



# SIGNAL DECOMPOSITION AND TIME-FREQUENCY REPRESENTATION USING VARIABLE-LENGTH SYMMETRIC FILTERS

Milton J. Porsani <sup>1,2\*</sup>, and Bjorn Ursin <sup>2,3</sup>

<sup>1</sup>Universidade Federal da Bahia - UFBA, Research Center in Geophysics and Geology, Salvador, BA, Brazil

<sup>2</sup>National Institute of Science and Technology of Petroleum Geophysics (INCT-GP/CNPq), Salvador, BA, Brazil

<sup>3</sup>Norwegian University of Science and Technology - NTNU, Dept. Electron. Sys., Acoustics Group, Trondheim, Norway

\*Corresponding author: [porsani@ufba.br](mailto:porsani@ufba.br)

**ABSTRACT.** We present a time-frequency decomposition method to represent a time signal into a 2D (time  $\times$  frequency) image, which describes how the frequency content varies along the time. This is done in two steps: firstly, by filtering the signal to obtain time-components; and secondly, by computing the average instantaneous frequency (AIF), which is used for moving the data components to the time-frequency plane. For the filtering process, we present an algorithm to generate a suite of symmetric filters that are computed recursively, starting with the high-frequency content of the signal, going down in frequency and leaving the lowest frequencies in the last filter component. This can be further decomposed by continuing the procedure. The symmetric impulse responses are zero-phase with positive frequency response, and they add up to a spike at the origin with a unit frequency response. The filtering procedure gives an exact decomposition of the signal and the traveltimes are preserved. Next, the analytic signal of each component is used for computing the AIF in sliding time windows, so that for each time sample, we have an associated AIF value. The 2D time-frequency plane is obtained by distributing and adding the data components along the frequency variable. Finally, by using the time  $\times$  frequency distribution, a time-frequency filtering may be performed by stacking data of sub-domains with similar features. The new technique has been applied to two synthetic signals which have previously been analyzed by many authors using a variety of algorithms. The new signal decomposition algorithm and the AIF computation are simple and produce effective results on the synthetic data.

**Keywords:** time-series analysis, time-frequency representation, seismic noise

## INTRODUCTION

Signal decomposition and time-frequency representation have numerous applications in geophysics (Chen et al., 2001; Lesage et al., 2002; Castagna et al., 2003; Lesage, 2008; Oropeza and Sacchi, 2011; Fomel, 2013; Han and van der Baan, 2013; Tary et al., 2013; Herrera et al., 2014; Mitrofanov and Priimenko, 2015; Liu et al., 2015, 2016; Cheng and Sacchi, 2016; Ursin and Porsani, 2021), biomedical signal analysis (Angelsen, 1981; Colominas et al., 2014; Wu et al., 2016; Hu et al., 2017), and in other fields. Discussions of different methods are given by Marple (1987), Cohen (1989), Mallat (2008), Auger et al. (2013), Tary et al. (2014, 2018), Iatsenko et al. (2015), and Fourer et al. (2017).

Three popular methods for signal decomposition are: empirical mode decomposition (EMD), variational mode decomposition (VMD), and singular spectrum analysis (SSA). EMD and its extensions (Huang et al., 1998, 1999; Torres et al., 2011; Han and van der Baan, 2013; Colominas et al., 2014) compute the modes recursively, starting with the most oscillatory one. The limitations are mode mixing and splitting, aliasing, end-point artifacts, and sensitivity to noise. VMD is more robust; the signal is decomposed into narrow-band modes with slowly varying amplitudes (Dragomiretskiy and Zosso, 2014; Liu et al., 2016). The center frequencies and amplitude functions of the intrinsic modes are estimated directly from the input signal.

Singular spectrum analysis (SSA) (Harris and Yuan, 2010; Oropeza and Sacchi, 2011; Cheng and Sacchi, 2016; Hu et al., 2017; Rodrigues et al., 2018; Golyandina and Zhigljavsky, 2020) is based on the singular value decomposition (SVD), (Golub and Loan, 1996) of the so-called trajectory matrix, which is equal to the filter matrix used in digital filtering (Robinson and Treitel, 2000). Harris and Yuan (2010) showed that the signal components are equal to the original data filtered with the autocorrelation functions of the eigenvectors of the covariance matrix. Since the autocorrelation function is symmetric, the signal phase is preserved in each component.

Porsani et al. (2019) extended the classical SSA by recursively computing the components of the signal using iterations in the dimension of the covariance matrix and, in each iteration, filtering with the autocorrelation of the first eigenvector. After a certain number of iterations, the result is a new signal component, which is subtracted from the signal, and the process is repeated. This process, termed recursive-iterative SSA (RI-SSA), is simplified here by assuming that the signal covariance matrix is a positive constant times an identity matrix, corresponding to a random signal. For the eigenvectors of the covariance matrix, we use a normalized vector with equal positive components. This is used recursively to compute filters that are applied in an iterative loop to compute a new signal component. The process is repeated for the remaining signal components, and the final remainder is the last signal component. The result is an exact signal decomposition with deterministic, variable-length symmetric filters which only depend on the number of iterative convolutions in the recursive signal estimation loop. It is an exact decomposition of the signal, keeping its phase, starting with the high-frequency components, going down in frequency content, and leaving the lowest frequency part in the residual (the last component). The new method is much simpler than the RI-SSA method proposed by Porsani et al. (2019), and on synthetic data examples it provides better results.

We shall consider the decomposition of a time signal into components depending on their frequency content. For each component, we shall compute a time-varying amplitude and average instantaneous frequency (AIF). The average frequency function of each signal component is computed from its analytic signal (Gabor, 1946; Taner et al., 1979) in a sliding time window using a one-step prediction-error operator (Porsani et al., 2019). Angelsen (1981) has shown that the average frequency is the average of the instantaneous frequency, which is being computed normally (Han and van der Baan, 2013; Vesnaver, 2017; Porsani et al., 2019). The real part of the complex signal value in the middle of the time window and the average instantaneous frequency (AIF) give the time-frequency representation of the signal component. The time-frequency distribution of the original

signal is then the sum of these time-varying spectral components.

The new technique has been applied to two synthetic signals which have previously been analyzed by many authors using a variety of algorithms. The first example (Tary et al., 2014) consists of three wavelets and several signals with variable frequency content. In the second example (Han and van der Baan, 2013) the signal consists of three wavelets and four harmonic components. Finally, a single seismic trace is analyzed and decomposed. Combining components in specific areas of the time-frequency domain results in a new signal-decomposition and filtering procedure.

## VARIABLE-LENGTH SYMMETRIC FILTERS

Porsani et al. (2019) proposed a recursive-iterative scheme for signal decomposition wherein the inner loop filter derived using SSA was applied to a data vector. The filter was a scaled version of the autocorrelation function of the normalized first eigenvector of the  $(k+1) \times (k+1)$  data covariance matrix. Now we let the covariance matrix be a  $(k+1) \times (k+1)$  identity matrix times a positive constant, corresponding to a random data vector. A normalized eigenvector of dimension  $(k+1) \times 1$  is

$$\mathbf{v}_k = \frac{1}{\sqrt{k+1}}(1, \dots, 1)^T = (v_k(0), \dots, v_k(k))^T \quad (1)$$

The symmetric filter becomes (see Porsani et al. (2019), equation (5), and the pseudo-code for their algorithm)

$$\mathbf{g}_k = \frac{1}{k+1} \mathbf{v}_k \oplus \mathbf{v}_k \quad (2)$$

where  $\oplus$  denotes correlation, and then

$$g_k(n) = \frac{k+1-|n|}{(k+1)^2}, \quad n = 0, \pm 1, \dots, \pm k. \quad (3)$$

From this, we may form the  $(N+1) \times (N+1)$  symmetric matrix

$$\mathbf{G}_k = \begin{bmatrix} g_k(0) & \dots & g_k(k) & \mathbf{O}_{N-k} \\ \vdots & g_k(0) & \ddots & g_k(k) \\ g_k(k) & \ddots & \ddots & \vdots \\ \mathbf{O}_{N-k} & g_k(k) & \dots & g_k(0) \end{bmatrix} \quad (4)$$

When a vector  $d(n)$  of dimension  $(N+1) \times 1$  is multiplied by this matrix it corresponds to convolution with the filter  $g_k$  and setting the components of the vector equal to zero outside their domain of definition. In the following we shall denote this operation by  $(g_k * d)(n)$ .

In the recursive filter loop we shall compute the

filters

$$f_k = g_k * f_{k-1} = g_k * g_{k-1} * \dots * g_1. \quad (5)$$

Figure 1 illustrates the filters  $g_k$ , the filters  $f_k$  in time and its corresponding frequency response.  $g_k$  and  $f_k$  were normalized before plotting. The signal is convolved  $J$  times with the filter  $(\delta - f_k)$  corresponding to the effective filter

$$c_k = [* (\delta - f_k)]^J \quad (6)$$

where  $\delta$  is the unit convolutional operator.

