Numerische
Mathematik

Numerical analysis for iterative filtering with new efficient implementations based on FFT

Antonio Cicone[1] · Haomin Zhou[2]

Abstract

The development of methods able to extract hidden features from non-stationary and non-linear signals in a fast and reliable way is of high importance in many research fields. In this work we tackle the problem of further analyzing the convergence of the Iterative Filtering method both in a continuous and a discrete setting in order to provide a comprehensive analysis of its behavior. Based on these results we provide a new efficient implementation of Iterative Filtering algorithm, called Fast Iterative Filtering, which reduces the original iterative algorithm computational complexity by utilizing, in a nontrivial way, Fast Fourier Transform in the computations.

1 Introduction

Non-stationary and non-linear signals are ubiquitous in applications. Standard techniques like Fourier or wavelet transform prove to be inadequate in capturing properly their hidden features [10]. For this reason Huang et al. proposed in 1998 a new kind of algorithm, called Empirical Mode Decomposition (EMD) [41], which allows to unravel the hidden features of a non-stationary signal $s(x)$, $x \in \mathbb{R}$, by iteratively decomposing it into a finite sequence of simple components, called Intrinsic Mode Functions (IMFs). Such IMFs fulfill two properties: the number of extrema and the number of zero crossings must either equal or differ at most by one; considering upper

✉ Antonio Cicone
  antonio.cicone@univaq.it

  Haomin Zhou
  hmzhou@math.gatech.edu

1 DISIM, Università degli Studi dell'Aquila, Via Vetoio n. 1, 67100 L'Aquila, Italy

2 School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA

and lower envelopes connecting respectively all the local maxima and minima of the function, their mean has to be zero at any point.

The EMD, over the last two decades, has attracted the attention of many researchers from a wide variety of applied fields. We mention here geophysical studies [6,42, 70], with applications in Seismology (data denoising and/or detrending [5,37,73], pre-seismic signal analysis [3,7,78], earthquake-induced co/post-seismic anomalies analysis [8]), Exploration Seismology (for improving signal-to-noise ratio in seismic data processing routines [4,9] or for seismic interpretation [76]), Geomagnetism [43, 63,91], Engineering Seismology (mainly for analysing ground motion data [39,52,87, 92,93]), climate, atmospheric and oceanographic sciences [21–25,45]. Its use is also common in Physics (for data analysis [32,40,56,71], data denoising and/or detrending [31,77], to assess causal relationships between two time series [86], or to extract information on multiple time scales [18]); Medicine and Biology [16,19,29,30,35, 49,82,85,97]; Engineering [2,46,51,59,65,81]; Economics and Finance [89,94,95]; Computer vision [1,47,84], just to mention a few. The impact of this technique is testified also by the high number of citations[1] that the original paper by Huang et al. [41] has received so far.

However, while the EMD method proves to be extremely powerful in extracting simple components from a given signal, it is unstable to perturbations [83] and susceptible to mode splitting and mode mixing [75]. These are the reasons why the Ensemble Empirical Mode Decomposition (EEMD) method [83] first, and then several alternative noise-assisted EMD-based methods (e.g. the complementary EEMD [88], the complete EEMD [72], the partly EEMD [96], the noise assisted multivariate EMD (NA-MEMD) [74], and fast multivariate EMD (FMEMD) [44]) have been proposed. While these newly developed methods are based on EMD, they all address the so called mode mixing problem and guarantee the stability of the decomposition with respect to noise [38]. However, mode splitting is still an open problem for all these methods and, more importantly, their mathematical analysis is by no means complete [38].

For all these reasons many alternative methods have been proposed recently, like, for instance, the sparse TF representation [33,34], the Geometric mode decomposition [90], the Empirical wavelet transform [28], the Variational mode decomposition [20], and similar techniques [54,62,64].

All of these methods are based on optimization. The only alternative technique proposed so far in the literature which is based on iterations, like the EMD-based methods, is the so called Iterative Filtering (IF) method proposed by Lin et al. in [50]. In recent years, this alternative iterative method has been used in a wide variety of applied fields like, for instance, in Engineering [48,55,65], Chemistry [15], Economy [60], Medicine [66], and Physics [26,27,53,57,58,61,67,68].

The mathematical analysis of IF has been tackled by several authors in the last few years [11,12,14,36,79,80], even for 2D or higher dimensional signals [17]. However several problems regarding this technique are still unsolved. In particular it is not yet clear how the stopping criterion used to discontinue the calculations of the IF algorithm

---

[1] The original work by Huang et al. [41] as received so far, by itself, more than 14600 unique citations, according to Scopus.

influences the decomposition. Furthermore all the aforementioned analyses focused on the convergence of IF when applied to the extraction of a single IMF from a given signal, the so called inner loop. Regarding the decomposition of all the IMFs contained in a signal, which is related to the outer loop convergence and potential finiteness of the decomposition itself, nothing has been said so far. In this work we further analyze the IF technique addressing these and other questions.

The rest of this work is organized as follows: in Sect. 2 we review the details and properties of the method in the continuous setting and we provide new results regarding its inner loop convergence in presence of a stopping criterion as well as the outer loop convergence and finiteness. In Sect. 3 we address the convergence analysis in the discrete setting for both the inner and outer loop of the algorithm. Based on these results in Sect. 4 we propose a new algorithm, called Fast Iterative Filtering (FIF), which produces the very same results as IF, but at consistently lower computational cost.

2 IF algorithm in the continuous setting

The key idea behind this decomposition technique is separating simple oscillatory components contained in a signal $s(x)$, $x \in \mathbb{R}$, the so called IMFs, by approximating the moving average of s and subtracting it from s itself. The approximated moving average is computed by convolution of s with a window/filter function w.

Definition 1 A filter/window w is a nonnegative and even function in C^0 ([$-L$, L]), $L > 0$, and such that $\int_{\mathbb{R}} w(z)\mathrm{d}z = \int_{-L}^{L} w(z)\mathrm{d}z = 1$.

We point out that the idea of iteratively subtracting the moving average comes from the Empirical Mode Decomposition (EMD) method [41] where the moving average was computed as a local average between an envelope connecting the maxima and one connecting the minima of the signal under study. The use of envelopes in an iterative way is the reason why the EMD algorithm is still lacking a rigorous mathematical framework.

The pseudocode of IF is given in Algorithm 1 where $w_m(t)$ is a nonnegative and compactly supported window/filter with area equal to one and support in [$-l_m$, l_m], where l_m is called filter length and represents the half support length.

The IF algorithm contains two loops: the inner and the outer loop, the second and first while loop in the pseudocode respectively. The former captures a single IMF, while the latter produces all the IMFs embedded in a signal.

Assuming $s_1 = s$, the key step of the algorithm consists in computing the moving average of s_m as

$$\mathcal{L}_m(s_m)(x) = \int_{-l_m}^{l_m} s_m(x + t) w_m(t) \mathrm{d}t, \tag{1}$$

which represents the convolution of the signal itself with the window/filter $w_m(t)$.

The moving average is then subtracted from s_m to capture the fluctuation part as

$$\mathcal{M}_m(s_m) = s_m - \mathcal{L}_m(s_m) = s_{m+1} \tag{2}$$

Algorithm 1 Iterative Filtering IMF = IF(s)

IMF = {}
while the number of extrema of $s \geq 2$ **do**
    $s_1 = s$
    while the stopping criterion is not satisfied **do**
        compute the filter length l_m for $s_m(x)$
        $s_{m+1}(x) = s_m(x) - \int_{-l_m}^{l_m} s_m(x+t)w_m(t)\mathrm{d}t$
        $m = m + 1$
    end while
    IMF = IMF $\cup \{s_m\}$
    $s = s - s_m$
end while
IMF = IMF $\cup \{s\}$

The first IMF, IMF_1, is computed repeating iteratively this procedure on the signal $s_m, m \in \mathbb{N}$, until a stopping criterion is satisfied, as described in the following section.

To produce the 2-nd IMF we apply the same procedure to the remainder signal $r = s - \text{IMF}_1$. Subsequent IMFs are produced iterating the previous steps.

The algorithm stops when r becomes a trend signal, meaning it has at most one local extremum.

We observe that, even thought the algorithm allows potentially to recompute the filter length l_m at every step of each inner loop, in practice we always compute the filter length only at the first step of an inner loop and then we keep it constant throughout the subsequent iterations. Hence $l_m = l_1 = l$ for every $m \geq 1$.

Following [50], one possible way of computing the filter length l is given by the formula

$$l := 2 \left\lfloor \nu \frac{N}{k} \right\rfloor \tag{3}$$

where N is the total number of sample points of a signal $s(x)$, k is the number of its extreme points, ν is a tuning parameter usually fixed around 1.6, and $\lfloor \cdot \rfloor$ rounds a positive number to the nearest integer closer to zero. In doing so we are computing some sort of average highest frequency contained in s.

Another possible way could be the calculation of the Fourier spectrum of s and the identification of its highest frequency peak. The filter length l can be chosen to be proportional to the reciprocal of this value.

The computation of the filter length l is an important step of the IF technique. Clearly, l is strictly positive and, more importantly, it is based solely on the signal itself. This last property makes the method nonlinear.

In fact, if we consider two signals p and q where $p \neq q$, assuming IMFs($\bullet$) represent the decomposition of a signal into IMFs by IF, the fact that we choose the half support length based on the signal itself implies that in general

$$\text{IMFs}(p+q) \neq \text{IMFs}(p) + \text{IMFs}(q)$$

Regarding the convergence analysis of the Iterative Filtering inner loop we recall here the following theorem

Theorem 1 (Convergence of the Iterative Filtering method [14,36]) *Given the filter function $w(t), t \in [-l, l]$ be L^2, symmetric, nonnegative, $\int_{-l}^{l} w(t)dt = 1$ and let $s(x) \in L^2(\mathbb{R})$.*
If $|1 - \widehat{w}(\xi)| < 1$ or $\widehat{w}(\xi) = 0$, where $\widehat{w}(\xi)$ is the Fourier transform of w computed at the frequency ξ,
  Then $\{\mathcal{M}^m(s)\}$ converges and

$$IMF_1 = \lim_{m \to \infty} \mathcal{M}^m(s)(x) = \int_{-\infty}^{\infty} \widehat{s}(\xi)\chi_{\{\xi \in \mathbb{R}|\widehat{w}(\xi)=0\}}e^{2\pi i \xi x}d\xi \qquad (4)$$

where $\chi_{\{A\}}$ represents the indicator function of the subset A of $\mathbb{R}$.

We observe here that given $h : [-\frac{l}{4}, \frac{l}{4}] \to \mathbb{R}, z \mapsto h(z)$, nonnegative, symmetric, with $\int_{\mathbb{R}} h(z)dz = \int_{-\frac{l}{4}}^{\frac{l}{4}} h(z)dz = 1$, if we construct the window w_1 as the convolution of h with itself and we fix $w_m = w_1$ throughout all the steps m of an inner loop, then the method converges for sure to the limit function (4) which depends only on the shape of the filter function chosen and the support length selected by the method [13,14].

