



Trend Filtering via Empirical Mode Decompositions

Azadeh Moghtaderi, Patrick Flandrin, Pierre Borgnat

► To cite this version:

Azadeh Moghtaderi, Patrick Flandrin, Pierre Borgnat. Trend Filtering via Empirical Mode Decompositions. 2011. ensl-00565293v1

HAL Id: ensl-00565293

<https://ens-lyon.hal.science/ensl-00565293v1>

Preprint submitted on 11 Feb 2011 (v1), last revised 15 Feb 2011 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Trend Filtering via Empirical Mode Decompositions

Azadeh Moghtaderi^{a,*}, Patrick Flandrin^b, Pierre Borgnat^b

^a*Department of Mathematics and Statistics, Queen's University
Kingston, Ontario, Canada K7L 3N6*

^b*École Normale Supérieure de Lyon, Laboratoire de Physique
46 allée d'Italie 69364 Lyon CEDEX 07, France*

Abstract

The present work is concerned with the problem of extracting low-frequency trend from a given time series. To solve this problem, the authors develop a nonparametric technique called empirical mode decomposition (EMD) trend filtering. A key assumption is that the trend is representable as the sum of intrinsic mode functions produced by the EMD. Based on an empirical analysis of the EMD, the authors propose an automatic procedure for selecting the requisite intrinsic mode functions. To illustrate the effectiveness of the technique, the authors apply it to simulated time series containing different types of trend, as well as real-world data collected from an environmental study (atmospheric carbon dioxide levels at Mauna Loa Observatory) and from a large-scale bicycle rental service (rental numbers of Grand Lyon Vélo'v).

Keywords: Empirical mode decomposition, trend filtering, adaptive data analysis, monthly mean carbon dioxide cycle, seasonality

1. Introduction

Many real-world time series exhibit a “composite” behavior, in the sense that such a time series can be decomposed into a superposition of two “components.” Typically one of these components can be classified as “trend,”

[☆]This document is a collaborative effort.

*Corresponding author

Email addresses: `azadeh@mast.queensu.ca` (Azadeh Moghtaderi),
`patrick.flandrin@ens-lyon.fr` (Patrick Flandrin), `pierre.borgnat@ens-lyon.fr`
(Pierre Borgnat)

while the other component is classified as “fluctuation.” (Note that the word “residual” is sometimes used instead of “fluctuation.” In this paper, however, the term residual will be used in the context of empirical mode decomposition; see Section 2.) The problem of effecting such a decomposition, and classifying the resulting components as trend or fluctuation, is called the *trend filtering problem* (or *trend estimation problem*). Solving this problem is desirable, since an analysis of the trend component of a time series can often yield valuable information, e.g., for prediction. An obvious initial barrier to solving the trend filtering problem is that the terms “decomposition,” “trend,” and “fluctuation” are context-dependent. Indeed, given a time series generated by a particular physical system, it may be clear (based on physical intuition) how to solve the trend filtering problem. In the absence of physical intuition, it may still be possible to solve the trend filtering problem, provided one makes an *ad hoc* definition of trend; see (Alexandrov et al., 2008). Such definitions may require extra assumptions concerning the nature of the time series.

A common *ad hoc* definition of trend is that of a “long-term change in the mean” (Chatfield, 1996; Alexandrov et al., 2008). This definition can lead to approaches which attempt to turn the trend filtering problem into one of regression. For example, it may be reasonable to assume that the time series has a trend component described by a low-degree polynomial. The coefficients of this polynomial can then be estimated by a standard polynomial regression; we again refer to (Alexandrov et al., 2008) for a more comprehensive discussion. Other approaches exist which do not impose such a strict model on the trend. For instance, nonparametric trend filtering assumes that the fluctuation possesses generic stationarity properties, and that the trend can be found by an *ad hoc* smoothing operation applied to the entire time series, e.g., using the Henderson filter (Henderson, 1916) or the Hodrick–Prescott filter (Hodrick and Prescott, 1997). Yet another possibility is to interpret the trend estimation problem in the frequency-domain sense—for instance, one can assume the trend is represented by a particular set of low-frequency (possibly polynomials or unit root) oscillations. This turns the trend filtering problem into a *bona fide* filtering problem. Viewed in this way, it may be profitable to use Wiener–Kolmogorov filtering (Pollock, 2006) to solve the trend filtering problem. Finally, it is worthwhile to mention that generalized “trend cycles,” defined as a “short-term trend [that] generally includes cyclical fluctuations,” have also been studied (Alexandrov et al., 2008). Deciding if a trend cycle should be considered as trend (e.g., in any of

the above senses) depends on the application and of course the observation scale.

In this paper, we introduce a novel approach to solving the trend filtering problem. We call this approach *empirical mode decomposition trend filtering*. It is philosophically similar to the “low-frequency approach” described in the preceding paragraph. Indeed, empirical mode decomposition trend filtering is based on the following definition: Trend is that component of a time series which is “slowly varying” in the sense that it is represented by the “slowest” intrinsic mode functions produced by the empirical mode decomposition (EMD). Recall (Huang et al., 1998) that the EMD is an algorithm which decomposes a time series into a finite additive superposition of “intrinsic mode functions,” or IMFs. The IMFs are computed in an iterative fashion—each iteration produces an IMF which is “rapidly varying” relative to the residual time series. Thus our decomposition into components is effected by the EMD. The remaining question is “Which of the IMFs produced by the EMD should be deemed the slowest?” It is precisely this question which is addressed by EMD trend filtering. In particular, we attempt to answer this question by examining certain properties of the IMFs’ energies and zero crossing numbers; these properties were first reported in (Flandrin et al., 2004b; Rilling et al., 2005). We give evidence which supports the fact that certain changes in these properties characterize the tipping point between trend and fluctuation.

It must be mentioned that the use of the EMD to solve the trend filtering problem has already been proposed in the literature. However, such work has either relied on an *a priori* model for the fluctuation (Flandrin et al., 2004a), or has considered the trend as being the final residual time series produced by the EMD (Wu et al., 2007). In a sense, using the EMD to solve the trend filtering problem shares common features with *singular-spectrum analysis* applied to the same problem (Vautard and Ghil, 1989; Ghil and Vautard, 1992; Vautard et al., 1991). This is because the SSA also effects a decomposition into oscillatory components. Like the EMD-based method proposed by Wu et al. (2007), a possible approach to solving the trend estimation problem using SSA is to identify the trend as the lowest-frequency oscillatory component. Other possibilities are to look for oscillatory components with prescribed smoothness or monotonicity properties; see (Alexandrov et al., 2008).

The rest of the paper is organized as follows. In Section 2, we briefly review some background material concerning the EMD. In Section 3 we state what trend means in the context of this paper. In Section 4, we describe the

details of EMD trend filtering. The performance of EMD trend filtering is demonstrated in Sections 5 and 6 through analyses of simulated and real-world time series. Finally, we make concluding remarks in Section 7.

2. The Empirical Mode Decomposition

The *empirical mode decomposition* (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 interpolation.) 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 $\mathbf{X} = \{X_t\}_{t \geq 0}$ be a (real, discrete-time, stochastic) process, and let $\mathbf{x} = (X_0, X_1, \dots, X_{N-1})$ be a realization of \mathbf{X} . (These assumptions illustrate a notational convention that is used throughout the rest of the paper, namely that time series of length N are written in bold typeface and are regarded as elements of the Euclidean space \mathbb{R}^N .) As an initialization step, set $i = 1$ and $\boldsymbol{\rho}^0 = \mathbf{x}$. The EMD computes the IMFs of \mathbf{x} using the following algorithm.

- (1) Identify the local maxima and local minima of $\boldsymbol{\rho}^{i-1}$.
- (2) Together with the chosen interpolation scheme, use the maxima and minima from step (1) to compute the upper and lower envelopes of $\boldsymbol{\rho}^{i-1}$.
- (3) Determine the *local trend*, denoted \mathbf{Q}^i , as the mean of the upper and lower envelopes from step (2).
- (4) Compute the *local fluctuation*, denoted $\mathbf{h} = \mathbf{x} - \mathbf{Q}^i$.
- (5) If \mathbf{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 $\boldsymbol{\rho}^{i-1} = \mathbf{h}$. (Huang et al. (1998) call this the *sifting process*; it is this process which depends on the stopping criterion.)

- (6) If \mathbf{h} is an IMF, in the sense that it satisfies the two requirements described at the beginning of this section, then the *i*th *intrinsic mode function* of \mathcal{X} is $\mathcal{M}^i = \mathbf{h}$, and the *i*th *residual* is $\boldsymbol{\rho}^i = \mathcal{X} - \mathcal{M}^i$. Increment *i*, then go to step (1).

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

$$\mathcal{X} = \sum_{i=1}^{\mathcal{I}} \mathcal{M}^i + \boldsymbol{\rho}^{\mathcal{I}}. \quad (1)$$

In this decomposition, \mathcal{M}^1 through $\mathcal{M}^{\mathcal{I}}$ can be thought of as containing a “spectrum” of local oscillations in \mathcal{X} , with the shortest-period (highest frequency) oscillations represented in \mathcal{M}^1 and the longest-period (lowest frequency) oscillations represented in $\mathcal{M}^{\mathcal{I}}$. The computational complexity of the algorithm depends on \mathcal{X} , 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., 2004a) that if \mathcal{X} is a broadband process (a broadband process includes a relatively wide range (or band) of frequencies), then the decomposition produced by the EMD has an almost dyadic filter-bank structure, typically with $\mathcal{I} \approx \log_2 N$. Moreover, it is known that the sifting process typically halts after some tens of iterations (Huang et al., 1998).

3. Trend in EMD

As discussed in Section 1, the term “trend” is meaningless and has to be made more precise in order to be useful. In this section we state what we mean by trend in this paper and in the context of EMD trend filtering. To begin with, let us introduce some notation.

Let $\mathcal{Y} = (Y_0, Y_1, \dots, Y_{N-1})$ be a realization of a process $\mathbf{Y} = \{Y_t\}_{t \geq 0}$, and let $\mathcal{C} = (C_0, C_1, \dots, C_{N-1}) \in \mathbb{R}^n$ be a *trend* component. Assume also that \mathbf{Y} is a broadband process with a continuous spectrum. From \mathcal{Y} and \mathcal{C} we may form two new time series: The first is $\mathcal{Y} + \mathcal{C}$, the *additive mix* of \mathcal{Y} and \mathcal{C} ; the second is $\mathcal{C}\mathcal{Y}$, the *multiplicative mix* of \mathcal{Y} and \mathcal{C} . (Here

the multiplication is being performed componentwise.) In either case, we say that \mathbf{Y} is *fluctuation* of the mix. Now let \mathbf{X} be the additive or multiplicative mix of \mathbf{Y} and \mathbf{C} . The question we wish to answer is: “Solely given \mathbf{X} as data, under what conditions should it be possible to accurately estimate \mathbf{C} from \mathbf{X} ?” To do so, we must constrain the trend and fluctuation of the mix in some fashion. We take the following pragmatic approach that is based on properties of EMD.

Recall 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). A loose “definition” of a trend in this paper is that \mathbf{C} is locally slowly varying as compared to \mathbf{Y} . 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 \mathbf{X} .

Let us now turn the attention to some properties of fluctuation of the mix which can also define (in contrast) the trend. First, and in agreement with Flandrin et al. (2004b) and Wu and Huang (2004), the mean frequency of the successive IMFs of broadband processes decrease, similarly to constant- Q filter-banks, with a factor near 2. This will be the first criterion studied in Section 4.1 by estimating the mean frequency from the number of zero crossings of IMFs. Second, the finding of Rilling et al. (2005) is that the “energy” of the IMFs of broadband processes decreases as the index of the IMFs increases. This has been first reported and demonstrated in Rilling et al. (2005) for fractional Gaussian noise processes (Embrechts and Maejima, 2002) (fractional Gaussian noise processes are convenient models for generic broadband processes), and we will provide more discussion in Section 4.2 for the validity of this characteristics. An explicit assumption in our work is that the fluctuation \mathbf{Y} contaminating the trend \mathbf{C} have such energy profile. In particular, situations where the spectrum of \mathbf{Y} experiences a substantial energy increase downwards low frequencies (e.g., processes with near unit-root behavior) are not considered in this paper.

