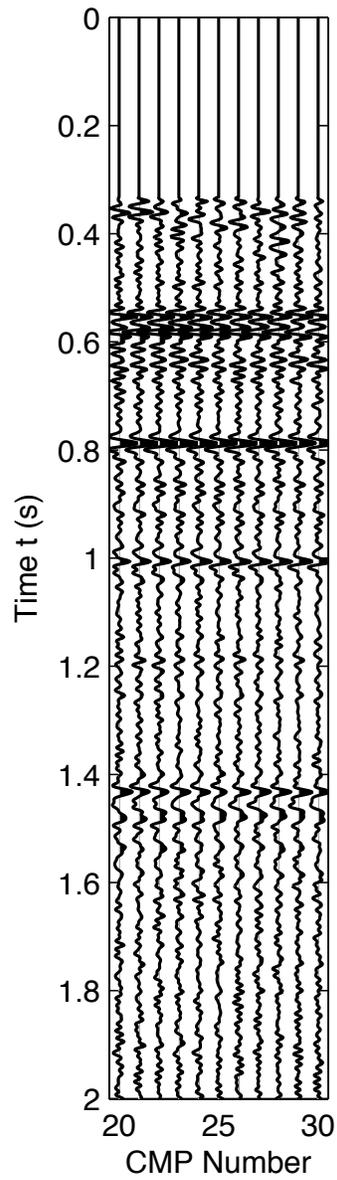


FIG. 15. Selection of traces from the pre-processed poststack section of NEBC land data.

