

TREND FILTERING: EMPIRICAL MODE DECOMPOSITIONS VERSUS ℓ_1 AND HODRICK–PRESCOTT

AZADEH MOGHTADERI*, PIERRE BORGNAT[†]
and PATRICK FLANDRIN[‡]

**Mathematics and Statistics Department,
University Avenue, Jeffery Hall, Queen’s University,
Kingston, Ontario, Canada, K7L 3N6*

*†Laboratoire de Physique, École Normale,
Supérieure de Lyon, 46 allée d’Italie,
Lyon, 69364/Cedex 07, France*

**azadeh@mast.queensu.ca*

†pierre.borgnat@ens-lyon.fr

‡patrick.flandrin@ens-lyon.fr

Considering the problem of extracting a trend from a time series, we propose a novel approach based on empirical mode decomposition (EMD), called EMD trend filtering. The rationale is that EMD is a completely data-driven technique, which offers the possibility of estimating a trend of arbitrary shape as a sum of low-frequency intrinsic mode functions produced by the EMD. Based on an empirical analysis of EMD, an automatic procedure is proposed to select the requisite intrinsic mode functions. The performance of the EMD trend filtering is evaluated on simulated time series containing different forms of trends. Comparing furthermore to two existing techniques (ℓ_1 -trend filtering and Hodrick–Prescott filtering), we observe that the EMD trend filtering performs very similarly, while it does not require assumptions on the form of the trend and it is free from estimation parameters. We also illustrate the performance of the technique on the S&P 500 index, as an example of real-world time series.

Keywords: Trend filtering; Hodrick–Prescott filtering; ℓ_1 -trend filtering; empirical mode decomposition.

1. Introduction

Many real-world time series can be characterized by a “composite” behavior where one part is considered as “fluctuation” and another part is considered (loosely speaking) as a “trend.” Since fluctuation and trend may carry different types of information, *trend filtering* is a desirable objective, either for trend *estimation* or, on the contrary, for trend *removal*.

Although a disentanglement “fluctuation versus trend” is often physically intuitive, it heavily relies on the ambiguous task of defining “trend,” which depends strongly on a number of side information — or assumptions — on the analyzed

time series. Generally speaking, a trend is viewed as a function with monotone variations. Furthermore, if a model is applicable, trend estimation can be turned into a well-posed problem of regression (e.g., linear models would use linear regression). However, if no such model is available, the problems of trend definition and trend estimation become more complicated since they may depend on the process under study and also on the observed time series. To make this point more explicit, let us consider a toy example where a realization of a given white noise process is superimposed on an oscillatory component with frequency f_0 . Depending on the observation scale N , such a component can be interpreted as an effective oscillation if $N \gg 1/f_0$, or as a trend in the above sense if $N \ll 1/f_0$. Therefore, deciding whether such a contribution is a trend or not appears as a matter of interpretation. In the following, we adopt the pragmatic approach of a *relative* definition, considering a trend in the signal to be any contribution that is *slowly varying* in comparison with the fluctuation.

Estimating a trend from the regression point of view has received a great deal of attention in the literature. The purpose of this paper, however, is not to review the corresponding contributions. The aim here is rather to focus on a recent publication detailing two such approaches [Kim *et al.* (2009)] and to contrast them with what can be obtained, on similar datasets, with the help of the model-free technique called *empirical mode decomposition (EMD) trend filtering*. The rationale is that, by construction, EMD sequentially extracts, at each step of the decomposition, a “fluctuation” part that is “quickly varying” as compared to the remaining residual, thus permitting to identify the trend with such a residual at some significant level of the decomposition that has to be adequately chosen. This kind of idea has already been pushed forward, either based on some assumed model for the background noise [Flandrin *et al.* (2004)], or by deliberately considering the trend as the final residual [Wu *et al.* (2007)] of the EMD. We revisit here the first of these two approaches by proposing a new, model-free, criterion aimed at identifying, in an automatic way, the level of the decomposition which delineates the fluctuation from the trend.

More precisely, the paper is organized as follows. In Sec. 2, we review some preliminaries which are necessary to the framework of this paper and a brief background on the Hodrick–Prescott (H–P) filtering in Sec. 2.1 and ℓ_1 -trend filtering in Sec. 2.2. In Sec. 3 we describe the EMD trend filtering approach by first reviewing the EMD in Sec. 3.1 and details on the proposed method in Sec. 3.2. The performance of the EMD trend filtering is evaluated in Sec. 4 via three simulated examples and one real-world example. The concluding remarks are made in Sec. 5.

2. Preliminaries and Background Material

Let $\mathbf{Y} = \{Y_t\}_{t \geq 0}$ be a discrete-time real-valued stochastic process. The process \mathbf{Y} is said to be broadband if its spectrum $S_{\mathbf{Y}}(f)$ is nonzero and purely continuous in the interval $f \in [0, 1/2]$. Let $\mathcal{Y} = (Y_0, Y_1, \dots, Y_{N-1})$ be a realization of \mathbf{Y} and let $\mathcal{C} = (C_0, C_1, \dots, C_{N-1}) \in \mathbb{R}^n$ be a *trend* component. From \mathcal{Y} and \mathcal{C} , we can make

two time series: The first is $\mathcal{Y} + \mathcal{C}$, the *additive mix* of \mathcal{Y} and \mathcal{C} ; the second is $\mathcal{Y}\mathcal{C}$, the *multiplicative mix* of \mathcal{Y} and \mathcal{C} . In either case, we say \mathcal{Y} is the fluctuation of the mix.

Now let \mathcal{X} be an additive or multiplicative mix of \mathcal{Y} and \mathcal{C} . The question we would like to answer in this paper is that solely given \mathcal{X} as data, how can one estimate \mathcal{C} from \mathcal{X} ? In order to answer this question, loosely speaking, we define the trend here to be that \mathcal{C} which is locally slowly varying as compared to \mathcal{Y} .

For the remaining of this section, we review two existing trend estimation techniques known as the H-P filtering and the ℓ_1 -trend filtering in Secs. 2.1 and 2.2, respectively.

2.1. Hodrick–Prescott filtering

Recall \mathcal{Y} and \mathcal{C} from above. Let $\mathcal{X} = (X_0, X_1, \dots, X_{N-1})$ be an additive mix of \mathcal{Y} and \mathcal{C} . The H-P filtering [Hodrick and Prescott (1997)] is a commonly used technique for trend estimation. In this technique, the trend estimate is chosen such that the weighted sum

$$\frac{1}{2} \sum_{t=0}^{N-1} (X_t - C_t)^2 + \lambda \sum_{t=1}^{N-2} (C_{t-1} - 2C_t + C_{t+1})^2 \quad (1)$$

is minimized. Here, $\lambda \geq 0$ is known as the *regularization parameter* and it controls the trade-off between the smoothness of C_t and the size of the residual $X_t - C_t$.

We can write Eq. (1) such that

$$\frac{1}{2} \|\mathcal{X} - \mathcal{C}\|_2^2 + \lambda \|\mathcal{D}\mathcal{C}\|_2^2, \quad (2)$$

where for $u = [\dots u_{-1} u_0 u_1 \dots]$, $\|u\|_2^2 = (\sum_i u_i^2)^{1/2}$ is the L^2 -norm of the vector u . Here, $\mathcal{D} \in \mathbb{R}^{(N-2) \times N}$ is the second-order difference matrix

$$\mathcal{D} = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix}.$$

The trend estimate \mathcal{C}^{HP} is the unique minimizer of Eq. (2) such that

$$\mathcal{C}^{\text{HP}} = (I + 2\lambda\mathcal{D}^T\mathcal{D})^{-1}\mathcal{X}, \quad (3)$$

where $I \in \mathbb{R}^{N \times N}$ is the identity matrix.

Some of the main properties of the H-P filtering are listed below:

- The number of steps necessary for computation of \mathcal{C}^{HP} is $O(N)$.
- \mathcal{C}^{HP} is a linear function of \mathcal{X} .
- As $\lambda \rightarrow 0$, \mathcal{C}^{HP} converges to \mathcal{X} .
- As $\lambda \rightarrow \infty$, \mathcal{C}^{HP} converges to the best affine fit to \mathcal{X} .

For further detail about H–P filtering see [Hodrick and Prescott (1997)] or [Kim *et al.* (2009)].

2.2. ℓ_1 -trend filtering

Let \mathcal{C} , \mathcal{Y} and \mathcal{X} be as in Sec. 2.1. The ℓ_1 -trend filtering [Kim *et al.* (2009)] is a variation of the H–P filtering where the trend is estimated such that the weighted sum

$$\frac{1}{2} \sum_{t=0}^{N-1} (X_t - C_t)^2 + \lambda \sum_{t=1}^{N-2} |C_{t-1} - 2C_t + C_{t+1}|, \quad (4)$$

is minimized. Here, $\lambda \geq 0$ is a regularization parameter as in the H–P filtering, used to control the trade-off between the smoothness of the trend and the size of the residual. Similarly to the H–P filtering, we write Eq. (4) in a matrix form

$$\frac{1}{2} \|\mathcal{X} - \mathcal{C}\|_2^2 + \lambda \|\mathcal{D}\mathcal{C}\|_1, \quad (5)$$

where $\|u\|_1 = (\sum_i |u_i|)$ is the L^1 -norm of the vector u . In the above, \mathcal{D} is as in Sec. 2.1. The minimizer of Eq. (5) is unique and is denoted by \mathcal{C}^{ℓ_1} . Note that, ℓ_1 -trend filtering does not provide a closed form expression for \mathcal{C}^{ℓ_1} analogous to Eq. (3) and the computation is entirely numerical. An efficient algorithm to solve this optimization problem (based on simple primal-dual method) is discussed in the paper by [Kim *et al.* (2009)].