In the presence of a trend, the prescription used in this paper is that the IMF index which shows a rupture in the two properties described above, explains how to separate the trend from the fluctuation. It follows from this prescription that a trend in the present work is neither restricted to be monotonic, nor some polynomial function, nor to have a (time-averaged) spectrum well separated from the one of the fluctuation. Also, the trend can have some oscillations while in Wu et al. (2007), only the residual of EMD

was deemed a trend, hence constraining it to have no oscillations at all.

In the following two sections, we will describe in more details the properties discussed above and their abilities in separation of the trend and fluctuation of a mix.

4. EMD Trend Filtering

Let \mathcal{X} be the additive mix of \mathcal{Y} and \mathcal{C} , where these entities are given as in the previous section. As described there, our goal is to accurately estimate \mathcal{C} from \mathcal{X} . This section is devoted to describing *EMD trend filtering* which can be used to obtain such an estimate.

The following notation and terminology will be employed throughout this section. Let \mathcal{M}^i be the IMFs of \mathcal{X} , where $1 \leq i \leq \mathcal{I}$, and let i_* be such that

$$\mathcal{C}_{i_*} = \sum_{i=i_*}^{\mathcal{I}} \mathcal{M}^i + \rho^{\mathcal{I}} \quad (2)$$

is the best approximation to \mathcal{C} in the Euclidean metric. We call i_* the *best index* and \mathcal{C}_{i_*} the *best approximation* of \mathcal{C} . Clearly, estimating \mathcal{C} is equivalent to estimating the best index. If \hat{i}_* is an estimate of i_* , then we denote by

$$\hat{\mathcal{C}}_{i_*} = \sum_{i=\hat{i}_*}^{\mathcal{I}} \mathcal{M}^i + \rho^{\mathcal{I}} \quad (3)$$

the corresponding estimate of \mathcal{C} . 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.

4.1. Ratio approach

In this subsection we describe the first approach to estimating i_* , which is based on an empirical property of the zero crossing numbers of IMFs.

To describe this property, we need to establish some additional notation. 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 $i \geq 2$. (This is well-defined since $Z^i \geq 1$;

see Section 2.) Of course, R^i depends fundamentally on the given time series; since the particular time series is always clear from context, we suppress this dependence. We call R^i the *i th ratio of the zero crossing numbers* (i th RZCN). It has been observed by Flandrin et al. (2004b) and Wu and Huang (2004) that if the time series under study is a realization of a generic broadband process, the approximation $R^i \approx 2$ holds.

This observation is also supported by the following data. We considered 20 broadband processes of the following types: 17 fractional Gaussian noise processes with Hurst exponents in the range $H = 0.1, 0.15, 0.2, \dots, 0.9$, two stationary AR(2) processes, and a nonstationary AR(2) process with time-dependent coefficients. For each process in the collection, we carried out the following procedure: We simulated $B = 10000$ realizations of length $N = 2000$, then computed the IMFs of each realization along with their zero crossing numbers. Denoting the i th RZCN of its b th realization by $R^{i,b}$, where $2 \leq i \leq \mathcal{I}^b$, and setting $\vec{R}^b = (R^{2,b} \ R^{3,b} \ \dots \ R^{\mathcal{I}^b,b})$, we then computed the empirical distribution of the elements of $\vec{R} = (\vec{R}^1 \ \vec{R}^2 \ \dots \ \vec{R}^B)$. The same procedure was 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} , and supports the contention that $R^i \approx 2$. In fact, these distributions appear to be approximately Gaussian with mean 2. Furthermore, it is evident from Figure 1 (particularly the left-hand plot) that apart from the expected peak at 2, we also observe several smaller but visible peaks at higher values. These peaks appear to be due to the presence of high-order IMFs; indeed, these slowly oscillating modes have small zero crossing numbers. Because RZCNs are calculated as the ratio of two integers, if the numerator is a small number, then the distribution of the elements of \vec{R} will have peaks at integer or rational values such as 2, $5/2$, 3, $4/3$, etc. Hence, RZCNs with integer or rational values for small denominators have slightly higher expected probabilities than neighbouring values.

Our key observation is that, generically, the approximation $R^i \approx 2$ fails for i near the best index i_* . This observation is supported by the following data. For each broadband process in the collection and using its realizations, we constructed 10000 additive and 10000 multiplicative mixes. Here the trend of the additive mix is assumed to be \mathcal{C}^4 , displayed in Figure 5, while the trend of the multiplicative mix is assumed to be $1 + \mathcal{C}^4$; this ensures that the logarithms are defined. We computed the IMFs of each mix along with

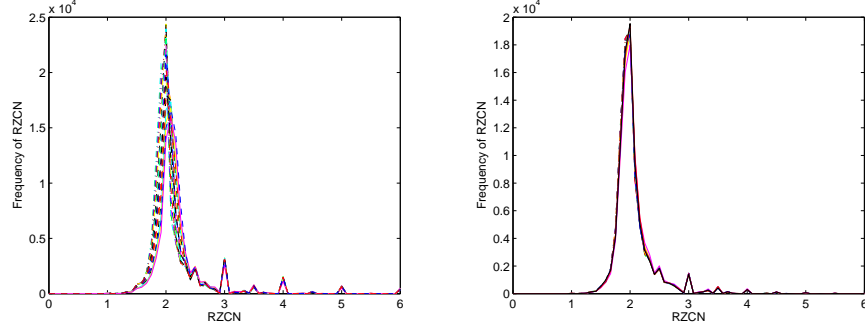


Figure 1: **Empirical distribution of the elements of \vec{R} computed for broadband processes:** The left-hand plot displays the empirical distribution of the elements of \vec{R} , for each broadband process in the collection. Different processes are represented by different line types and colors. The right-hand plot is similar, but concerns the log-transformed versions of each broadband process in the collection.

their RZCNs and took two approaches. For the first approach, we set \vec{R} and computed the empirical distribution of its elements for both additive and multiplicative mixes shown in the top plots in Figure 2. These plots indicate that the empirical distribution of the elements of \vec{R} for both additive and multiplicative mixes are non-Gaussian as the side peaks in both distributions grow taller in comparison with the empirical distributions displayed in Figure 1. Clearly at this point we can conclude that the approximation $R^i \approx 2$ fails but it is not yet clear whether it fails at or around the best index i_* or any other index. To clarify this, we proceeded to the second approach. For the second approach we used the IMFs obtained for each additive mix (resp. multiplicative mix) and used the knowledge of \mathcal{C}^4 (resp. $\log(1 + \mathcal{C}^4)$) to evaluate the best index i_* (details on how one can evaluate i_* is described in an algorithm in Section 5.) Using the knowledge of i_* , we then evaluated the fluctuation of each mix by eliminating those IMFs whose indices are greater than or equal to i_* . We call this *detrending the mix* and refer to the fluctuation obtained from detrending as the *best approximation of the fluctuation*. For all the mixes, we now use the RZCNs of the remaining IMFs in order to set \vec{R} . The bottom plots in Figure 2 display the empirical distributions of the elements of \vec{R} computed for the best approximation of the fluctuations. We can see that after detrending the mixes, the empirical distribution of the elements of \vec{R} appears to be strongly Gaussian with mean 2 as was the case with broadband processes. This result indicates that what

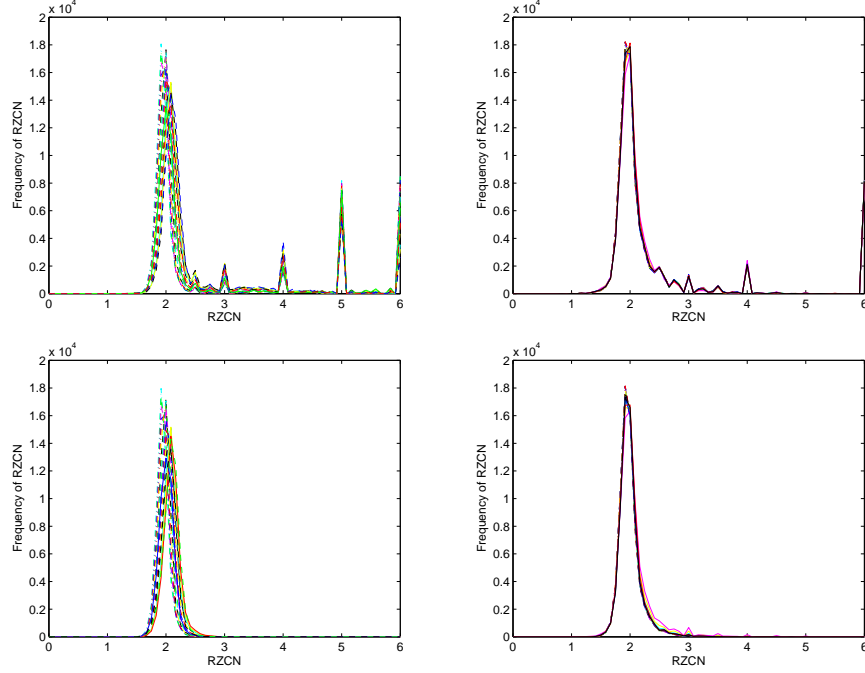


Figure 2: **Empirical distribution of the elements of \vec{R} computed for mixes:** The top-left (resp. top-right) plot displays the empirical distribution of the elements of \vec{R} for additive mixes (resp. multiplicative mixes) obtained from 20 broadband processes in the collection. The bottom-left (resp. bottom-right) plot displays the empirical distribution of the elements of \vec{R} for the best approximation of the fluctuations obtained after detrending the additive mixes (resp. multiplicative mixes). Different processes in the collection are represented by different line types and colors.

makes the empirical distribution of the elements of \vec{R} to be non-Gaussian, is the existing trends in the mixes.

Finally, using the result of our simulation, we estimate i_* by choosing \hat{i}_* to be the smallest index i for which R^i is significantly different from 2. We refer to this as the *ratio approach*. The results of our simulations for broadband processes suggest that a common threshold test could be used to conclude whether or not R^i is significantly different from 2. For $0 \leq p \leq 100$, we therefore compute $p\%$ and $(100 - p)\%$ significance level of the empirical distributions from Figure 1 as the *left threshold* and the *right threshold* respectively. Table 1 reports the averaged left and right thresholds computed for a few different values of p and for both distributions shown in Figure 1. At the end, any RZCN which is outside of the appropriate right

| p | T_l | T_r | T_l^{\log} | T_r^{\log} | p | T_l | T_r | T_l^{\log} | T_r^{\log} |
|-----|-------|-------|--------------|--------------|-----|-------|-------|--------------|--------------|
| 1 | 1.503 | 4.976 | 1.507 | 3.532 | 18 | 1.905 | 2.305 | 1.849 | 2.199 |
| 3 | 1.671 | 3.809 | 1.665 | 3.024 | 19 | 1.911 | 2.283 | 1.854 | 2.181 |
| 5 | 1.738 | 3.043 | 1.723 | 2.703 | 20 | 1.918 | 2.259 | 1.859 | 2.169 |
| 7 | 1.783 | 3.000 | 1.754 | 2.520 | 21 | 1.925 | 2.244 | 1.864 | 2.154 |
| 8 | 1.801 | 2.924 | 1.771 | 2.500 | 22 | 1.931 | 2.225 | 1.869 | 2.140 |
| 9 | 1.816 | 2.737 | 1.783 | 2.462 | 23 | 1.936 | 2.212 | 1.874 | 2.127 |
| 10 | 1.830 | 2.645 | 1.794 | 2.412 | 24 | 1.941 | 2.199 | 1.878 | 2.114 |
| 11 | 1.842 | 2.266 | 1.801 | 2.365 | 25 | 1.946 | 2.185 | 1.883 | 2.103 |
| 12 | 1.854 | 2.520 | 1.810 | 2.337 | 26 | 1.950 | 2.173 | 1.887 | 2.094 |
| 13 | 1.864 | 2.480 | 1.818 | 2.312 | 27 | 1.954 | 2.162 | 1.891 | 2.086 |
| 14 | 1.873 | 2.445 | 1.826 | 2.279 | 28 | 1.959 | 2.153 | 1.896 | 2.079 |
| 15 | 1.882 | 2.400 | 1.832 | 2.254 | 29 | 1.963 | 2.144 | 1.899 | 2.071 |
| 16 | 1.890 | 2.368 | 1.837 | 2.234 | 30 | 1.966 | 2.136 | 1.904 | 2.063 |
| 17 | 1.898 | 2.334 | 1.843 | 2.209 | 31 | 1.970 | 2.129 | 1.908 | 2.055 |

Table 1: **The left and right thresholds for different values of p :** T_l and T_r (resp. T_l^{\log} and T_r^{\log}) are the left and right thresholds associated with $p\%$ and $(100 - p)\%$ significance level of the empirical distribution shown in the left-hand (resp. right-hand) plot in Figure 1.

and left thresholds is considered *significantly different from 2*.