In general we can assume that the filter functions $w_m(u)$ are defined as some scaling of an a priori fixed filter shape $w : [-1, 1] \to \mathbb{R}$. In particular we define the scaling function

$$g_m : [-1, 1] \to [-l_m, l_m], \qquad t \mapsto u = g_m(t), \qquad (5)$$

where g_m is assumed to be invertible and monotone, such that $w_m(u) = C_m w(g_m^{-1}(u)) = C_m w(t)$, where $t = g_m^{-1}(u)$, $u = g_m(t)$ and C_m is a scaling coefficient which is required to ensure that $\int_{\mathbb{R}} w_m(u)du = \int_{-l_m}^{l_m} w_m(u)du = 1$.

We point out that, from Theorem 1, it is clear that the choice of the filter function $w(t)$ can have an important influence in the decomposition produced by the IF algorithm. This is an important direction of research, which has never been explored in the literature so far, to best of our knowledge. However its analysis is out of the scope of this work. We plan to study it in a future work.

Regarding the computation of the scaling coefficient C_m, from the observation that $du = g_m'(t)dt$, it follows that

$$\int_{-l_m}^{l_m} w_m(u)du = \int_{-l_m}^{l_m} C_m w(g_m^{-1}(u))du = C_m \int_{-1}^{1} w(t)|g_m'(t)|dt \qquad (6)$$

hence

$$C_m = \frac{1}{\int_{-1}^{1} w(t)|g_m'(t)|dt} \qquad (7)$$

and

$$w_m(u) = C_m w(g_m^{-1}(u)) = \frac{w(g_m^{-1}(u))}{\int_{-1}^{1} w(t)|g_m'(t)|dt} \qquad (8)$$

As an example of a scaling function we can consider, for instance, linear or quadratic scalings: $g_m(t) = l_m t$ and $g_m(t) = l_m t^2$ respectively.

In the case of linear scaling we have that $g_m^{-1}(u) = \frac{u}{l_m}$, $g_m'(t) = l_m \geq 0$, for every $t \in \mathbb{R}$, and $C_m = \frac{1}{l_m}$. Hence

$$w_m(u) = \frac{w\left(\frac{u}{l_m}\right)}{l_m}. \tag{9}$$

2.1 IF inner loop convergence in presence of a stopping criterion

In Algorithm 1 the inner loop has to be iterated infinitely many times. In numerical computations, however, some stopping criterion has to be introduced. One possible stopping criterion follows from the solution of

Problem 1 *For a given $\delta > 0$ we want to find the value $N_0 \in \mathbb{N}$ such that*

$$\|\mathcal{M}^N(s)(x) - \mathcal{M}^{N+1}(s)(x)\|_{L^2} < \delta \quad \forall N \geq N_0$$

Applying the aforementioned stopping criterion, the inner loop of Algorithm 1 converges in finite steps to an IMF whose explicit form is given in the following theorem where $\widehat{s}(\xi)$ represents the Fourier transform of s at frequency ξ.

Theorem 2 *Given $s \in L^2(\mathbb{R})$ and w obtained as the convolution $\widetilde{w} * \widetilde{w}$, where $\widetilde{w}$ is a filter/window, Definition 1, and fixed $\delta > 0$.*
Then, for the minimum $N_0 \in \mathbb{N}$ such that the following inequality holds true

$$\frac{N_0^{N_0}}{(N_0 + 1)^{N_0+1}} < \frac{\delta}{\|\widehat{s}(\xi)\|_{L^2}} \quad \forall \xi \in \mathbb{R} \tag{10}$$

we have that $\left\|\mathcal{M}^N(s)(x) - \mathcal{M}^{N+1}(s)(x)\right\|_{L^2} < \delta \ \ \forall N \geq N_0$ and the first IMF is given by

$$IMF_1^{SC} = \mathcal{M}^N(s)(x) = \int_{\mathbb{R}} (1 - \widehat{w}(\xi))^N \widehat{s}(\xi) e^{2\pi i \xi x} d\xi \quad \forall N \geq N_0 \tag{11}$$

Proof From the hypotheses on the filter w it follows that its Fourier transform is in the interval $[0, 1]$, see [14]. Furthermore from the linearity of the Fourier transform it follows that

$$\widehat{\mathcal{M}^N(s)(x)}(\xi) = (1 - \widehat{w}(\xi))^N \widehat{s}(\xi) = \begin{cases} \widehat{s}(\xi) & \text{if } \widehat{w}(\xi) = 0 \\ (1 - \widehat{w}(\xi))^N \widehat{s}(\xi) & \text{if } |1 - \widehat{w}(\xi)| < 1 \end{cases}$$

since the Fourier Transform is a unitary operator, by the Parseval's Theorem, it follows that

$$\left\|\mathcal{M}^N(s)(x) - \mathcal{M}^{N+1}(s)(x)\right\|_{L^2} = \left\|\widehat{\mathcal{M}^N(s)(x)}(\xi) - \widehat{\mathcal{M}^{N+1}(s)(x)}(\xi)\right\|_{L^2}$$

$$= \left\|(1 - \widehat{w}(\xi))^N \left[1 - (1 - \widehat{w}(\xi))\right]\widehat{s}(\xi)\right\|_{L^2}$$

$$= \left\| (1 - \widehat{w}(\xi))^N \, \widehat{w}(\xi) \widehat{s}(\xi) \right\|_{L^2}$$

We point out that this formula can also be interpreted as the L^2-norm of the moving average of $\mathcal{M}^N$ which is given by the convolution $\mathcal{M}^N * w$.

For a fixed N we can compute the maximum of the function $(1 - \widehat{w}(\xi))^N \widehat{w}$, for $\widehat{w} \in [0, \, 1]$, that is attained for $\widehat{w}(\xi) = \frac{1}{N+1}$. Therefore

$$\left\| (1 - \widehat{w}(\xi))^N \widehat{w}(\xi) \widehat{s}(\xi) \right\|_{L^2} \leq \left\| \left(1 - \frac{1}{N+1} \right)^N \frac{1}{N+1} \widehat{s}(\xi) \right\|_{L^2}$$

$$= \left\| \frac{N^N}{(N+1)^{N+1}} \widehat{s}(\xi) \right\|_{L^2} < \delta$$

Hence we consider the smallest $N_0 \in \mathbb{N}$ such that

$$\frac{N_0^{N_0}}{(N_0+1)^{N_0+1}} < \frac{\delta}{\|\widehat{s}(\xi)\|_{L^2}}.$$

□

Equation (11) provides a valuable insight on how the implemented algorithm is actually decomposing a signal into IMFs. We recall that without any stopping criterion each IMF of a signal s is given by the inverse Fourier transform of $\widehat{s}$ computed at the frequencies corresponding to zeros of $\widehat{w}$, as stated in (4).

Therefore, from the observation that $\widehat{w}$ is a function not compactly supported and with isolated zeros, the IMFs produced with IF are given by the summation of pure and well separated tones.

Whereas, when we enforce a stopping criterion, we end up producing IMFs containing a much richer spectrum. In fact from (11) we discover that an IMF is now given by the inverse Fourier transform of $\widehat{s}$ computed at every possible frequency in $\mathbb{R}$, each multiplied by the coefficient $(1 - \widehat{w}(\xi))^N$. Since, by construction, $0 \leq \widehat{w}(\xi) \leq 1$, $\forall \xi \in \mathbb{R}$, then $(1 - \widehat{w}(\xi))^N$ is equal to 1 if and only if $\widehat{w}(\xi) = 0$, whereas for all the other frequencies it is smaller than 1 and it tends to zero as N grows. The $(1 - \widehat{w}(\xi))^N$ quantity represents in practice the percentage with which each frequency is contained in the reconstruction of an IMF from the Fourier transform of the original signal. The higher is the number of iterations N the narrower are the intervals of frequencies that are almost completely captured in each IMF. And as $N \to \infty$ such intervals coalesce into isolated points corresponding to the zeros of $\widehat{w}$.

2.1.1 Convergence with a threshold

We start recalling a few properties regarding the filter functions w. Assuming $w(x)$, $x \in \mathbb{R}$, is a filter function supported on $(-1, \, 1)$, if we use the linear scaling described in (9), then we can construct

$$w^a(x) = \frac{1}{a} w \left(\frac{x}{a} \right) \tag{12}$$

where $w^a(x)$ is supported on $(-a, \ a)$.

If we define $\widehat{w}(\xi) = \int_{-\infty}^{+\infty} w(x) e^{-i\xi x 2\pi} \, dx$, then

$$\widehat{w^a}(\xi) = \int_{-\infty}^{+\infty} \frac{1}{a} w\left(\frac{x}{a}\right) e^{-i\xi \frac{x}{a} a 2\pi} \, dx = \widehat{w}(a\xi) \tag{13}$$

Therefore, if ξ_0 is a root of $\widehat{w}(\xi) = 0$, then $\frac{\xi_0}{a}$ is a root of $\widehat{w^a}(\xi) = 0$ because $\widehat{w^a}\left(\frac{\xi_0}{a}\right) = \widehat{w}\left(a\frac{\xi_0}{a}\right) = \widehat{w}(\xi_0) = 0$.

We remind that, since w are compactly supported functions, their Fourier transform are defined on $\mathbb{R}$ and they have zeros which are isolated points.

Given $0 < \gamma < 1$, we identify the set

$$I_{w,\gamma,N} = \left\{ \xi \in \mathbb{R} \ : \ \widehat{w}(\xi) \leq 1 - \sqrt[N]{1-\gamma} \right\}. \tag{14}$$

As $N \to \infty$ the quantity $1 - \sqrt[N]{1-\gamma} \to 0$, therefore $I_{w,\gamma,N}$ coalesces into isolated points corresponding to the zeros of $\widehat{w}$.

If we consider filters like the Fokker-Planck filters [14] or any filter with smooth finite support properties we must have that, for a fixed $N \in \mathbb{N}$ and $\gamma > 0$, there exists $\varXi_0 > 0$ such that

$$\widehat{w}(\xi) \leq 1 - \sqrt[N]{1-\gamma} < 1 \qquad \text{for all } |\xi| \geq \varXi_0 \tag{15}$$

In fact, since $\int |w(x)|^2 dx < +\infty$ with $w(x)$ smooth function, then $\int |\widehat{w}(\xi)|^2 d\xi < +\infty$ which implies that $\widehat{w}(\xi)$ decays as $|\xi| \to \infty$.

So for a filter w with smooth finite support properties the set $I_{w,\gamma,N}$ is made up of a finite number of disjoint compact intervals, containing zeros of $\widehat{w}$, together with the intervals $(-\infty, \ -\varXi_0]$ and $[\varXi_0, \ \infty)$.

Furthermore if we scale these filters using a linear scaling with coefficient $a > 1$ it follows from the previous observations that $\varXi_0 \to 0$ and, as a consequence, $I_{w,\gamma,N}$ converges to $\mathbb{R}\backslash\{0\}$.