Some of the basic properties of ℓ_1 -trend filtering are as follows:

- The number of steps necessary to numerically compute \mathcal{C}^{ℓ_1} is $O(N)$.
- In contrast with \mathcal{C}^{HP} , \mathcal{C}^{ℓ_1} is not a linear function of \mathcal{X} .
- As $\lambda \rightarrow 0$, \mathcal{C}^{ℓ_1} converges to \mathcal{X} .
- As $\lambda \rightarrow \infty$, \mathcal{C}^{ℓ_1} converges to the best affine fit to \mathcal{X} . This convergence, however, happens for finite value of λ .

In general the ℓ_1 -trend filtering is advantageous to the H–P filtering when the trend is a piecewise linear function. For further detail about ℓ_1 -trend filtering see [Kim *et al.* (2009)].

3. Trend Estimation Via EMD

Estimation of the trend in the H–P filtering and the ℓ_1 -trend filtering depends strongly on the choice of the regularization parameter λ , for which there exists no deciding rule. We propose a novel estimator of the trend which does not require any free parameter and uses the EMD in its development. This novel estimator is called the EMD trend filtering.

In this section, we review the EMD in Sec. 3.1 and introduce the EMD trend filtering in Sec. 3.2.

3.1. The Empirical mode decomposition

The EMD is an algorithm which decomposes a time series into a finite additive superposition of oscillatory components, each of which is called an *intrinsic mode function* (IMF); see [Huang *et al.* (1998)]. The EMD does not rely on any technical assumptions concerning the nature of the time series; note that this includes modelling assumptions. The basic idea is that IMFs are computed subject to two requirements: First, the number of local extrema and number of zero crossings of each IMF vary by at most one. Second, the mean of the upper and lower envelopes of each IMF should be identically equal to zero, where the envelopes are computed by means of a fixed interpolation scheme. (In the numerical results presented in this paper, we have confined ourselves to the use of cubic spline and linear interpolations.) The IMFs are computed by means of an iterative scheme. This scheme however depends on a stopping criterion which guarantees that the requirements above are satisfied within a given tolerance while at the same time each extracted IMF is meaningful in both its amplitude and frequency modulations; we again refer to [Huang *et al.* (1998)] for details.

To make this intuitive description more precise, let $\mathbf{Z} = \{Z_t\}_{t \geq 0}$ be a (real-valued, discrete-time, stochastic) process, and let $\mathcal{Z} = (Z_0, Z_1, \dots, Z_{N-1})$ be a realization of \mathbf{Z} . As an initialization step, set $i = 1$ and $\mathcal{R}^0 = \mathcal{Z}$. The EMD computes the IMFs of \mathcal{Z} using the following algorithm.

- (1) Identify the local maxima and local minima of \mathcal{R}^{i-1} .
- (2) Together with the chosen interpolation scheme, use the maxima and minima from step (1) to compute the upper and lower envelopes of \mathcal{R}^{i-1} .
- (3) Determine the *local trend*, denoted \mathcal{Q}^i , as the mean of the upper and lower envelopes from step (2).
- (4) Compute the *local fluctuation*, denoted $\mathcal{H} = \mathcal{Z} - \mathcal{Q}^i$.
- (5) If \mathcal{H} is not an IMF, in the sense that it does not satisfy the two requirements described in the beginning of this section, then increment i and go to step (1) with $\mathcal{R}^{i-1} = \mathcal{H}$. (Authors in [Huang *et al.* (1998)] call this the *sifting process*; it is this process which depends on the stopping criterion.)
- (6) If \mathcal{H} is an IMF, in the sense that it satisfies the two requirements described at the beginning of this section, then the *ith intrinsic mode function* of \mathcal{Z} is $\mathcal{M}^i = \mathcal{H}$, and the *ith residual* is $\mathcal{R}^i = \mathcal{Z} - \mathcal{M}^i$. Increment i and go to step (1).

The algorithm halts when the *ith* residual has no further oscillations, in the sense that it has no local maxima or local minima. We denote by i_m the largest index for which \mathcal{M}^i is defined. Then

$$\mathcal{Z} = \sum_{i=1}^{i_m} \mathcal{M}^i + \mathcal{R}^{i_m}. \quad (6)$$

In this decomposition, \mathcal{M}^1 through \mathcal{M}^{i_m} can be thought of as containing a “spectrum” of local oscillations in \mathcal{Z} , with the shortest-period (highest frequency)

oscillations represented in \mathcal{M}^1 and the longest-period (lowest frequency) oscillations represented in \mathcal{M}^{i_m} . The computational complexity of the algorithm depends on \mathcal{Z} , the chosen interpolation scheme, and the stopping criterion. However, the algorithm usually halts in a reasonably small number of steps. For example, it is known [Flandrin *et al.* (2004)] that if \mathcal{Z} is a broadband process, then the decomposition produced by the EMD has an almost dyadic filter-bank structure, typically with $i_m \approx \log_2 N$. Moreover, it is known that the sifting process typically halts after some tens of iterations [Huang *et al.* (1998)].

3.2. EMD trend filtering

In this section, we introduce a novel technique for trend estimation which we call the EMD trend filtering. The EMD trend filtering generalizes and modifies a technique, which was introduced by the authors in [Moghtaderi *et al.* (2010)]. In the following, we describe the details of this technique.

Recall \mathcal{Y} and \mathcal{C} from Sec. 2. Let \mathcal{X} be an additive mix of \mathcal{Y} and \mathcal{C} . We apply EMD to \mathcal{X} in order to extract its IMFs, denoted $\mathcal{M}^i = (M_0^i, M_1^i, \dots, M_{N-1}^i)$ for $1 \leq i \leq i_m$. Recall the definition of trend from Sec. 2. Also recall from Sec. 3.1 that in EMD the successive IMFs are oscillations going from high frequency to low frequency, and that this property is valid locally in time (there is not necessarily a global separation of spectrum of successive IMFs) [Huang *et al.* (1998)]. Hence, a pragmatic way of satisfying this is that the trend should be obtained as the sum of the last few IMFs and the residual extracted from \mathcal{X} such that

$$\mathcal{C} \approx \sum_{i=i_*}^{i_m} \mathcal{M}^i + \mathcal{R}^{i_m}. \quad (7)$$

Clearly, in order to estimate \mathcal{C} , all we need to do is to estimate i_* . In the following, we denote the estimate of \mathcal{C} by $\widehat{\mathcal{C}}$, and the estimate of i_* by \widehat{i}_* .

Note that if the mix is multiplicative and the elements of \mathcal{C} are positive, then the situation reduces to the additive case. Indeed, one can take logarithms to obtain $\log |\mathcal{X}| = \log \mathcal{C} + \log |\mathcal{Y}|$, where the logarithm and absolute value functions are being applied elementwise.

EMD trend filtering, described over the course of the next three subsections, actually consists of three approaches to estimating i_* . These are called, respectively, the *ratio*, *energy*, and *energy-ratio* approaches.

3.2.1. Energy approach

In this section, we describe the first approach for evaluating \widehat{i}_* .

Let $\{Z_t\}_{t \geq 0}$ be an arbitrary process. For a given time series which is a realization of $\{Z_t\}$, we define the *energy of its i th IMF*, denoted G^i , by

$$G^i \triangleq \sum_{t=0}^{N-1} |M_t^i|^2, \quad 1 \leq i \leq i_m.$$

Assume now that we have B different time series obtained from $\{Z_t\}$. Given the b th time series, $1 \leq b \leq B$, if $G^{i,b}$ denotes the energy of its i th IMF, the *averaged energy of its i th IMF* is defined by

$$\overline{G^i} \triangleq \frac{1}{B} \sum_{b=1}^B G^{i,b}. \quad (8)$$

It is shown in [Rilling *et al.* (2005)] that if the time series under study are realizations of a generic broadband process, then $\overline{G^i}$ is a decreasing sequence in i . The authors in [Rilling *et al.* (2005)] concluded this by studying fractional Gaussian noise processes, which provide a general model for broadband processes with no particular frequency band. The power spectral density of a fractional Gaussian noise process at frequency f can be approximated by $K|f|^{1-2h}$ when $|f| \rightarrow 0$. Here, K is a constant and $0 < h < 1$ is known as the Hurst exponent. As a result, fractional Gaussian noise processes are convenient models for processes with broadband or power-law spectra. For details about fractional Gaussian noise processes, see [Flandrin *et al.* (2004)] or [Rilling *et al.* (2005)].

Keeping what we described above in mind, our key observation is that, generically, $\overline{G^i}$ increases for i near the index i_* [Moghtaderi *et al.* (2010)]. Identifying the smallest index $i \geq 2$ such that $G^i > G^{i-1}$ evaluates \hat{i}_* . This approach is called the *energy approach*.

The limitation with the energy approach is that one is often given a single time series to use in the trend estimation procedure. Clearly, Computation of energy, based on only one time series, may cause false increases in G^i at indices which do not associate with the trend.

3.2.2. Ratio approach

In this section, we describe the second approach to evaluate \hat{i}_* .