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 p , the smallest i for which R^i appears significantly different from 2 is a false detection. In Section 5, we will use extended simulations to decide on the “best” p .

4.2. Energy approach

In this subsection we describe the second approach to estimating i_* , which is based on an empirical property of the so-called “energy” of the IMFs.

To describe this property, we need to establish some additional notation. 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

$$G^i \triangleq \sum_{t=0}^{N-1} |\mathcal{M}_t^i|^2, \quad 1 \leq i \leq \mathcal{I}.$$

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

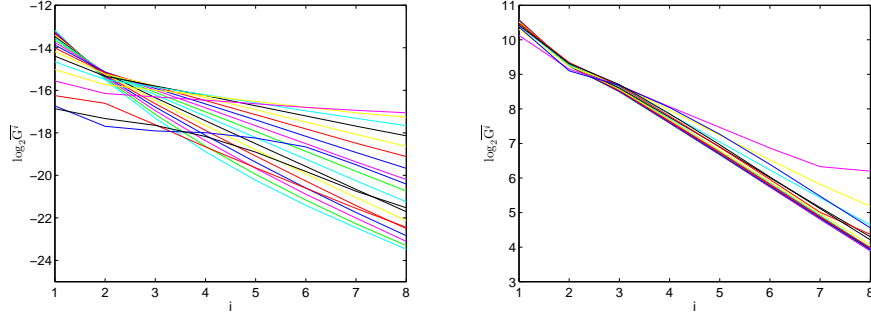


Figure 3: **$\log_2 \overline{G}^i$ for broadband processes:** The left-hand plot displays $\log_2 \overline{G}^i$ for each broadband process in the collection. Different processes are represented by different line types and colors. The right-hand plot is similar, but concerns the log-transformed versions of each broadband process in the collection.

averaged energy of its i th IMF is defined by

$$\overline{G}^i \triangleq \frac{1}{B} \sum_{b=1}^B G^{i,b}. \quad (4)$$

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. This observation is also supported by the following data. We considered the same 20 broadband processes introduced in Section 4.1. For each process in the collection, we simulated $B = 10000$ realizations of length $N = 2000$, and computed the IMFs for each realization along with $G^{i,b}$ and \overline{G}^i . Figure 3 displays $\log_2 \overline{G}^i$ for 20 broadband processes and their log-transformed versions respectively. The result of this simulation supports the idea that the averaged energy of the IMFs of a broadband process is a decreasing sequence in i .

Our key observation is that, generically, \overline{G}^i increases for i near the best index i_* . This observation is supported by the following data. We used the exact same additive and multiplicative mixes introduced in Section 4.1. For each mix, we computed its IMFs and the energy of its IMFs and took two approaches. For the first approach, we computed \overline{G}^i for the additive and multiplicative mixes. The top plots in Figure 4 display $\log_2 \overline{G}^i$ computed for additive and multiplicative mixes. The result of this simulation supports the idea that \overline{G}^i increase at some i for both additive and multiplicative mixes.

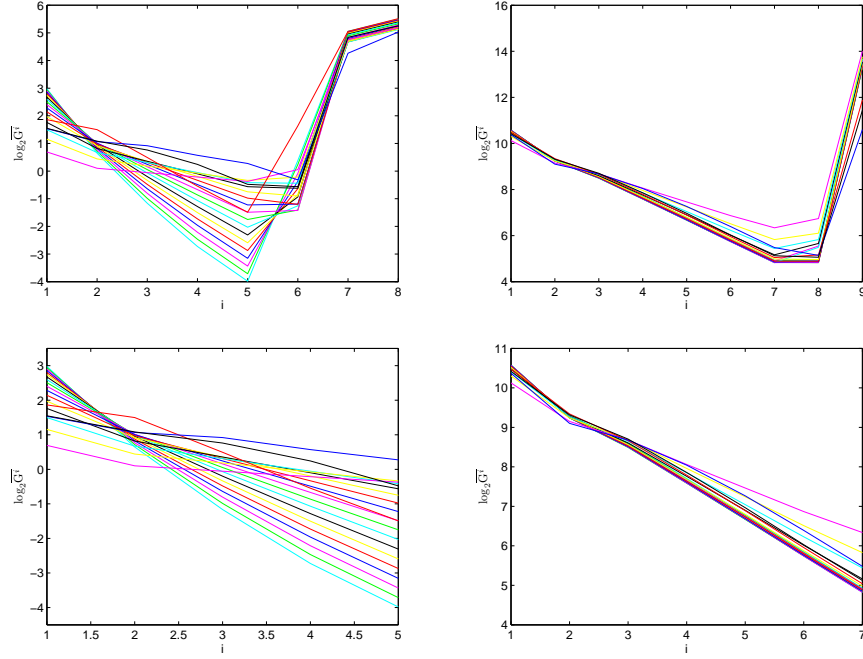


Figure 4: $\log_2 \overline{G^i}$ for mixtures: The top-left (resp. top-right) plot displays $\log_2 \overline{G^i}$ for additive mixtures (resp. multiplicative mixtures) obtained from 20 broadband processes in the collection. The bottom-left (resp. bottom-right) plot displays $\log_2 \overline{G^i}$ for the best approximation of the fluctuations obtained after detrending the additive mixtures (resp. multiplicative mixtures). Different mixtures or the best approximation of the fluctuations are represented by different line types and colors.

However, this result does not provide any information on whether or not these increases have occurred at or around the best index i_* . To show this we proceeded to the second approach. For the second approach, we computed $\overline{G^i}$ for the best approximation of the fluctuations obtained after detrending the mixtures as described in Section 4.1. The bottom plots in Figure 4 display $\log_2 \overline{G^i}$, computed for the best approximation of the fluctuations. The result of this simulation supports the idea that $\overline{G^i}$ increases at the best index i_* . This is clear from the comparison between the top and bottom plots of Figure 4.

Based on the above discussion, identifying the smallest index $i \geq 2$ such that $\overline{G^i} > \overline{G^{i-1}}$ evaluates \hat{i}_* . This approach is called the *energy approach*. As for the ratio approach, one could think of looking for *significant* increases which would be based on some statistical information about the dispersion

of energy of each IMF. This viewpoint has been considered first for white Gaussian noise in Huang et al. (2003) and further generalized in Flandrin and Gonçalves (2004) and Flandrin et al. (2004a), even in a detrending perspective. The limitation however is that the confidence intervals associated with this approach depend strongly on some prior knowledge about the spectra of broadband processes. This is mainly the reason for which we do not follow such direction, as we are merely interested in a procedure which is model-free.

The limitation with the energy approach is that one is often given a single time series to use in the trend estimation procedure. Clearly, the energy approach in this case is to identify the smallest index $i \geq 2$ such that $G^i > G^{i-1}$. Computation of energy, based on only one realization, may cause false increases in G^i at indices which do not associate with trend.

4.3. Energy-ratio approach

In this section, we introduce the last and most important approach to estimating i_* .

As we 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 detections can be reduced by combining these two approaches.

To be more precise, for each $2 \leq i \leq \mathcal{I}$, 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 2. We then evaluate \hat{i}_* to be the smallest common index in both approaches. This approach is called the *energy-ratio approach*. In the next section, we demonstrate the overall performance of this approach in comparison with the other two.

5. Performance Evaluation of the EMD Trend Filtering; Evaluation of the best p

We follow two main goals in this section. The first goal is to evaluate the overall performance of the EMD trend filtering. The second goal is to empirically evaluate the best p using which we can improve the performance of the energy-ratio approach by reducing the number of false detections in the energy and ratio approaches.

In order to do the above, we constructed 10 simulated examples structured as follows. Let $\mathbf{Y}^k = \{Y_t^k\}_{t \geq 0}$, $1 \leq k \leq 6$, be 6 generic broadband processes where for $k = \{1, 4, 5, 6\}$, \mathbf{Y}^k are fractional Gaussian noise processes with Hurst exponents 0.5, 0.7, 0.15, and 0.75 respectively and for $k = \{2, 3\}$,

$$\begin{aligned} Y_t^2 &= 0.8Y_{t-1}^2 - 0.4Y_{t-2}^2 + \zeta_t, \quad \text{and} \\ Y_t^3 &= 0.2Y_{t-1}^3 + 0.5Y_{t-2}^3 + \xi_t, \end{aligned}$$

where $\{\zeta_t\}$ and $\{\xi_t\}$ are two independent white noise processes with variance 10^4 . Let $\mathbf{y}^k = (Y_0^k, Y_1^k, \dots, Y_{N-1}^k)$ be a realization of \mathbf{Y}^k . Now, let us assume that $\mathbf{C}^k = (C_0^k, C_1^k, \dots, C_{N-1}^k)$, $1 \leq k \leq 6$, are 6 trends where for $k = \{1, 2, 3, 4\}$, \mathbf{C}^k are randomly constructed trends using piecewise linear and cubic spline techniques and for $k = \{5, 6\}$

$$\begin{aligned} C_t^5 &= 2 - e^{\frac{-(t-1000)^2}{2 \times 400^2}}, \quad \text{and} \\ C_t^6 &= 1.5 + \cos(2\pi f_s t), \quad f_s = 0.002. \end{aligned}$$

Figure 5 displays \mathbf{C}^k for $1 \leq k \leq 6$ and for $N = 2000$.

We constructed 6 additive mix and 4 multiplicative mix \mathbf{X}^k , $1 \leq k \leq 10$, such that

$$\mathbf{X}^k = \begin{cases} \mathbf{C}^k + \mathbf{y}^k, & 1 \leq k \leq 6 \\ (1 + \mathbf{C}^{k-6})\mathbf{y}^{k-6}, & 7 \leq k \leq 8 \\ \mathbf{C}^{k-4}\mathbf{y}^{k-4}, & 9 \leq k \leq 10. \end{cases} \quad (5)$$

For each k , we created $B = 10000$ realizations of length $N = 2000$ of \mathbf{Y}^k and constructed the mixes appropriately using Eq. (5). In the following we denote the b th realization of the k th example by b_k .

Recall from earlier that the first goal in this section is to evaluate the overall performance of the EMD trend filtering. In order to achieve such goal, we first proceed with the following algorithm and make a discussion afterwards.