As an example of a compactly supported filter we can consider the triangular filter function

$$w(x) = \begin{cases} \frac{1}{L} - \frac{1}{L^2}|x| & \text{for } |x| \leq L \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

whose Fourier transform is

$$\widehat{w}(\xi) = \frac{1}{L} \frac{\sin^2 (L\pi\xi)}{(\pi\xi)^2}. \tag{17}$$

The triangular filter and its Fourier transform are depicted in Fig. 1.

Given the threshold value $1 - \sqrt[N]{1-\gamma}$ depicted in the right panel of Fig. 1 and the triangular filter (16) with $L = 1$, the set $I_{w,\gamma,N}$ is made up of four intervals: two compactly supported and centered around $1/2$ and $-1/2$, and other two starting around 0.8 and -0.8 and ending at infinity and minus infinity, respectively.
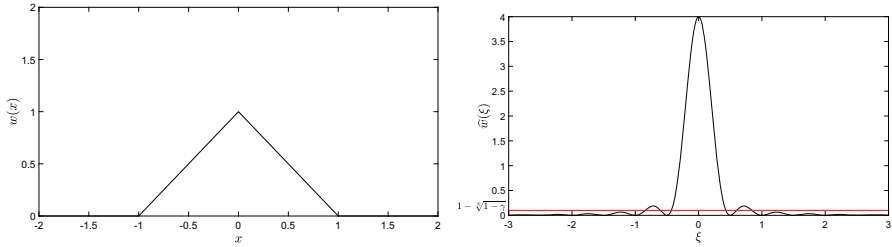
Fig. 1 Left panel, triangular filter (16) with $L = 1$. Right panel, in black the Fourier transform (17) and in red the threshold value $1 - \sqrt[N]{1-\gamma}$

We can use the threshold value $1 - \sqrt[N]{1-\gamma}$ in the computation of an IMF as follows: given (11), whenever $(1 - \widehat{w}(\xi))^N \geq 1 - \gamma$, we substitute $\widehat{w}(\xi)$ with zero. This is equivalent to setting $\widehat{w}(\xi) = 0$ whenever $\xi \in I_{w,\gamma,N}$.

Therefore, using the previously described thresholding and based on Theorem 2, Algorithm 1 converges to

$$\text{IMF}_1^{\text{TH}} = \int_{\mathbb{R}\setminus I_{w,\gamma,N}} (1 - \widehat{w}(\xi))^N \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi + \int_{I_{w,\gamma,N}} \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi \quad \forall N \geq N_0 \tag{18}$$

where $I_{w,\gamma,N}$ is defined in (14).

We are now ready to prove the following

Proposition 1 *Assuming that all the hypotheses of Theorem 2 are fulfilled, then for every $\epsilon > 0$ there exist a stopping criterion value $\delta > 0$ and a threshold $0 < \gamma < 1$ such that*

$$\left\| IMF_1 - IMF_1^{TH} \right\| \leq \frac{\epsilon}{2}, \qquad \left\| IMF_1^{TH} - IMF_1^{SC} \right\| \leq \frac{\epsilon}{2} \tag{19}$$

and

$$\left\| IMF_1 - IMF_1^{SC} \right\| \leq \epsilon \tag{20}$$

where IMF_1, IMF_1^{SC}, and IMF_1^{TH} are defined in (4), (11), and (18) respectively.

Proof First of all we have that

$$\left\| \text{IMF}_1 - \text{IMF}_1^{\text{SC}} \right\| \leq \left\| \text{IMF}_1 - \text{IMF}_1^{\text{TH}} \right\| + \left\| \text{IMF}_1^{\text{TH}} - \text{IMF}_1^{\text{SC}} \right\|$$

where

$$\left\| \text{IMF}_1 - \text{IMF}_1^{\text{TH}} \right\|$$
$$\leq \left\| \int_{\mathbb{R}} \widehat{s}(\xi) \chi_{\{\xi \in \mathbb{R} \ | \ \widehat{w}(\xi)=0\}} e^{2\pi i \xi x} \mathrm{d}\xi \right.$$
$$\left. - \int_{\mathbb{R}\setminus I_{w,\gamma,N}} (1 - \widehat{w}(\xi))^N \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi - \int_{I_{w,\gamma,N}} \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi \right\|$$

$$\leq \left\| \int_{\mathbb{R}\setminus I_{w,\gamma,N}} (1 - \widehat{w}(\xi))^N \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi \right\| + \left\| \int_{I_{w,\gamma,N}\setminus\{\xi\in\mathbb{R} \mid \widehat{w}(\xi)=0\}} \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi \right\| \tag{21}$$

and

$$\left\| \mathrm{IMF}_1^{\mathrm{TH}} - \mathrm{IMF}_1^{\mathrm{SC}} \right\| \leq \left\| \int_{\mathbb{R}\setminus I_{w,\gamma,N}} (1 - \widehat{w}(\xi))^N \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi \right.$$

$$\left. + \int_{I_{w,\gamma,N}} \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi - \int_{\mathbb{R}} (1 - \widehat{w}(\xi))^N \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi \right\|$$

$$\leq \left\| \int_{I_{w,\gamma,N}} \left[1 - (1 - \widehat{w}(\xi))^N \right] \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi \right\| \tag{22}$$

From (14) and the fact that $\int_{I_{w,\gamma,N}} \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi \to \int_{\{\xi\in\mathbb{R} \mid \widehat{w}(\xi)=0\}} \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi$ as $\gamma \to 0$ or $N \to \infty$, it follows that there exist $N_1 \in \mathbb{N}$ big enough and $0 < \gamma_1 < 1$ small enough such that

$$\left\| \int_{I_{w,\gamma_1,N_1}\setminus\{\xi\in\mathbb{R} \mid \widehat{w}(\xi)=0\}} \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi \right\| \leq \frac{\epsilon}{4}$$

Furthermore there exist $0 < \gamma_2 < 1$ small enough and a $N_2 \in \mathbb{N}$ so that

$$\left\| \int_{I_{w,\gamma_2,N_2}} \left[1 - (1 - \widehat{w}(\xi))^{N_2} \right] \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi \right\| \leq \frac{\epsilon}{2}$$

in fact as $\gamma_2 \to 0$ the interval I_{w,γ_2,N_2} tends to the set of frequencies corresponding to the zeros of $\widehat{w}(\xi)$. Given $\gamma = \min\{\gamma_1, \gamma_2\}$, then there exists $N_3 \in \mathbb{N}$ big enough such that $(1 - \widehat{w}(\xi))^{N_3}$ is small enough in order to have

$$\left\| \int_{\mathbb{R}\setminus I_{w,\gamma,N}} (1 - \widehat{w}(\xi))^{N_3} \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi \right\| \leq \frac{\epsilon}{4}$$

If we consider $N_0 = \max\{N_1, N_2, N_3\}$ there exists $\delta > 0$ such that (10) holds true for every $N \geq N_0$. $\square$

This proposition implies that $\mathrm{IMF}_1^{\mathrm{TH}}$ can be as close as we like to both $\mathrm{IMF}_1^{\mathrm{SC}}$ and IMF_1 if we choose wisely the stopping criterion value δ and the threshold γ.

2.2 IF outer loop convergence

We do have now all the tools needed to study the Iterative Filtering outer loop convergence.

Definition 2 (Significant IMFs with respect to $\eta > 0$) Fixed $\eta > 0$ and given a signal s and its decomposition in IMFs obtained using Algorithm 1, then we define *significant IMFs* with respect to η all the IMFs whose L^∞-norm is bigger than η.

Theorem 3 *Given a signal $s \in L^2(\mathbb{R}) \cap L^\infty(\mathbb{R})$, whose continuous frequency spectrum is compactly supported with upper limit $B > 0$ and lower limit $b > 0$, and such that $\|\widehat{s}\|_\infty = c < \infty$, chosen a filter w produced as convolution of a filter with itself, fixed $\delta > 0$ and $\eta > 0$.*
Then the inner loop of Algorithm 1 converges to (11) and the outer loop produces only a finite number $M \in \mathbb{N}$ of significant IMFs whose norm is bigger than η.

Proof Let us consider the Fourier transform of the signal s. From the hypotheses it follows that $|\widehat{s}(\xi)| = 0$ for every $\xi \geq B$.

We can assume that Algorithm 1 in the first step of its outer loop starts selecting a filter w_1 such that the zero of $\widehat{w_1}$ with smallest frequency is at B. We recall in fact that one of the possible way to choose the filter length is based on the Fourier transform of s, as explained in Sect. 2. Given $\delta > 0$ we can identify $N_1 \in \mathbb{N}$ such that (10) is fulfilled for every $N \geq N_1$.

Now, from the hypothesis that $\|\widehat{s}\|_\infty = c < \infty$ it follows there exists the upper bound c on $\widehat{s}(\xi)$ uniformly on $\xi \in \mathbb{R}$. From the hypotheses on the filter function it follows that $0 < \widehat{w_1} < 1$, ref. end of Sect. 2 in [14]. Furthermore, from the assumption on the lower bound b and upper bound B of the continuous frequency spectrum of s, the fact that $\left\| e^{2\pi i \xi x} \right\|_\infty \leq 1$ for every $x, \xi \in \mathbb{R}$, by definition of the interval $I_{w_1, \gamma, \widetilde{N}_1}$, and for every $\widetilde{N}_1 \geq N_1$ and $0 < \gamma < \frac{\eta}{2c(B-b)}$, it follows that

$$
\begin{aligned}
&\left\| \mathrm{IMF}_1^{SC} - \mathrm{IMF}_1^{TH} \right\|_\infty \\
&\leq \left\| \int_{[b,\,B] \cap I_{w_1,\gamma,\widetilde{N}_1}} \left[1 - (1 - \widehat{w_1}(\xi))^{\widetilde{N}_1} \right] \widehat{s}(\xi) e^{2\pi i \xi x} \mathrm{d}\xi \right\|_\infty \\
&\leq \int_{[b,\,B] \cap I_{w_1,\gamma,\widetilde{N}_1}} \left\| \left[1 - (1 - \widehat{w_1}(\xi))^{\widetilde{N}_1} \right] \right\|_\infty \|\widehat{s}(\xi)\|_\infty \left\| e^{2\pi i \xi x} \right\|_\infty \mathrm{d}\xi \\
&\leq c \int_{[b,\,B] \cap I_{w_1,\gamma,\widetilde{N}_1}} \left\| \left[1 - (1 - \widehat{w_1}(\xi))^{\widetilde{N}_1} \right] \right\|_\infty \mathrm{d}\xi \leq c\gamma(B-b) < \frac{\eta}{2} \quad (23)
\end{aligned}
$$

In particular we point out that $I_{w_1, \gamma, \widetilde{N}_1}$, defined as in (14), covers the interval of frequencies $[B - r_1,\ B + r_1]$, for some $r_1 > \varepsilon > 0$.