For a given time series, we denote the zero crossing number of its i th IMF by Z^i , and define $R^i = Z^{i-1}/Z^i$ for $i \geq 2$. We call R^i the *i th ratio of the zero-crossing numbers* (i th RZCN). We observe that if the time series under study is a realization of a generic broadband process, the approximation $R^i \approx 2$ holds. This number might be slightly different based on the type of interpolation used in the EMD but always remains the same within a given interpolation scheme.

In order to assess the truth of this approximation, we construct 21 broadband processes, including (1) 17 fractional Gaussian noise processes with Hurst exponents $h = 0.1, 0.15, 0.2, 0.25, \dots, 0.9$, (2) two autoregressive process of order 2 (or AR(2)), and (3) two non-stationary processes, the first being AR(2) with time-dependent coefficients and the second being frequency modulated. For each process, we create 10,000 realizations of length $N = 2,000$, and compute the IMFs and RZCNs for each realization. We denote the i th RZCN of the b th realization by $R^{i,b}$, where $2 \leq i \leq i_m^b$ and write $\vec{R}^b = [R^{2,b} \ R^{3,b} \ \dots \ R^{i_m^b,b}]$. We then compute the empirical distribution of the elements of $\vec{R} = [\vec{R}^1 \ \vec{R}^2 \ \dots \ \vec{R}_{10,000}^2]$. The same procedure was

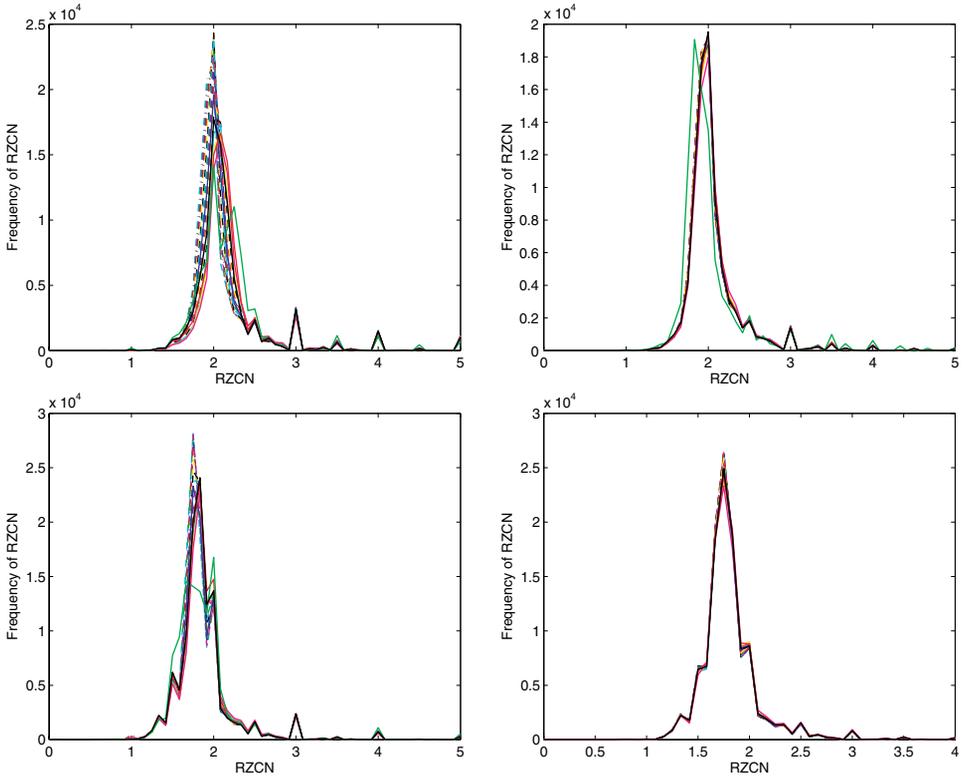


Fig. 1. Empirical distribution of the elements of \vec{R} : The top-left (resp. top-right) plot displays the empirical distribution of the elements of \vec{R} for 21 broadband (log-transformed broadband) processes when cubic spline interpolation is used in the EMD. The bottom-left (resp. bottom-right) plot displays the empirical distribution of the elements of \vec{R} for 21 broadband (resp. log-transformed broadband) processes when linear interpolation is used in the EMD. In all plots, different processes are represented by different line types and colors.

also applied to the log-transformed version of each process in the collection. Figure 1 displays the empirical distributions of the elements of \vec{R} computed for each process and each log-transformed process when cubic spline and linear interpolations are used in the EMD. The result of our simulation encourages the idea that $R^i \approx 2$. In fact, these distributions appear to be approximately Gaussian with mean 2. We also observe that apart from the expected peak at the center of the distribution, there exist several smaller but visible peaks. They appear in the presence of high order IMFs which have small values of zero crossing numbers.

Our key observation is that, generically, the approximation $R^i \approx 2$ fails for i near the index i_* . We therefore evaluate \hat{i}_* by looking for the smallest index i for which R^i is “significantly different from two.” The result of simulation suggests that a common threshold test could be used to conclude whether or not R^i is significantly different from two. In order to do so, we compute α and $1-\alpha$ ($0 \leq \alpha \leq 1$) significance level of the empirical distribution of the elements of \vec{R} as the *right threshold* and

Table 1. Right and left thresholds for a few values of α .

α	T_r^{SP}	T_l^{SP}	T_r^{ln}	T_l^{ln}	L_r^{SP}	L_l^{SP}	L_r^{ln}	L_l^{ln}
0.01	4.9762	1.5033	4.9762	1.5033	3.5317	1.5073	3.0000	1.2665
0.03	3.8095	1.6708	3.0000	1.3804	3.0238	1.6647	2.5079	1.3487
0.05	3.0433	1.7377	2.5179	1.4990	2.7030	1.7232	2.3359	1.4692
0.08	2.9244	1.8014	2.2912	1.5000	2.5000	1.7708	2.1433	1.5000
0.1	2.6454	1.8300	2.1912	1.5534	2.4117	1.7941	2.0542	1.5050
0.2	2.2589	1.9183	2.0000	1.6863	2.1687	1.8587	1.9281	1.6456

Note: T_r^{SP} and T_l^{SP} (resp. T_r^{ln} and T_l^{ln}) are the right and left thresholds of the top-left (resp. bottom-left) plot in Fig. 1. Similarly, L_r^{SP} and L_l^{SP} (resp. L_r^{ln} and L_l^{ln}) are the right and left thresholds of the top-right (resp. bottom-right) plot in Fig. 1.

the *left threshold*, respectively. Precisely, we compute the left and right thresholds for each process and then average them to obtain the final value. Table 1 reports the right and left thresholds for all the empirical distributions shown in Fig. 1 and for a few values of α .

Finally, any RZCN which is outside of the appropriate right and left thresholds is considered *significantly different from two*.

The problem with the ratio approach is that, since selection of the left and right thresholds is entirely based on empirical results, it is always possible that for a given α , the smallest i for which R^i appears significantly different from 2 is a false detect.

3.2.3. Energy–ratio approach

As described in the previous two sections, the energy and ratio approaches are confronted with possible false detections of the smallest index, which does not associate with the trend. Since the criteria proposed by the energy approach and ratio approach to evaluate \hat{i}_* are independent, the number of false detects can be reduced by combining these two approaches.

To be more precise, for each $2 \leq i \leq i_m$, we compute each index i such that $G^i > G^{i-1}$. We also evaluate every index i where R^i is significantly different from two. We then evaluate \hat{i}_* to be the smallest common index in both approaches. This approach is called the energy–ratio approach. The performance of the energy–ratio approach and its ability to reduce the number of false detects is discussed in the paper by [Moghtaderi *et al.* (2010)] via simulations.

4. Examples

In this section, we demonstrate the performance of the EMD trend filtering by applying it to three simulated examples in Sec. 4.1 and one real-world example in Sec. 4.2. For each example, we compare the performance of the EMD trend filtering via the energy-ratio approach with ℓ_1 -trend filtering and H–P filtering. The comparison method we use here is based on the Euclidean distance between

the theoretical trend and the estimated trend obtained from a given method as

$$E = \left(\sum_{t=0}^{N-1} |C_t - \widehat{C}_t|^2 \right)^{1/2}. \quad (9)$$

Clearly, if any given example follows a multiplicative mix, we replace C_t in Eq. (9) with $\log C_t$ and \widehat{C}_t with $\widehat{\log C}_t$.

In the following, we use $\alpha = 0.08$ in the energy-ratio approach (this choice of α is due to its good performance in a large number of simulations.)

4.1. Simulated examples

In this section, we introduce three simulated examples. For each example, we construct a time series of length $N = 2,000$ and apply EMD to it once using cubic spline interpolation and once using linear interpolation. For each example and each method of interpolation, we extract the IMFs and apply the energy-ratio approach to them in order to evaluate \widehat{i}_* . We then estimate the trend using \widehat{i}_* as described in Sec. 3.2.

For each example, we use the same time series used in the EMD trend filtering in order to estimate the trend using ℓ_1 -trend filtering and H-P filtering. We select the regularization parameter λ to be the nonnegative integers satisfying $0 \leq \lambda \leq \lambda_{\max}$ where $\lambda_{\max} \in \mathbb{N}$. We then compute the Euclidean distance for each λ and find the value λ which associates with the minimum Euclidean distance. We denote this value by λ_* .