1. Apply EMD to \mathbf{X}^{b_k} in order to extract its IMFs. Denote \mathcal{I} , \mathcal{M}^i , and $\rho^{\mathcal{I}}$ by \mathcal{I}^{b_k} , $\mathcal{M}^{i^{b_k}}$, and $\rho^{\mathcal{I}^{b_k}}$ respectively.
2. Use knowledge of \mathbf{C}^k and compute the Euclidean distance

$$E_{t_{\dagger}}^{\mathbf{C}^{b_k}} \triangleq \left(\sum_{t=0}^{N-1} \left| C_t^k - C_{t, t_{\dagger}}^{b_k} \right|^2 \right)^{1/2}, \quad (6)$$

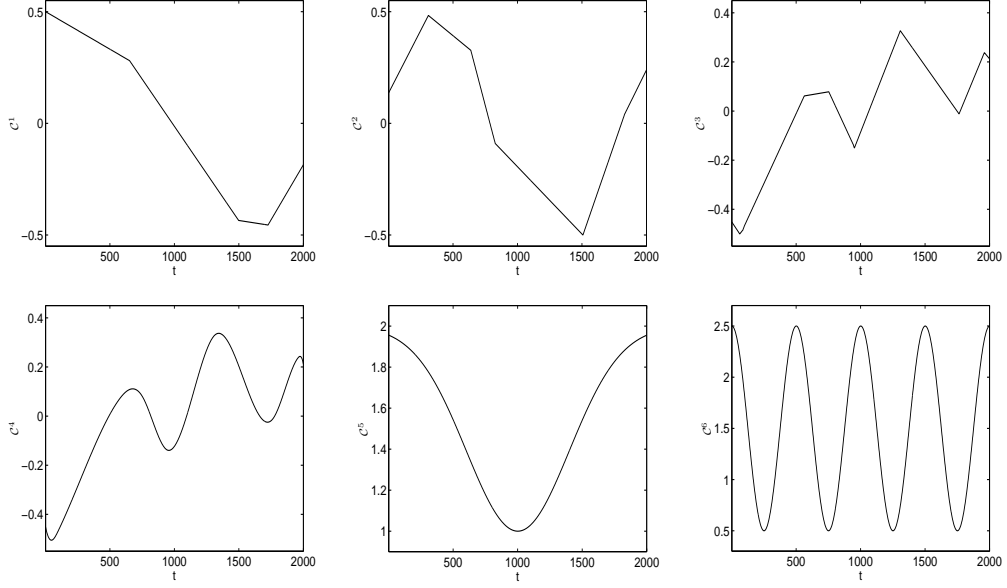


Figure 5: **Trends used in simulated examples:** \mathcal{C}^k for $1 \leq k \leq 6$ are six trends used in simulated examples 1 to 10

where $\mathcal{C}_{i_{\dagger}}^{b_k} = (C_{0,i_{\dagger}}^{b_k}, C_{1,i_{\dagger}}^{b_k}, \dots, C_{N-1,i_{\dagger}}^{b_k})$ is

$$\mathcal{C}_{i_{\dagger}}^{b_k} = \sum_{i=i_{\dagger}}^{\mathcal{I}^{b_k}} \mathcal{M}^{i^{b_k}} + \rho^{\mathcal{I}^{b_k}}, \quad i_{\dagger} \in \{1, 2, \dots, \mathcal{I}^{b_k}\}.$$

3. The best index i_* is that i_{\dagger} which results in minimum $E_{i_{\dagger}}^{\mathcal{C}^{b_k}}$, denoted $E_{i_*}^{\mathcal{C}^{b_k}}$. Clearly, $\mathcal{C}_{i_*}^{b_k}$ is the best approximation of \mathcal{C}^k .
4. Compute the Euclidean norm

$$E_{i_*}^{\mathcal{Y}^{b_k}} \triangleq \left(\sum_{t=0}^{N-1} |\mathcal{Y}_{t,i_*}^{b_k}|^2 \right)^{1/2}, \quad (7)$$

where $\mathcal{Y}_{i_*}^{b_k} = \mathcal{X}^{b_k} - \mathcal{C}_{i_*}^{b_k}$ is the best approximation of \mathcal{Y}^{b_k} .

5. Compute the Euclidean norm of the fluctuation \mathcal{Y}^{b_k} by

$$E^{\mathcal{Y}^{b_k}} = \left(\sum_{t=0}^{N-1} |\mathcal{Y}_t^{b_k}|^2 \right)^{1/2}.$$

6. Apply the Hodrick–Prescott (H-P) filter (Hodrick and Prescott, 1997) to \mathbf{x}^{b_k} to estimate \mathbf{C}^k , denoted $\hat{\mathbf{C}}_{\text{hp}}^{b_k}$. Compute the Euclidean distance between \mathbf{C}^k and $\hat{\mathbf{C}}_{\text{hp}}^{b_k}$ and denote it by $E_{\text{hp}}^{\hat{\mathbf{C}}^{b_k}}$.
7. Apply the singular-spectrum analysis (SSA) (Vautard et al., 1991) to \mathbf{x}^{b_k} to estimate \mathbf{C}^k , denoted $\hat{\mathbf{C}}_{\text{ssa}}^{b_k}$. Compute the Euclidean distance between \mathbf{C}^k and $\hat{\mathbf{C}}_{\text{ssa}}^{b_k}$ and denote it by $E_{\text{ssa}}^{\hat{\mathbf{C}}^{b_k}}$.

Note that the above algorithm is designed for examples with additive mixes. The extension of this algorithm to examples with multiplicative mixes via log-transformation is straight forward.

For each k , we now average $E_{i_*}^{\mathbf{C}^{b_k}}$, $E_{i_*}^{\mathbf{y}^{b_k}}$, $E_{\text{hp}}^{\hat{\mathbf{C}}^{b_k}}$, and $E_{\text{ssa}}^{\hat{\mathbf{C}}^{b_k}}$ and denote them, in order, by $\bar{E}_{i_*}^{\mathbf{C}^k}$, $\bar{E}_{i_*}^{\mathbf{y}^k}$, $\bar{E}^{\mathbf{y}^k}$, $\bar{E}_{\text{hp}}^{\hat{\mathbf{C}}^k}$, and $\bar{E}_{\text{ssa}}^{\hat{\mathbf{C}}^k}$. Table 2 reports all these averaged Euclidean distances/norms.

The parameters used for the H-P filter are 10^5 and 5×10^5 (associated with the second and third columns of Table 2 respectively) and the window lengths used for the SSA are 100 and 200 (associated with the fourth and fifth columns of Table 2 respectively). To evaluate the performance of the EMD trend filtering, we make the following discussion. The first column of Table 2 reports the average of the Euclidean distances between \mathbf{C}^k and the b_k th best approximation of \mathbf{C}^k for different values of k . Looking solely at these values, we cannot say anything about the performance of the EMD trend filtering unless we make some type of comparisons. The first attempt we make is to compare the best approximation of \mathbf{C}^k obtained from EMD trend filtering with the estimates obtained from H-P filter and the SSA. This is done by comparing the first column of Table 2 with the second to fifth columns for each k . It is clear from the reported values that all three trend filtering methods are comparable since their reported averaged Euclidean distances are comparable. We should note however that both H-P filter and SSA are dependent on free parameters. The value of the free parameter in either method can change the performance of the trend filtering in comparison with EMD trend filtering. This is clear where i.e., the majority of the values reported in the second column of Table 2 are larger than those from the first column of Table 2 whereas this situation reverses for the third column of Table 2. A similar case is applied for the SSA. The second attempt is an alternative way to determine the performance of EMD trend filtering by comparing the fluctuation of each mix, \mathbf{y}^{b_k} , with the best approximation of

| k | $\bar{E}_{i_*}^{\mathcal{C}^k}$ | $\bar{E}_{hp}^{\hat{\mathcal{C}}^k}$ | $\bar{E}_{hp}^{\tilde{\mathcal{C}}^k}$ | $\bar{E}_{ssa}^{\hat{\mathcal{C}}^k}$ | $\bar{E}_{ssa}^{\tilde{\mathcal{C}}^k}$ | $\bar{E}_{i_*}^{\mathcal{Y}^k}$ | $\bar{E}^{\mathcal{Y}^k}$ |
|----|---------------------------------|--------------------------------------|--|---------------------------------------|---|---------------------------------|---------------------------|
| 1 | 0.752 | 0.898 | 0.733 | 0.871 | 0.624 | 7.392 | 7.400 |
| 2 | 0.822 | 0.840 | 0.697 | 0.818 | 0.627 | 5.475 | 5.493 |
| 3 | 0.887 | 0.930 | 0.785 | 0.920 | 0.729 | 2.854 | 2.922 |
| 4 | 0.642 | 0.655 | 0.581 | 0.652 | 0.542 | 2.334 | 2.398 |
| 5 | 0.631 | 0.393 | 0.286 | 1.651 | 0.211 | 17.32 | 17.31 |
| 6 | 4.594 | 4.314 | 3.850 | 4.273 | 3.715 | 12.47 | 13.18 |
| 7 | 4.808 | 6.014 | 4.898 | 3.924 | 3.070 | 49.73 | 49.63 |
| 8 | 6.369 | 7.028 | 5.732 | 3.924 | 3.070 | 49.84 | 49.66 |
| 9 | 4.803 | 6.396 | 5.212 | 4.144 | 2.863 | 49.75 | 49.65 |
| 10 | 8.513 | 7.315 | 6.123 | 6.265 | 13.75 | 49.45 | 49.61 |

Table 2: **Average Euclidean distances/norms computed for three trend filtering methods:** For each k , the first column of this table reports the averaged Euclidean distances between C^k and the b_k th best approximation of C^k . The second and third columns report the averaged Euclidean distances between C^k and $\hat{C}_{hp}^{b_k}$ using parameters 10^5 and 5×10^5 respectively. The fourth and fifth columns of this table report the averaged Euclidean distances between C^k and $\hat{C}_{ssa}^{b_k}$ using window lengths 100 and 200 respectively. The sixth column is the averaged Euclidean norms of the best approximation of \mathcal{Y}^{b_k} and the seventh column is the averaged Euclidean norms of the fluctuation \mathcal{Y}^{b_k} .

\mathcal{Y}^{b_k} . This is done by comparing the averaged Euclidean norms reported in the sixth and seventh columns of Table 2. The fact that the Euclidean norms from column 6 can be compared well with those in the seventh column is an indication that the EMD trend filtering has performed well.

Recall now that the second goal is to empirically evaluate the best p in order to obtain the best estimate of i_* . In order to achieve this goal, we first proceed with the following algorithm and make a discussion afterwards.

- Apply EMD to \mathcal{X}^{b_k} in order to extract its IMFs. Denote \mathcal{I} , \mathcal{M}^i , and $\rho^{\mathcal{I}}$ by \mathcal{I}^{b_k} , $\mathcal{M}^{i^{b_k}}$, and $\rho^{\mathcal{I}^{b_k}}$ respectively.
- Estimate the best index, \hat{i}_* , using the energy approach and compute an estimate of \mathcal{C}^k by

$$\hat{\mathcal{C}}_g^{b_k} = \sum_{i=\hat{i}_*}^{\mathcal{I}^{b_k}} \mathcal{M}^{i^{b_k}} + \rho^{\mathcal{I}^{b_k}}. \quad (8)$$

- Fix the value p in the interval $0 \leq p \leq 100$ and estimate the best index, \hat{i}_* , using the ratio approach and the energy-ratio approach. Compute

the estimates of \mathcal{C}^k using ratio and energy-ratio approaches by $\hat{\mathcal{C}}_r^{b_k,p}$ and $\hat{\mathcal{C}}_{gr}^{b_k,p}$, respectively.

- Compute the Euclidean distance between \mathcal{C}^k and estimates $\hat{\mathcal{C}}_g^{b_k}$, $\hat{\mathcal{C}}_r^{b_k,p}$, and $\hat{\mathcal{C}}_{gr}^{b_k,p}$ and denote them by $E_g^{\hat{\mathcal{C}}^{b_k}}$, $E_r^{\hat{\mathcal{C}}^{b_k,p}}$, and $E_{gr}^{\hat{\mathcal{C}}^{b_k,p}}$, respectively.
- Compute the Euclidean distance between \mathcal{X}^{b_k} and estimates $\hat{\mathcal{C}}_g^{b_k}$, $\hat{\mathcal{C}}_r^{b_k,p}$, and $\hat{\mathcal{C}}_{gr}^{b_k,p}$ and denote them by $E_g^{\hat{\mathcal{Y}}^{b_k}}$, $E_r^{\hat{\mathcal{Y}}^{b_k,p}}$, and $E_{gr}^{\hat{\mathcal{Y}}^{b_k,p}}$, respectively.