This last inequality follows from the fact that if we scale linearly the filter function w to enlarge its support, as in (12) for $a > 1$, its Fourier transform is proportionally shrunk (13). However the signal s does have a lower bound b in the continuous frequency spectrum which implies that the filter function w_1 cannot have a too wide support and as a consequence its Fourier transform cannot be too much squeezed. Therefore it does exist $\varepsilon > 0$ which lower bounds the radius r_1.

If $\left\| \mathrm{IMF}_1^{TH} \right\|_\infty < \frac{\eta}{2}$ then we can for sure regard this component as not significant because $\left\| \mathrm{IMF}_1^{SC} \right\|_\infty \leq \left\| \mathrm{IMF}_1^{SC} - \mathrm{IMF}_1^{TH} \right\|_\infty + \left\| \mathrm{IMF}_1^{TH} \right\|_\infty < \eta$. Otherwise, assum-

ing $\left\| \text{IMF}_1^{\text{TH}} \right\|_\infty \geq \frac{\eta}{2}$, if $\left\| \text{IMF}_1^{\text{SC}} \right\|_\infty \geq \eta$, then IMF_1^{SC} represents the first significant IMF in the decomposition. This conclude the first step of the outer loop in Algorithm 1.

In the second step of the outer loop Algorithm 1 iterates the previous passages using now the remainder signal $s_2 = s - \text{IMF}_1^{SC}$ and selecting a filter w_2 such that the zero of $\widehat{w_2}$ with smallest frequency is at $B - r_1$.

Also in this case, given $\delta > 0$, we can identify $N_2 \in \mathbb{N}$ such that (10) is fulfilled for every $N \geq N_2$. Furthermore $\widehat{s_2}(\xi) = \widehat{s}(\xi) - \widehat{\text{IMF}}_1^{SC}(\xi) = \left[1 - (1 - \widehat{w_2}(\xi))^{\widetilde{N}_1} \right] \widehat{s}(\xi)$, $\forall \xi \in \mathbb{R}$ which implies that

$$\|\widehat{s_2}\|_\infty \leq \left\| \left[1 - (1 - \widehat{w_2}(\xi))^{\widetilde{N}_1} \right] \right\|_\infty \|\widehat{s}(\xi)\|_\infty \leq \|\widehat{s}(\xi)\|_\infty \qquad (24)$$

since $\widehat{w_2}(\xi) \in [0, 1]$, $\forall \xi \in \mathbb{R}$ [14]. Hence $\widehat{s_2}$ has the same uniform upper bound c over all $\xi \in \mathbb{R}^+$ as $\widehat{s}(\xi)$.

Therefore

$$\left\| \text{IMF}_2^{\text{SC}} - \text{IMF}_2^{\text{TH}} \right\|_\infty \leq \left\| \int_{I_{w_2, \gamma, \widetilde{N}_2}} \left[1 - (1 - \widehat{w_2}(\xi))^{\widetilde{N}_2} \right] \widehat{s_2}(\xi) e^{2\pi i \xi x} d\xi \right\|_\infty$$

$$\leq c\gamma(B - b) < \frac{\eta}{2} \qquad (25)$$

for every $\widetilde{N}_2 \geq N_2$ and $0 < \gamma < \frac{\eta}{2c(B-b)}$.

Furthermore $I_{w_2, \gamma, \widetilde{N}_2}$ covers the interval of frequencies $[B - r_2, \ B + r_2]$, for some $r_2 > \varepsilon > 0$. This last inequality follows from the same reasoning as before and the fact that the lower bound on the continuous frequency spectrum of s_2 is again b, by construction of s_2, the fact that γ is fixed for every IMF and the Fourier transform of the scaled filter w_2 is a squeezed version of $\widehat{w}$, ref. equation (13).

If $\left\| \text{IMF}_2^{\text{TH}} \right\|_\infty < \frac{\eta}{2}$ then we can regard this component as not significant. If instead $\left\| \text{IMF}_2^{\text{TH}} \right\|_\infty \geq \frac{\eta}{2}$ and $\left\| \text{IMF}_2^{\text{SC}} \right\|_\infty \geq \eta$, then IMF_2^{SC} represents another significant IMF in the decomposition.

The subsequent outer loop steps follow similarly. The existence of the lower limit ε for all $r_k > 0$, $k \geq 1$, ensures that we can have a finite coverage of the interval of frequencies $[b, \ B]$. In particular the algorithm generates a set $\{r_k\}_{k=1}^R$ such that $\sum_{k=1}^R r_k = B - b$ and there exists a natural number $0 \leq M \leq R$ which represents the number of significant IMFs with respect to η. $\qquad \square$

We point out that this theorem holds true also if we consider the L^2-norm instead of the L^∞-norm thanks to the inclusion of L^p spaces on a finite measure space.

From this Theorem it follows that IF with a stopping criterion allows to decompose a signal into a finite number of components given by (11) each of which contains frequencies of the original signal filtered in a smart way.

We observe also that this theorem, together with Theorems 1 and 2 , allow to conclude that the IF method can not produce fake oscillations. Each IMF is in fact containing part of the oscillatory content of the original signal, as described in (4) and (11).

3 IF algorithm in the discrete setting

Real life signals are discrete and compactly supported, therefore we want to analyze the IF algorithm discretization and study its properties.

Consider a signal $s(x)$, $x \in \mathbb{R}$, we assume for simplicity it is supported on $[0, \ 1]$, sampled at n points $x_j = \frac{j}{n-1}$, with $j = 0, \ldots, n - 1$, with a sampling rate which allows to capture all its fine details, so that aliasing will not play any role. The goal is to decompose the vector $\left[s(x_j) \right]_{j=0}^{n-1}$ into vectorial IMFs. Without loosing generality we can assume that $\| \left[s(x_j) \right] \|_2 = 1$.

From now on, to simplify the formulas, we use the notation $s = \left[s(x_j) \right]_{j=0}^{n-1}$. Furthermore, if not specified differently, we consider as matrix norm the so called Frobenius norm $\|A\|_2 = \sqrt{\sum_{i, \ j=0}^{n-1} |a_{ij}|^2}$ which is unitarily invariant.

Definition 3 A vector $w \in \mathbb{R}^n$, n odd number, is called a **filter** if its values are symmetric with respect to the middle, nonnegative, and $\sum_{p=1}^{n} w_p = 1$.

We assume that a filter shape has been selected a priori, like one of the Fokker-Planck filters described in [14], and that some invertible and monotone scaling function g_m has been chosen so that $w_m(\xi)$ can be computed as described in (8). Therefore, assuming $s_1 = s$, the main step of the IF method becomes

$$s_{m+1}(x_i) = s_m(x_i) - \int_{x_i-l_m}^{x_i+l_m} s_m(y) w_m(x_i - y) \mathrm{d}y \approx s_m(x_i)$$

$$- \sum_{x_j=x_i-l_m}^{x_i+l_m} s_m(x_j) w_m(x_i - x_j) \frac{1}{n}, \quad j = 0, \ldots, n - 1 \qquad (26)$$

In matrix form we have

$$s_{m+1} = (I - W_m)s_m \qquad (27)$$

where

$$W_m = \left[w_m(x_i - x_j) \cdot \frac{1}{n} \right]_{i, \ j=0}^{n-1} = \left[\frac{w(g_m^{-1}(x_i - x_j))}{\sum_{z_r=-1}^{1} w(z_r)|g_m'(z_r)|\Delta z_r} \cdot \frac{1}{n} \right]_{i, \ j=0}^{n-1} \qquad (28)$$

Algorithm 2 provides the discrete version of Algorithm 1.

We remind that the first while loop is called outer loop, whereas the second one inner loop.

The first IMF is given by $\mathrm{IMF}_1 = \lim_{m \to \infty} (I - W_m)s_m$, where we point out that the matrix $W_m = [w_m(x_i - x_j)]_{i, \ j=0}^{n-1}$ depends on the half support length l_m at every step m.

However in the implemented code the value l_m is usually computed only in the first iteration of each inner loop and then kept constant in the subsequent steps, so that the matrix W_m is equal to W for every $m \in \mathbb{N}$. So the first IMF is given by

$$\mathrm{IMF}_1 = \lim_{m \to \infty} (I - W)^m s \qquad (29)$$

Algorithm 2 Discrete Iterative Filtering IMF = DIF(s)

IMF = {}
while the number of extrema of $s \geq 2$ **do**
    $s_1 = s$
    while the stopping criterion is not satisfied **do**
        compute the function $w_m(\xi)$, whose half support length l_m is based on the signal $\left[s_m(x_i)\right]_{i=0}^{n-1}$
        $s_{m+1}(x_i) = s_m(x_i) - \sum_{j=0}^{n-1} s_m(x_j) w_m(|x_i - x_j|) \frac{1}{n}, \quad i = 0, \ldots, n-1$
        $m = m + 1$
    end while
    IMF = IMF $\cup$ {s_m}
    $s = s - s_m$
end while
IMF = IMF $\cup$ {s}

Furthermore in the implemented algorithm we do not let m to go to infinity, instead we use a stopping criterion as described in Sect. 2.1. For instance, we can define the following quantity

$$SD := \frac{\|s_{m+1} - s_m\|_2}{\|s_m\|_2} \tag{30}$$

and we can stop the process when the value SD reaches a certain threshold. Another possible option is to introduce a limit on the maximal number of iterations for all the inner loops. It is always possible to adopt different stopping criteria for different inner loops.

If we consider the case of linear scaling, making use of (9), the matrix W_m becomes

$$W_m = \left[\frac{w\left(\frac{x_i - x_j}{l_m}\right)}{l_m} \cdot \frac{1}{n}\right]_{i,\ j=0}^{n-1} = \left[\frac{w\left(\frac{i-j}{(n-1)l_m}\right)}{l_m} \cdot \frac{1}{n}\right]_{i,\ j=0}^{n-1} \tag{31}$$

We point out that the previous formula represent an ideal W_m, however we need to take into account the quadrature formula we use to compute the numerical convolution in order to build the appropriate W_m to be used in the DIF algorithm.

For instance, if we use the rectangle rule, we need to substitute the exact value of $w(y)$ at y with its average value in the interval of length $\frac{1}{n}$ centered in y and multiply this value for the length of interval itself. Furthermore we should handle appropriately the boundaries of the support of $w(y)$, in fact the half length of the support is, in general, a non integer value. This can be done by handling separately the first and last interval in the quadrature formula. In fact we can scale the value of the integral on these two intervals proportionally to the actual length of the intervals themselves.

If we take into account all the aforementioned details we can reproduce a matrix W_m which is row stochastic.

We observe that in the implemented code we simply scale each row of W_m by its sum so that the matrix becomes row stochastic.

3.1 Spectrum of W_m

Since $W_m \in \mathbb{R}^{n \times n}$ represents the discrete convolution operator, it can be a circulant matrix, Toeplitz matrix or it can have a more complex structure. Its structure depends on the way we extend the signal outside its boundaries.