4.1.1. Simulated example 1

Let $\mathbf{Y}^{(1)} = \{Y_t^{(1)}\}_{t \geq 0}$ be an AR(2) process such that

$$Y_t^{(1)} = 0.8Y_{t-1}^{(1)} - 0.4Y_{t-2}^{(1)} + \zeta_t.$$

Here $\{\zeta_t\}$ is a white noise process with variance 10^4 . Now, let $\mathcal{Y}^{(1)} = (Y_0^{(1)}, Y_1^{(1)}, \dots, Y_{N-1}^{(1)})$ be a realization of $\mathbf{Y}^{(1)}$ and $\mathcal{C}^{(1)}$ be a trend component which is randomly constructed using piecewise linear spline technique. Using $\mathcal{Y}^{(1)}$ and $\mathcal{C}^{(1)}$, we can now construct the additive mix $\mathcal{X}^{(1)} = \mathcal{C}^{(1)} + \mathcal{Y}^{(1)}$. The left hand side plot in Fig. 2 displays $\mathcal{C}^{(1)}$ and the right hand side plot displays $\mathcal{C}^{(1)}$ superimposed on $\mathcal{X}^{(1)}$.

We apply EMD to $\mathcal{X}^{(1)}$ using cubic spline interpolation (resp. linear interpolation), and obtain $i_m = 9$ (resp. $i_m = 11$). The energy-ratio approach evaluates $\widehat{i}_* = 7$ (resp. $\widehat{i}_* = 9$) which we use to estimate $\mathcal{C}^{(1)}$. Figure 3 displays the energy-ratio approach for both interpolation schemes and Fig. 4 displays the estimated trends based on the energy-ratio approach.

We now use $\mathcal{X}^{(1)}$ to estimate $\mathcal{C}^{(1)}$ using ℓ_1 -trend filtering. Recall from Sec. 2 that the best affine fit in the ℓ_1 -trend filtering is obtained for a finite value of λ . The Euclidean distance for the best affine fit to $\mathcal{X}^{(1)}$ is $E = 9.956$. We therefore

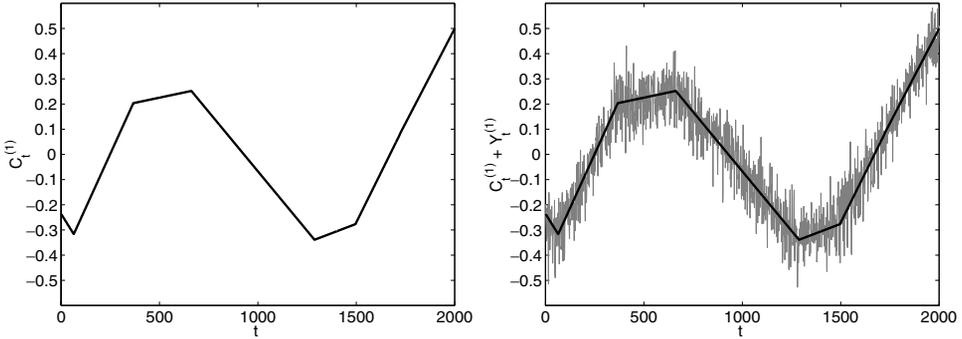


Fig. 2. Trend in example 1: The left-hand side plot displays $\mathcal{C}^{(1)}$ and the right-hand side plot displays $\mathcal{C}^{(1)}$ superimposed on $\mathcal{X}^{(1)}$.

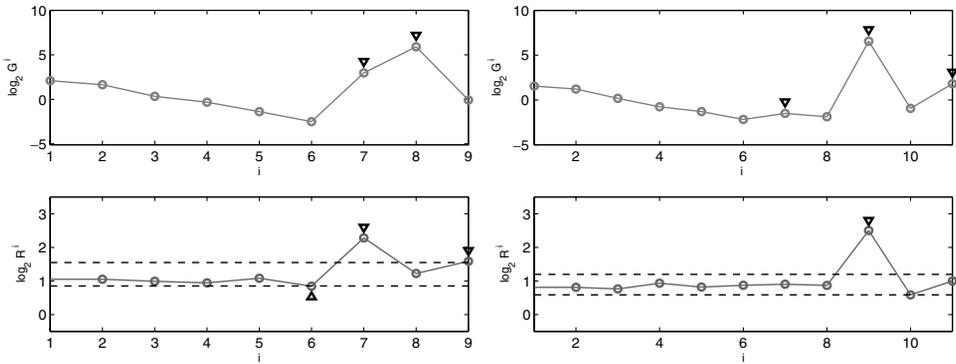


Fig. 3. Energy-ratio approach in example 1: The top-left (resp. top-right) plot displays the energy approach when using cubic spline interpolation (resp. linear interpolation) in EMD. The small circles are $\log_2 G^i$ for $1 \leq i \leq 9$ (resp. $2 \leq i \leq 11$) and the small triangles mark those indices $i \geq 2$ where $G^i > G^{i-1}$. The bottom-left (resp. bottom-right) plot displays the ratio approach when using cubic spline interpolation (resp. linear interpolation) in EMD. The small circles are $\log_2 R^i$ for $2 \leq i \leq 9$ (resp. $2 \leq i \leq 11$), the dashed lines are T_l^{sp} and T_r^{sp} (resp. T_l^{ln} and T_r^{ln}) for $\alpha = 0.08$ in Table 1, and the small triangles mark those indices i where R^i is significantly different from two. Here, the energy approach evaluates $\hat{i}_* = 7$ (resp. $\hat{i}_* = 7$), the ratio approach evaluates $\hat{i}_* = 6$ (resp. $\hat{i}_* = 9$), and the energy-ratio approach evaluates $\hat{i}_* = 7$ (resp. $\hat{i}_* = 9$) which we use to estimate $\mathcal{C}^{(1)}$.

increase λ_{\max} until the Euclidean distance associated with λ_{\max} reaches 9.956. This happens at $\lambda \approx 32 \times 10^3$. Here, the minimum Euclidean distance is $E = 0.858$, which associates with $\lambda_* = 15$. Figure 5 displays the Euclidean distance as a function of λ and the estimated trend using $\lambda_* = 15$.

Similarly, we can use $\mathcal{X}^{(1)}$ to estimate $\mathcal{C}^{(1)}$ using H-P filtering. Recall from Sec. 2 that the best affine fit in the H-P filtering is obtained when $\lambda \rightarrow \infty$. Computationally speaking, we cannot choose λ_{\max} to reach the best affine fit to $\mathcal{X}^{(1)}$. We therefore select λ_{\max} large enough to see the increase in the Euclidean distance and find an approximate range for λ_* . We select $\lambda_{\max} = 3 \times 10^8$ and observe that the

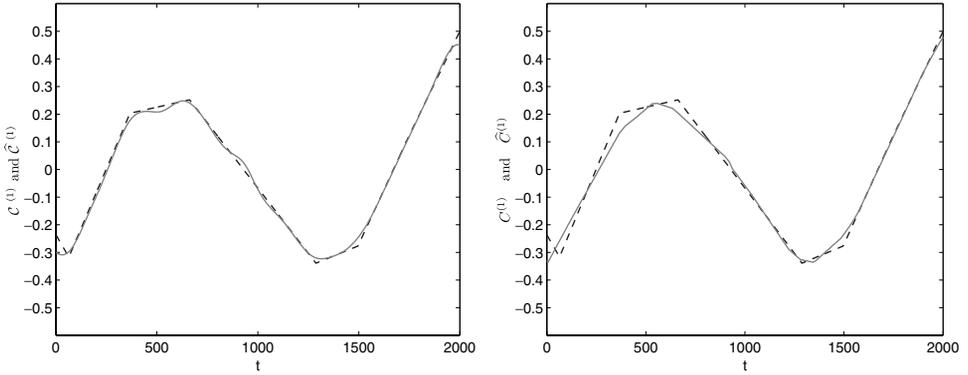


Fig. 4. EMD trend filtering in example 1: The left-hand side (resp. right-hand side) plot displays $C^{(1)}$ (dashed line) versus $\hat{C}^{(1)}$ (solid line) using the energy-ratio approach when $i_* = 7$ (resp. $i_* = 9$). The Euclidean distance for this estimate is $E = 0.559$ (resp. $E = 1.078$).

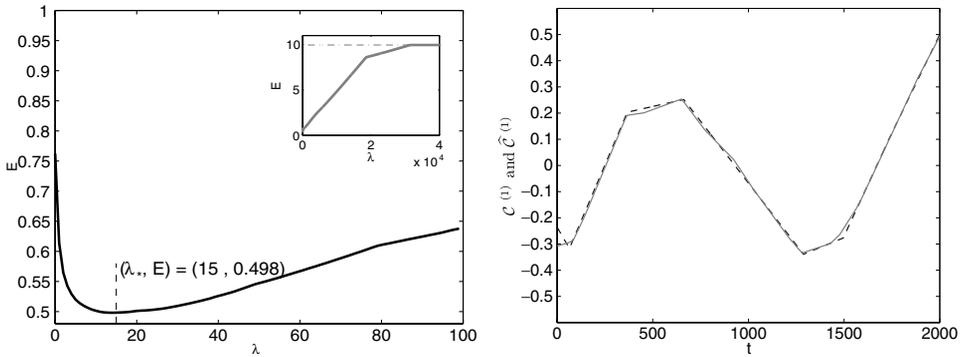


Fig. 5. The best estimate of $C^{(1)}$ using ℓ_1 -trend filtering in example 1: The left-hand side plot displays the Euclidean distance as a function of λ . The larger plot shows the Euclidean distance only up to $\lambda = 100$ so the minimum Euclidean distance $E = 0.498$ at $\lambda_* = 15$ can be seen clearly. The smaller plot shows the Euclidean distance for $\lambda_{\max} = 4 \times 10^4$. The Euclidean distance clearly reaches the best affine fit to $\mathcal{X}^{(1)}$, which is marked with a dashed-dotted line. The right-hand side plot displays $C^{(1)}$ (dashed line) versus $\hat{C}^{(1)}$ (solid line) for $\lambda_* = 15$.

minimum Euclidean distance appears at $\lambda_* \approx 545 \times 10^3$, where we obtain $E = 0.538$. Figure 6 displays the Euclidean distance as a function of λ and the estimated trend using $\lambda_* = 545 \times 10^3$.