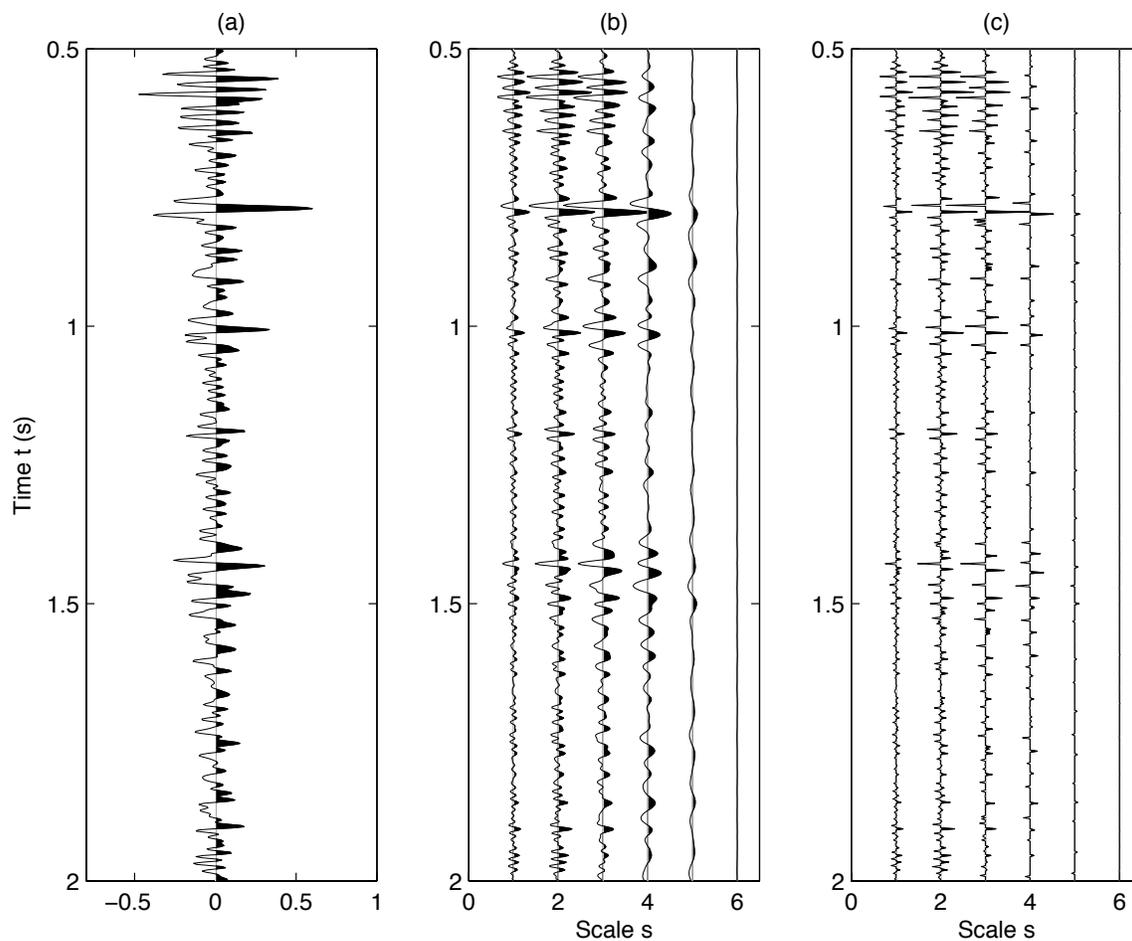


FIG. 16. Transformation of the leftmost trace in Figure 15. (a) Original trace $d(t)$; (b) continuous wavelet transform, i.e., $Wd(t)$; (c) maxima of the continuous wavelet transform, i.e., $MWd(t)$.

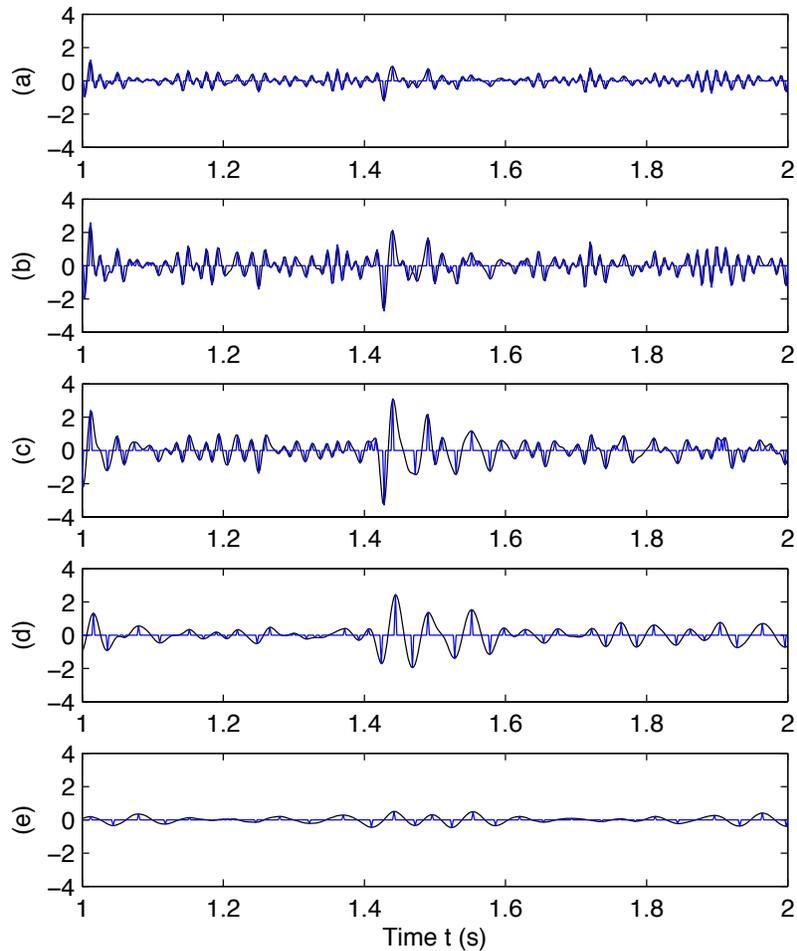


FIG. 17. The five smallest scales of the data trace, zoomed in between 1 and 2 seconds. (a)–(e) scales 1 through 5. In blue is the CWT, and in black all but the CWT maxima have been culled.