From now on we assume for simplicity that n is an odd natural number, and that we have periodic extension of signals outside the boundaries, therefore W_m is a circulant matrix given by

$$W_m = \begin{bmatrix} c_0 & c_{n-1} & \cdots & c_1 \\ c_1 & c_0 & \cdots & c_2 \\ \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & \cdots & c_0 \end{bmatrix} \tag{32}$$

where $c_j \geq 0$, for every $j = 0, \ldots, n-1$, and $\sum_{j=0}^{n-1} c_j = 1$. Each row contains a circular shift of the entries of a chosen vector filter w_m. For the non periodic extension case we refer the reader to [12].

Denoting by $\sigma(W_m)$ the spectrum of the matrix, in the case of a circulant matrix it is well known that the eigenvalues $\lambda_j \in \sigma(W_m)$, $j = 0, \ldots, n-1$ are given by the formula

$$\lambda_j = c_0 + c_{n-1}\omega_j + \cdots + c_1\omega_j^{n-1}, \quad \text{for} \quad j = 0, \ldots, n-1 \tag{33}$$

where $i = \sqrt{-1}$, and $\omega_j = e^{\frac{2\pi ij}{n}}$ j-th power of the n-th root of unity, for $j = 0, \ldots, n-1$.

Since we construct the matrices W_m using symmetric filters w_m, we have that $c_{n-j} = c_j$ for every $j = 1, \ldots, \frac{n-1}{2}$. Hence W_m is circulant, symmetric and

$$\lambda_j = c_0 + c_1\left(\omega_j + \omega_j^{n-1}\right) + c_2\left(\omega_j^2 + \omega_j^{n-2}\right)\cdots + c_{\frac{n-1}{2}}\left(\omega_j^{\frac{n-1}{2}} + \omega_j^{\frac{n+1}{2}}\right)$$

$$= c_0 + \sum_{k=1}^{\frac{n-1}{2}} c_k\left(\omega_j^k + \omega_j^{n-k}\right) = c_0 + \sum_{k=1}^{\frac{n-1}{2}} c_k\left(e^{\frac{2\pi ij}{n}k} + e^{\frac{2\pi ij}{n}(n-k)}\right)$$

$$= c_0 + \sum_{k=1}^{\frac{n-1}{2}} c_k\left(e^{\frac{2\pi ij}{n}k} - e^{\frac{2\pi ij}{n}k}e^{2\pi ij}\right) \tag{34}$$

Therefore

$$\lambda_j = c_0 + 2\sum_{k=1}^{\frac{n-1}{2}} c_k \cos\left(\frac{2\pi jk}{n}\right), \quad \text{for} \quad j = 0, \ldots, n-1 \tag{35}$$

It is evident that, for any $j = 0, \ldots, n-1$, λ_j is real and $\sigma(W_m) \subseteq [-1, 1]$ since W_m is a stochastic matrix.

Furthermore, if we make the assumption that the filter half supports length is always $l_m \leq \frac{n-1}{2}$, then the entries c_j of the matrix W_m are going to be zero at least for any $j \in [\frac{n-1}{4}, \frac{3}{4}(n-1)]$.

We observe that the previous assumption is reasonable since it implies that we can study oscillations with periods at most equal to half of the length of a signal.

Theorem 4 *Considering the circulant matrix W_m given in (32), assuming that $n > 1$, $\sum_{j=0}^{n-1} c_j = 1$, $c_j \geq 0$, and $c_{n-j} = c_j$, for every $j = 1, \ldots, n - 1$.*

Then W_m is non-defective, diagonalizable and has real eigenvalues.

Furthermore, if the filter half supports length l_m is small enough so that $c_0 = 1$ and $c_j = 0$, for every $j = 1, \ldots, n - 1$, then we have n eigenvalues λ_j all equal 1.

Otherwise, if the filter half supports length l_m is big enough so that $c_0 < 1$ and the values c_k correspond to the discretization of a function with compact and connected support, then there is one and only one eigenvalue equal to 1, which is λ_0, all the other eigenvalues λ_j are real and strictly less than one in absolute value. So they belong to the interval $(-1, 1)$.

Proof First of all we recall that symmetric matrices are always non-defective, diagonalizable and with a real spectrum.

In the case of $c_0 = 1$ the conclusion follows immediately from the observation that W_m reduces to an identity matrix.

When $c_0 < 1$ from (35) it follows that $\lambda_0 = 1$ and all the other eigenvalues belong to the interval $[-1, 1]$. Let us assume, by contradiction, that there exists another eigenvalue $\lambda_d = 1$ for some $d \in \{1, 2, \ldots, n - 1\}$. We assume for simplicity that n is odd. The proof in the even case works in a similar way.

From (35) and the fact that $c_{n-j} = c_j$, for every $j = 1, \ldots, \frac{n-1}{2}$, it follows that

$$\lambda_d = c_0 + 2 \sum_{k=1}^{\frac{n-1}{2}} c_k \cos\left(\frac{2\pi dk}{n}\right), \quad \text{for} \quad d \in \{1, 2, \ldots, n - 1\} \qquad (36)$$

In the right hand side we have among the terms c_k, which by themselves would add up to 1, at least $c_1 > 0$ which is multiplied by $\cos\left(\frac{2\pi d}{n}\right) < 1$ for any $d \in \{1, 2, \ldots, n - 1\}$. Therefore the right hand side will never add up to 1. Hence we have a contradiction.

From (35) it follows also that $\lambda_d \neq -1$ for any $d \in \{1, 2, \ldots, n - 1\}$ because λ_d is given by a convex combination of cosines and $+1$.

So all the eigenvalues of W_m except λ_0 are real and strictly less than one in modulus. $\qquad \square$

We observe that in the discrete iterative filtering algorithm the entries c_k derive from the discretization of a filter function which is by Definition 1 compactly supported. Furthermore, since the filter function is used to compute the moving average of a signal, it is reasonable to require its support to be connected.

Form this theorem it follows that

Corollary 1 *Considering the matrix W_m given in the previous theorem, assuming $c_0 < 1$ and that W_m is constructed using a filter w_m that is produced as convolution of a*

symmetric filter h_m with itself, then there is one and only one eigenvalue equal to 1, all the other eigenvalues belong to the interval $[0, 1)$.

Proof The proof follows directly from the previous theorem and the fact that the matrix $W_m = \widetilde{W}_m^T * \widetilde{W}_m = \widetilde{W}_m^2$, where $\widetilde{W}_m$ is a circulant symmetric convolution matrix associated with the filter $\widetilde{w}_m$. □

Corollary 2 *Assuming $c_0 < 1$, the eigenvector of W_m corresponding to $\lambda_0 = 1$ is a basis for the kernel of the matrix $(I - W_m)$, which has dimension one.*

Before presenting the main proposition we recall that, given a circulant matrix $C = \left[c_{pq}\right]_{p, q=0,\dots,n-1}$, its eigenvalues are

$$\lambda_p = \sum_{q=0}^{n-1} c_{1q} e^{-2\pi i p \frac{q}{n}} \quad p = 0, \dots, n-1 \tag{37}$$

and the corresponding eigenvectors are

$$u_p = \frac{1}{\sqrt{n}} \left[1, \ e^{-2\pi i p \frac{1}{n}}, \dots, \ e^{-2\pi i p \frac{n-1}{n}}\right]^T \quad p = 0, \dots, n-1 \tag{38}$$

which form an orthonormal set.

We recall that an eigenvalue of a matrix is called semisimple whenever its algebraic multiplicity coincides with its geometric multiplicity.

Proposition 2 *Given a matrix W_m, assuming that all the assumptions of Theorem 4 and Corollary 1 hold true, and assuming that $W_m = W$ for any $m \geq 1$. Given $\{\lambda_p\}_{p=0,\dots,n-1}$, semisimple eigenvalues of W, and the corresponding eigenvectors $\{u_p\}_{p=0,\dots,n-1}$, we define the matrix U having as columns the eigenvectors u_p. Assuming that W has k zero eigenvalues, where k is a number in the set $\in \{0, 1, \dots, n-1\}$,*

Then

$$\lim_{m\to\infty} (I - W)^m = U Z U^T \tag{39}$$

where U is unitary and Z is a diagonal matrix with entries all zero except k elements in the diagonal which are equal to one.

Proof From Theorem 4 we know that W is diagonalizable, therefore the matrix U is orthogonal and all the eigenvalues of W are semisimple. Furthermore, since the eigenvectors of W are orthonormal, it follows that U is a unitary matrix. Hence $W = UDU^T$, where D is a diagonal matrix containing in its diagonal the eigenvalues of W. From the assumption that W is associated with a double convolved filter it follows that the spectrum of W is contained in $[0, 1]$, ref. Corollary 1. Therefore also the spectrum of $(I - W)$ is contained in $[0, 1]$. Furthermore

$$(I - W) = U(I - D)U^T$$

and $I - D$ is a diagonal matrix whose diagonal entries are in the interval $(0, 1)$ except the first one which equals 0, ref. Corollary 2, and k entries that are equal to 1. Hence

$$\lim_{m \to \infty} (I - W)^m = \lim_{m \to \infty} U(I - D)^m U^T = U Z U^T$$

where Z is a diagonal matrix with entries all zero except k elements in the diagonal which are equal to one.   □

From the previous proposition it follows

Corollary 3 *Given a signal $s \in \mathbb{R}^n$, assuming that we are considering a doubly convolved filter, and the half filter support length is constant throughout all the steps of an inner loop,*
Then the first outer loop step of the DIF method converges to

$$IMF_1 = \lim_{m \to \infty} (I - W)^m s = U Z U^T s \qquad (40)$$

So the DIF method in the limit produces IMFs that are projections of the given signal s onto the eigenspace of W corresponding to the zero eigenvalue which has algebraic and geometric multiplicity $k \in \{0, 1, \ldots, n - 1\}$. Clearly, if W has only a trivial kernel then the method converges to the zero vector. We point out that since (37) is also the Discrete Fourier Transform (DFT) formula of the sequence $\{c_{1q}\}_{q=0,\ldots,n-1}$, where $C = [c_{pq}]$ is a circulant matrix, it follows that the eigenvalues of W, can be computed directly as the DFT of the sequence $\{w_{1q}\}_{q=0,\ldots,n-1}$, by means of the Fast Fourier Transform (FFT). If we regard the DFT as a discretization of the Fourier Transform of the filter function w it becomes clear that, since the latter has only isolated zeros, in many cases we will not have eigenvalues exactly equal to zero. So in general W has only a trivial kernel and (40) converges to the zero vector. In order to ensure that the method produces a non zero vector we need to discontinue the calculation introducing some stopping criterion.

3.2 DIF inner and outer loop convergence in presence of a stopping criterion

If we assume that the half support length l_m is computed only in the beginning of each inner loop, then the first IMF is given by (29) and (40).

In order to have a finite time method we may introduce a stopping criterion in the DIF algorithm, like the condition

$$\|s_{m+1} - s_m\|_2 < \delta \qquad \forall m \geq N_0 \qquad (41)$$

for some fixed $\delta > 0$.