It is clear from the outcome of our simulating example that all trend filtering methods discussed here perform similarly: the magnitude of the Euclidean distances is comparable and the general appearance of all the estimated trends does not differ significantly.

4.1.2. Simulated example 2

Let $\mathbf{Y}^{(2)} = \{Y_t^{(2)}\}_{t \geq 0}$ be a fractional Gaussian noise process with Hurst exponent $h = 0.7$. Let $\mathcal{Y}^{(2)} = (Y_0^{(2)}, Y_1^{(2)}, \dots, Y_{N-1}^{(2)})$ be a realization of $\mathbf{Y}^{(2)}$ and $\mathcal{C}^{(2)}$ be

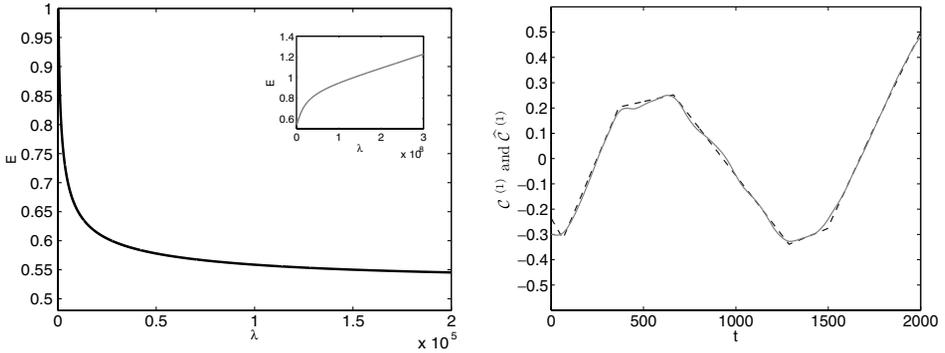


Fig. 6. The best estimate of $\mathcal{C}^{(1)}$ using H-P filtering in example 1: The left-hand side plot displays the Euclidean distance as a function of λ . The larger plot displays the Euclidean distance as a purely decreasing function of λ up to $\lambda = 2 \times 10^5$. The smaller plot shows the Euclidean distance for $\lambda_{\max} = 3 \times 10^8$, which shows that the Euclidean distance increases but will not reach the best affine fit for a finite λ . Here, $\lambda_* \approx 545 \times 10^3$, where $E = 0.538$. The right-hand side plot displays $\mathcal{C}^{(1)}$ (dashed line) versus $\widehat{\mathcal{C}}^{(1)}$ (solid line) for $\lambda_* = 545 \times 10^3$.

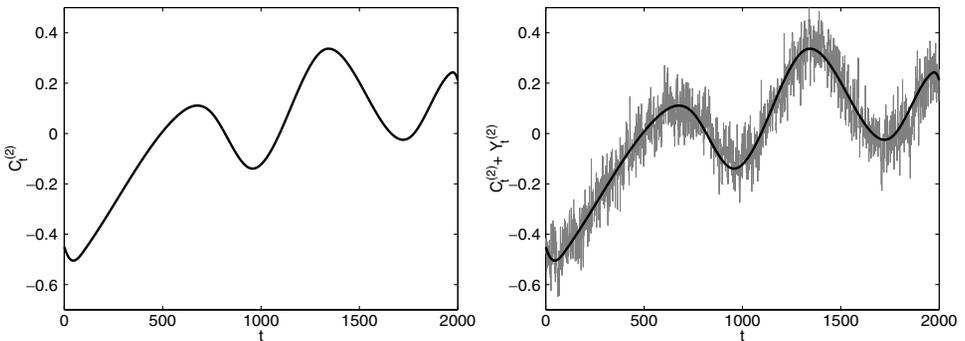


Fig. 7. Trend in example 2: The left-hand side plot displays $\mathcal{C}^{(2)}$ and the right-hand side plot displays $\mathcal{C}^{(2)}$ superimposed on $\mathcal{X}^{(2)}$.

a trend component which is randomly constructed using piecewise cubic spline technique. Using $\mathcal{Y}^{(2)}$ and $\mathcal{C}^{(2)}$, we now construct the additive mix $\mathcal{X}^{(2)} = \mathcal{C}^{(2)} + \mathcal{Y}^{(2)}$. The left hand side plot in Fig. 7 displays $\mathcal{C}^{(2)}$ and the right hand side plot displays $\mathcal{C}^{(2)}$ superimposed on $\mathcal{X}^{(2)}$.

We apply EMD to $\mathcal{X}^{(2)}$ using cubic spline interpolation (resp. linear interpolation) and obtain $i_m = 8$ (resp. $i_m = 12$). In this case, the energy-ratio approach evaluates $\widehat{i}_* = 7$ (resp. $\widehat{i}_* = 9$) which we use to estimate $\mathcal{C}^{(2)}$. Figure 8 displays the energy-ratio approach for both interpolation schemes and Fig. 9 displays the estimated trends based on the energy-ratio approach.

Now, we use $\mathcal{X}^{(2)}$ to estimate $\mathcal{C}^{(2)}$ using ℓ_1 -trend filtering and H-P filtering. For the ℓ_1 -trend filtering, we select $\lambda_{\max} = 4 \times 10^4$ and obtain $\lambda_* = 15$ with minimized Euclidean distance $E = 0.876$. For the H-P filtering, we select $\lambda_{\max} = 5 \times 10^7$ and

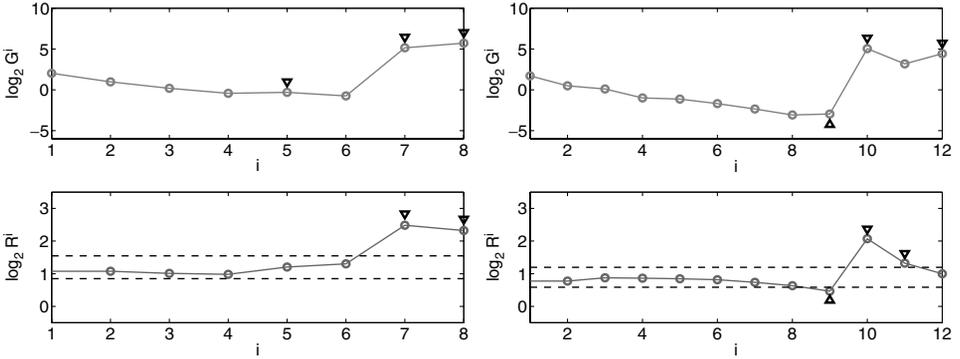


Fig. 8. Energy-ratio approach in example 2: The top-left (resp. top-right) plot displays the energy approach when cubic spline interpolation (resp. linear interpolation) is used in the EMD. The small circles are $\log_2 G^i$ for $1 \leq i \leq 8$ (resp. $2 \leq i \leq 12$) and the small triangles mark those indices $i \geq 2$ where $G^i > G^{i-1}$. The bottom-left (resp. bottom-right) plot displays the ratio approach when cubic spline interpolation (resp. linear interpolation) is used in the EMD. The small circles are $\log_2 R^i$ for $2 \leq i \leq 8$ (resp. $2 \leq i \leq 12$), the dashed lines are T_l^{SP} and T_r^{SP} (resp. T_l^{ln} and T_r^{ln}) for $\alpha = 0.08$ in Table 1, and the small triangles mark those indices i where R^i is significantly different from two. Here, the energy approach evaluates $\hat{i}_* = 5$ (resp. $\hat{i}_* = 9$), the ratio approach evaluates $\hat{i}_* = 7$ (resp. $\hat{i}_* = 9$), and the energy-ratio approach evaluates $\hat{i}_* = 7$ (resp. $\hat{i}_* = 9$) which we use to estimate $\mathcal{C}^{(2)}$.

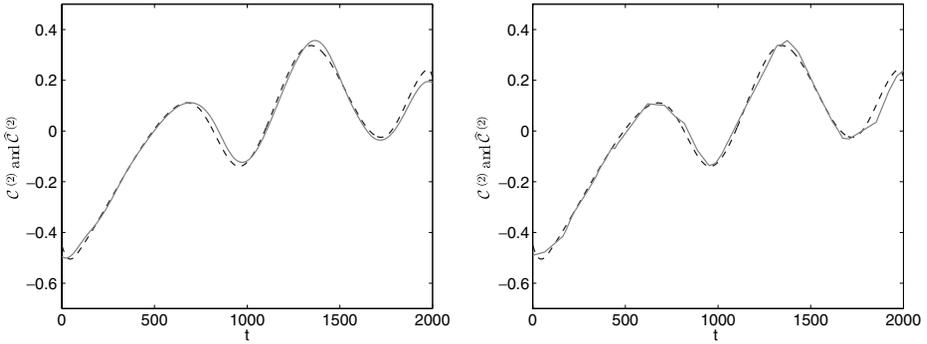


Fig. 9. EMD trend filtering in example 2: The left-hand side (resp. right-hand side) plot displays $\mathcal{C}^{(2)}$ (dashed line) versus $\widehat{\mathcal{C}}^{(2)}$ (solid line) using the energy-ratio approach when $\hat{i}_* = 7$ (resp. $\hat{i}_* = 9$). The Euclidean distance for this estimate is $E = 0.850$ (resp. $E = 0.886$).

obtain $\lambda_* \approx 2197 \times 10^3$ with Euclidean distance $E = 0.829$. Figure 10 display the best estimated trend for ℓ_1 -trend filtering and H-P filtering using λ_* in each case.

