Anomalous volatility scaling in high frequency financial data

Noemi Nava\textsuperscript{1,3}, T. Di Matteo\textsuperscript{2} Tomaso Aste\textsuperscript{1,3}

\textsuperscript{1} Department of Computer Science, University College London, Gower Street, London, WC1E 6BT, UK
\textsuperscript{2} Department of Mathematics, King’s College London, The Strand, London, WC2R 2LS, UK
\textsuperscript{3} Systemic Risk Centre, London School of Economics and Political Sciences, London, WC2A2AE, UK

Abstract

Volatility of intra-day stock market indices computed at various time horizons exhibits a scaling behaviour that differs from what would be expected from fractional Brownian motion (fBm). We investigate this anomalous scaling by using Empirical Mode Decomposition (EMD), a method which separates time series into a set of cyclical components at different time-scales. By applying EMD to fBm, we retrieve a scaling law that relates the variance of the components to a power law of the oscillating period. In contrast, when analysing 22 different stock market indices, we observe deviations from the fBm and Brownian motion scaling behaviour. These deviations are associated to the characteristics of financial markets, with the larger deviations corresponding to the less developed markets.

Keywords: Empirical Mode Decomposition, Hurst exponent, Multi-scaling, Market efficiency.

1. Introduction

Over the last few years, financial markets have witnessed the availability and widespread use of data sampled at high frequencies. The study of these data allows to identify the intra-day structure of financial markets (Dacorogna et al., 2001; Bartolozzi et al., 2007). Data at these frequencies have dynamic properties which are not generated by a single process but by several
components that are superimposed on top of each other. These components are not immediately apparent, but once identified, they can be meaningfully categorized as noise, cycles at different time-scales and trends (Dacorogna et al., 2001).

Since the early work of Mandelbrot (1963), it was recognized that different time-scales contribute to the complexity of financial time series in a self-similar (fractal) manner. Empirical properties of financial data at various frequencies have been observed in a number of studies, see for example Muller et al. (1990), Cont (2001), Di Matteo et al. (2003), Lillo and Farmer (2004), Glattfelder et al. (2011).

According to the Efficient Market Hypothesis (Fama, 1965), financial market dynamics can be described by a random walk, a self-similar process with scaling exponent (Hurst exponent) $H = 0.5$. Opposing this theory, Peters (1994) introduced the Fractal Market hypothesis (FMH) and described financial market dynamics by fractional Brownian motion (fBm), a self-similar process with scaling exponent $0 < H < 1$. The focus of this last theory is on the interaction of agents with various investment horizons and differing interpretations of information.

In self-similar uni-scaling process, such as fBm, all time-scales contribute in a proportional manner and there is a specific relation that links statistical properties at different time-scales (Feller, 1966). However, real financial time series have more complex scaling patterns, with some time-scales contributing disproportionately. These patterns characterize multi-scaling processes whose statistical properties vary at each time-scale (Di Matteo et al., 2003, Barunik et al., 2012).

The knowledge of scaling laws helps us to understand market dynamics and to construct efficient and profitable trading strategies. In this paper, we use Empirical Mode Decomposition (EMD), an algorithm introduced by Huang et al. (1998), to decompose intra-day financial data into a trend and a finite set of simple oscillations. These oscillations, called Intrinsic Mode Functions (IMFs), are associated with the time-scale of cycles latent in the time series. EMD provides a tool for an exploratory analysis that takes into account both the fine and coarse structure of data. This decomposition has been widely used in many fields, including the analysis of financial time series (Huang et al., 2003, Nava et al., 2014), Guhathakurta et al. (2008), river flow fluctuations (Huang et al., 2009), wind speed (Guo et al., 2012), hear rate variability (Balocchi et al., 2004), etc.