Then, based on Corollary 3, we produce an approximated first IMF given by

$$\overline{\text{IMF}}_1 = (I - W)^{N_0} s = U(I - D)^{N_0} U^T s. \tag{42}$$

Theorem 5 *Given $s \in \mathbb{R}^n$, we consider the convolution matrix W defined in (32), associated with a filter vector w given as a symmetric filter h convolved with itself. Assuming that W has k zero eigenvalues, where k is a number in the set $\in \{0, 1, \ldots, n-1\}$, and fixed $\delta > 0$,*
   Then, calling $\tilde{s} = U^T s$, for the minimum $N_0 \in \mathbb{N}$ such that it holds true the inequality

$$\frac{N_0^{N_0}}{(N_0 + 1)^{N_0+1}} < \frac{\delta}{\|\tilde{s}\|_\infty \sqrt{n - 1 - k}} \tag{43}$$

we have that $\|s_{m+1} - s_m\|_2 < \delta \quad \forall m \geq N_0$ and the first IMF is given by

$$\overline{IMF}_1 = U(I - D)^{N_0} U^T s = UP \begin{bmatrix} 0 & & & & & & \\ & (1 - \lambda_1)^{N_0} & & & & & \\ & & \ddots & & & & \\ & & & (1 - \lambda_{n-1-k})^{N_0} & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} P^T U^T s \tag{44}$$

where P is a permutation matrix which allows to reorder the columns of U, which correspond to eigenvectors of W, so that the corresponding eigenvalues $\{\lambda_p\}_{p=1,\ldots, n-1}$ are in decreasing order.

Proof

$$\|s_{m+1} - s_m\|_2 = \|(I - W)^{m+1} - (I - W)^m\|_2 = \|U(I - D)^m (I - D - I) U^T s\|_2$$
$$= \|(I - D)^m (I - D - I) U^T s\|_2 = \|(I - D)^m (I - D - I) \tilde{s}\|_2 \tag{45}$$

since U is a unitary matrix and where $\tilde{s} = U^T s$.
   Given a permutation matrix P such that the entries of the diagonal PDP^T are the eigenvalues of W in decreasing order of magnitude, starting from $\lambda_0 = 1$, and assuming that W has k zero eigenvalues, where k is a number in the set $\in \{0, 1, \ldots, n-1\}$, then

$$\|(I - D)^m (I - D - I)\widetilde{s}\|_2$$

$$\leq \left\| P \begin{bmatrix} 0 & & & & & & \\ & (1-\lambda_1)^m \lambda_1 & & & & & \\ & & \ddots & & & & \\ & & & (1-\lambda_{n-1-k})^m \lambda_{n-1-k} & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} P^T \begin{bmatrix} \|\widetilde{s}\|_\infty \\ \vdots \\ \|\widetilde{s}\|_\infty \end{bmatrix} \right\|_2$$

$$\leq \sqrt{n-1-k} \left(1 - \frac{1}{m+1}\right)^m \frac{1}{m+1} \|\widetilde{s}\|_\infty = \sqrt{n-1-k} \frac{m^m}{(m+1)^{m+1}} \|\widetilde{s}\|_\infty$$

$$(46)$$

because the function $(1 - \lambda)^m \lambda$ achieves its maximum at $\lambda = \frac{1}{m+1}$ for $\lambda \in [0, \ 1]$.

Hence the stopping criterion (41) is fulfilled for N_0 minimum natural number such that 43 holds true. □

We observe that, as we mentioned earlier, since (37) is also the Discrete Fourier Transform (DFT) formula of the sequence $\{c_{1q}\}_{q=0,\dots,n-1}$, it follows that the eigenvalues of $W = [w_{pq}]_{p, \ q=0,\dots,n-1}$, can be computed directly as the DFT of the sequence $\{w_{1q}\}_{q=0,\dots,n-1}$, by means of the Fast Fourier Transform (FFT). This calculation can be done "off line", in fact, once the filter shape w has been fixed, we can compute and store its FFT for different values of the size of its support. This fact, together with other previous results, can be used to improve the efficiency of the method as explained in the following section.

It is interesting to notice that each IMF is generated as a linear combination of elements in an orthonormal basis. Therefore we can regard the IMFs as elements of a frame which allows to decompose a given signal into a few significant components. From this prospective the IF algorithm can be viewed as a method that automatically produces elements of a frame associated with a signal. The possible connections between IF and the frame theory are fascinating, but out of the scope of the present work. We plan to follow this direction of research in a future work.

Regarding the DIF outer loop convergence they hold true the same results described in Sect. 2.2 for the continuous setting. In fact, while the inner loop of the IF algorithm requires a discretization to deal with discrete signals, the outer loop does not require any form of discretization and it works the very same as in the continuous setting.

4 Efficient implementation of the DIF algorithm

In this section we present an efficient implementation of the DIF algorithm applied to the decomposition of a signal s of length n, named Fast Iterative Filtering algorithm. This approach assumes implicitly that the signal at the boundaries is periodical. This is clearly a limitation, since most signals are far from being periodical at the edges,

but it can be treated by pre-extending the signal and making it periodical at the new boundaries [12,69].

We start from Theorem 5 which allows to compute each IMF as fast as the FFT. In fact, we can compute the IMF using (44) where the eigenvalues $\{\lambda_k\}_{k=1, 2,..., n}$ can be evaluated using (37), or by means of the Fast Fourier Transform since (37) is equivalent to the Discrete Fourier Transform of the sequence $\{w_{1q}\}_{q=0,...,n-1}$. Furthermore we recall that $U^T s$ is the DFT of s that can be computed using the FFT algorithm, whose computational complexity is $n \log(n)$, and that multiplying on the left by the matrix $U = [u_k]$ is equivalent to computing the inverse DFT (iDFT) which can be done using the inverse FFT. Hence the following corollary holds true and provides an explicit formula for the calculation of each IMF.

Corollary 4 *Given a signal $s \in \mathbb{R}^n$, and a filter vector w derived from a symmetric filter h convolved with itself. Fixed $\delta > 0$, then, for the minimum $N_0 \in \mathbb{N}$ such that the stopping criterion $SD = \frac{\|s_{N_0+1} - s_{N_0}\|_2}{\|s_{N_0}\|_2} < \delta$ is satisfied, the first IMF contained in s is given by*

$$IMF = \sum_{k=0}^{n-1} u_k (1 - \lambda_k)^{N_0} \sigma_k = iDFT\left((I - D)^{N_0} DFT(s)\right) \tag{47}$$

where σ_k represents the k-th element of the DFT of the signal s, and D is a diagonal matrix containing as entries the eigenvalues of the convolution matrix W defined in (32), associated with the filter vector w.

We exploit this theoretical new result in the so called Fast Iterative Filtering (FIF) method implemented for Matlab and available online.[2] The pseudocode of FIF is given in Algorithm 3.

Algorithm 3 Fast Iterative Filtering IMF = FIF(s)

IMF = {}
while the number of extrema of $s \geq 2$ **do**
    compute the filter length l for $s(x)$ and the corresponding filter w
    compute the DFT of the signal s and of the filter w
    while the stopping criterion is not satisfied **do**
        $\widehat{s}_{m+1} = (I - \text{diag}\,(\text{DFT}(w)))^m \text{DFT}(s)$
        $m = m + 1$
    end while
    IMF = IMF $\cup$ {iDFT $(\widehat{s}_m)$}
    $s = s - $ iDFT $(\widehat{s}_m)$
end while
IMF = IMF $\cup$ {s}

In FIF method we speed up the calculations significantly by computing the convolution as product in the frequency domain via FFT, as suggested by Corollary 4. In the recently published paper [11] several tests and a comparison of the standard DIF and

---

Table 1 Computational times versus the number of sample points n for the decomposition of both artificial and real life signals via DIF and FIF method

n (N_{IMF})	DIF (ave. num. it.)	FIF (ave. num. it.)
$1.6 \cdot 10^3$ (1)	0.0448 (8)	0.0049 (8)
$1.6 \cdot 10^4$ (5)	5.58 (6)	0.079 (6)
$2 \cdot 10^4$ (22)	81.77 (17)	0.512 (17)
$1.2 \cdot 10^7$ (5)	3329.7 (4)	40.72 (4)

In between parentheses the number of IMFs N_{IMF}, and the average number of iterations needed to compute each IMF. All tests were conducted on a laptop equipped with an Intel Core i7-8550U CPU, 1.80 GHz, 16.0 GB RAM, Windows 10 Pro, Matlab R2018a. The DIF and FIF Matlab codes are freely available at www.cicone.com. Details on the signals and their decompositions can be found in [11]

the newly proposed FIF method have been conducted on both artificial and real life signals. The results obtained in those examples show, as expected, that the two techniques produces decompositions that are identical, up to machine precision. However, based on those tests, the newly proposed FIF method proves to be roughly three order of magnitudes faster, from a computational time prospective, than the standard DIF method, when applied to signals containing 10^4 sample points or more, Table 1.

We point out that other approaches can be implemented and explored to further speed up the computational cost of the FIF method. For instance, one possibility is the precomputation of the number of iterations needed to achieve the required accuracy δ in the computation of a certain IMF. This number of iterations can be approximated by the minimum $N_0 \in \mathbb{N}$ satisfying the inequality (43).

Another idea that can be implemented is the identification of the appropriate number of iteration via a bisection approach. The FIF algorithm inner loop can be initiated, in fact, with a reasonable big value of m and then, via a bisection approach, the minimal value of m which allows to satisfy the stopping criterion can be identified.

To speed up the FIF computation it is also possible to precompute the eigenvalues λ_k corresponding to any possible scaling of a filter w. In doing so we can reduce even more the computational time of the algorithm.

The FIF algorithm further acceleration is, however, out of the scope of this paper, whose focus is the numerical analysis of the IF method. The interested reader can find more ideas and details in this direction of research in the recently published paper [11].

5 Conclusions and outlook

In this work we tackle the problem of a complete analysis of the IF algorithm both in the continuous and discrete setting. In particular in the continuous setting we show how IF can decompose a signal into a finite number of so called IMFs and that each IMF contains frequencies of the original signal filtered in a "smart way".

In the discrete setting we prove that the DIF method is also convergent and, in the case of periodic extension at the boundaries of the given signal, we provide an explicit

formula for the a priori calculation of each IMF. From this equation it follows that each IMF is a smart summation of eigenvectors of a circulant matrix.

We show that no fake oscillations can be produced neither in the continuous nor in the discrete setting.

From the properties of the DIF algorithm and the explicit formula for the IMFs produced by this method and derived in this work, we propose new ideas that has been directly incorporated in the implemented algorithm in order to increase its efficiency and reduce its computational complexity. The result is the so called FIF method which allows to quickly decompose a signal by means of the FFT. This is an important result in this area of research which opens the doors to an almost instantaneous analysis of non stationary signals.