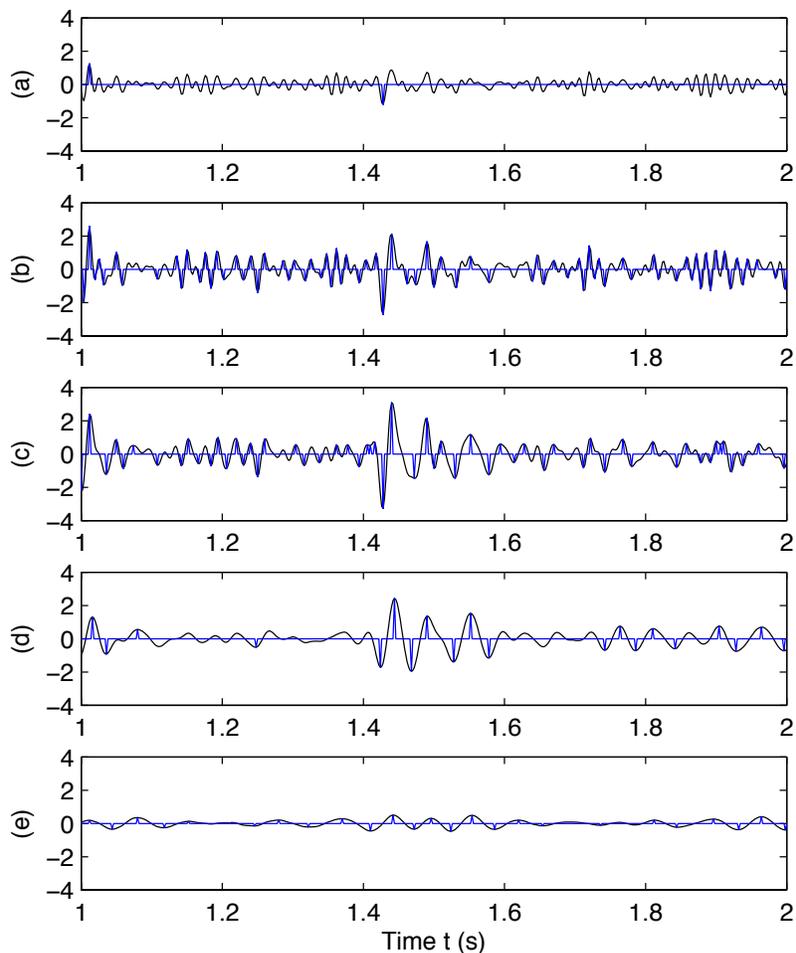


FIG. 18. Similar to Figure 17, but with the sub-threshold maxima (in black) set to zero. The matrix of numbers plotted in black is, in operator terms, $\mathbf{TMW}d(t)$.

These maxima are now subject to the iterative Mallat-Zhong reconstruction $(\mathbf{MW})^{-1}$. In Figure 19a the results, namely $\tilde{d}(t) = [\mathbf{MW}]^{-1}[\mathbf{TMW}d(t)]$, are displayed in contrast to the original trace in Figure 19b. The gains are modest—the component of (b) which has been removed is displayed in Figure 19c.

