Empirical Mode Decomposition

November 9, 2023

1 Introduction

EMD is a relatively new technique used for the spectral analysis of astrophysical signals. It is fundamentally different from Fourier and Wavelet based techniques and primarily relies on the time domain decomposition of the different modes present in the signal. It is not a replacement for fourier and wavelet methods, but simply another tool to testify the obtained results.
EMD in python

2 Initial Processing

The given signal is de-meaned and standard deviation is normalised to unity.

$$X \to X - \bar{X} \tag{1}$$

$$X \to \frac{X}{\sigma^2} \tag{2}$$

where $\bar{X}$ is the mean and σ^2 is the Standard deviation

The signal is decomposed into various Intrinsic Mode Functions called as IMFs, which represent a Band limited signal. The sifting processes has a Chi-squared criterion, and stops when it's met. Summing up the IMFs we get back the original signal.

This is one of the key advantages of the EMD analysis over Fourier and Wavelets methods as it allows for Non stationary and non linear data to be analysed.

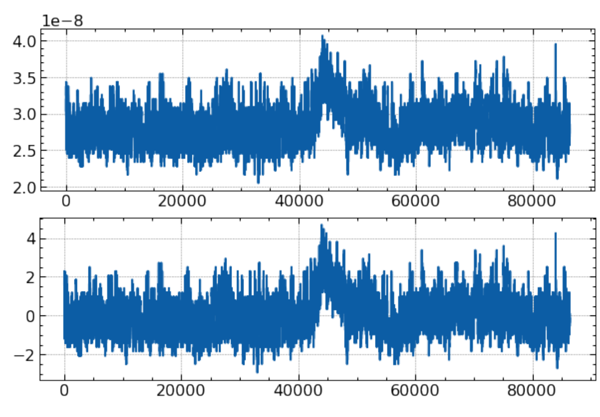
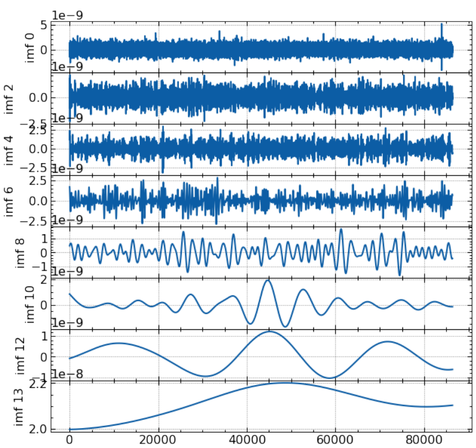
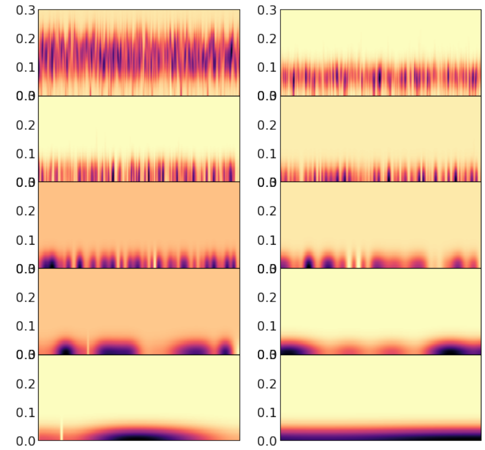


Figure 1: Upper panel :The mean is non zero, and standard deviation is not normalised
Lower panel: From the label, its indicative that the signal has unit energy.

3 The sifting process



(a) label 1



(b) label 2

Figure 2: Left panel:Some of the IMFs are shown as a result of sifting. Clearly the frequency content decreases with the IMFs, the first IMF has the highest frequency while the last one is nearly monotone and aperiodic. Right Panel :We can see the decreasing frequency content with increasing the IMF number. The hilbert holopsectrum band occupies a lower value for subsequent IMF.

4 Modal Period and Energy

We define Modal Energy and period as follows

$$E_m^j = \frac{1}{N}\sum_i^N |C_{j_{i,m}}|^2 \tag{3}$$

$$P_m^j = \frac{2N}{b_m} \tag{4}$$

where $C_{j_{i,m}}$ is the amplitudes of the j^{th} IMF and b_m as the total no of extremas in the IMF

5 The EMD Spectrum

The EMD spectrum is constructed by plotting the points (E_m^j, P_m^j) on a Modal energy vs Modal Period graph where each point represents the modal energy and period of an Intrinsic Mode. This is known as the EMD spectrum and is used for Colored Noise Analysis, for detecting Periodic Signatures in signals.