There are several open problems that remain unsolved. First of all from the proposed analysis it is clear that different filter functions have different Fourier transform and hence the decomposition produced by IF and DIF algorithms is directly influenced by this choice. In a future work we plan to study more in details the connections between the shape of the filters and the quality of the decomposition produced by these methods.

In the current work we analyzed the DIF assuming a periodic extension of the signals at the boundaries. We plan to study in a future work the behavior of the DIF method in the case of reflective, anti-reflective and other boundaries extensions of a signal.

Based on the numerical evidence [14,17] we claim that the Iterative Filtering method is stable under perturbations of the signal. We plan to study rigorously such stability in a future work.

The results about the DIF algorithm convergence suggest that the method allows, in general, to automatically generate a frame associated with a given signal. We plan to further analyze this connection in a future work.

Finally we recall that it is still an open problem how to extend all the results obtained for the Iterative Filtering technique to the case of the Adaptive Local Iterative Filtering method, whose convergence and stability analysis is still under investigation [13,14,61].

References

1. Abdelouahad, A.A., El Hassouni, M., Cherifi, H., Aboutajdine, D.: Reduced reference image quality assessment based on statistics in empirical mode decomposition domain. SIViP **8**(8), 1663–1680 (2014)

2. An, N., Zhao, W., Wang, J., Shang, D., Zhao, E.: Using multi-output feedforward neural network with empirical mode decomposition based signal filtering for electricity demand forecasting. Energy **49**, 279–288 (2013)
3. Barman, C., Ghose, D., Sinha, B., Deb, A.: Detection of earthquake induced radon precursors by hilbert huang transform. J. Appl. Geophys. **133**, 123–131 (2016)
4. Battista, B.M., Knapp, C., McGee, T., Goebel, V.: Application of the empirical mode decomposition and Hilbert–Huang transform to seismic reflection data. Geophysics **72**(2), H29–H37 (2007)
5. Baykut, S., Akgül, T., İnan, S., Seyis, C.: Observation and removal of daily quasi-periodic components in soil radon data. Radiat. Meas. **45**(7), 872–879 (2010)
6. Bowman, D.C., Lees, J.M.: The Hilbert–Huang transform: a high resolution spectral method for nonlinear and nonstationary time series. Seismol. Res. Lett. **84**(6), 1074–1080 (2013)
7. Chen, C.H., Yeh, T.K., Liu, J.Y., Wang, C.H., Wen, S., Yen, H.Y., Chang, S.H.: Surface deformation and seismic rebound: implications and applications. Surv. Geophys. **32**(3), 291 (2011)
8. Chen, C.H., Wang, C.H., Liu, J.Y., Liu, C., Liang, W.T., Yen, H.Y., Yeh, Y.H., Chia, Y.P., Wang, Y.: Identification of earthquake signals from groundwater level records using the HHT method. Geophys. J. Int. **180**(3), 1231–1241 (2010)
9. Chen, Y.: Dip-separated structural filtering using seislet transform and adaptive empirical mode decomposition based dip filter. Geophys. J. Int. **206**(1), 457–469 (2016)
10. Cicone, A.: Nonstationary signal decomposition for dummies. In: Singh, V., Gao, D., Fischer, A. (eds.) Advances in Mathematical Methods and High Performance Computing, pp. 69–82. Springer, Cham (2019)
11. Cicone, A.: Iterative filtering as a direct method for the decomposition of nonstationary signals. Numer. Algorithms **85**(3), 811–827 (2020)
12. Cicone, A., Dell'Acqua, P.: Study of boundary conditions in the iterative filtering method for the decomposition of nonstationary signals. J. Comput. Appl. Math. **373**, 112248 (2020)
13. Cicone, A., Garoni, C., Serra-Capizzano, S.: Spectral and convergence analysis of the discrete ALIF method. Linear Algebra Appl. **580**, 62–95 (2019)
14. Cicone, A., Liu, J., Zhou, H.: Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis. Appl. Comput. Harmon. Anal. **41**(2), 384–411 (2016)
15. Cicone, A., Liu, J., Zhou, H.: Hyperspectral chemical plume detection algorithms based on multidimensional iterative filtering decomposition. Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. **374**(2065), 20150196 (2016)
16. Cicone, A., Wu, H.-T.: How nonlinear-type time-frequency analysis can help in sensing instantaneous heart rate and instantaneous respiratory rate from photoplethysmography in a reliable way. Front. Physiol. **8**, 701 (2017)
17. Cicone, A., Zhou, H.: Multidimensional iterative filtering method for the decomposition of high-dimensional non-stationary signals. Numer. Math. Theory Methods Appl. **10**(2), 278–298 (2017)
18. Costa, M., Goldberger, A.L., Peng, C.K.: Broken asymmetry of the human heartbeat: loss of time irreversibility in aging and disease. Phys. Rev. Lett. **95**(19), 198102 (2005)
19. Cummings, D.A., Irizarry, R.A., Huang, N.E., Endy, T.P., Nisalak, A., Ungchusak, K., Burke, D.S.: Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. Nature **427**(6972), 344 (2004)
20. Dragomiretskiy, K., Zosso, D.: Variational mode decomposition. IEEE Trans. Signal Process. **62**(3), 531–544 (2013)
21. Duffy, D.G.: The application of Hilbert–Huang transforms to meteorological datasets. In: Huang, N.E., Shen, S.P. (eds.) Hilbert–Huang Transform and Its Applications, pp. 203–221. World Scientific, Singapore (2014)
22. Ezer, T., Atkinson, L.P., Corlett, W.B., Blanco, J.L.: Gulf stream's induced sea level rise and variability along the U.S. mid-Atlantic coast. J. Geophys. Res. Oceans **118**(2), 685–697 (2013)
23. Ezer, T., Corlett, W.B.: Is sea level rise accelerating in the Chesapeake bay? A demonstration of a novel new approach for analyzing sea level data. Geophys. Res. Lett. **39**(19), 6 (2012)
24. Franzke, C.: Multi-scale analysis of teleconnection indices: climate noise and nonlinear trend analysis. Nonlinear Process. Geophys. **16**(1), 65–76 (2009)
25. Franzke, C.: Nonlinear trends, long-range dependence, and climate noise properties of surface temperature. J. Clim. **25**(12), 4172–4183 (2012)

26. Ghobadi, H., Spogli, L., Alfonsi, L., Cesaroni, C., Cicone, A., Linty, N., Romano, V., Cafaro, M.: Disentangling ionospheric refraction and diffraction effects in GNSS raw phase through fast iterative filtering technique. GPS Solut. **24**, 85 (2020)

27. Hossein, G., Caner, S., Luca, S., Fabio, D., Antonio, C., Massimo, C..: A comparative study of different phase detrending algorithms for scintillation monitoring. In 2020 XXXIIIrd General Assembly and Scientific Symposium of the International Union of Radio Science, pp. 1–4. IEEE

28. Gilles, J.: Empirical wavelet transform. IEEE Trans. Signal Process. **61**(16), 3999–4010 (2013)

29. Gregoriou, G.G., Gotts, S.J., Desimone, R.: Cell-type-specific synchronization of neural activity in FEF with V4 during attention. Neuron **73**(3), 581–594 (2012)

30. Hassan, A.R., Bhuiyan, M.I.H.: Automatic sleep scoring using statistical features in the EMD domain and ensemble methods. Biocybern. Biomed. Eng. **36**(1), 248–255 (2016)

31. Hillier, A., Morton, R.J., Erdélyi, R.: A statistical study of transverse oscillations in a quiescent prominence. Astrophys. J. Lett. **779**(2), L16 (2013)

32. Hofmann-Wellenhof, B., Lichtenegger, H., Wasle, E.: GNSS-Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and More. Springer, Berlin (2007)

33. Hou, T.Y., Shi, Z.: Adaptive data analysis via sparse time-frequency representation. Adv. Adapt. Data Anal. **3**(01–02), 1–28 (2011)

34. Hou, T.Y., Yan, M.P., Wu, Z.: A variant of the EMD method for multi-scale data. Adv. Adapt. Data Anal. **1**(04), 483–516 (2009)

35. Hu, K., Lo, M.T., Peng, C.K., Liu, Y., Novak, V.: A nonlinear dynamic approach reveals a long-term stroke effect on cerebral blood flow regulation at multiple time scales. PLoS Comput. Biol. **8**(7), e1002601 (2012)

36. Huang, C., Yang, L., Wang, Y.: Convergence of a convolution-filtering-based algorithm for empirical mode decomposition. Adv. Adapt. Data Anal. **1**(04), 561–571 (2009)

37. Huang, J.Y., Wen, K.L., Li, X.J., Xie, J.J., Chen, C.T., Su, S.C.: Coseismic deformation time history calculated from acceleration records using an EMD-derived baseline correction scheme: a new approach validated for the 2011 Tohoku earthquake. Bull. Seismol. Soc. Am. **103**(2B), 1321–1335 (2013)

38. Huang, N.E.: Introduction to the Hilbert–Huang Transform and Its Related Mathematical Problems. World Scientific, SIngapore (2014)

39. Huang, N.E., Chern, C.C., Huang, K., Salvino, L.W., Long, S.R., Fan, K.L.: A new spectral representation of earthquake data: Hilbert spectral analysis of station TCU129, Chi-Chi, Taiwan, 21 September 1999. Bull. Seismol. Soc. Am. **91**(5), 1310–1338 (2001)

40. Huang, N.E., Shen, Z., Long, S.R.: A new view of nonlinear water waves: the Hilbert spectrum. Annu. Rev. Fluid Mech. **31**(1), 417–457 (1999)

41. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **454**(1971), 903–995 (1998)

42. Huang, N.E., Wu, Z.: A review on Hilbert–Huang transform: method and its applications to geophysical studies. Rev. Geophys. **46**(2), RG2006 (2008)

43. Jackson, L.P., Mound, J.E.: Geomagnetic variation on decadal time scales: what can we learn from empirical mode decomposition? Geophys. Rese. Lett. **37**(14), L14307 (2010)

44. Lang, X., Zheng, Q., Zhang, Z., Lu, S., Xie, L., Horch, A., Su, H.: Fast multivariate empirical mode decomposition. IEEE Access **6**, 65521–65538 (2018)

45. Lee, T., Ouarda, T.B.M.J.: Prediction of climate nonstationary oscillation processes with empirical mode decomposition. J. Geophys. Res: Atmos.**116**(D6), D06107 (2011)

46. Lei, Y., Lin, J., He, Z., Zuo, M.J.: A review on empirical mode decomposition in fault diagnosis of rotating machinery. Mech. Syst. Signal Process. **35**(1), 108–126 (2013)

47. Li, X., Su, J., Yang, L.: Building detection in SAR images based on bi-dimensional empirical mode decomposition algorithm. IEEE Geosci. Remote Sens. Lett. **17**(4), 641–645 (2019)

48. Li, Y., Wang, X., Liu, Z., Liang, X., Si, S.: The entropy algorithm and its variants in the fault diagnosis of rotating machinery: a review. IEEE Access **6**, 66723–66741 (2018)