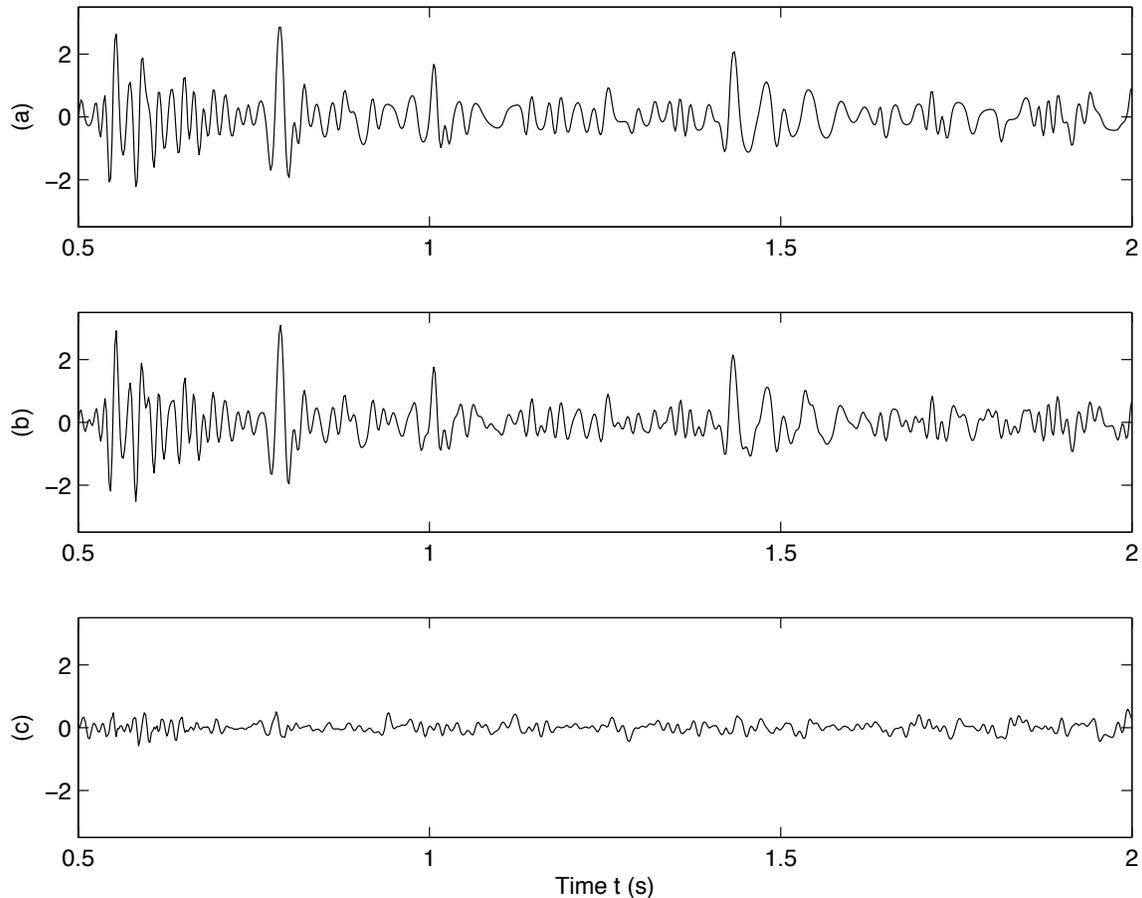


FIG. 19. Reconstruction $\tilde{d}(t) = [\mathbf{MW}]^{-1}[\mathbf{TMW}d(t)]$. (a) Reconstructed signal; (b) original signal; (c) difference.

Denosing is difficult—probably impossible—to assess, and it is unclear whether we would call Figure 19a “better” or “worse” than b. However, we can examine the noise which has been removed to see if it is largely white, or if it shares any autocorrelation features with the input. In Figure 20 the short lags of the autocorrelations of the input and the subtracted noise are compared after both are normalized. For nonzero lags, the subtracted noise has significantly reduced values relative to the input data, which is suggestive that not too much signal has been lost to this denosing.

“Aggressive” noise removal with a threshold derived from coherency across scales

Some modest improvement was detected in threshold denosing in the previous sections. However, the payoff is quite small considering the trouble gone to. Thresholding seems to hold some potential for real meaningful extraction of signal components, and we would like to investigate this also.