4.1.3. Simulated example 3

Let $\mathbf{Y}^{(3)} = \{Y_t^{(3)}\}_{t \geq 0}$ be a fractional Gaussian noise process with Hurst exponent $h = 0.4$. Let $\mathcal{Y}^{(3)} = (Y_0^{(3)}, Y_1^{(3)}, \dots, Y_{N-1}^{(3)})$ be a realization of $\mathbf{Y}^{(3)}$. Now, let $\mathcal{C}^{(3)}$ and $\mathcal{C}^{(4)}$ be two trend components where $\mathcal{C}^{(4)}$ is randomly constructed using peacewise

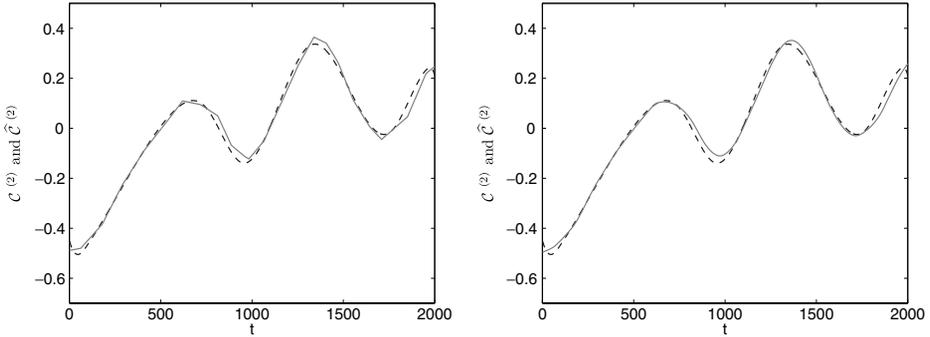


Fig. 10. The best estimate of $\mathcal{C}^{(2)}$ using ℓ_1 -trend filtering and H-P filtering in example 2: The left-hand side (resp. right-hand side) plot displays $\mathcal{C}^{(2)}$ (dashed line) versus $\widehat{\mathcal{C}}^{(2)}$ (solid line) using ℓ_1 -trend filtering (resp. H-P filtering) for $\lambda_* = 15$ (resp. $\lambda_* = 2197 \times 10^3$). The Euclidean distance is $E = 0.876$ (resp. $E = 0.829$).

linear spline technique and

$$C_t^{(3)} = 100 + \left(\frac{t + 10}{500} \right)^2.$$

We first construct a multiplicative mix $\mathcal{X}^{(3)} = \mathcal{C}^{(3)}\mathcal{Y}^{(3)}$ and use it to construct an additive mix $\mathcal{X}^{(4)} = \mathcal{X}^{(3)} + \mathcal{C}^{(4)}$. The top-left plot in Fig. 11 displays $\log \mathcal{C}^{(3)}$ and the top-right plot displays $\log \mathcal{C}^{(3)}$ superimposed on $\log |\mathcal{X}^{(3)}|$. The bottom-left plot in Fig. 11 displays $\mathcal{C}^{(4)}$ and the bottom-right plot displays $\mathcal{C}^{(4)}$ superimposed on $\mathcal{X}^{(4)}$.

We first apply EMD to $\mathcal{X}^{(4)}$ using cubic spline interpolation (resp. linear interpolation) and obtain $i_m = 8$ (resp. $i_m = 12$). The energy-ratio approach evaluates $\widehat{i}_* = 7$ (resp. $\widehat{i}_* = 10$) which we use to estimate $\mathcal{C}^{(4)}$. Figure 12 displays the energy-ratio approach for both interpolation schemes and Fig. 13 displays the estimated $\mathcal{C}^{(4)}$ based on the energy-ratio approach.

We now use $\mathcal{X}^{(4)}$ to estimate $\mathcal{C}^{(4)}$ using ℓ_1 -trend filtering and H-P filtering. For the ℓ_1 -trend filtering, we select $\lambda_{\max} = 4 \times 10^4$ and obtain $\lambda_* = 8$ with minimized Euclidean distance $E = 0.284$. For the H-P filtering, we select $\lambda_{\max} = 5 \times 10^7$ and obtain $\lambda_* \approx 2 \times 10^5$ with Euclidean distance $E = 0.307$. Figure 14 displays the best estimated trend for ℓ_1 -trend filtering and H-P filtering using λ_* in each case.

We now turn into estimating $\log \mathcal{C}^{(3)}$. In order to do that, let $\mathcal{Z} = \mathcal{X}^{(4)} - \widehat{\mathcal{C}}^{(4)}$ be an estimate of $\mathcal{X}^{(3)}$ where $\widehat{\mathcal{C}}^{(4)}$ denotes the estimated $\mathcal{C}^{(4)}$. Since $\mathcal{X}^{(3)} = \mathcal{C}^{(3)}\mathcal{Y}^{(3)}$ is a multiplicative mix, we apply EMD to $\log |\mathcal{Z}|$ using cubic spline interpolation (resp. linear interpolation) and obtain $i_m = 10$ (resp. $i_m = 12$). In this case, the energy-ratio approach evaluates $\widehat{i}_* = 8$ (resp. $\widehat{i}_* = 10$) which we use to estimate $\log \mathcal{C}^{(3)}$. Figure 15 displays the energy-ratio approach for both interpolation schemes and Fig. 16 displays the estimated log-transformed trends using the energy-ratio approach.

We now use $\log |\mathcal{Z}|$ to estimate $\log \mathcal{C}^{(3)}$ using ℓ_1 -trend filtering and H-P filtering. For the ℓ_1 -trend filtering, we select $\lambda_{\max} = 4 \times 10^4$ and obtain $\lambda_* = 487$ with

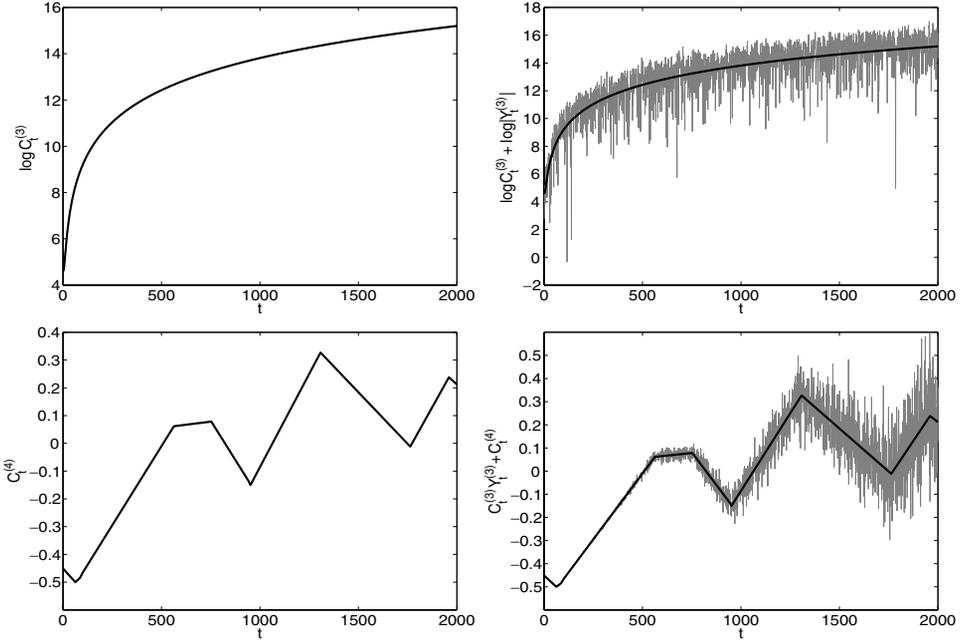


Fig. 11. Trends in example 3: The top-left and right plots display $\log \mathcal{C}^{(3)}$ and $\log \mathcal{C}^{(3)}$ superimposed on $\log |\mathcal{X}^{(3)}|$ respectively. The bottom-right and left plots display $\mathcal{C}^{(4)}$ and $\mathcal{C}^{(4)}$ superimposed on $\mathcal{X}^{(4)}$ respectively.