Similarly to the previous algorithm, the above algorithm is designed for the examples with additive mixes. The extension to examples with multiplicative mixes via log-transformation is straight forward.

For each k , we now average $E_g^{\hat{\mathcal{C}}^{b_k}}$, $E_r^{\hat{\mathcal{C}}^{b_k,p}}$, $E_{gr}^{\hat{\mathcal{C}}^{b_k,p}}$, $E_g^{\hat{\mathcal{Y}}^{b_k}}$, $E_r^{\hat{\mathcal{Y}}^{b_k,p}}$, and $E_{gr}^{\hat{\mathcal{Y}}^{b_k,p}}$ denoted, in order, by $\bar{E}_g^{\hat{\mathcal{C}}^k}$, $\bar{E}_r^{\hat{\mathcal{C}}^k}$, $\bar{E}_{gr}^{\hat{\mathcal{C}}^k}$, $\bar{E}_g^{\hat{\mathcal{Y}}^k}$, $\bar{E}_r^{\hat{\mathcal{Y}}^k}$, and $\bar{E}_{gr}^{\hat{\mathcal{Y}}^k}$.

Tables 3 and 4 report the averaged Euclidean distances for the ratio and energy-ratio approaches respectively when using 22 values of p for each example. In order to evaluate the best p , one may think of selecting p so that it associates with the minimum averaged Euclidean distances reported in Tables 3 and 4. Looking at the reported values in Table 3, we can see that the minimum averaged Euclidean distances in examples 1 to 10 are obtained when $p = (5, 11, 11, 11, 1, 13, 1, 5, 3, 7)$. A similar search in Table 4 gives $p = (24, 24, 15, 18, 5, 22, 9, 13, 9, 24)$. Clearly, this thought cannot be used since the values of p associated with the minimum Euclidean distances appear to be strongly dependent on the estimation approach and also the type of example. In the following, we make an argument using which we can evaluate the best p . Note that we make this evaluation for additive and multiplicative mixes independently.

- One of the main discussions in this paper is that the energy-ratio approach should reduce the false detections in the energy and ratio approaches. The reported values in Tables 3 and 4 show that for all values of p , except for a few, all examples show a smaller averaged Euclidean distance for the energy-ratio approach than for the ratio approach. This mainly means that the energy-ratio approach reduces the false detections of the ratio approach regardless of p . In order for the energy-ratio approach to reduce the false detections of the energy

| p | $\overline{E}_r^{\widehat{C}^1_p}$ | $\overline{E}_r^{\widehat{C}^2_p}$ | $\overline{E}_r^{\widehat{C}^3_p}$ | $\overline{E}_r^{\widehat{C}^4_p}$ | $\overline{E}_r^{\widehat{C}^5_p}$ | $\overline{E}_r^{\widehat{C}^6_p}$ | $\overline{E}_r^{\widehat{C}^7_p}$ | $\overline{E}_r^{\widehat{C}^8_p}$ | $\overline{E}_r^{\widehat{C}^9_p}$ | $\overline{E}_r^{\widehat{C}^{10}_p}$ |
|-----|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|---------------------------------------|
| 1 | 1.524 | 5.346 | 5.819 | 3.441 | 0.655 | 21.43 | 5.421 | 9.033 | 6.350 | 21.62 |
| 3 | 1.092 | 2.987 | 3.400 | 1.807 | 0.663 | 10.86 | 5.457 | 8.262 | 5.877 | 17.44 |
| 5 | 1.020 | 1.709 | 1.618 | 0.917 | 0.863 | 6.927 | 6.037 | 7.972 | 6.115 | 13.20 |
| 7 | 1.083 | 1.694 | 1.601 | 0.916 | 1.237 | 6.847 | 7.206 | 8.607 | 7.185 | 11.72 |
| 8 | 1.109 | 1.474 | 1.341 | 0.838 | 1.514 | 6.558 | 8.150 | 9.297 | 8.031 | 12.31 |
| 9 | 1.137 | 1.379 | 1.137 | 0.770 | 1.792 | 5.977 | 9.142 | 9.913 | 8.805 | 12.29 |
| 10 | 1.191 | 1.366 | 1.101 | 0.759 | 2.101 | 5.752 | 10.18 | 10.73 | 9.822 | 12.76 |
| 11 | 1.253 | 1.336 | 1.065 | 0.750 | 2.419 | 5.550 | 10.94 | 11.50 | 10.63 | 12.99 |
| 12 | 1.335 | 1.371 | 1.068 | 0.760 | 2.766 | 5.506 | 11.88 | 12.37 | 11.54 | 13.73 |
| 13 | 1.404 | 1.412 | 1.067 | 0.770 | 3.122 | 5.456 | 12.88 | 13.28 | 12.59 | 14.45 |
| 14 | 1.491 | 1.453 | 1.078 | 0.782 | 3.482 | 5.463 | 13.84 | 14.16 | 13.62 | 15.23 |
| 15 | 1.577 | 1.513 | 1.091 | 0.797 | 3.815 | 5.465 | 14.72 | 14.95 | 14.58 | 15.94 |
| 16 | 1.671 | 1.574 | 1.112 | 0.814 | 9.125 | 5.508 | 15.44 | 15.61 | 15.35 | 16.54 |
| 17 | 1.772 | 1.639 | 1.156 | 0.837 | 4.438 | 5.559 | 16.37 | 16.63 | 16.32 | 17.35 |
| 18 | 1.884 | 1.706 | 1.211 | 0.858 | 4.739 | 5.629 | 17.24 | 17.43 | 17.14 | 18.08 |
| 19 | 1.986 | 1.772 | 1.277 | 0.880 | 5.019 | 5.720 | 18.10 | 18.26 | 18.02 | 18.94 |
| 20 | 2.097 | 1.856 | 1.379 | 0.912 | 5.314 | 5.832 | 18.74 | 18.98 | 18.75 | 19.65 |
| 22 | 2.319 | 2.011 | 1.580 | 0.983 | 5.861 | 6.135 | 20.31 | 20.72 | 20.43 | 21.25 |
| 24 | 2.503 | 2.174 | 1.763 | 1.059 | 6.264 | 6.535 | 21.91 | 22.49 | 22.08 | 22.97 |
| 26 | 2.695 | 2.375 | 1.943 | 1.156 | 6.599 | 7.083 | 23.48 | 24.26 | 23.67 | 31.54 |
| 28 | 2.907 | 2.597 | 2.076 | 1.258 | 6.931 | 7.639 | 24.98 | 25.85 | 25.33 | 26.15 |
| 30 | 3.111 | 2.812 | 2.178 | 1.350 | 7.195 | 8.149 | 26.50 | 27.58 | 26.97 | 27.76 |

Table 3: **Averaged Euclidean distances computed for the ratio approach:** For each k and a fixed $1 \leq p \leq 30$, this table reports the average of the Euclidean distances between C^k and $\widehat{C}_r^{b_k, p}$. The appropriate left and right thresholds for a given p are taken from Table 1. For each p , the minimum averaged Euclidean distances in examples 1 to 10 are marked in bold.

| p | $\overline{E}_{\text{gr}}^{\hat{\mathcal{C}}^1_p}$ | $\overline{E}_{\text{gr}}^{\hat{\mathcal{C}}^2_p}$ | $\overline{E}_{\text{gr}}^{\hat{\mathcal{C}}^3_p}$ | $\overline{E}_{\text{gr}}^{\hat{\mathcal{C}}^4_p}$ | $\overline{E}_{\text{gr}}^{\hat{\mathcal{C}}^5_p}$ | $\overline{E}_{\text{gr}}^{\hat{\mathcal{C}}^6_p}$ | $\overline{E}_{\text{gr}}^{\hat{\mathcal{C}}^7_p}$ | $\overline{E}_{\text{gr}}^{\hat{\mathcal{C}}^8_p}$ | $\overline{E}_{\text{gr}}^{\hat{\mathcal{C}}^9_p}$ | $\overline{E}_{\text{gr}}^{\hat{\mathcal{C}}^{10}_p}$ |
|-----|--|--|--|--|--|--|--|--|--|---|
| 1 | 1.222 | 3.377 | 4.346 | 2.500 | 0.654 | 21.42 | 5.429 | 9.101 | 6.359 | 21.59 |
| 3 | 1.086 | 2.275 | 3.174 | 1.729 | 0.633 | 10.89 | 5.300 | 8.353 | 5.729 | 17.44 |
| 5 | 0.993 | 1.483 | 1.584 | 0.921 | 0.633 | 6.930 | 5.204 | 7.636 | 5.229 | 12.99 |
| 7 | 0.992 | 1.482 | 1.582 | 0.920 | 0.633 | 6.846 | 5.231 | 7.423 | 5.246 | 10.89 |
| 8 | 0.963 | 1.328 | 1.355 | 0.832 | 0.634 | 6.528 | 5.238 | 7.429 | 5.252 | 10.89 |
| 9 | 0.927 | 1.209 | 1.137 | 0.751 | 0.634 | 5.897 | 5.202 | 7.178 | 5.103 | 10.16 |
| 10 | 0.912 | 1.141 | 1.092 | 0.724 | 0.635 | 5.621 | 5.228 | 7.162 | 5.133 | 9.909 |
| 11 | 0.893 | 1.071 | 1.047 | 0.702 | 0.635 | 5.370 | 5.250 | 7.132 | 5.173 | 9.511 |
| 12 | 0.890 | 1.045 | 1.037 | 0.697 | 0.636 | 5.265 | 5.266 | 7.146 | 5.191 | 9.444 |
| 13 | 0.875 | 1.032 | 1.024 | 0.691 | 0.636 | 5.158 | 5.288 | 7.078 | 5.229 | 9.330 |
| 14 | 0.875 | 1.007 | 1.024 | 0.686 | 0.637 | 5.088 | 5.322 | 7.088 | 5.270 | 9.236 |
| 15 | 0.871 | 0.976 | 1.020 | 0.681 | 0.638 | 5.008 | 5.348 | 7.099 | 5.300 | 9.172 |
| 16 | 0.857 | 0.955 | 1.026 | 0.680 | 0.639 | 4.962 | 5.369 | 7.097 | 5.317 | 9.054 |
| 17 | 0.856 | 0.952 | 1.052 | 0.679 | 0.640 | 4.916 | 5.398 | 7.109 | 5.356 | 9.022 |
| 18 | 0.853 | 0.925 | 1.085 | 0.677 | 0.642 | 4.887 | 5.414 | 7.096 | 5.389 | 8.999 |
| 19 | 0.847 | 0.916 | 1.135 | 0.677 | 0.642 | 4.878 | 5.433 | 7.120 | 5.412 | 8.982 |
| 20 | 0.846 | 0.909 | 1.217 | 0.679 | 0.643 | 4.866 | 5.447 | 7.123 | 5.429 | 8.967 |
| 22 | 0.842 | 0.906 | 1.370 | 0.681 | 0.645 | 4.862 | 5.491 | 7.141 | 5.484 | 8.949 |
| 24 | 0.835 | 0.899 | 1.516 | 0.685 | 0.647 | 4.873 | 5.523 | 7.170 | 5.537 | 8.919 |
| 26 | 0.836 | 0.904 | 1.668 | 0.690 | 0.648 | 4.901 | 5.550 | 7.211 | 5.578 | 9.013 |
| 28 | 0.838 | 0.920 | 1.790 | 0.695 | 0.649 | 4.938 | 5.584 | 7.244 | 5.620 | 8.922 |
| 30 | 0.839 | 0.933 | 1.880 | 0.698 | 0.650 | 4.957 | 5.616 | 7.281 | 5.666 | 8.937 |

Table 4: **Averaged Euclidean distance computed for the energy-ratio approach:** For each k and a fixed $1 \leq p \leq 30$, this table reports the average of the Euclidean distances between C^k and $\hat{\mathcal{C}}_{\text{gr}}^{b_{k,p}}$. The appropriate left and right thresholds for a given p are taken from Table 1. For each p , the minimum averaged Euclidean distances in examples 1 to 10 are marked in bold. The selection of these values is based on the four digit decimal points.

approach, we select p such that at least on average, the energy-ratio approach overperforms the energy approach. For this to happen, we have to select p such that the averaged Euclidean distance for the energy-ratio approach is smaller than the one for the energy approach (see the first row of Table 5) for all or majority of the proposed examples. We observe that for $p = 13$, 5 out of 6 examples with additive mix have a Euclidean distance smaller than the one from the energy approach. This value changes to $p = 9$ where 3 out of 4 examples with multiplicative mix follow such property. We now have to select the best p to be larger or equal to these values.