In this paper, we first apply EMD to fBm, uncovering a power law scaling
between the period and variance of the IMFs with scaling exponent related to the Hurst exponent. We then apply EMD to 22 different stock market indices whose prices are sampled at 30 sec intervals over a time span of 6 months. In this case, we encounter more complex scaling laws than in fBm. The deviations from the fBm behaviour are quantified and interpreted as an anomalous multi-scaling behaviour. We discuss the deviations from Brownian motion and fBm as measures of “market (in)efficiency”.

This paper is organized as follows. In section 2, we introduce EMD. In section 3, we present the variance scaling properties of fBm. In section 4, we present an application to high frequency financial data. We finally conclude in section 5.

2. Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is a fully data-driven decomposition that can be applied to non-stationary and non-linear data \cite{Huang1998}. Differently from Fourier and wavelet transform, EMD does not require any a priori filter function \cite{Peng2005}. The purpose of the method is to identify a finite set of oscillations with scale defined by the local maxima and minima of the data itself. Each oscillation is empirically derived from data and is referred as an Intrinsic Mode Function (IMF). An IMF must satisfy two criteria:

1. The number of extrema and the number of zero crossings must either be equal or differ at most by one.
2. At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

IMFs are obtained through a process that makes use of local extrema to separate oscillations starting with the highest frequency. Hence, given a time series \( x(t), t = 1, 2, ..., T \), the process decomposes it into a finite number of Intrinsic Mode Functions denoted as \( IMF_k(t), k = 1, ..., n \) and a residue \( r_n(t) \). If the decomposed data consist of uniform scales in the frequency space, the EMD acts as a dyadic filter and the total number of IMFs is close to \( n = \log_2(T) \) \cite{Flandrin2004}. The residue is the non-oscillating drift of the data. At the end of the decomposition process, the original time series can be reconstructed as:

\[
x(t) = \sum_{k=1}^{n} IMF_k(t) + r_n(t). \tag{1}
\]
The EMD comprises the following steps:

1. Initialize the residue to the original time series \( r_0(t) = x(t) \) and set the IMF index \( k = 1 \).

2. Extract the \( k^{th} \) IMF:
   
   (a) initialize \( h_0(t) = r_{k-1}(t) \) and the iteration counter \( i = 1 \);
   
   (b) find the local maxima and local minima of \( h_{i-1}(t) \);
   
   (c) create the upper envelope \( E_u(t) \) by interpolating between the maxima (lower envelope \( E_l(t) \) for minima, respectively);
   
   (d) calculate the mean of both envelopes as \( m_{i-1}(t) = \frac{E_u(t) + E_l(t)}{2} \);
   
   (e) subtract the envelope mean from input time series, obtaining \( h_i(t) = h_{i-1}(t) - m_{i-1}(t) \);
   
   (f) verify if \( h_i(t) \) satisfies the IMF’s conditions:
   
   • if \( h_i(t) \) does not satisfy the IMF’s conditions, increase \( i = i + 1 \) and repeat the sifting process from step (b);
   
   • if \( h_i(t) \) satisfies the IMF’s conditions, set \( IMF_k(t) = h_i \) and define \( r_k(t) = r_{k-1}(t) - IMF_k(t) \).

3. When the residue \( r_k(t) \) is either a constant, a monotonic slope or contains only one extrema stop the process, otherwise continue the decomposition from step 2 setting \( k = k + 1 \).

The IMFs are nearly orthogonal to each other and the total variance is approximated by the sum of the variance of the components and the variance of the residue. However, it must be noted that EMD is based on the time-scale and not on the demand of orthogonality. For some non-linear data, orthogonality may not be satisfied implying that in general the sum of the variance of the components and the residue is different than the variance of the original time series \( \text{[Huang et al. 1998]} \).

3. Self-similar scaling exponent

Self-similarity or scale invariance is an attribute of many laws of nature and is the underlying concept of fractals. Self-similarity is related to the occurrence of similar patterns at different time-scales. In this sense, probabilistic properties of self-similar processes remain invariant when the process is viewed at different time-scales \( \text{[Mandelbrot 1982, Mandelbrot et al. 1997, Calvet and Fisher 2002]} \).
A stochastic process \( X(t) \) is statistically self-similar, with Hurst exponent \( 0 < H < 1 \), if for any real \( a > 0 \) it follows the scaling law:

\[
X(at) \overset{d}{=} a^H X(t) \quad t \in \mathbb{R},
\]

where the equality \( (\overset{d}{=}) \) is in probability distribution. The parameter \( H \) is called the scaling exponent of the process.