In the component recursive loop, the data remainder is convolved with this filter to obtain a new component. This is removed from the data remainder, and the process is repeated with a new filter. The filters do not depend on the data, so they can be computed in a separate recursive scheme.

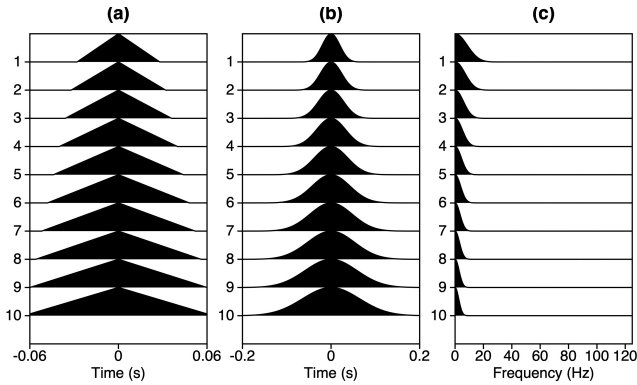


Figure 1: The normalized filters  $g_k$  and  $f_k$  in (a) and (b), respectively, and the frequency responses of  $f_k$  in (c).

The result is the pseudo-code to compute the signal components:

- Initial values:  $d_1 = d$ ,  $f_0 = \delta$
- **DO**  $k = 1, \dots, K$ 
  - $f_k = g_k * f_{k-1}$  Filter loop
  - $c_k = [* (\delta - f_k)]^J$
  - $x_k = c_k * d_k$  Component loop
  - $d_{k+1} = d_k - x_k = (\delta - c_k) * d_k$
- **END DO**
- $x_{K+1} = d_{K+1}$

The signal decomposition is

$$d = \sum_{k=1}^K x_k + d_{K+1} = \sum_{k=1}^K h_k * d + h_{K+1} * d. \quad (7)$$

The last term,  $d_{K+1}$ , is the signal residual which may be further decomposed by applying filters  $c_{K+1}$

and higher-order. The filters  $h_k$  represent the total impulse responses of the system, given by

$$\begin{aligned} h_1 &= c_1, \\ h_k &= c_k * (\delta - c_{k-1}) * \dots * (\delta - c_1) \quad k = 2, 3, \dots \end{aligned} \quad (8)$$

and

$$h_{K+1} = \delta - \sum_{k=1}^K h_k. \quad (9)$$

In the algorithm we compute  $x_k = c_k * d_k$  equivalent to multiplication with a matrix as defined in Equation 4. Both  $d_k$  and  $x_k$  are of length  $N + 1$ , and  $c_k$  is cut off to length about  $N/2$ . The filters  $c_k$  are independent of the data, and further filters may be computed to obtain more signal components with lower frequency content. All filters  $c_k$  depend on  $J$ , the number of factors in the convolutional product in Equation 6.

In Appendix A we have further analyzed the impulse responses. The results are:

1. The signal decomposition starts with the high-frequency parts and low-frequency components may be added by computing higher-order components that are independent of the previous ones.
2. The filters are symmetric, zero-phase and the sum of the impulse response elements is a unit delta function (see Equation A19).
3. The Fourier spectra are positive and less than one, and their sum is one.

This means that the signal is exactly represented by the decomposition and that the signal phase is preserved in each component. We may decompose a signal into a large number of components and, after that, we sum the components with similar characteristics to obtain a smaller number of components. A classification algorithm, based on the SSA method, to provide a fully automatic component extraction is presented by Harmouche et al. (2018).

In order to investigate the influence of the number of convolutional factors,  $J$ , we have computed the impulse responses  $h_k$  which we normally do not compute because they become very long (see Appendix A). They are shown in Figures 2 to 6 for  $K=9$  components, and  $J = 1, 15, 30, 50$  and 100 convolutions in the filter loop. For  $J = 1$  there is large overlap between the different frequency responses, and the low-frequency response is narrow. At  $J = 15$ , there is already improved frequency separation, and uniform low-frequency response (it can be further decomposed by adding more components). For  $J = 30, 50$  and 100, the responses at medium frequencies become narrower; however, there is mode mixing between the first three high-frequency components.

It is seen that the convolutional product is necessary to separate different frequency components of the

signal. Large  $J$  gives a longer impulse response, but the frequency response does not change significantly.

$J$  and  $K$  are user-defined parameters where  $J$  does not need to be very large and  $K$  should be such that the meaningful signal components appear between 1 and  $K$ . If there are still separate low-frequency components in the remaining part,  $K + 1, K$  may be increased to compute further components.

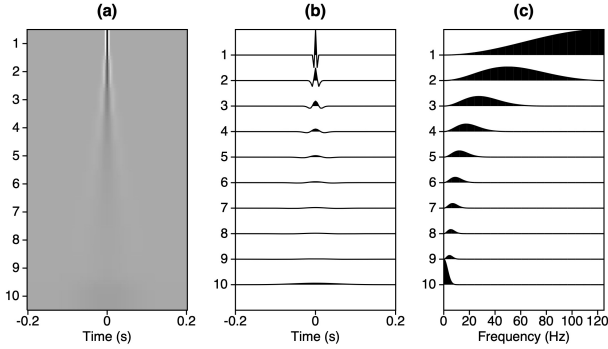


Figure 2: Filter impulse responses for  $K = 9$  and  $J = 1$  (one single convolution). (a) and (b) time responses, (c) frequency responses.

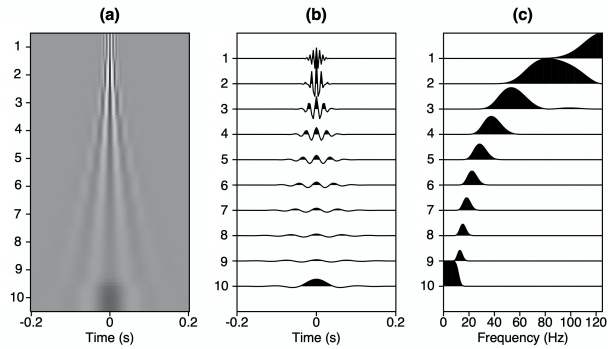


Figure 3: Filter impulse responses for  $K = 9$  and  $J = 15$  (filter convolutions). (a) and (b) time responses, (c) frequency responses.

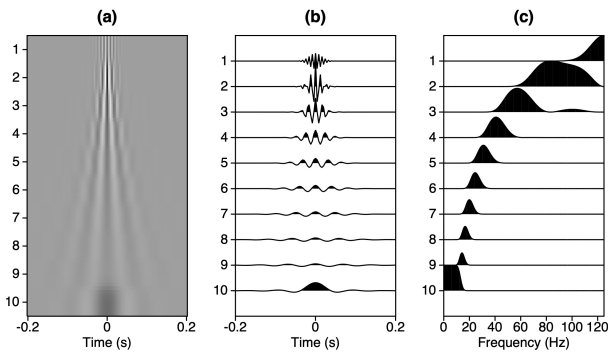


Figure 4: Filter impulse responses for  $K = 9$  and  $J = 30$  filter convolutions. (a) and (b) time responses, (c) frequency responses.

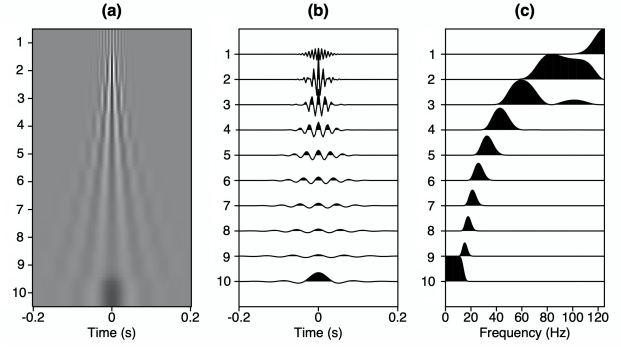


Figure 5: Filter impulse responses for  $K = 9$  and  $J = 50$  filter convolutions. (a) and (b) time responses, (c) frequency responses.

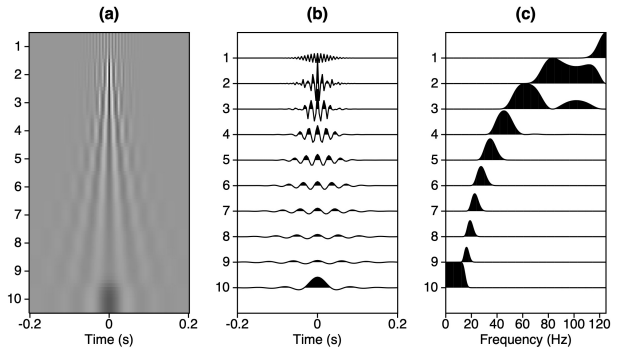


Figure 6: Filter impulse responses for  $K = 9$  and  $J = 100$  filter convolutions. (a) and (b) time responses, (c) frequency responses.