- For a given k , we now compute the absolute difference between the averaged Euclidean distance reported in the first row of Table 5 and the averaged Euclidean distances reported in Table 4. For each p , we then average these differences over all examples with additive or multiplicative mixes. We select the best p to be the one associated with the minimum averaged difference.

Following the above two steps, the best p for the additive and the multiplicative examples are 18 and 16 respectively. We denote the best p by p_* .

Finally using p_* , a comparison between energy, ratio and energy-ratio approaches is shown in Table 5 by reporting the averaged Euclidean distances $\overline{E}_g^{\hat{\mathcal{C}}^k}$, $\overline{E}_r^{\hat{\mathcal{C}}^k}$, $\overline{E}_{gr}^{\hat{\mathcal{C}}^k}$, $\overline{E}_g^{\hat{\mathcal{Y}}^k}$, $\overline{E}_r^{\hat{\mathcal{Y}}^k}$, and $\overline{E}_{gr}^{\hat{\mathcal{Y}}^k}$. For all k , the energy-ratio approach performs better than energy and ratio approaches. This can be concluded in two ways. The first observation is that $\overline{E}_{gr}^{\hat{\mathcal{C}}^k}$ is always smaller than or equal to $\overline{E}_g^{\hat{\mathcal{C}}^k}$ and $\overline{E}_r^{\hat{\mathcal{C}}^k}$. Second observation is that for each k , $\overline{E}_{gr}^{\hat{\mathcal{Y}}^k}$ is closer to the averaged Euclidean distance reported in the seventh column of Table 2 than $\overline{E}_r^{\hat{\mathcal{Y}}^k}$ and $\overline{E}_g^{\hat{\mathcal{Y}}^k}$. The second observation indicates that if we detrend the mixes using the estimates obtained from the energy-ratio approach, the estimated fluctuation is comparable with the actual fluctuation.

6. Examples

In this section, we demonstrate the performance of the EMD trend filtering by applying it to two simulated examples and two real-world examples.

| | k=1 | k=2 | k=3 | k=4 | k=5 | k=6 | k=7 | k=8 | k=9 | k=10 |
|-------------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $\overline{\mathcal{E}}_g^k$ | 0.853 | 1.072 | 2.220 | 0.734 | 0.659 | 5.223 | 5.813 | 7.419 | 5.906 | 9.590 |
| $\overline{\mathcal{E}}_r^{p_*}$ | 1.884 | 1.706 | 1.211 | 0.858 | 4.739 | 5.629 | 15.45 | 15.61 | 15.35 | 16.54 |
| $\overline{\mathcal{E}}_{gr}^{p_*}$ | 0.853 | 0.925 | 1.085 | 0.677 | 0.642 | 4.887 | 5.369 | 7.097 | 5.317 | 9.054 |
| $\overline{\mathcal{Y}}_g^k$ | 7.383 | 5.338 | 1.918 | 2.297 | 17.32 | 12.15 | 49.72 | 49.87 | 49.72 | 49.34 |
| $\overline{\mathcal{Y}}_r^{p_*}$ | 7.154 | 5.230 | 2.705 | 2.252 | 16.86 | 12.14 | 47.29 | 47.45 | 47.34 | 47.08 |
| $\overline{\mathcal{Y}}_{gr}^{p_*}$ | 7.407 | 5.503 | 2.805 | 2.337 | 17.32 | 12.56 | 49.64 | 49.73 | 49.65 | 49.57 |

Table 5: **Comparison between energy, ratio, and energy-ratio approaches by means of p_* :** This table reports in its first to third rows the average of the Euclidean distances between C_k and the b_k th trend estimates from the energy, ratio and energy-ratio approaches respectively. The values reported in rows fourth to sixth are the average of the Euclidean distances between \mathcal{X}^{b_k} and the b_k th trend estimates from the energy, ratio and energy-ratio approaches respectively. For the computations using ratio and energy-ratio approaches we have used $p_* = 18$ for examples 1 to 6 and $p_* = 16$ for examples 7 to 10.

6.1. Simulated examples

In this section, we recall two of the examples introduced in Section 5 for further analysis. The examples are \mathcal{X}^2 which is an additive mix and \mathcal{X}^9 which is a multiplicative mix. We use EMD trend filtering via energy, ratio, and energy-ratio approaches for both examples to estimate the best index using each approach.

The notation used in this section is exactly the same as in Section 5 except that since we only work with one time series of each mix, we replace b_k in the notation with k . For the ratio and energy-ratio approaches, we use $p_* = 18$ and $p_* = 16$ for the additive and multiplicative mixes respectively.

6.1.1. Simulated example 1

Recall \mathbf{Y}^2 and \mathbf{C}^2 from Section 5. Let $\mathbf{Y}^2 = \{Y_0^2, Y_1^2, \dots, Y_{N-1}^2\}$ be a realization of \mathbf{Y}^2 and set the additive mix $\mathcal{X}^2 = \mathbf{Y}^2 + \mathbf{C}^2$ for $N = 2000$. We apply EMD to \mathcal{X}^2 and extract its IMFs and obtain $\mathcal{I} = 10$.

Looking at the energy of the IMFs of \mathcal{X}^2 , we observe that the IMF indices for which $G^i > G^{i-1}$ are $i = \{6, 8, 9, 10\}$. Based on the energy approach, we evaluate $\hat{i}_* = 6$ which is the smallest observed index in this case. Using $\hat{i}_* = 6$, we obtain the trend estimate $\hat{\mathcal{C}}_g^2$ and using the knowledge of \mathbf{C}^2 , we compute the Euclidean distance $E_g^{\hat{\mathcal{C}}^2} = 1.197$. Looking at the RZCN of each IMF on

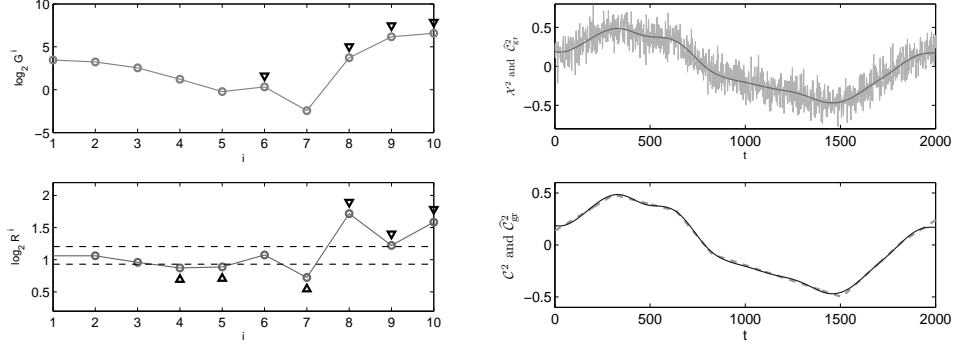


Figure 6: **EMD trend filtering for simulated example 1:** The top-left plot displays the energy approach. The small circles are $\log_2 G^i$ for $1 \leq i \leq 10$ and the small triangles mark those indices $i \geq 2$ where $G^i > G^{i-1}$. Based on only the energy approach, we evaluate $\hat{i}_* = 6$. The bottom-left plot displays the ratio approach. The small circles are $\log_2 R^i$ for $2 \leq i \leq 10$, the dashed lines are T_l and T_r for $p = 18$ in Table 1, and the small triangles mark those indices i where R^i is significantly different from 2. Based on the ratio approach, we evaluate $\hat{i}_* = 4$. The energy-ratio approach evaluates $\hat{i}_* = 8$ which we use to estimate \mathcal{C}^2 . The top-right plot displays \mathcal{X}^2 and $\hat{\mathcal{C}}_{\text{gr}}^2$ while the bottom-right plot displays \mathcal{C}^2 (dashed line) and $\hat{\mathcal{C}}_{\text{gr}}^2$ (solid line).

the other hand, we observe that the IMF indices for which R^i is significantly different from two are $i = \{4, 5, 7, 8, 9, 10\}$. Based on the ratio approach, we evaluate $\hat{i}_* = 4$ which is the smallest observed index in this case as well. Using $\hat{i}_* = 4$, we obtain the trend estimate $\hat{\mathcal{C}}_r^2$ and using the knowledge of \mathcal{C}^2 , we compute the Euclidean distance $E_r^{p*} = 2.191$. Finally, the energy-ratio approach evaluates $\hat{i}_* = 8$ as the smallest common IMF index between the energy and ratio approaches. Using $\hat{i}_* = 8$ we obtain $\hat{\mathcal{C}}_{\text{gr}}^2$ and by means of \mathcal{C}^2 , we obtain $E_{\text{gr}}^{p*} = 0.628$. Figure 6 displays the energy and ratio approaches together with the estimated trend using the energy-ratio approach. It is clear from the above discussion and the left-hand plots in Figure 6 that the energy-ratio approach outperforms the energy and ratio approaches by eliminating the false detections in those methods.

In order to determine the performance of the energy-ratio approach in estimation of the best index we do the following. We first recall Eq. (6) and compute the Euclidean distances $E_{i_{\dagger}}^{\mathcal{C}^2}$ for each $1 \leq i_{\dagger} \leq 10$. For all

| i_{\dagger} | $E_{i_{\dagger}}^{\mathcal{C}^2}$ | $E_{i_{\dagger}}^{\mathcal{Y}^2}$ | $E_{i_{\dagger}}^{\log \mathcal{C}^5}$ | $E_{i_{\dagger}}^{\log \mathcal{Y}^5 }$ |
|---------------|-----------------------------------|-----------------------------------|--|--|
| 1 | 5.493 | 2.05e-15 | 47.38 | 7.47e-14 |
| 2 | 4.381 | 3.320 | 36.56 | 34.81 |
| 3 | 3.214 | 4.644 | 27.89 | 40.34 |
| 4 | 2.191 | 5.173 | 19.71 | 43.90 |
| 5 | 1.575 | 5.315 | 14.22 | 45.47 |
| 6 | 1.197 | 5.378 | 10.37 | 46.41 |
| 7 | 0.691 | 5.483 | 8.254 | 46.90 |
| 8 | 0.628 | 5.496 | 5.047 | 47.17 |
| 9 | 3.800 | 6.606 | 3.682 | 47.18 |
| 10 | 9.746 | 11.187 | 2.492 | 47.24 |
| 11 | - | - | 10.41 | 48.69 |

Table 6: **Search for the best index of \mathcal{X}^2 and \mathcal{X}^9** : The first and second columns of this table report the Euclidean distances $E_{i_{\dagger}}^{\mathcal{C}^2}$ and $E_{i_{\dagger}}^{\mathcal{Y}^2}$ for simulated example 1 where $1 \leq i_{\dagger} \leq 10$. The third and fourth columns of this table reports the Euclidean distances $E_{i_{\dagger}}^{\log \mathcal{C}^5}$ and $E_{i_{\dagger}}^{\log |\mathcal{Y}^5|}$ for simulated example 2 where $1 \leq i_{\dagger} \leq 11$.