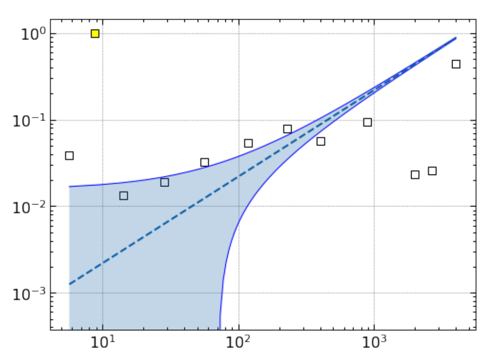


Figure 3

EMD Spectrum. Each point in this scatter plot represents an IMF. See that, a linear relationship can be observed indicating a power law like noise profile. The yellow dot represents the overall signal while the 1^{st} IMF lies above the significance intervals, candidate for QPO.

6 Fourier Spectra of IMFs

spectral shape and area coverage for each spectrum are identical, we can have an integral expression to represent, to the first order of approximation, the functional form of Fourier spectrum for any IMF (except the first one) as

$$\int S_{\ln T,n}\, d\ln T = const \tag{5}$$

$S_{lnT,n}$ is the Fourier Spectrum of the n^{th} IMF as a function of lnT

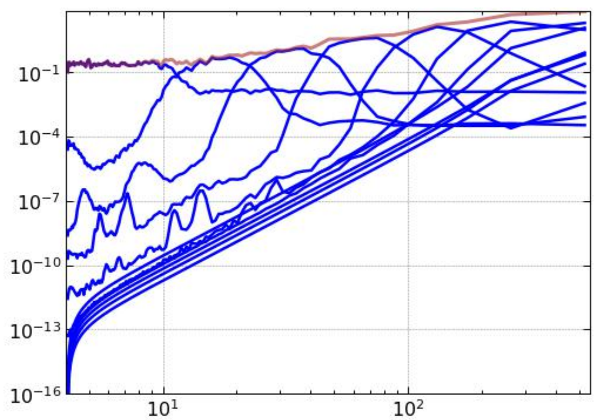


Figure 4: Welch's Periodogram of the IMFs showing their spectral content. As predicted the peaks of the spectrums are at increasing periods with IMFs. The first IMF has the left most peak. Higher IMFs do not show any particular peak spectrum and are purely noisy.

Energy of the n^{th} IMF is given by

$$NE_n = \int S_{\omega,n}\, d\omega \tag{6}$$

Through a series of variable changes, from frequency to period to logarithmic

$$NE_n = \int S_{\omega,n}\, d\omega = \int S_{T,n}\, \frac{dT}{T^2} = \int S_{\ln T,n}\, \frac{d\ln T}{T} = \frac{\int S_{\ln T,n}\, d\ln T}{\bar{T_n}} \tag{7}$$

Mean Period

$$\bar{T_n} = \left(\int S_{\ln T,n}\, d\ln T\right)\left(\int S_{\ln T,n}\, \frac{d\ln T}{T}\right)^{-1} \tag{8}$$

$\bar{T_n}$ is the definition of the average period calculated from any given spectrum identical to one obtained from zero crossings
Energy density E_n and averaged Period $\bar{T_n}$ relate as

$$E_n T_n = const \tag{9}$$

$$\ln E_n + \ln T_n = const \tag{10}$$

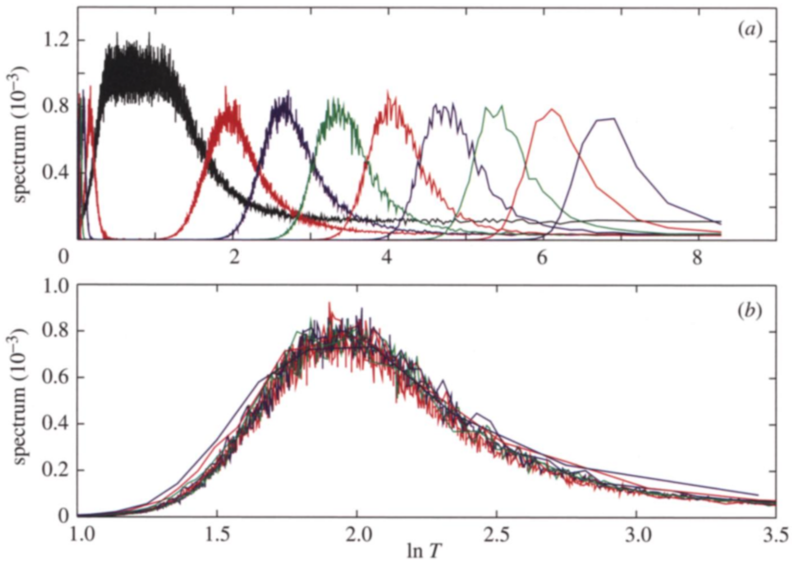


Figure 5

Energy density of the n^{th} IMF is given by

$$E_n = \frac{1}{N}\sum_{j=1}^{N}|C_n(j)|^2 \tag{11}$$

Probaility Distribution of NE_n is the χ^2 distribution with NE_n degrees of freedom