### TIME-FREQUENCY DECOMPOSITION

In this section, we present the equation used to compute the average instantaneous frequency (AIF) for the signal components, and we show how to obtain a time-frequency matrix representation for the original signal.

From each estimated signal component  $x_k(t)$  we form the analytic signal

$$z_k(t) = x_k(t) + iy_k(t) = A_k(t)e^{i\phi_k(t)} \quad (10)$$

where  $y_k(t)$  is the Hilbert transform of  $x_k(t)$ , and for each complex component we consider a time window,  $z_k(t + j)$ ,  $j = -L, \dots, L$ , of length  $2L + 1$  centered at  $t$ . The instantaneous frequency, which is often used to characterize a signal, is given by

$$f_k(t) = \frac{1}{2\pi} \phi'_k(t). \quad (11)$$

Computing this time derivative, as shown in Appendix B, is an unstable operation, and Porsani et al. (2019) used a local averaging procedure to stabilize the estimate of the instantaneous frequency.

Here we compute the AIF for each signal compo-



nent  $k$  (see [Appendix B](#)),

$$\begin{aligned}\bar{f}_k(t) &= \frac{1}{\pi\Delta t} \text{Im}\{\log[1 + r_k(t)]\} \\ &= \frac{1}{\pi\Delta t} \arctan\left\{\frac{\text{Im}\{r_k(t)\}}{1 + \text{Re}\{r_k(t)\}}\right\}\end{aligned}\quad (12)$$

where  $\Delta t$  is the time-sample interval, and  $r_k(t)$  is the reflection coefficient associated with the prediction error operator (PEO) of order 1 ([Burg, 1975](#)).

The AIF value computed using [Equation 12](#) is attributed to the complex data,  $z_k(t)$ , at the center of the time window. By computing the AIF for every signal component and for all time sample we obtain the time-frequency distribution,

$$\{b[t, \bar{f}_k(1)], \dots, b[t, \bar{f}_k(K)]\} = \sum_{k=1}^K z_k(t) \delta[f - \bar{f}_k(t)] \quad (13)$$

To obtain a matrix representation from [Equation 13](#) we may discretize the AIF values by using a frequency-sample interval,  $\Delta f$ ,

$$b[t, \bar{f}_k(t)] \approx b[t, j_k(t)\Delta f] = z_k(t) \delta\{[j - j_k(t)]\Delta f\} \quad (14)$$

where,

$$j_k(t) = \text{int}\left\{\frac{|\bar{f}_k(t)|}{\Delta f}\right\} \quad (15)$$

is a integer positive number,  $0 \leq j_k(t) \leq f_{Ny}/\Delta f$ , ( $f_{Ny} = 1/2\Delta t$ , the Nyquist frequency).

For each signal component we use [Equation 14](#) and [Equation 15](#) to compute the elements  $b[t, j_k(t)\Delta f]$ ,  $\{t = 0, \dots, N\}$  and  $\{j_k(t) = 0, \dots, f_{Ny}/\Delta f\}$ , to form a time-frequency matrix,  $\mathbf{B}_k$ . By adding these matrices we obtain the time-frequency matrix representation of the original signal,

$$\mathbf{B} = \sum_{k=1}^K \mathbf{B}_k. \quad (16)$$

From [Equation 13](#) and [Equation 14](#) we note that,

$$z(t) = \sum_{k=1}^K b[t, \bar{f}_k(t)] = \sum_{k=1}^K b[t, j_k(t)\Delta f] = \sum_{k=1}^K \mathbf{B}_k(t). \quad (17)$$

$\mathbf{B}_k(t)$  is a frequency slice of the time-frequency matrix  $\mathbf{B}_k$ . [Equation 17](#) provides an exact time-frequency decomposition of the original signal. Adding all frequency components for a given time, the original signal will be recovered. Also, it is possible to use the method as a 2D (time  $\times$  frequency) filtering procedure, by adding groups of components with similar features in subdomains of the 2D time-frequency plane, to obtain fewer components with cleaner characteristics.

A pseudo-code for the time-frequency decomposi-

tion is:

Initial parameters:

$2L + 1$ : window length to compute AIF values,

$K$ : number of components,

$\Delta f$ : frequency sample interval

- DO  $k = 1, \dots, K$  (Component loop)
  - obtain component,  $x_k(t)$ , using the previous algorithm
  - obtain the complex component,  $z_k(t)$ , (eq. 10),
    - \* DO  $t = 0, \dots, N$ 
      - compute the AIF,  $\bar{f}_k(t)$ , (eq. 12)
      - compute the frequency index,  $j_k(t)$ , (eq. 15),
      - add component,  $z_k(t)$ , to the time-frequency matrix (eq. 17)
    - \* END DO
- END DO

We remark that the length of the time window to compute the AIF values may be adjusted to adapt to the frequency variation in the signal. In the numerical examples we have chosen  $2L+1 = k(k+1)+1$ , equal to the length of the filter  $f_k$  (eq. A-9), where  $k$  is the index of the component being analyzed.

## APPLICATIONS

To test the proposed algorithms we analyze two noise-free synthetic signals. [Tary et al. \(2014\)](#) defined a composite signal consisting of a Morlet wavelet and several components with varying frequency. They used 8 different methods to compute the time-frequency representation of this signal, and [Andrade et al. \(2018\)](#) used complex auto-regressive time-frequency analysis. [Porsani et al. \(2019\)](#) showed both a signal decomposition and the corresponding time-frequency representation, including amplitude information.

In the second synthetic data example we decompose a signal analyzed by [Han and van der Baan \(2013\)](#), consisting of a Morlet wavelet, two Ricker wavelets, and several harmonic components, using a short-time Fourier transform (STFT), the continuous wavelet transform (CWT) and several EMD algorithms. It has also been analyzed by [Herrera et al. \(2014\)](#) using the CWT, and EMD algorithms and the synchrosqueezing transform (SST). It was also analyzed by [Liu et al. \(2016\)](#) using two EMD algorithms and VMD. Furthermore, [Andrade et al. \(2018\)](#) and [Porsani et al. \(2019\)](#) also produced time-frequency representation for this signal.

### Synthetic signal I

The first synthetic signal we analyse is a sum of five elements ([Tary et al., 2014](#)):

$$\begin{aligned}
 s_1(t) &= 3 e^{-1250(t-2)^2} \cos[710(t-2)] & 0 \text{ s} \leq t \leq 10 \text{ s} \\
 s_2(t) &= \sin\left\{\frac{8\pi 100^{t/8}}{\log(100)}\right\} & 6 \text{ s} \leq t \leq 10 \text{ s} \\
 s_3(t) &= 0.7 \cos(130\pi t) + 5 \sin(2\pi t) & 4 \text{ s} \leq t \leq 8 \text{ s} \\
 s_4(t) &= 0.6 \cos(70\pi t) & 0 \text{ s} \leq t \leq 6 \text{ s} \\
 s_5(t) &= 0.8 \cos(30\pi t) & 0 \text{ s} \leq t \leq 6 \text{ s}
 \end{aligned}$$

It is composed of two harmonic components with frequencies of 15 and 35 Hz, a frequency-modulated harmonic around 65 Hz, a sliding harmonic from 35 to 158 Hz, and a Morlet wavelet with a central frequency of approximately 113 Hz.

Figure 7 shows the 5 components and the synthetic signal at the top. Its corresponding time-frequency representations are shown in Figure 8.

The signal was analyzed with our algorithm with  $K = 9$  and  $J = 50$  filter convolution. The composite signal is shown at the top of Figure 9, and its 10 components are shown below. The time-frequency representations of the first 10 components are shown in Figure 10, and their sum, a composite time-frequency representation of the signal, is shown on the top to the left. These 10 components are summed in groups to form four signal components as shown in Figure 11, and the time-frequency representations of these signals are shown in Figure 12.

These results are slightly better than the ones obtained by Porsani et al. (2019) with a more complicated algorithm. Our composite time-frequency representation in Figure 12 can also be compared to the results in Figure 3 in Tary et al. (2014), and Figure 2 in Andrade et al. (2018). In both cases, the authors provide time-frequency representations of the same signal using various algorithms.

Tary et al. (2014) obtained the best result with basis pursuit, Chen et al. (2001) used a decomposition of Morlet wavelets, and Andrade et al. (2018) obtained the best result with complex short-term autoregressive modeling using the Marple (1980) or Morf et al. (1977) algorithm for computing the complex prediction error operator.