$1 \leq i_{\dagger} \leq 10$, we then compute

$$E_{i_{\dagger}}^{\mathcal{Y}^2} \triangleq \left(\sum_{t=0}^{N-1} \left| \mathbf{x}_t^2 - \mathbf{c}_{t,i_{\dagger}}^2 \right|^2 \right)^{1/2}.$$

Using the mix \mathcal{X}^2 , we also compute

$$\begin{aligned} E^{\mathcal{Y}^2} &= \left(\sum_{t=0}^{N-1} \left| \mathbf{x}_t^2 - \mathbf{c}_t^2 \right|^2 \right)^{1/2} \\ &= 5.493. \end{aligned}$$

The Euclidean distances $E_{i_{\dagger}}^{\mathcal{C}^2}$ and $E_{i_{\dagger}}^{\mathcal{Y}^2}$ are reported in the first and second columns of Table 6.

Following the result in the first two columns of Table 6, we can see that since the index $i_{\dagger} = 8$ results in the minimum Euclidean distance $E_{i_{\dagger}}^{\mathcal{C}^2}$, we then conclude that $i_* = 8$. This agrees with what we obtain for $\hat{i}_* = 8$ from the energy-ratio approach. Also, it is interesting to see that $E_{i_{\dagger}}^{\mathcal{Y}^2}$ for $i_{\dagger} = 8$ is the closest value to $E^{\mathcal{Y}^2} = 5.493$. Putting all the above observations together, we conclude that EMD trend filtering has performed excellently in estimation of C_t^2 .

6.1.2. Simulated example 2

Recall \mathbf{Y}^5 and \mathbf{C}^5 from Section 5. Let $\mathbf{Y}^5 = \{Y_0^5, Y_1^5, \dots, Y_{N-1}^5\}$ be a realization of \mathbf{Y}^5 and set the multiplicative mix $\mathbf{X}^9 = \mathbf{C}^5 \mathbf{Y}^5$ for $N = 2000$. We apply EMD to $\log |\mathbf{X}^9|$ and extract its IMFs and obtain $\mathcal{I} = 11$.

Looking at the energy of the IMFs of $\log |\mathbf{X}^9|$, we observe that the IMF indices for which $G^i > \hat{G}^{i-1}$ are $i = \{7, 10, 11\}$. Based on the energy approach, we evaluate $\hat{i}_* = 7$ which is the smallest observed index in this case. Using $\hat{i}_* = 7$, we obtain the log-transformed trend estimate $\log \hat{\mathbf{C}}_{\text{g}}^5$ and using the knowledge of $\log \mathbf{C}^5$, we compute the Euclidean distance $E_{\text{g}}^{\log \hat{\mathbf{C}}_{\text{g}}^5} = 8.254$. Looking at the RZCN of each IMF on the other hand, we observe that the IMF indices for which R^i is significantly different from two are $i = \{2, 8, 9, 10, 11\}$. Based on the ratio approach, we evaluate $\hat{i}_* = 2$ which is the smallest observed index in this case as well. Using $\hat{i}_* = 2$, we obtain the trend estimate $\log \hat{\mathbf{C}}_{\text{r}}^{5, p^*}$ and using the knowledge of $\log \mathbf{C}^5$, we compute the Euclidean distance $E_{\text{r}}^{\log \hat{\mathbf{C}}_{\text{r}}^{5, p^*}} = 36.56$. Finally, the energy-ratio approach evaluates $\hat{i}_* = 10$ as the smallest common IMF index between the energy and ratio approaches. Using $\hat{i}_* = 10$ we obtain $\log \hat{\mathbf{C}}_{\text{gr}}^{5, p^*}$ and by means of $\log \mathbf{C}^5$, we obtain $E_{\text{gr}}^{\log \hat{\mathbf{C}}_{\text{gr}}^{5, p^*}} = 2.492$. Figure 7 displays the energy and ratio approaches together with the estimated log-transformed trend using the energy-ratio approach. It is clear from the above discussion and the left-hand plots in Figure 7 that the energy-ratio approach outperforms the energy and ratio approaches by eliminating the false detections in those methods.

In order to determine the performance of the energy-ratio approach in estimation of the best index we do the following. We first recall Eq. (6) and compute the Euclidean distances $E_{i_{\dagger}}^{\log \mathbf{C}^5}$ For each $1 \leq i_{\dagger} \leq 11$. For all $1 \leq i_{\dagger} \leq 11$, we then compute

$$E_{i_{\dagger}}^{\log |\mathbf{Y}^5|} \triangleq \left(\sum_{t=0}^{N-1} \left| \log |\mathbf{X}_t^9| - \log \mathbf{C}_{t, i_{\dagger}}^5 \right|^2 \right)^{1/2}.$$

Using the mix $\log |\mathbf{X}^9|$, we also compute

$$\begin{aligned} E^{\log |\mathbf{Y}^5|} &= \left(\sum_{t=0}^{N-1} \left| \log |\mathbf{X}_t^9| - \log \mathbf{C}_t^5 \right|^2 \right)^{1/2} \\ &= 47.38. \end{aligned}$$

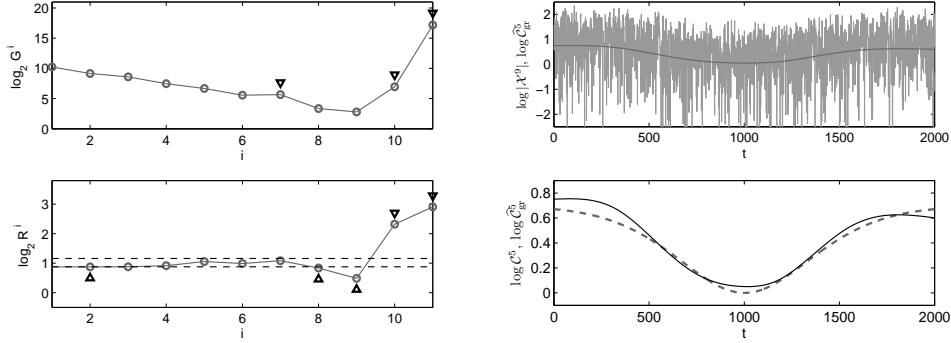


Figure 7: **EMD trend filtering for simulated example 2:** The top left plot displays the energy approach. The small circles are $\log_2 G^i$ for $1 \leq i \leq 11$ and the small triangles mark those indices $i \geq 2$ where $G^i > G^{i-1}$. Based on only the energy approach, we evaluate $\hat{i}_* = 7$. The bottom left plot displays the ratio approach. The small circles are $\log_2 R^i$ for $2 \leq i \leq 11$, the dashed lines are T_l^{\log} and T_r^{\log} for $p = 16$ in Table 1, and the small triangles mark those indices i where R^i is significantly different from two. Based on the ratio approach, we evaluate $\hat{i}_* = 2$. The energy-ratio approach evaluates $\hat{i}_* = 10$ which we use to estimate $\log \mathcal{C}^5$. The top-right plot displays $\log |\mathcal{X}^9|$ and $\log \hat{\mathcal{C}}^5$ while the bottom right plot displays $\log \mathcal{C}^5$ (dashed line) and $\log \hat{\mathcal{C}}^5$ (solid line).

The Euclidean distances $E_{i_{\dagger}}^{\log \mathcal{C}^5}$ and $E_{i_{\dagger}}^{\log |\mathcal{Y}^5|}$ are reported in the third and fourth columns of Table 6.

Following the result in the last two columns of Table 6, we can see that since the index $i_{\dagger} = 10$ results in the minimum Euclidean distance $E_{i_{\dagger}}^{\log \mathcal{C}^5}$, we then conclude that $i_* = 10$. This agrees with what we obtain for $\hat{i}_* = 10$ from the energy-ratio approach. Also, it is interesting to see that $E_{i_{\dagger}}^{\log |\mathcal{Y}^5|}$ for $i_{\dagger} = 10$ is the closest value to $E^{\log |\mathcal{Y}^5|} = 47.38$. Putting all the above observations together, we conclude that EMD trend filtering has performed excellently in estimation of $\log \mathcal{C}_t^5$.

6.2. Real-world examples

In this section we introduce two real world examples. The first example is the monthly mean carbon dioxide (CO_2) data from Mauna Loa and the second example is the Grand Lyon-Vélo'v bicycle rental data from the city of Lyon in France.

For each example, we estimate the underlying trend using the EMD trend filtering.

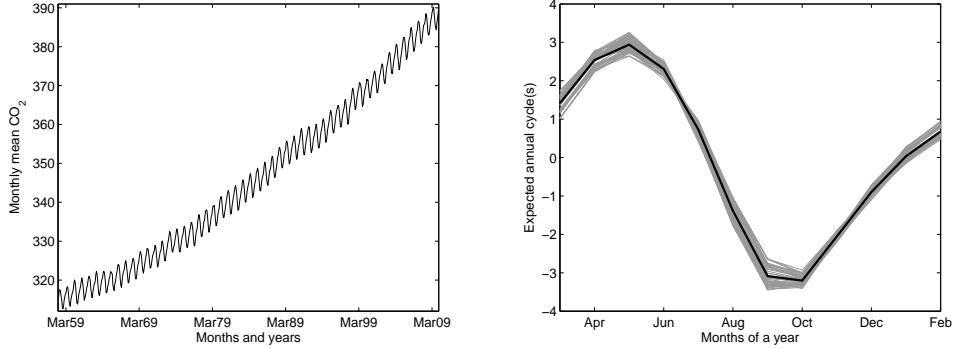


Figure 8: **Monthly mean CO₂ data and the expected annual cycle:** The left-hand plot displays the monthly mean CO₂ data from March 1958 to March 2010. The right-hand plot displays the yearly cycles of the detrended monthly mean CO₂ data using the expected trend together with the expected annual cycle

6.2.1. Monthly mean CO₂ at Mauna Loa

In this section, we analyze the monthly mean CO₂ data collected from March 1958 to March 2010 and measured at Mauna Loa observatory in Hawaii (Available via FTP:ftp://ftp.cmdl.noaa.gov/ccg/co2/trends/co2_mm_mlo.txt. The authors have received permission from Dr. Pieter Tans in order to use this data.) The left-hand plot in Figure 8 displays the monthly mean CO₂ data at Mauna Loa. After removing the averaged seasonal cycle expected in the monthly mean CO₂ data, a trend is obtained. This trend is given at the URL together with the data, and it will serve as a reference for a comparison with the result from EMD trend filtering. For more information on the known seasonal cycle and trend calculation see the URL provided above. The right-hand plot in Figure 8 displays the one year cycles of the monthly mean CO₂ data after removing the expected trend together with their average. We call this average the *expected annual cycle*.

We now use EMD trend filtering for monthly mean CO₂ data in order to estimate its underlying trend. Applying EMD to this data, we obtain $\mathcal{I} = 3$ and following the energy-ratio approach, we evaluate $\hat{i}_* = 3$. The left-hand plot in Figure 9 displays the estimated trend plotted together with the expected trend obtained from removing the seasonal cycle. Since these two trends look very similar, the smaller plot is made to display only a small portion of these trends. It is clear that the estimated trend from the EMD trend filtering is only a smoother version of the expected trend.