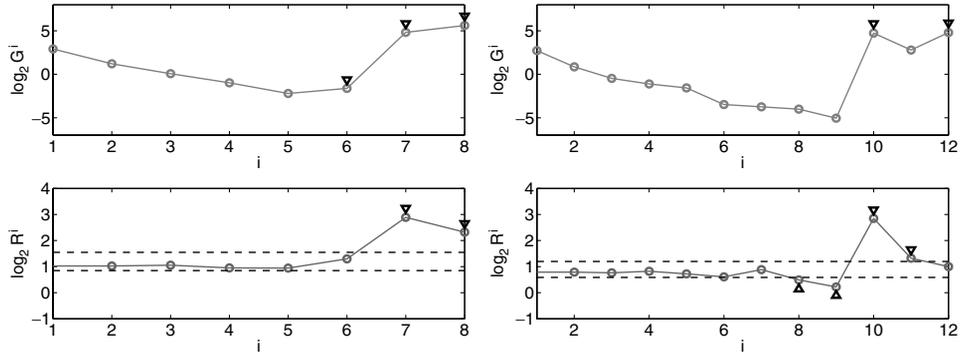


Fig. 12. Energy-ratio approach to estimate $\mathcal{C}^{(4)}$ in example 3: The top-left (resp. top-right) plot displays the energy approach when cubic spline interpolation (resp. linear interpolation) is used in the EMD. The small circles are $\log_2 G^i$ for $1 \leq i \leq 8$ (resp. $2 \leq i \leq 12$) and the small triangles mark those indices $i \geq 2$ where $G^i > G^{i-1}$. The bottom-left (resp. bottom-right) plot displays the ratio approach when cubic spline interpolation (resp. linear interpolation) is used in the EMD. The small circles are $\log_2 R^i$ for $2 \leq i \leq 8$ (resp. $2 \leq i \leq 12$), the dashed lines are T_l^{SP} and T_r^{SP} (resp. T_l^{ln} and T_r^{ln}) for $\alpha = 0.08$ in Table 1, and the small triangles mark those indices i where R^i is significantly different from two. Here, The energy approach evaluates $\hat{i}_* = 6$ (resp. $\hat{i}_* = 10$), the ratio approach evaluates $\hat{i}_* = 7$ (resp. $\hat{i}_* = 8$), and the energy-ratio approach evaluates $\hat{i}_* = 7$ (resp. $\hat{i}_* = 10$) which we use to estimate $\mathcal{C}^{(4)}$.

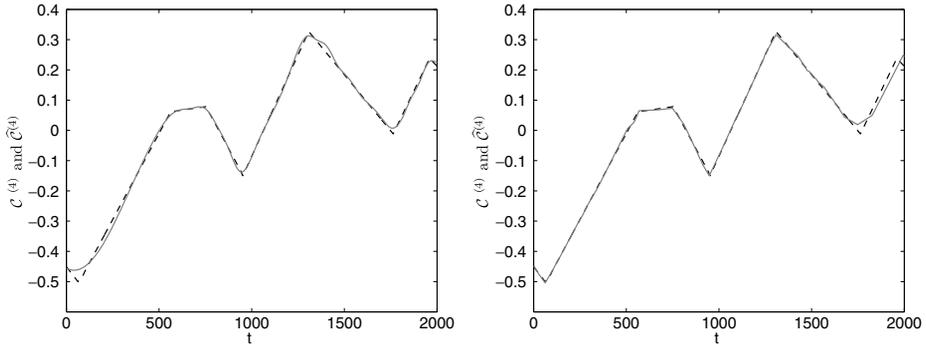


Fig. 13. EMD trend filtering to estimate $C^{(4)}$ in example 3: The left-hand side (resp. right-hand side) plot displays $C^{(4)}$ (dashed line) versus $\widehat{C}^{(4)}$ (solid line) using the energy-ratio approach when $\widehat{i}_* = 7$ (resp. $\widehat{i}_* = 10$). The Euclidean distance for this estimate is $E = 0.471$ (resp. $E = 0.415$).

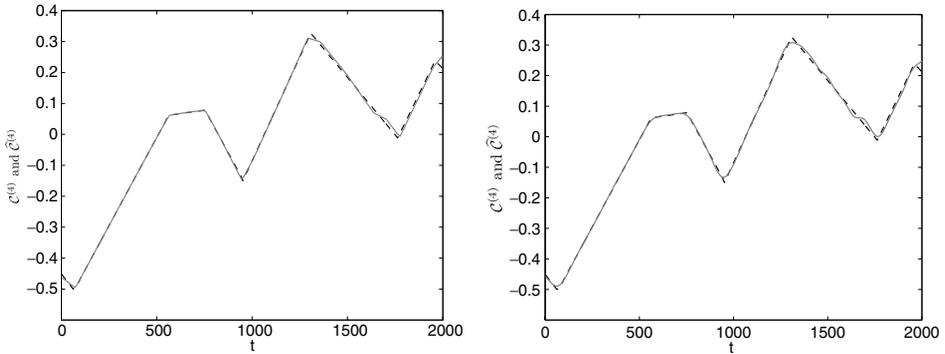


Fig. 14. The best estimate of $C^{(4)}$ using ℓ_1 -trend filtering and H-P filtering in example 3: The left-hand side (resp. right-hand side) plot displays $C^{(4)}$ (dashed line) versus $\widehat{C}^{(4)}$ (solid line) using ℓ_1 -trend filtering (resp. H-P filtering) for $\lambda_* = 8$ (resp. $\lambda_* = 2 \times 10^5$). The Euclidean distance is $E = 0.284$ (resp. $E = 0.307$).

minimized Euclidean distance $E = 3.894$. For the H-P filtering, we select $\lambda_{\max} = 5 \times 10^8$ and obtain $\lambda_* \approx 1271 \times 10^3$ with Euclidean distance $E = 5.287$. Figure 17 displays the best estimated trend for ℓ_1 -trend filtering and H-P filtering using λ_* in each case.

4.2. S&P 500 index data

In this section, we use the S&P 500 index data used in [Kim *et al.* (2009)] to demonstrate the performance of the EMD trend filtering in comparison with the ℓ_1 -trend filtering and H-P filtering. Fig. 18 displays the S&P 500 index data from March 25th, 1999 to March 9th, 2007.

The authors in [Kim *et al.* (2009)] estimated the trend of the log-transformed S&P 500 index data using the ℓ_1 -trend filtering with parameter $\lambda = 100$ and H-P

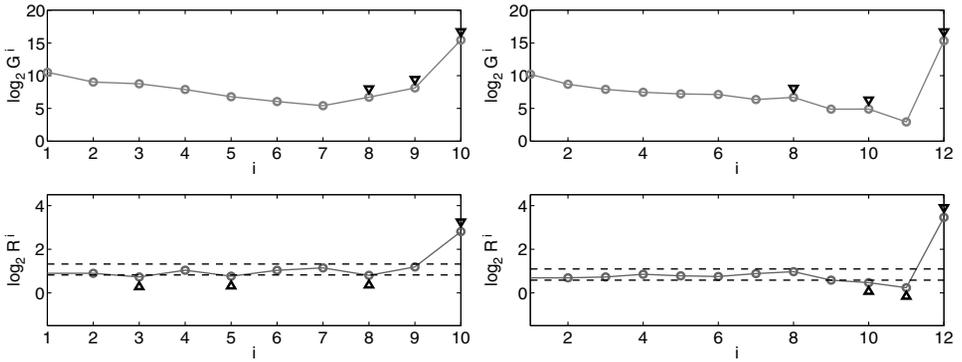


Fig. 15. Energy-ratio approach to estimate $\log \mathcal{C}^{(3)}$ in example 3: The top-left (resp. top-right) plot displays the energy approach when cubic spline interpolation (resp. linear interpolation) is used in the EMD. The small circles are $\log_2 G^i$ for $1 \leq i \leq 10$ (resp. $2 \leq i \leq 12$) and the small triangles mark those indices $i \geq 2$ where $G^i > G^{i-1}$. The bottom-left (resp. bottom-right) plot displays the ratio approach when cubic spline interpolation (resp. linear interpolation) is used in the EMD. The small circles are $\log_2 R^i$ for $2 \leq i \leq 10$ (resp. $2 \leq i \leq 12$), the dashed lines are L_l^{SP} and L_r^{SP} (resp. L_l^{ln} and L_r^{ln}) for $\alpha = 0.08$ in Table 1, and the small triangles mark those indices i where R^i is significantly different from two. Here, The energy approach evaluates $\hat{i}_* = 8$ (resp. $\hat{i}_* = 8$), the ratio approach evaluates $\hat{i}_* = 3$ (resp. $\hat{i}_* = 10$), and the energy-ratio approach evaluates $\hat{i}_* = 8$ (resp. $\hat{i}_* = 10$) which we use to estimate $\log \mathcal{C}^{(3)}$.

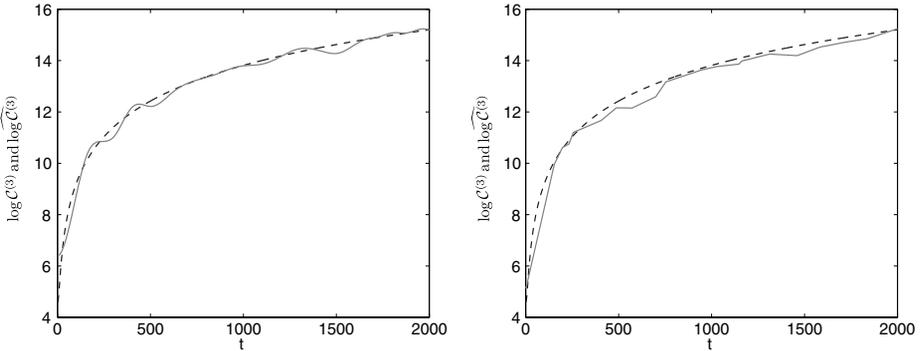


Fig. 16. EMD trend filtering to estimate $\log \mathcal{C}^{(3)}$ in example 3: The left-hand side (resp. right-hand side) plot displays $\log \mathcal{C}^{(3)}$ (dashed line) versus $\widehat{\log \mathcal{C}^{(3)}}$ (solid line) using the energy-ratio approach when $\hat{i}_* = 8$ (resp. $\hat{i}_* = 10$). The Euclidean distance for this estimate is $E = 10.156$ (resp. $E = 14.764$).

filtering with parameter λ so to obtain the same fitting error. Figure 19 displays these trend estimates.