In the latter time-frequency representation, there are blanks in the representations of the 15 Hz and 35 Hz signals at 2 s due to the influence of the Morlet wavelet. Only the basis pursuit algorithm produces better time-frequency representation than our new algorithm, but there is no decomposition into signal components.

### Synthetic signal II

The second synthetic signal is taken from Han and van der Baan (2013). It is composed of an initial 20 Hz cosine wave, a 100 Hz Morlet wavelet at 0.3 s, two 30 Hz Ricker wavelets at 1.07 s and 1.1 s, and three different frequency components between 1.3 s and 1.7

s of 7, 30 and 40 Hz, respectively. The 7 Hz component is not continuous and has three separate less than one-period portions which appear at 1.37, 1.51, and 1.65 s.

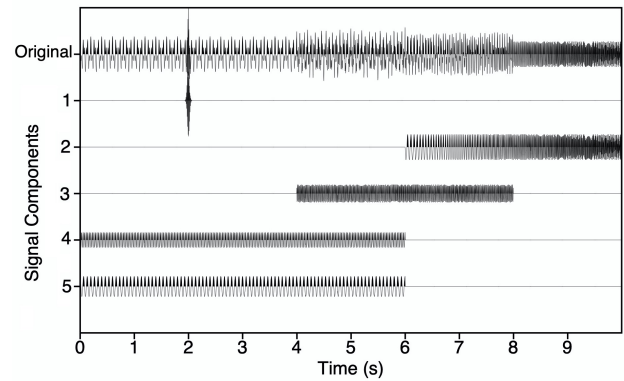


Figure 7: The 5 signal components and the synthetic signal (Tary et al., 2014) at the top.

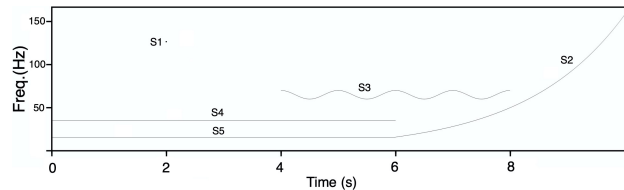


Figure 8: The time-frequency representation of the signal components in Figure 7.

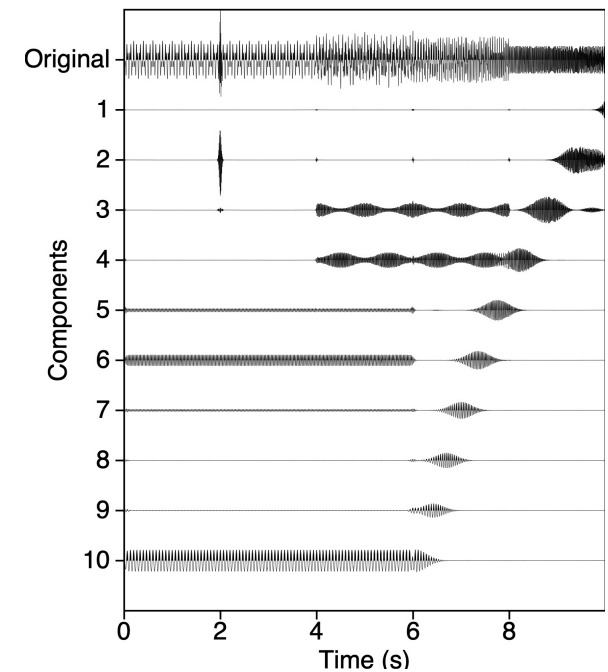


Figure 9: Decomposition of the original signal (Tary et al., 2014) shown on the top.  $K = 9$  and  $J = 50$  filter convolutions.

The composite signal is shown in Figure 17a. It has been decomposed into 25 components and 4 different values  $J$ . In Figure 13 is shown the results for  $J = 1$ , in (a),  $J = 5$ , in (b),  $J = 15$ , in (c),

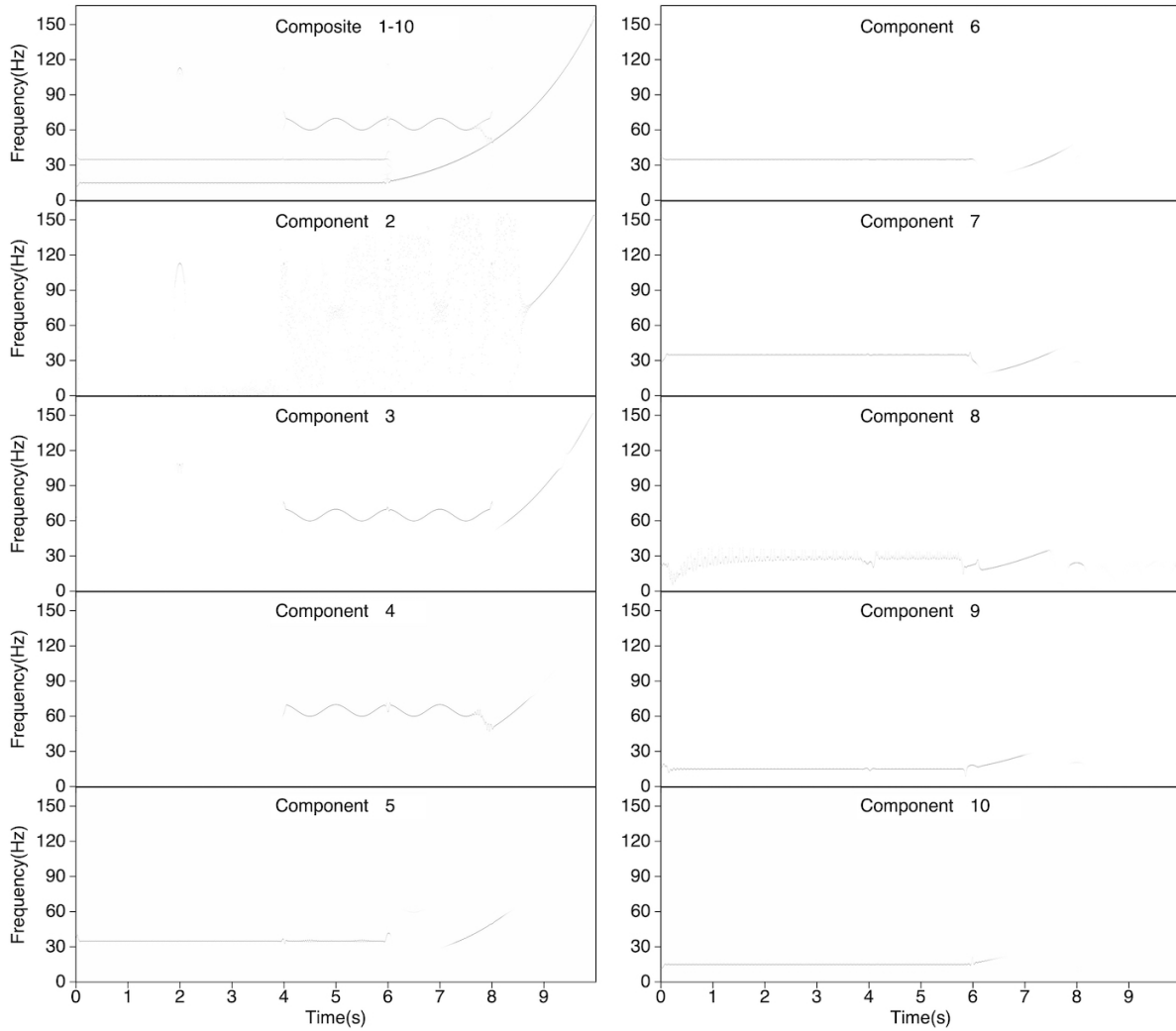


Figure 10: Time-frequency representation ( $K=9, J=50$ ) of the first 10 components of the signal shown in Figure 7. The composite time-frequency representation is shown on the top left.

and  $J = 50$  in (d). As shown in the previous example, by increasing the number  $J$ , of convolution, the meaningful features of the components move to the components of higher-order.

Figure 14 shows the time-frequency representation for the components in Figure 13c. The boxes indicate areas where the time-frequency components will be extracted. The resulting six modes are shown in Figure 15 and the composite time-frequency representation is shown in Figure 16. The resulting signal decomposition (Figure 15) is very clean with well-separated components. They are summed to form the composite signal in Figure 17b. It is very similar to the original in Figure 17a, however there is a small error signal as shown in Figure 17c. We do not have access to the individual components of the input signal but from Figure 17c we can see that there is a small end effect for the three parts of the 20 Hz signal. The two Ricker wavelets are almost perfectly recovered.