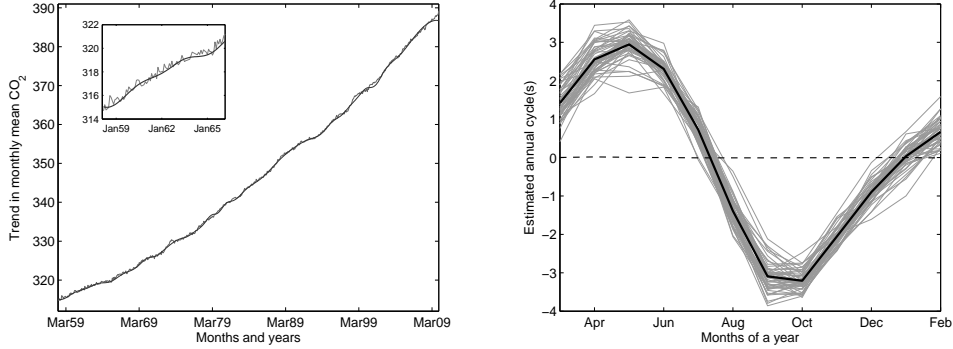


Figure 9: **Estimated trend and annual cycle for the monthly mean CO₂ data:** The left-hand plot displays the expected trend together with the estimated trend using EMD trend filtering. The estimated trend is obtained using $\mathcal{I} = 3$ and $\hat{i}_* = 3$. Since these two trends look very similar, the smaller plot is made to display only a small portion of these trends. The right-hand plot displays the yearly cycles of the detrended monthly mean CO₂ data using the estimated trend together with the estimated annual cycle. The dashed line displays the difference between the expected and estimated annual cycles

After subtracting the estimated trend from the data, we divide the detrended data into one year cycles and then average over all cycles to obtain the estimated annual cycle. The right-hand plot in Figure 9 displays all the one year cycles of the monthly mean CO₂ data after removing the estimated trend together with the estimated annual cycle. The dashed line in Figure 9 displays the difference between the expected and the estimated annual cycles. This difference confirms the strong similarities between the two annual cycles.

6.2.2. Grand Lyon-Vélo'v

In this section we analyze the data from Vélo'v, the community shared bicycle program that started in Lyon in May 2005 (For more information, see <http://www.velov.grandlyon.com>.) The program Vélo'v is a major initiative in public transportation, in which bicycles are proposed to rental by anyone at fully automated stations in many places all over the city, to be returned at any other station. Such a community shared system offers both a new and versatile option of public transportation, and a way to look into the movements of people across the city. In order to understand the dynamics of this system, a question is to estimate and model the evolution in time of the number of rentals made throughout the city (Borgnat et al., 2009). The

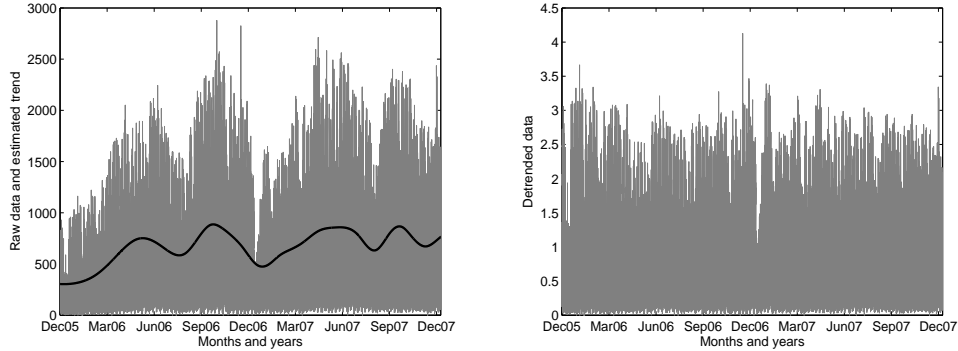


Figure 10: **Vélo’v raw and detrended data:** The left-hand plot displays the raw Vélo’v data together with the estimated trend using EMD trend filtering. The estimated trend is obtained using $\mathcal{I} = 12$ and $\hat{i}_* = 10$. The right-hand plot displays the detrended Vélo’v data by dividing the raw data with the estimated trend

left-hand plot in Figure 10 displays the raw data which is the number of hourly rentals for two years of activity of the Vélo’v system (from December 2005 to December 2007). (The authors would like to thank JCDecaux for providing access to this data.)

The number of rentals is both driven by cyclic patterns over the day (more activity during day, mostly at specific rush hours, than during night) and the week (more activities during week-days than week-ends) to which are superimposed fluctuations due to external contingencies (such as rain, holidays) and a general multiplicative trend over the months (Borgnat et al., 2009). We apply EMD trend filtering to this data in order to estimate the underlying multiplicative trend. We obtain $\mathcal{I} = 12$ and using the energy-ratio approach we evaluate $\hat{i}_* = 10$ which we use to estimate the trend. The estimated trend is displayed in the left-hand plot in Figure 10 where superimposed over the raw data.

This trend is meaningful for the data, and can be related to, and explained by, two effects: (i) the system was expanded in 2005 and 2006 at the same time it was already in exploitation, hence, there is a long-term increase of the hourly rentals over the two years of data, (ii) because of seasonal effects, the use of Vélo’v is smaller during winter, and also during the main summer holidays; this causes several drops of the trend, during winter times and during summer holidays.

Using the estimated trend, detrended Vélo’v data are obtained by divid-

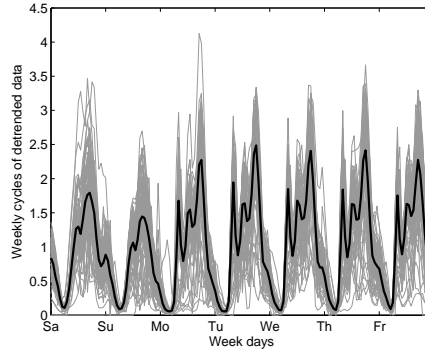


Figure 11: **Weekly cycle of detrended Vélo'v data:** This figure displays the weekly detrended Vélo'v data together with its average

ing the number of hourly locations by the estimated trend. This is displayed in the right-hand plot in Figure 10. The detrended data is, visually, more stationary than the raw data. This allows a good estimation of the cyclic pattern over the week of the number of hourly rentals. Figure 11 displays the weekly cycles of the Vélo'v data after removing the estimated trend, and the average over all the weeks.

This estimate of the average use of Vélo'v bicycles as a function of the time of the week, is meaningful in that it reveals the main features of the Vélo'v activity: during week-days, there are three sharp peaks of use in the morning, at noon and at the end of the afternoon; during week-end, there is a small peak at noon, and a smooth and large peak during the afternoon.

Finally, let us note that here the multiplicative trend estimation procedure was applied to a case where the underlying process that the trend multiplies to is not actually a broadband process: it is more specifically a periodic process (with clear periods of one week and one day) with added fluctuations. Nevertheless, the procedure is able to find the relevant multiplicative trend describing the evolutions at the scale of the seasons, and that is used to detrend the data. This is believed to be due mostly to the fact that fluctuations have typical periodic scales (one day or one week) which are much smaller than the typical scale of evolution (several months) of the trend, making of this scale separation a prerequisite that might be more important than the existence of a broadband spectrum in a stricter sense.

7. Conclusion

An automated method has been proposed to filter the trend in a time series, whose principle is to extract the lowest frequency intrinsic mode functions (IMFs) via empirical mode decomposition (EMD). The core of the method is to decide which IMFs belong to the trend, on the rationale that a trend causes both a departure of the ratio of zero crossing numbers from 2, and an increase of the energy contained in the low-frequency IMFs, as compared to the expected behavior of broadband processes. Combining both criteria, the procedure was shown to work well on several examples. Also, it works both for the trend filtering of additive and multiplicative trends. Let us emphasize that the approach is fully data-driven (as is EMD) and, besides the parameters of the decomposition itself, the trend filtering method has only one parameter, the level of significance p for the test of estimation of what is the trend. From the numerical experiments that are reported in this paper, $p = 18$ and $p = 16$ are optimum parameters for relevant estimations of the additive and multiplicative trends respectively. The method is not based on a specific model of the data and, once the best p fixed, has no other free parameter.

Many numerical examples were reported to illustrate the robustness of this EMD trend filtering and its potential interests have been further illustrated on two real-world examples: the CO₂ data which displays an additive trend, and the Vélo'v data which shows a multiplicative trend. In both cases, filtering of the trends allows us to propose an estimation of the cycle inside the data (annual cycle for the CO₂ data, weekly and daily cycles for the Vélo'v data) that compares favorably to existing methods both for the extracted trends and estimated cycles. A strength of the method is that it works, even if the fluctuations above the trend do not follow exactly a priori behaviors for the fluctuations that were used to design empirically the test (displaying for instance oscillatory behaviors more than the assumed broadband behavior.) This is related to its character as a fully data-driven and model-free approach.

A perspective of this work would be to go beyond trend-filtering and use the same type of approach to group together IMFs obtained by EMD in several signals describing a trend, then the major cycles, and finally the rapid fluctuations. This would be an interesting asset for the model-free decomposition of processes.

8. Acknowledgment

The authors would like to thank the anonymous reviewers and Associate Editor for their helpful comments and suggestions. Most of the work reported here was completed during the postdoctoral stay of Azadeh Moghtaderi at ENS Lyon, which was supported by ANR Grant ANR-07-BLAN-0191-01 StaRAC.

References

- Alexandrov, T., Bianconcini, S., Dagum, E. B., Maass, P., McElroy, T., March 2008. A review of some modern approaches to the problem of trend extraction. Research Report Series, *Statistics #2008-3*, Statistical Research Division, U.S. Census Bureau, Washington. To appear in *Econometric Reviews*.
- Borgnat, P., Abry, P., Flandrin, P., Rouquier, J.-B., 2009. Studying Lyon's Vélo'v: A statistical cyclic model. In: *Proceedings of ECCS'09 (European Conference of Complex Systems)*. Warwick, United Kingdom.
- Chatfield, C., 1996. *The analysis of time series: An introduction*. Chapman and Hall/CRC.
- Embrechts, P., Maejima, M., 2002. *Selfsimilar Processes*. Princeton University Press, Princeton, NJ.
- Flandrin, P., Gonçalves, P., 2004. Empirical mode decompositions as data-driven wavelet-like expansions. *International Journal of Wavelets, Multiresolution and Information Proceedings* 2 (4), 477–496.
- Flandrin, P., Gonçalves, P., Rilling, G., 2004a. Detrending and denoising with empirical mode decompositions. In: *Proceedings of EUSIPCO-04*. Vienna, Austria, pp. 1581–1584.
- Flandrin, P., Rilling, G., Gonçalves, P., 2004b. Empirical mode decomposition as a filter bank. *IEEE Signal Processing Letters* 11 (2), 112–114.
- Ghil, M., Vautard, R., 1992. Interdecadal oscillations and the warming trend in global temperature time series. *Nature* 58, 95–126.

- Henderson, R., 1916. Note on graduation by adjusted average. *Transactions on the Actuarial Society of America* 17, 43–48.
- Hodrick, R. J., Prescott, E. C., 1997. Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit, and 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., Liu, H. H., 1998. The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 454, 903–995.
- Huang, N. E., Wu, M.-L., Long, S., Shen, S., Qu, W., Gloersen, P., Fan, K., 2003. A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. *Proceedings of the Royal Society of London A* 459 (2037), 2317–2345.
- Pollock, D. S. G., 2006. Wiener–Kolmogorov filtering frequency-selective filtering and polynomial regression. *Econometric Theory* 23, 71–83.
- Rilling, G., Flandrin, P., Gonçalves, P., 2005. Empirical mode decomposition, fractional Gaussian noise, and Hurst exponent estimation. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 489–492.
- Vautard, R., Ghil, M., 1989. Singular-spectrum analysis in nonlinear dynamics with applications to paleoclimatic time series. *Physica D* 35, 395–424.
- Vautard, R., Yiou, P., Ghil, M., 1991. Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica D* 350, 324–327.
- Wu, Z., Huang, N. E., 2004. A study of the characteristics of white noise using the empirical mode decomposition method. *Proceedings of the Royal Society of London A* 460, 1597–1611.
- Wu, Z., Huang, N. E., Long, S., Peng, C.-K., 2007. On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proceedings of the National Academy of Sciences* 4 (38), 14889–14894.