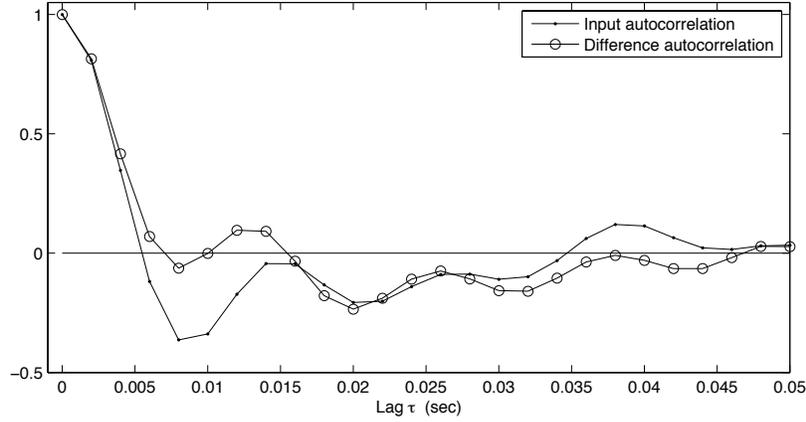


FIG. 20. Autocorrelations of the input data and the noise removed by continuous wavelet transform thresholding. Both autocorrelations have been normalized to 1 at zero lag.

Having decomposed the signal into scales with the CWT, we should consider devising a processing regimen which works on the assumption that signal and noise can be separated via their scaling behaviour. We might postulate, for instance, that *signal* be defined as those parts of the input which persist across all scales, whereas *noise* be defined as those parts without such coherency.

We will use this idea to derive a data-dependent thresholding operator. The wavelet transform $\mathbf{W}d(t)$, we recall (see e.g., Figure 2), is a matrix,

$$\mathbf{W}d(t) = \begin{bmatrix} W_{2^1}d(t) \\ W_{2^2}d(t) \\ W_{2^3}d(t) \\ \vdots \end{bmatrix}, \quad (9)$$

whose rows $W_{2^n}d(t)$ are convolutions of the input $d(t)$ with the wavelet at scales $s = 2^n$. If a portion of the input is signal, it should be coherent across the scales. This is the part of the input we would like to retain through thresholding—the incoherent part can be suppressed. We design a thresholding operator \mathbf{T} as follows. At each scale we square the continuous wavelet transform, and normalize it to unity. We then sum those normalized squares, and once more, normalize to unity. The result is a positive function of time, whose larger values coincide with regions of the continuous wavelet transform with coherence across scales:

$$\mathbf{WSQ}_n(t) = \frac{W_{2^n}^2(t)}{\max \{W_{2^n}^2(t)\}} \quad (10)$$

$$\mathbf{TH}(t) = \frac{\sum_{n=1}^N \mathbf{WSQ}_n(t)}{\max \left\{ \sum_{n=1}^N \mathbf{WSQ}_n(t) \right\}}. \quad (11)$$

The thresholding operator \mathbf{T} we propose amounts to multiplying the maxima of the continuous wavelet transform by this window. The window created by the land data trace we studied in the previous section is plotted in Figure 21.

Applying this operator to $\mathbf{M}\mathbf{W}d(t)$ and calculating the Mallat-Zhong reconstruction, and a very weak thresholding of small values, $\{0.2, 0.2, 0.2, 0.2, 0.2, 0.2\}$, we recover the

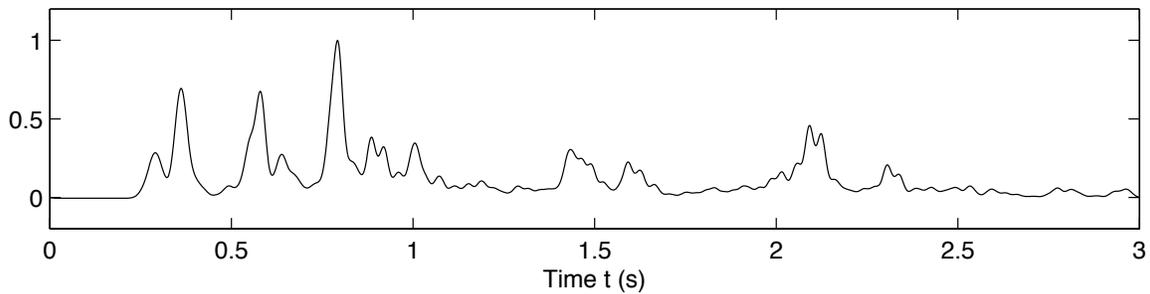


FIG. 21. Threshold operator designed from coherency of $Wd(t)$ across scales.

signal in Figure 22b, which is plotted between the original trace, Figure 22a, and a synthetic trace, Figure 22c.