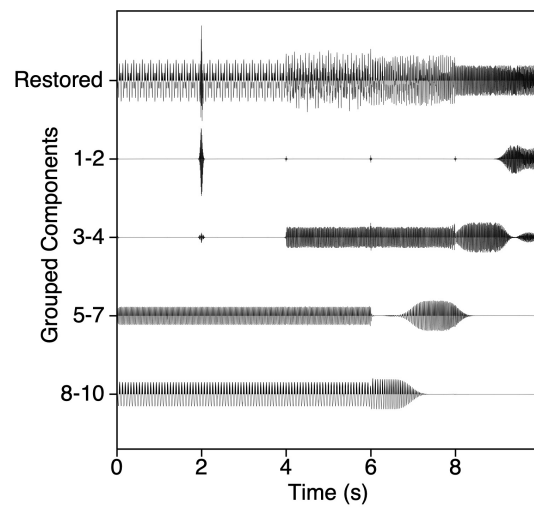


Figure 11: Decomposition of the signal in Figure 7 by combining the elements in Figure 9 in common signal groups.

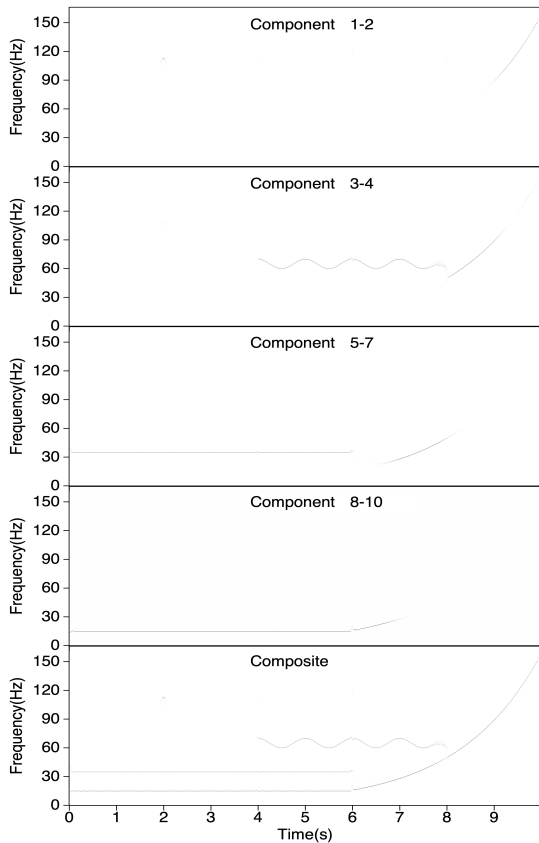


Figure 12: Time-frequency representation of the signal components shown in Figure 11.

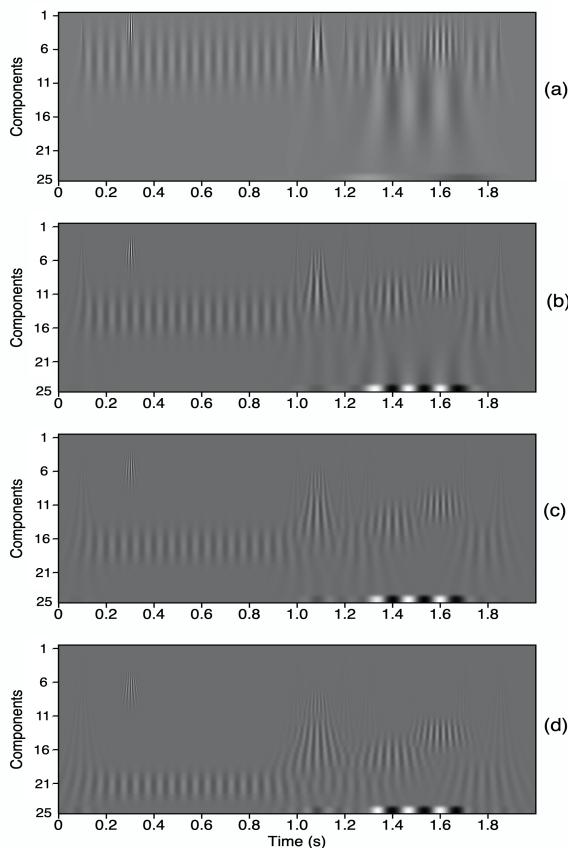


Figure 13: Decomposition of the original signal (Han and van der Baan, 2013) shown in Figure 17a.  $K=24$  and convolutions varying from  $J=1$  in (a),  $J=5$  in (b),  $J=15$  in (c) and  $J=50$  in (d).

This signal has also been analyzed previously by Han and van der Baan (2013), Herrera et al. (2014), Liu et al. (2016), Andrade et al. (2018), and Por-sani et al. (2019) using a large variety of methods. Our overall impression is that VMD (Liu et al., 2016) gives the best-combined mode separation and time-frequency representation. For signal decomposition we can compare our Figs. 15 and 16 to Fig. 8c in Liu et al. (2016). Our mode separation is much cleaner than the one obtained with VMD. In particular, the two Ricker wavelets are very well represented in Figure 15. The time-frequency representation obtained with VMD is shown in Fig. 8c in Liu et al. (2016). When this is compared to our Figure 15, we believe that our result is slightly better, having fewer variations in the constant-frequency parts of the signal.

### Analysis of a seismic trace

The seismic trace shown in Figure 20a is decomposed in 20 components with  $K=19$  and  $J=15$  as shown in Figure 18. The time-frequency representation is shown in Figure 19 where it is also shown two zones in which the signal will be decomposed. Adding the components in these two zones results in the two components shown in Figure 20b and Figure 20c. It is clearly seen that the high-frequency noise is separated from the low-frequency signal. The average amplitude spectra are shown in Figure 21.

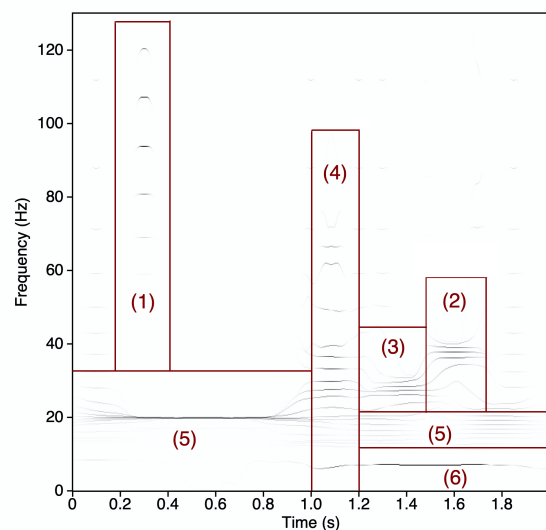


Figure 14: The time-frequency representation for the (Han and van der Baan, 2013) signal, shown in Figure 17a, for  $K=24$  components and  $J=15$  filter convolution. The boxes indicate areas from where the time-frequency data will be summed.

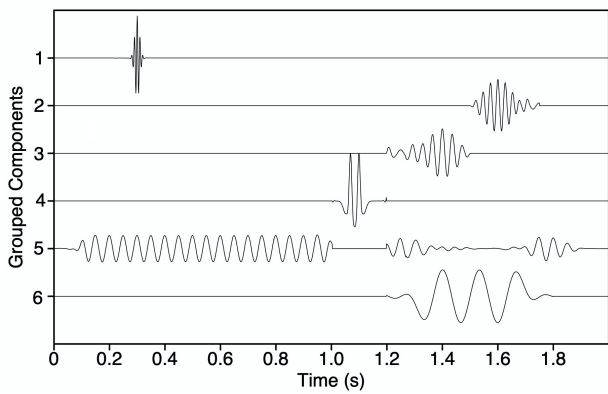


Figure 15: Signal components obtained by stacking time-frequency data inside the boxes in Figure 14.

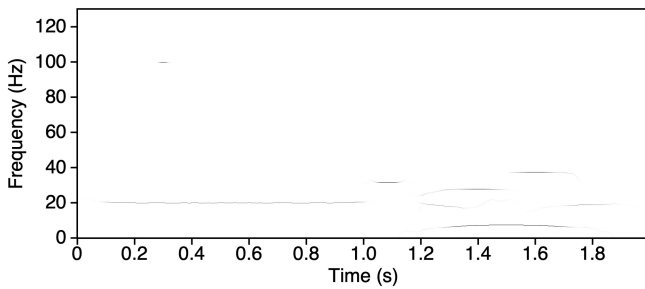


Figure 16: Absolute value of the time-frequency distribution of the signal components in Figure 15.