49. Liang, H., Bressler, S.L., Buffalo, E.A., Desimone, R., Fries, P.: Empirical mode decomposition of field potentials from macaque V4 in visual spatial attention. Biol. Cybern. **92**(6), 380–392 (2005)

50. Lin, L., Wang, Y., Zhou, H.: Iterative filtering as an alternative algorithm for empirical mode decomposition. Adv. Adapt. Data Anal. **1**(4), 543–560 (2009)

51. Liu, H., Chen, C., Tian, H.Q., Li, Y.F.: A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. Renew. Energy **48**, 545–556 (2012)

52. Loh, C.H., Wu, T.C., Huang, N.E.: Application of the empirical mode decomposition-Hilbert spectrum method to identify near-fault ground-motion characteristics and structural responses. Bull. Seismol. Soc. Am. **91**(5), 1339–1357 (2001)

53. Materassi, M., Piersanti, M., Consolini, G., Diego, P., D'Angelo, G., Bertello, I., Cicone, A.: Stepping into the Equatorward Boundary of the Auroral Oval: preliminary results of multi scale statistical analysis. Ann. Geophys. **61**, 55 (2019)

54. Meignen, S., Perrier, V.: A new formulation for empirical mode decomposition based on constrained optimization. IEEE Signal Process. Lett. **14**(12), 932–935 (2007)

55. Mitiche, I., Morison, G., Nesbitt, A., Hughes-Narborough, M., Stewart, B.G., Boreham, P.: Classification of partial discharge signals by combining adaptive local iterative filtering and entropy features. Sensors **18**(2), 406 (2018)

56. Morton, R.J., Erdélyi, R., Jess, D.B., Mathioudakis, M.: Observations of sausage modes in magnetic pores. Astrophys. J. Lett. **729**(2), L18 (2011)

57. Papini, E., Cicone, A., Piersanti, M., Franci, L., Hellinger, P., Landi, S., Verdini, A.: Multidimensional iterative filtering: a new approach for investigating plasma turbulence in numerical simulations. J. Plasma Phys. **86**(5) (2020)

58. Papini, E., Piersanti, M., Cicone, A., Franci, L., Landi, S.: Multidimentional iterative filtering: a new approach for investigating plasma turbulence in Hall-MHD and Hybrid-PIC simulations. In: Geophysical Research Abstracts, vol. 21 (2019)

59. Parey, A., El Badaoui, M., Guillet, F., Tandon, N.: Dynamic modelling of spur gear pair and application of empirical mode decomposition-based statistical analysis for early detection of localized tooth defect. J. Sound Vib. **294**(3), 547–561 (2006)

60. Piersanti, G., Piersanti, M., Cicone, A., Canofari, P., Di Domizio, M.: An inquiry into the structure and dynamics of crude oil price using the fast iterative filtering algorithm. Energy Econ. **92**, 104952 (2020)

61. Piersanti, M., Materassi, M., Cicone, A., Spogli, L., Zhou, H., Ezquer, R.G.: Adaptive local iterative filtering: a promising technique for the analysis of nonstationary signals. J. Geophys. Res. Space Phys. **123**(1), 1031–1046 (2018)

62. Pustelnik, N., Borgnat, P., Flandrin, P.: A multicomponent proximal algorithm for empirical mode decomposition. In: 2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO), pp. 1880–1884. IEEE (2012)

63. Roberts, P.H., Yu, Z.J., Russell, C.T.: On the 60-year signal from the core. Geophys. Astrophys. Fluid Dyn. **101**(1), 11–35 (2007)

64. Selesnick, I.W.: Resonance-based signal decomposition: a new sparsity-enabled signal analysis method. Signal Process. **91**(12), 2793–2809 (2011)

65. Sfarra, S., Cicone, A., Yousefi, B., Ibarra-Castanedo, C., Perilli, S., Maldague, X.: Improving the detection of thermal bridges in buildings via on-site infrared thermography: the potentialities of innovative mathematical tools. Energy Build. **182**, 159–171 (2019)

66. Sharma, R., Pachori, R.B., Upadhyay, A.: Automatic sleep stages classification based on iterative filtering of electroencephalogram signals. Neural Comput. Appl. **28**(10), 2959–2978 (2017)

67. Spogli, L., Piersanti, M., Cesaroni, C., Materassi, M., Cicone, A., Alfonsi, L., Romano, V., Ezquer, R.G.: Role of the external drivers in the occurrence of low-latitude ionospheric scintillation revealed by multi-scale analysis. J. Space Weather Space Clim. **9**, A35 (2019)

68. Spogli, L., Piersanti, M., Cesaroni, C., Materassi, M., Cicone, A., Alfonsi, L., Romano, V., Ezquer, R.G.: Role of the external drivers in the occurrence of low-latitude ionospheric scintillation revealed by multi-scale analysis. In: 2019 URSI Asia-Pacific Radio Science Conference (AP-RASC), number 8738254, pp. 1–1 (2019)

69. Stallone, A., Cicone, A., Materassi, M.: New insights and best practices for the successful use of empirical mode decomposition, iterative filtering and derived algorithms. Sci. Rep. **10**, 15161 (2020)

70. Tary, J.B., Herrera, R.H., Han, J., van der Baan, M.: Spectral estimation—what is new? What is next? Rev. Geophys. **52**(4), 723–749 (2014)

71. Terradas, J., Oliver, R., Ballester, J.L.: Application of statistical techniques to the analysis of solar coronal oscillations. Astrophys. J. **614**(1), 435 (2004)

72. Torres, M.E., Colominas, M.A., Schlotthauer, G., Flandrin, P.: A complete ensemble empirical mode decomposition with adaptive noise. In: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4144–4147. IEEE (2011)

73. Tsolis, G.S., Xenos, T.D.: A qualitative study of the seismo-ionospheric precursors prior to the 6 April 2009 earthquake in l'aquila, Italy. Nat. Hazards Earth Syst. Sci. **10**(1), 133–137 (2010)

74. Ur Rehman, N., Mandic, D.P.: Filter bank property of multivariate empirical mode decomposition. IEEE Trans. Signal Process. **59**(5), 2421–2426 (2011)

75. ur Rehman, N., Park, C., Huang, N.E., Mandic, D.P.: EMD via MEMD: multivariate noise-aided computation of standard EMD. Adv. Adapt. Data Anal. **5**(02), 1350007 (2013)

76. Vasudevan, K., Cook, F.A.: Empirical mode skeletonization of deep crustal seismic data: theory and applications. J. Geophys. Res. Solid Earth **105**(B4), 7845–7856 (2000)

77. Wang, C., Choi, H.J., Kim, S.J., Desai, A., Lee, N., Kim, D., Bae, Y., Lee, K.: Deconvolution of subcellular protrusion heterogeneity and the underlying actin regulator dynamics from live cell imaging. Nat. Commun. **9**, 1–17 (2018)

78. Wang, D., Hwang, C., Shen, W.: Investigations of anomalous gravity signals prior to 71 large earthquakes based on a 4-years long superconducting gravimeter records. Geod. Geodyn. **8**(5), 319–327 (2017)

79. Wang, Y., Wei, G.-W., Yang, S.: Iterative filtering decomposition based on local spectral evolution kernel. J. Sci. Comput. **50**(3), 629–664 (2012)

80. Wang, Y., Zhou, Z.: On the Convergence of Iterative Filtering Empirical Mode Decomposition. Excursions in Harmonic Analysis, vol. 2, pp. 157–172. Birkhäuser, Boston (2013)

81. Wei, Y., Chen, M.C.: Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. Transp. Res. Part C Emerg. Technol. **21**(1), 148–162 (2012)

82. Wu, C.H., Chang, H.C., Lee, P.L., Li, K.S., Sie, J.J., Sun, C.W., Yang, C.Y., Li, P.H., Deng, H.T., Shyu, K.K.: Frequency recognition in an ssvep-based brain computer interface using empirical mode decomposition and refined generalized zero-crossing. J. Neurosci. Methods **196**(1), 170–181 (2011)

83. Wu, Z., Huang, N.E.: Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv. Adapt. Data Anal. **1**(01), 1–41 (2009)

84. Xia, Y., Zhang, B., Pei, W., Mandic, D.P.: Bidimensional multivariate empirical mode decomposition with applications in multi-scale image fusion. IEEE Access **7**, 114261–114270 (2019)

85. Yang, A.C., Huang, N.E., Peng, C.K., Tsai, S.J.: Do seasons have an influence on the incidence of depression? The use of an internet search engine query data as a proxy of human affect. PLOS ONE **5**(10), e13728 (2010)

86. Yang, A.C., Peng, C.K., Huang, N.E.: Causal decomposition in the mutual causation system. Nat. Commun. **9**(1), 3378 (2018)

87. Yang, J.N., Lei, Y., Lin, S., Huang, N.: Identification of natural frequencies and dampings of in situ tall buildings using ambient wind vibration data. J. Eng. Mech. **130**(5), 570–577 (2004)

88. Yeh, J.R., Shieh, J.S., Huang, N.E.: Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method. Adv. Adapt. Data Anal. **2**(02), 135–156 (2010)

89. Yu, L., Wang, S., Lai, K.K.: Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Econ. **30**(5), 2623–2635 (2008)

90. Yu, S., Ma, J., Osher, S.: Geometric mode decomposition. Inverse Probl. Imaging **12**(4), 831–852 (2018)

91. Yu, Z.G., Anh, V., Wang, Y., Mao, D., Wanliss, J.: Modeling and simulation of the horizontal component of the geomagnetic field by fractional stochastic differential equations in conjunction with empirical mode decomposition. J. Geophys. Res. Space Phys. **115**(A10), 1–11 (2010)

92. Zhang, R.R., Ma, S., Hartzell, S.: Signatures of the seismic source in EMD-based characterization of the 1994 Northridge, California, earthquake recordings. Bull. Seismol. Soc. Am. **93**(1), 501–518 (2003)

93. Zhang, R.R., Ma, S., Safak, E., Hartzell, S.: Hilbert–Huang transform analysis of dynamic and earthquake motion recordings. J. Eng. Mech. **129**(8), 861–875 (2003)

94. Zhang, X., Lai, K.K., Wang, S.Y.: A new approach for crude oil price analysis based on empirical mode decomposition. Energy Econ. **30**(3), 905–918 (2008)

95. Zhang, X., Yu, L., Wang, S., Lai, K.K.: Estimating the impact of extreme events on crude oil price: an EMD-based event analysis method. Energy Econ. **31**(5), 768–778 (2009)

96. Zheng, J., Cheng, J., Yang, Y.: Partly ensemble empirical mode decomposition: an improved noise-assisted method for eliminating mode mixing. Signal Process. **96**, 362–374 (2014)

97. Zheng, Y., Wang, G., Li, K., Bao, G., Wang, J.: Epileptic seizure prediction using phase synchronization based on bivariate empirical mode decomposition. Clin. Neurophysiol. **125**(6), 1104–1111 (2014)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.