CONCLUSIONS

In signal analysis some hold that the maxima of continuous wavelet transforms contain much—effectively all—of the information carried by the signal. It is possible, following some popular signal models from the early 1990s, to process as well as analyze signals in the continuous wavelet transform maxima domain. The reconstruction of a signal from its CWTM data is iterative, and its exactitude lacks proofs. However, anecdotally the quality of the reconstructions is very high, to within a few percent in its worst areas of error, as we establish here. A wide range of possible applications in exploration seismology exist, and we analyze the reduction of random noise using a data-driven CWTM thresholding operator as an example. It must be classified as “aggressive”, and thus may impact signal as well as noise, and so must be applied with caution. In a companion paper we recognize the suitability of denoised data of this kind in the construction of internal multiple prediction operators in highly challenging land environments.

ACKNOWLEDGEMENTS

This work was funded by CREWES and NSERC. We thank in particular the sponsors of CREWES for continued support. CREWES thanks Nexen Inc. for permission to publish results involving the NEBC seismic and well data.

REFERENCES

- Canny, J., 1986, A computational approach to edge detection: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **PAMI-8**, 679–698.
- Chui, C. K., 1992, *Wavelets: a tutorial in theory and application*: Academic Press, New York, first edn.
- Daubechies, I., 1992, *Ten Lectures on Wavelets*: Society for Industrial and Applied Mathematics, Philadelphia, first edn.
- Hernandez, M. J., and Innanen, K. A., 2013, Identifying internal multiples using 1D prediction: physical modelling and land data examples: submitted to *Geophysics*.
- Innanen, K. A., 2003, Local signal regularity and Lipschitz exponents as a means to estimate Q : *Journal of Seismic Exploration*, **12**, 53–74.