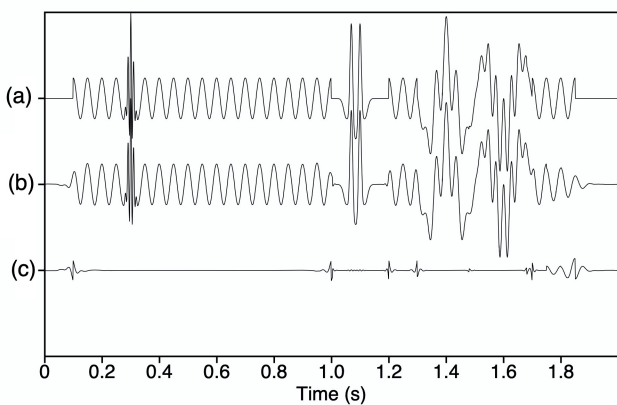


Figure 17: The original synthetic signal (Han and van der Baan, 2013) in (a), the reconstructed signal in (b) and the difference in (c).

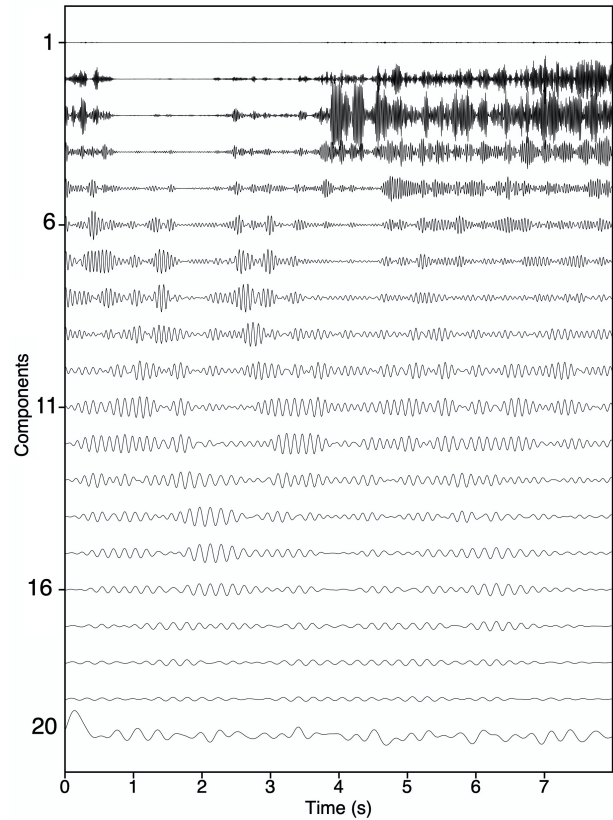


Figure 18: Decomposition of the seismic signal shown in Figure 20a, for  $K = 19$  and  $J = 15$  filter convolutions.

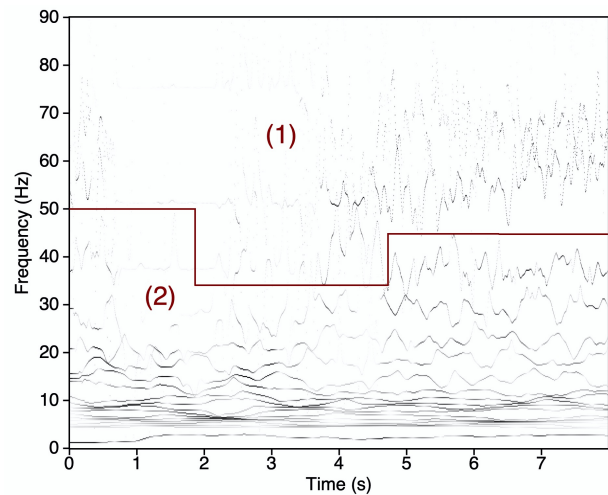


Figure 19: The time-frequency representation for the signal shown in Figure 20a, for  $K = 19$  and  $J = 15$  filter convolution. The two zones indicate areas from where the time-frequency data will be added.



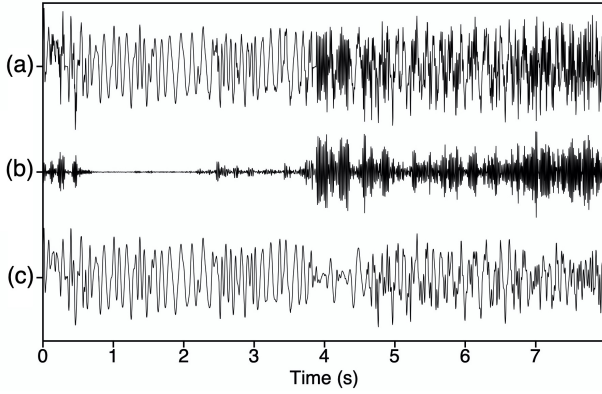


Figure 20: The original seismic signal in (a), the residue component in (b), obtained by adding the data into the area (1) of Figure 19, and the difference (filtered signal) in (c), associated with data in area (2) of Figure 19.

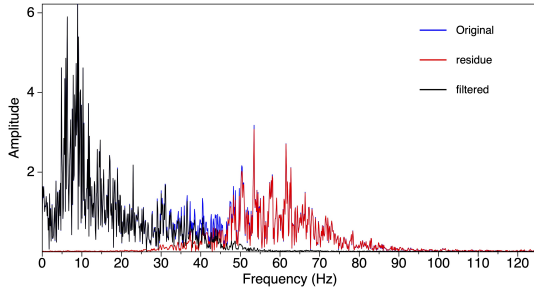


Figure 21: Average amplitude spectra of the signals in Figure 20a, Figure 20b and Figure 20c.

## CONCLUSION

The new recursive method for signal decomposition does not depend on the data, only on the number of the components to be obtained and the number of filter convolutions in the recursive loop. The recursive signal decomposition method is exact and phase-preserving, but non-unique. It starts with the high-frequency component, then providing new components with lower frequency, and the final residual consisting of the lowest frequency content of the signal. Further low-frequency components may be obtained by applying higher-order filters to the residual from the initial decomposition. This procedure provides an exact decomposition of the signal with filters which are zero-phase with positive frequency response, thus preserving arrival time information in the signal. A stable method for time-frequency analysis is obtained using a first-order prediction error operator to compute the average instantaneous frequency in sliding time windows. Applying this to each signal component gives a time-frequency representation of the input signal. The number of signal components may be reduced by summing similar modes, again resulting in an exact representation of the input signal. An alter-

native method is to sum components in an area in the time-frequency domain, resulting in a cleaner, but not exact, representation of the input signal. The results for two synthetic signals demonstrate that the time-frequency decomposition method is very effective to reproduce the data, even though the decomposition does not exactly reproduce the input signal.

## ACKNOWLEDGMENTS

The authors would like to thank INCT-GP/CNPq/MCTIC, CAPES, ANP, FINEP, FAPESB Brazil for financial support. Bjorn Ursin received financial support from CNPq, Brazil through the project INCT-GP and from the Norwegian Research Council through the Center for Geophysical Forecasting at NTNU (Grant No. 309960).

## APPENDIX A. PROPERTIES OF THE FILTER IMPULSE RESPONSES

In order to analyze the impulse responses of the symmetric filters in the proposed algorithm, we shall use two properties of the convolution,  $c = a * b$ , of two symmetric filters of length or dimension  $2k + 1$  and  $2\sigma + 1$ , respectively. They are

$$\dim\{c\} = 2(k + \sigma) + 1 \quad (\text{A1})$$

and

$$\sum_{n=-\sigma-k}^{\sigma+k} c(n) = \sum_{n=-k}^k a(n) \sum_{n=-\sigma}^{\sigma} b(n) \quad (\text{A2})$$

The starting point for the algorithm is the normalized vector

$$v_k(n) = \begin{cases} \frac{1}{\sqrt{k+1}}, & n = 0, 1, \dots, k \\ 0, & \text{otherwise} \end{cases} \quad (\text{A3})$$

of length  $k + 1$ , with

$$\sum_{n=0}^k v_k(n) = \sqrt{k+1} \quad (\text{A4})$$

The Fourier transform is

$$\begin{aligned} V_k(f) &= \frac{1}{\sqrt{k+1}} \sum_{n=0}^k e^{-2\pi i f n \Delta t} \\ &= \frac{1}{\sqrt{k+1}} \frac{1 - e^{-2\pi i f (k+1) \Delta t}}{1 - e^{-2\pi i f \Delta t}} \end{aligned} \quad (\text{A5})$$

where  $\Delta t$  is the sampling interval of the time series.