$$\rho(NE_n) = (NE_n)^{\frac{N\bar{E_n}}{2}-1}e^{\frac{-NE_n}{2}} \tag{12}$$

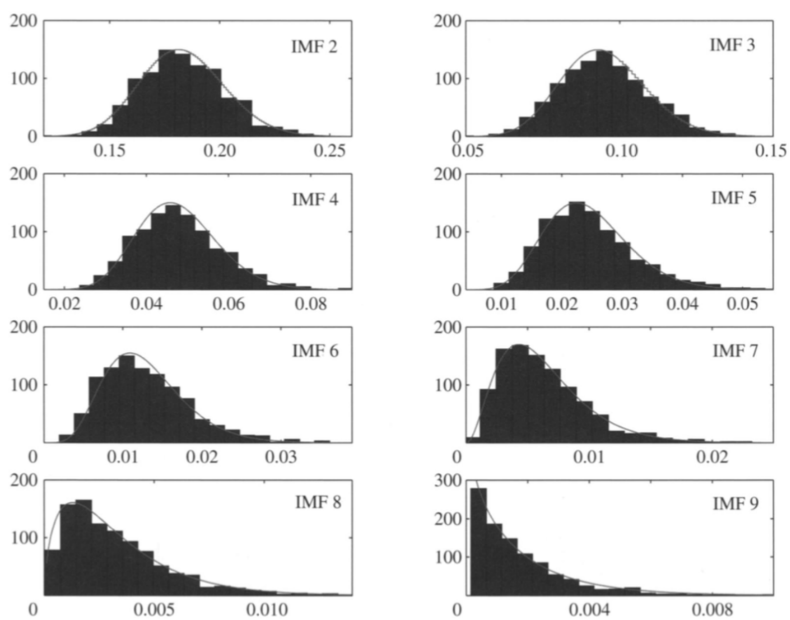


Figure 6: Histograms of the energy density for IMFs (modes) 2-9 for a white-noise sample with 50 000 data points. The superimposed grey lines are the X2 fits for each IMF.

Probability of E_n is given by

$$\rho(E_n) = N(NE_n)^{\frac{N\bar{E_n}}{2}-1}e^{\frac{-NE_n}{2}} \tag{13}$$

7 Spread of Energy

We introduce a new variable y as $y = \ln E$, hence the distribution of y :

$$\rho(y) = N(Ne^y)^{\frac{N\bar{E}}{2}-1} e^{-\frac{NE}{2}} e^y \tag{14}$$

$$= C \exp\left(\frac{1}{2} y N E_n - \frac{1}{2} N E\right) \tag{15}$$

$$= C \exp\left(-\frac{N\bar{E}}{2}\left(\frac{E}{\bar{E}-y}\right)\right) \tag{16}$$

where $C = N^{\frac{N\bar{E}}{2}}$
Since $E = e^y$

$$\frac{E}{\bar{E}} = e^{y-\bar{y}} = 1 + (y-\bar{y}) + \frac{1}{2!}(y-\bar{y})^2 + \frac{1}{3!}(y-\bar{y})^3 + ... \tag{17}$$

$$= C' \exp\left(-\frac{N\bar{E}}{2}\left[\frac{1}{2!}(y-\bar{y})^2 + \frac{1}{3!}(y-\bar{y})^3 + ...\right]\right) \tag{18}$$

$$\sigma^2 = \frac{2\bar{T}_n}{N} \tag{19}$$

The distribution of E_n is approximately a Gaussian with a standard deviation 0
The spread lines are defined as

$$y = x \pm k\sqrt{\frac{2}{N}} e^{\frac{x}{2}} \tag{20}$$

where $x = \ln T$, and k is a constant determined by percentiles of a standard normal distribution. For example, we will have k equal to $-2.326, -0.675, -0.0$ and 0.675 for the first, $25^{th}, 50^{th}, 75^{th}$ and 99^{th} percentiles, respectively.

8 Colored Noise Analysis

We assume power law fits for the background stochastic noise processes. And hence fit linear functions to the EMD spectrum using least squares. Most of the IMFs will be composed of pure noisy elements, however those lying above the confdence intervals can be considered to contain significant oscillatory signatures.

Thus,, we construct the 95% Confidence levels around the power law fits and thus can say with enough certainty that the modes lying outside this interval are significant.