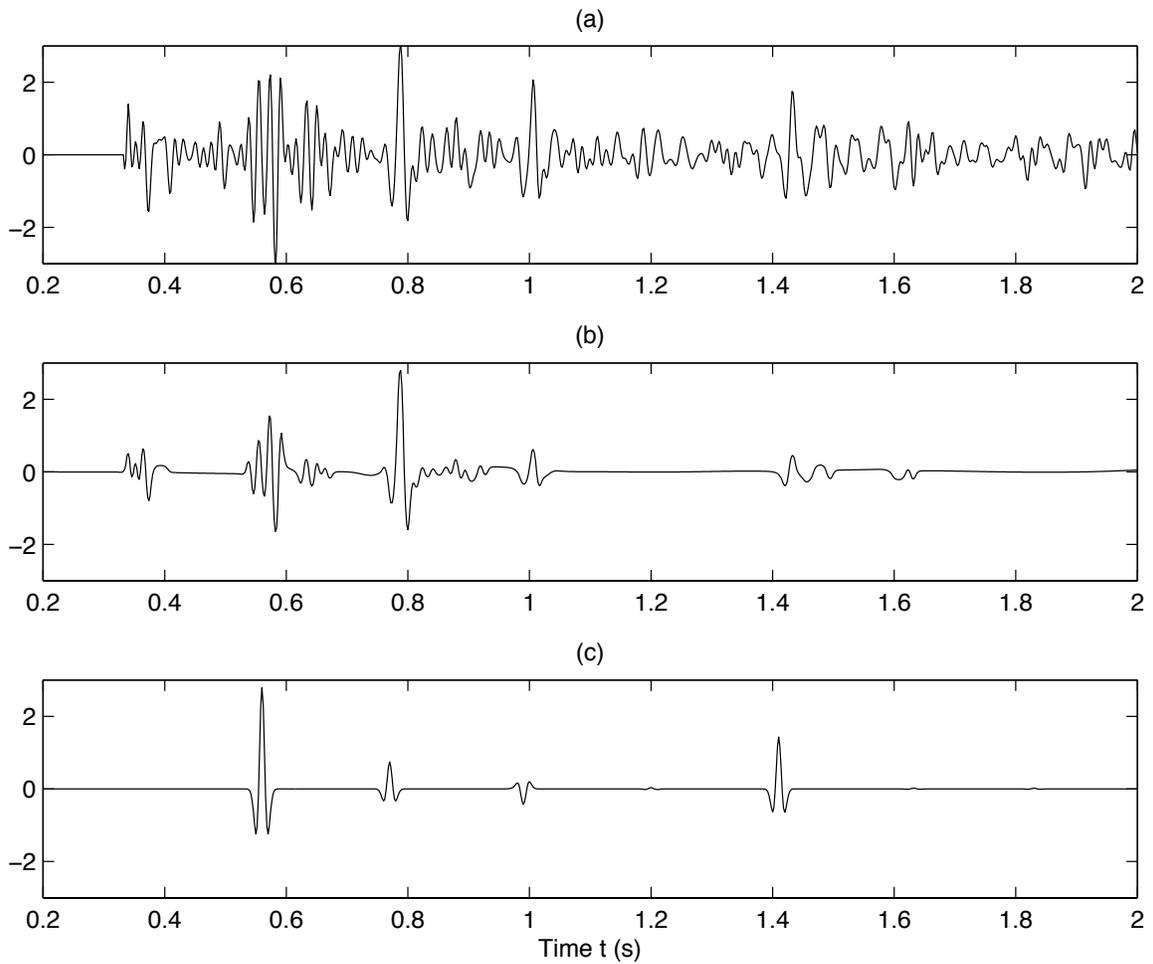


FIG. 22. “Aggressive” denoising carried out with a scale-coherency threshold. (a) Input trace; (b) denoised trace (the result of a scale-coherency designed threshold applied to the maxima of the continuous wavelet transform, followed by the Mallat-Zhong reconstruction); (c) synthetic trace.

- Innanen, K. A., 2013, Internal multiple prediction in the continuous wavelet transform maxima domain: CREWES Annual Report, **25**.
- Izadi, H., Innanen, K. A., and Lamoureux, M. P., 2011a, Continuous wavelet transforms and Lipschitz exponents as a means for analyzing seismic data: CREWES Annual Report, **23**.
- Izadi, H., Innanen, K. A., and Lamoureux, M. P., 2011b, The relationship between Lipschitz exponents and Q: synthetic data tests: CREWES Annual Report, **23**.
- Izadi, H., Innanen, K. A., and Lamoureux, M. P., 2012, Q estimation via continuous wavelet transform: CREWES Annual Report, **24**.
- Izadi, H., Innanen, K. A., and Lamoureux, M. P., 2013, Local signal regularity and smoothness as a means for seismic Q estimation: Submitted to Geophysics.
- Mallat, S. G., 1989, A theory for multiresolution signal decomposition: the wavelet representation: IEEE Transactions on Pattern Analysis and Machine Intelligence, **11**, 674–693.
- Mallat, S. G., 1998, A Wavelet Tour of Signal Processing, Geophysics reprint series: Academic Press.
- Mallat, S. G., and Hwang, W. L., 1992, Singularity detection and processing with wavelets: IEEE Transactions on Information Theory, **38**, 607–643.
- Mallat, S. G., and Zhong, S., 1992, Characterization of signals from multiscale edges: IEEE Transactions on Pattern Analysis and Machine Intelligence, **14**, 710–732.
- Zuleta, L., 2012, Near-surface characterization and V_P/V_S analysis of a shale gas basin, MSc Thesis, University of Calgary.