The first filter is

$$g_k = \frac{1}{k+1} v_k \oplus v_k \quad (\text{A6})$$

where  $\oplus$  denotes correlation. The filter coefficients

are

$$g_k(n) = \begin{cases} \frac{k+1-|n|}{(k+1)^2} & n = 0, \pm 1, \dots, \pm k \\ 0, & \text{otherwise} \end{cases} \quad (\text{A7})$$

and

$$\begin{aligned} \dim\{g_k\} &= 2k+1, \\ g_k(n) &> 0, \\ \sum_{n=-k}^k g_k(n) &= 1, \\ G_k(f) &= \frac{1}{k+1} |V_k(f)|^2, \\ 0 < G_k(f) &\leq 1, \quad G_k(0) = 1 \end{aligned} \quad (\text{A8})$$

The next filter is

$$f_k = g_k * \dots * g_1 = g_k * f_{k-1} \quad (\text{A9})$$

with properties

$$\begin{aligned} \dim\{f_k\} &= 2 \sum_{k=1}^k k+1 = k(k+1)+1, \\ f_k(n) &> 0, \\ \sum_n f_k(n) &= 1, \end{aligned} \quad (\text{A10})$$

$$\begin{aligned} F_k(f) &= \frac{1}{(k+1)!} \prod_{k=1}^k |V_k(f)|^2, \\ 0 < F_k(f) &\leq 1, \quad F_k(0) = 1 \end{aligned}$$

The filter

$$a_k = \delta - f_k \quad (\text{A11})$$

has properties

$$\begin{aligned} \dim\{a_k\} &= \dim\{f_k\}, \\ a_k(0) &> 0, \\ a_k(n) &< 0 \quad k \neq 0, \\ \sum_n a_k(n) &= 0, \\ A_k(f) &= 1 - F_k(f), \\ 0 \leq A_k(f) &< 1, \quad A_k(0) = 0 \end{aligned} \quad (\text{A12})$$

From this we define the convolution product

$$c_k = (*a_k)^J = [* (\delta - f_k)]^J \quad (\text{A13})$$

of  $J$  equal factors. Then

$$\begin{aligned} \dim\{c_k\} &= J \dim\{f_k\} = Jk(k+1)+1, \\ \sum_n c_k(n) &= 0, \\ C_k(f) &= A_k(f)^J = [1 - F_k(f)]^J, \\ 0 \leq C_k(f) &< 1, \quad C_k(0) = 0 \end{aligned} \quad (\text{A14})$$

The impulse response filter is, for  $J \geq 2$ ,

$$\begin{aligned} h_1 &= c_1, \\ h_k &= c_k * (\delta - c_{k-1}) * \dots * (\delta - c_1), \quad k = 2, 3, \dots \end{aligned} \quad (\text{A15})$$

with

$$\begin{aligned} \dim\{h_k\} &= J \sum_{j=1}^k j(j+1) + 1 \\ &= J \frac{k(k+1)(k+2)}{3} + 1 \\ \sum_n h_k(n) &= 0 \end{aligned} \quad (\text{A16})$$

and

$$\begin{aligned} H_1(f) &= C_1(f), \\ H_k(f) &= C_k(f)[1 - C_k(f)] \dots [1 - C_1(f)], \\ k &= 2, 3, \dots \end{aligned}$$

$$0 \leq H_k(f) < 1, \quad H_k(0) = 0$$

For  $J = 1$ ,  $c_k = \delta - f_k$ , and

$$\begin{aligned} h_1 &= \delta - f_1, \\ h_k &= (1 - f_k) * f_{k-1} * \dots * f_1, \quad k = 2, 3, \dots \end{aligned} \quad (\text{A17})$$

with

$$\begin{aligned} \dim\{h_k\} &= \frac{k(k+2)(k+3)}{3} + 1, \\ \sum_n h_k(n) &= 0, \end{aligned} \quad (\text{A18})$$

$$\begin{aligned} H_k(f) &= [1 - F_k(f)]F_{k-1}(f) \dots F_1(f) \\ 0 < H_k(f) &\leq 1, \quad H_k(0) = 0 \end{aligned}$$

Finally, the last component is computed from

$$h_{K+1} = \delta - \sum_{k=1}^K h_k \quad (\text{A19})$$

and

$$\begin{aligned} \dim\{h_{K+1}(n)\} &= \dim\{h_K\} = \frac{K(K+1)(K+3)}{3} + 1, \\ \sum_n h_{K+1}(n) &= 1, \\ H_{K+1}(f) &= 1 - \sum_{k=1}^K H_k(f), \\ 0 < H_{K+1}(f) &\leq 1, \quad H_{K+1}(0) = 1 \end{aligned} \quad (\text{A20})$$

Equation A18 and Equation A20 show that the remaining low-frequency part of the signal, which may be further decomposed, always remains in component number  $K + 1$ .

## APPENDIX B. AVERAGE INSTANTANEOUS FREQUENCY

We consider a complex signal (not necessarily analytic)

$$z(t) = x(t) + iy(t) = A(t)e^{i\phi(t)} \quad (\text{B21})$$

where

$$\begin{aligned} A(t) &= [x(t)^2 + y(t)^2]^{1/2} \\ \phi(t) &= \arctan \left\{ \frac{y(t)}{x(t)} \right\} \end{aligned} \quad (\text{B22})$$

The instantaneous frequency is given by the time derivative of the phase function (Angelsen, 1981; Han and van der Baan, 2013)

$$\hat{f}(t) = \frac{1}{2\pi} \phi'(t) = \frac{1}{2\pi} \frac{y'(t)x(t) - x'(t)y(t)}{x(t)^2 + y(t)^2} \quad (\text{B23})$$

The derivative is an unstable operation, so Porsani et al. (2019) used a local averaging technique to obtain a more stable estimate.

Here we shall use the average frequency Angelsen (1981)

$$\begin{aligned} \bar{f} &= \frac{\int_{-\infty}^{\infty} f S_z(f) df}{\int_{-\infty}^{\infty} S_z(f) df} \\ &= \frac{1}{2\pi} \text{Im}\{\rho'_z(0)\} \end{aligned} \quad (\text{B24})$$

where

$$\rho_z(k) = \frac{R_z(k)}{R_z(0)} \quad (\text{B25})$$

is the normalized autocorrelation function of the signal and  $S_z(f)$  is the power spectrum. It can be shown that the average frequency is the average value of the instantaneous frequency (Angelsen, 1981).

The average frequency is a constant computed for the input signal. In order to obtain a time-varying average frequency, we consider data in a moving time-window centered at  $t$ , and compute the average instantaneous frequency (AIF) (Andrade et al., 2018). In this window we compute a one-step prediction error operator (PEO),  $[1, -r(t)]$  using the Burg (1975) algorithm. The data in the time-window may then be represented by the inverse of the PEO, which is a minimum-delay wavelet ( $|r(t)| < 1$ ),

$$w_t(n) = r(t)^n, \quad n = 0, 1, \dots \quad (\text{B26})$$

The autocorrelation function is

$$R_t(n) = \begin{cases} R_t(0)r(t)^n, & n = 1, 2, \dots \\ R_t(0) = \frac{1}{1 - |r(t)|^2}, & n = 0 \\ R_t(0)r^*(t)^n, & n = -1, -2, \dots \end{cases} \quad (\text{B27})$$

where  $r^*(t)$  is the complex conjugate. We also need

the discrete time-derivative operator

$$\frac{\partial}{\partial n} \approx \frac{1}{\Delta t} \begin{cases} q(n), & n > 0 \\ 0, & n = 0 \\ -q(n), & n < 0 \end{cases} \quad (\text{B28})$$

where

$$q(n) = \begin{cases} \frac{1}{n}, & n = 1, 3, \dots \\ -\frac{1}{n}, & n = 2, 4, \dots \end{cases} \quad (\text{B29})$$

Using equation B27 to B29 in equation B24 gives

$$\begin{aligned} \bar{f}(t) &= \frac{1}{2\pi\Delta t} \text{Im} \left\{ \sum_{k=1}^{\infty} [r(t)^k q(k) - r^*(t)^k q(k)] \right\} \\ &= \frac{1}{\pi\Delta t} \text{Im} \left\{ \sum_{k=1}^{\infty} r(t)^k q(k) \right\} \\ &= \frac{1}{\pi\Delta t} \text{Im} \left\{ \frac{r(t)}{1} - \frac{r(t)^2}{2} + \frac{r(t)^3}{3} - \dots \right\}. \end{aligned} \quad (\text{B30})$$

A standard Taylor-series expression gives the AIF

$$\begin{aligned} \bar{f}(t) &= \frac{1}{\pi\Delta t} \text{Im} \{ \log [1 + r(t)] \} \\ &= \frac{1}{\pi\Delta t} \arctan \left\{ \frac{\text{Im}[r(t)]}{1 + \text{Re}[r(t)]} \right\}. \end{aligned} \quad (\text{B31})$$