Here, we use EMD trend filtering to estimate the trend of the log-transformed S&P 500 index data using both cubic spline and linear interpolation schemes. Figure 20 displays these estimates. We can clearly see that the estimates obtained from this method are comparable with those obtained in [Kim *et al.* (2009)].

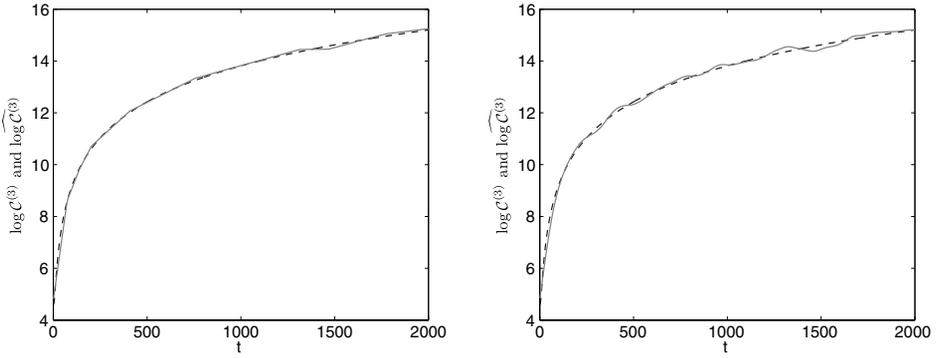


Fig. 17. The best estimate of $\log \mathcal{C}^{(3)}$ using ℓ_1 -trend filtering and H-P filtering in example 3: The left-hand side (resp. right-hand side) plot displays $\log \mathcal{C}^{(3)}$ (dashed line) versus $\log \mathcal{C}^{(3)}$ (solid line) using ℓ_1 -trend filtering (resp. H-P filtering) for $\lambda_* = 487$ (resp. $\lambda_* = 1271 \times 10^3$). The Euclidean distance is $E = 3.894$ (resp. $E = 5.287$).

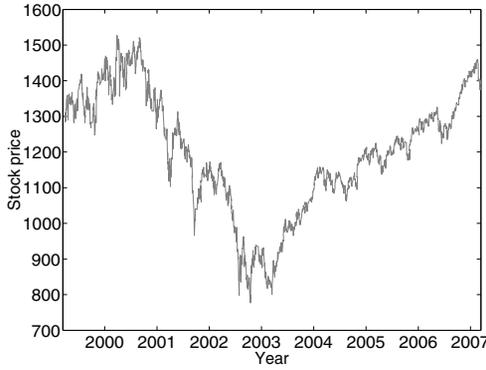


Fig. 18. S&P 500 index data: The S&P 500 index from March 25th, 1999 to March 9th, 2007.

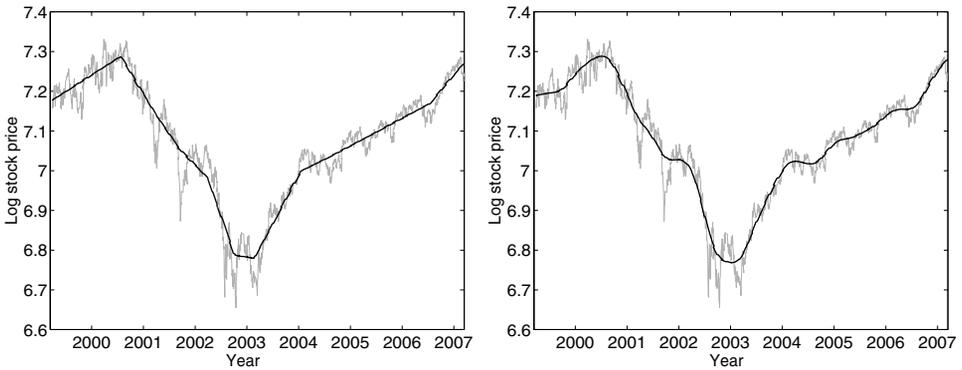


Fig. 19. Estimating trend for the log-transformed S&P 500 index data using ℓ_1 -trend filtering and H-P filtering: The left-hand side plot displays the trend estimate using ℓ_1 -trend filtering for $\lambda = 100$ and the right-hand side plot displays the trend estimate using H-P filtering with $\lambda = 2 \times 10^6$.

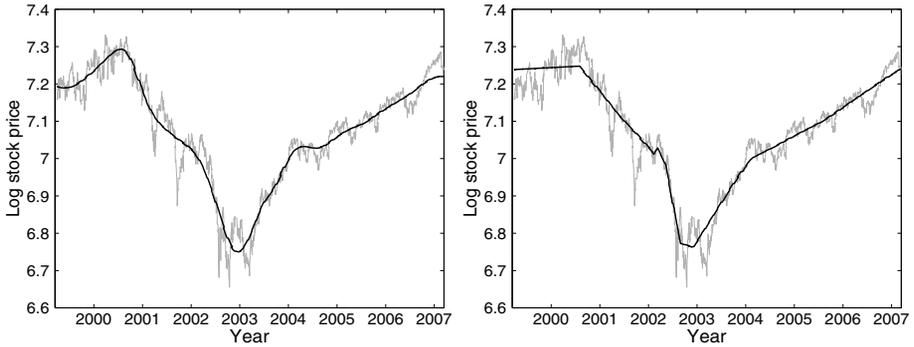


Fig. 20. Estimating trend for the log-transformed S&P 500 index data using EMD trend filtering: The left-hand side (resp. right-hand side) plot displays the trend estimate when cubic spline (resp. linear) interpolation is used in the EMD. In the left-hand side (resp. right-hand side) plot, we have used $\hat{i}_* = 7$ (resp. $\hat{i}_* = 11$) using the energy-ratio approach with $i_m = 9$ (resp. $i_m = 13$).

5. Conclusion

In this paper, we have considered EMD as an alternative to more classical approaches to trend filtering, namely H–P filtering and a recent ℓ_1 -variation upon it. The EMD rationale for what is considered here as trend filtering is to disentangle the fluctuating part of a signal from a slowly varying contribution made of a number of IMFs beyond some critical order. A mixed criterion, based on the properties of the IMFs’ energy and zero-crossing numbers, has been proposed for the automatic selection of this critical order, leading to an EMD trend filtering strategy that has been compared to H–P and ℓ_1 via extensive and well-controlled simulation experiments. As a general result of this study, it turns out that none of the three methods outperform significantly the two other ones in all cases, differences lying mostly in prior assumptions and/or effective implementation.

It is clear that, if some prior assumption about the trend is available, making use of it could help when choosing the estimation method: this is, e.g., the case with a piecewise linear model that is implicit in the use of the ℓ_1 approach. On the other hand, one strength of EMD is to be data driven rather than model based, and it is worth noticing that its performance compares favorably with ℓ_1 in the piecewise linear case while being more versatile in other situations since the solution it offers is not *a priori* constrained.

Another issue is that EMD trend filtering only depends on basic degrees of freedom of EMD (interpolation scheme, stopping criterion) and on detection thresholds that are fixed once for all, whereas both H–P and ℓ_1 call for a regularization parameter λ whose variation can result in dramatically different performance. In this respect, an encouraging feature of the EMD trend filtering is that the performance it achieves is generally close to that of H–P and/or ℓ_1 when λ is chosen in an optimal way (minimum error), a situation that is not feasible in practice when the actual trend is of course unknown.

Even if, because it is based on EMD which still lacks strong theoretical grounds, EMD trend filtering would require further studies, it is believed that it could be an interesting addition to the existing toolkit of trend filtering methods, to be used in conjunction with them.

Acknowledgment

Most of the work reported here was completed during the first author's postdoctoral stay at ENS Lyon, which was supported by ANR Grant ANR-07-BLAN-0191-01 StaRAC.

References

- Flandrin, P., Gonçalves, P. and Rilling, G. (2004). Detrending and denoising with empirical mode decompositions. *Proceeding of the XIIth European Signal Processing Conference (EUSIPCO-04)*, pp. 1581–1584.
- Flandrin, P., Rilling, G. and Gonçalves, P. (2004). Empirical mode decomposition as a filter bank. *IEEE Sig. Proc. Lett.*, **11**(2): 112–114.
- Hodrick, R. J. and Prescott, E. C. (1997). Postwar U.S. business cycles: an empirical investigation. *J. Money Credit Banking*, **29**(1): 1–16.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. L., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C. and Liu, H. H. (1998). The empirical mode decomposition and Hilbert spectrum for non-linear and non-stationary time series analysis. *Proc. R. Soc. Lond. A: Math. Phys. Eng. Sci.*, **454**: 903–995.
- Kim, S. J., Koh, K., Boyd, S. and Gorinevsky, D. (2009). ℓ_1 trend filtering. *SIAM Rev.*, **51**(2): 339–360.
- Moghtaderi, A., Flandrin, P. and Borgnat, P. (2010). Time-varying spectrum estimation of uniformly modulated processes by means of surrogate data and empirical mode decomposition. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 3678–3681.
- Rilling, G., Flandrin, P. and Gonçalves, P. (2005). Empirical mode decomposition, fractional Gaussian noise and Hurst exponent estimation. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 489–492.
- Wu, Z., Huang, N. E., Long, S.R. and Peng, C.-K. (2007). On the trend, detrending, and variability of non-linear and non-stationary time series. *Proc. Natl. Acad. Sci.*, **4**(38): 14889–14894.