## REFERENCES

- Andrade, M. C., M. J. Porsani, and B. Ursin, 2018, Complex autoregressive time-frequency analysis: estimation of time-varying periodic signal components: *IEEE Signal Processing Magazine*, **35**, 142–153, doi: 10.1109/MSP.2017.2783942.
- Angelsen, B. A., 1981, Instantaneous frequency, mean frequency, and variance of mean frequency estimators for ultrasonic blood velocity Doppler signals: *IEEE Transactions on Bio-medical Engineering*, **28**, 733–741, doi: 10.1109/TBME.1981.324853.
- 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.
- Burg, J. P., 1975, Maximum entropy spectral analysis.: PhD thesis, Stanford University.
- Castagna, J. P., S. Sun, and R. W. Siegfried, 2003, Instantaneous spectral analysis: Detection of low-frequency shadows associated with hydrocarbons: *The Leading Edge*, **22**, 120–127, doi: 10.1190/1.1559038.
- Chen, S. S., D. L. Donoho, and M. A. Saunders, 2001, Atomic decomposition by basis pursuit: *SIAM Review*, **43**, 129–159, doi: 10.1137/S003614450037906X.
- Cheng, J., and M. Sacchi, 2016, Fast and memory-

- efficient singular spectrum analysis for seismic data reconstruction and denoising: SEG Technical Program Expanded Abstracts 2016, Society of Exploration Geophysicists, 4064–4068. doi: 10.1190/segam2016-13955076.1.
- Cohen, L., 1989, Time–frequency distributions – a review: *Proceedings of the IEEE*, **77**, 941–981, doi: 10.1109/5.30749.
- Colominas, M. A., G. Schlotthauer, and M. E. Torres, 2014, Improved complete ensemble EMD: A suitable tool for biomedical signal processing: *Biomedical Signal Processing and Control*, **14**, 19–29, doi: 10.1016/j.bspc.2014.06.009.
- Dragomiretskiy, K., and D. Zosso, 2014, Variational mode decomposition: *IEEE Transactions on Signal Processing*, **62**, 531–544, doi: 10.1109/TSP.2013.2288675.
- Fomel, S., 2013, Seismic data decomposition into spectral components using regularized nonstationary autoregression: *Geophysics*, **78**, O69–O76, doi: 10.1190/geo2013-0221.1.
- Fourer, D., J. Harmouche, J. Schmitt, T. Oberlin, S. Meignen, F. Auger, and P. Flandrin, 2017, The ASTRES toolbox for mode extraction of non-stationary multicomponent signals: 2017 25th European Signal Processing Conference (EUSIPCO), 1130–1134. doi: 10.23919/EUSIPCO.2017.8081384.
- 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.
- Golub, G. H., and C. F. V. Loan, 1996, *Matrix computations*, 3rd ed.: Johns Hopkins University Press.
- Golyandina, N., and A. Zhigljavsky, 2020, *Singular spectrum analysis for time series*, 2nd ed.: Springer.
- Han, J., and M. van der Baan, 2013, Empirical mode decomposition for seismic time-frequency analysis: *Geophysics*, **78**, O9–O19, doi: 10.1190/geo2012-0199.1.
- Harmouche, J., D. Fourer, F. Auger, P. Borgnat, and P. Flandrin, 2018, The sliding singular spectrum analysis: A data-driven nonstationary signal decomposition tool: *IEEE Transactions on Signal Processing*, **66**, 251–263, doi: 10.1109/TSP.2017.2752720.
- Harris, T. J., and H. Yuan, 2010, Filtering and frequency interpretations of singular spectrum analysis: *Physica D: Nonlinear Phenomena*, **239**, 1958–1967, doi: 10.1016/j.physd.2010.07.005.
- Herrera, R. H., J. Han, and M. van der Baan, 2014, Applications of the synchrosqueezing transform in seismic time-frequency analysis: *Geophysics*, **79**, V55–V64, doi: 10.1190/geo2013-0204.1.
- Hu, H., S. Guo, R. Liu, and P. Wang, 2017, An adaptive singular spectrum analysis method for extracting brain rhythms of electroencephalography: *PeerJ*, **5**, e3474, doi: 10.7717/peerj.3474.
- Huang, N. E., Z. Shen, and S. R. Long, 1999, A new view of nonlinear water waves: The Hilbert spectrum: *Annual Review of Fluid Mechanics*, **31**, 417–457, doi: 10.1146/annurev.fluid.31.1.417.
- 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 nonstationary time series analysis: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, **454**, 903–995, doi: 10.1098/rspa.1998.0193.
- 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*, **42**, 1–26, doi: 10.1016/j.dsp.2015.03.004.
- Lesage, P., 2008, Automatic estimation of optimal autoregressive filters for the analysis of volcanic seismic activity: *Natural Hazards and Earth System Sciences*, **8**, 369–376, doi: 10.5194/nhess-8-369-2008.
- Lesage, P., F. Glangaud, and J. Mars, 2002, Applications of autoregressive models and time–frequency analysis to the study of volcanic tremor and long-period events: *Journal of Volcanology and Geothermal Research*, **114**, 391–417, doi: 10.1016/S0377-0273(01)00298-0.
- Liu, W., S. Cao, and Y. Chen, 2015, Seismic time–frequency analysis via empirical wavelet transform: *IEEE Geoscience and Remote Sensing Letters*, **13**, 28–32, doi: 10.1109/LGRS.2015.2493198.
- Liu, W., S. Cao, and Y. Chen, 2016, Applications of variational mode decomposition in seismic time-frequency analysis: *Geophysics*, **81**, V365–V378, doi: 10.1190/geo2015-0489.1.
- Mallat, S., 2008, *A wavelet tour of signal processing: The sparse way*, 3rd ed.: Academic Press.
- Marple, L., 1980, A new autoregressive spectrum analysis algorithm: *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **28**, 441–454, doi: 10.1109/TASSP.1980.1163429.
- Marple, S. L., 1987, *Digital spectral analysis: With applications*: Prentice-Hall.
- Mitrofanov, G., and V. Priimenko, 2015, Prony filtering of seismic data: *Acta Geophysica*, **63**, 652–678, doi: 10.1515/acgeo-2015-0012.
- Morf, M., B. Dickinson, T. Kailath, and A. Vieira, 1977, Efficient solution of covariance equations for linear prediction: *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **25**, 429–433, doi: 10.1109/TASSP.1977.1162989.
- Oropeza, V., and M. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: *Geophysics*, **76**, V25–V32, doi: 10.1190/1.3552706.

- Porsani, M. J., B. Ursin, and M. G. Silva, 2019, Signal decomposition and time–frequency representation using iterative singular spectrum analysis: *Geophysical Journal International*, **217**, 748–765, doi: 10.1093/gji/ggz046.
- Robinson, E. A., and S. Treitel, 2000, *Geophysical signal analysis*: Society of Exploration Geophysicists.
- Rodrigues, P. C., P. G. S. E. Tuy, and R. Mahmoudvand, 2018, Randomized singular spectrum analysis for long time series: *Journal of Statistical Computation and Simulation*, **88**, 1921–1935, doi: 10.1080/00949655.2018.1462810.
- Taner, M. T., F. Koehler, and R. E. Sheriff, 1979, Complex seismic trace analysis: *Geophysics*, **44**, 1041–1063, doi: 10.1190/1.1440994.
- Tary, J. B., R. Herrera, and M. Baan, 2013, Time-varying autoregressive model for spectral analysis of microseismic experiments and long-period volcanic events: *Geophysical Journal International*, **196**, 600–611, doi: 10.1093/gji/ggt400.
- 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.
- Tary, J. B., R. H. Herrera, and M. van der Baan, 2018, Analysis of time-varying signals using continuous wavelet and synchrosqueezed transforms: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **376**, 20170254, doi: 10.1098/rsta.2017.0254.
- Torres, M. E., M. A. Colominas, G. Schlotthauer, and P. Flandrin, 2011, A complete ensemble empirical mode decomposition with adaptive noise: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4144–4147. doi: 10.1109/ICASSP.2011.5947265.
- Ursin, B., and M. J. Porsani, 2021, Signal time–frequency representation and decomposition using partial fractions: *Geophysical Journal International*, **226**, 617–626, doi: 10.1093/gji/ggab115.
- Vesnaver, A., 2017, Instantaneous frequency and phase without unwrapping: *Geophysics*, **82**, F1–F7, doi: 10.1190/geo2016-0185.1.
- Wu, H.-T., G. F. Lewis, M. I. Davila, I. Daubechies, and S. W. Porges, 2016, Optimizing estimates of instantaneous heart rate from pulse wave signals with the synchrosqueezing transform: *Methods of Information in Medicine*, **55**, 463–472, doi: 10.3414/ME16-01-0026.

**Porsani, M. J.:** was responsible for the development, implementation, and testing of the algorithms. Also generated the figures and final results presented in the article; **Ursin, B.:** contributed to the critique and evaluation of the results. Also helped in the writing and final organization of the article.

Received on December 13, 2021 / Accepted on January 27, 2022



- Creative Commons attribution CC BY