An example of self-similar process is fractional Brownian motion (fBm), characterized by a positive scaling exponent \( 0 < H < 1 \). fBm can be considered a generalization of Brownian motion with independent increment process and \( H = \frac{1}{2} \). When \( 0 < H < \frac{1}{2} \), fBm is said to be anti-persistent with increments negatively auto-correlated. For the case \( \frac{1}{2} < H < 1 \) fBm reflects a persistent behaviour and the increments are positive auto-correlated.

### 3.1. EMD based scaling exponent

Flandrin and Gonçalves (2004) empirically showed that when decomposing fractional Gaussian noise (fGn) of Hurst exponent \( H > \frac{1}{2} \), EMD can be used to estimate this exponent. The authors ascertained that the variance progression across IMFs satisfies

\[
\text{var}(IMF_k^{fGn}) \propto \tau_k^{2(H-1)},
\]

where the function \( \tau_k \) denotes the period of the \( k^{th} \)-IMF.

In this paper we follow a similar approach to Flandrin and Gonçalves (2004), but instead of applying EMD to fGn, we considered its integrated process, fBm. We empirically showed that a similar scaling law holds for the variance of IMFs:

\[
\text{var}(IMF_k^{fBm}) \propto \tau_k^{2H}. \tag{3}
\]

Therefore, for fBm and fGn the IMF variance follows a power law scaling behaviour with respect to its particular period of oscillation, and the scaling parameter is related to the Hurst exponent. The EMD estimator of \( H \) can be determined by the slope of a linear regression fit on the logarithmic of the variance as a function of the logarithmic of the period,

\[
\log \left( \text{var}(IMF_k) \right) = 2H \log(\tau_k) + \log(c_0). \tag{4}
\]

\(^1\)The periods \( \tau_k \) can be approximated as the total number of data points divided by the total number of zero crossings of each IMF.
In the following section we will provide the simulation results supporting these findings.

3.2. fBm simulation analysis

In order to verify the empirical scaling law of Equation 3, we generated \( N = 100 \) fBm processes for different values of Hurst exponent \( H = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 \) and \( 0.9 \). For each exponent \( H \), we generated processes with two different lengths \( T_1 = 10,000 \) and \( T_2 = 100,000 \).

We applied EMD to each fBm and to its respective difference process, fGn. Figure 1 shows the mean of the estimated exponent \( H^* \) computed over the 100 simulations, left: length 10,000 and right: length 100,000. The error bars represent the Root Mean Square Error (RMSE) of the estimators, 

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{N} (H^*_i - H)^2}{N}}.
\]

We observe that the longer the analysed time series, the smaller the RMSE is. For length \( T = 10,000 \), the estimator \( H^* \) obtained through fBm is indeed very close to the Hurst exponent \( H \) (for all values of \( H \)). Conversely, the estimator through the decomposition of fGn deviates from the theoretical values when \( H < 0.5 \). This is consistent with the original observations made by Flandrin and Gonçalves (2004).

Figure 1: Mean of the scaling exponent \( H^* \) over 100 simulations of fBm (fGn) with parameter \( H = 0.1, 0.2, \ldots, 0.9 \) and length: left \( T_1 = 10,000 \) and right: \( T_2 = 100,000 \).

To visualize the linear relationship of Equation 4 in Figure 2 we explicitly show the relation between \( \log(Var(IMF_{k}^{fBm})) \) and \( \log(\tau_k) \) for a fBm.

---

\(^2\)All the fBm paths were generated using MATLAB\textsuperscript{\textregistered} wavelet toolbox.
simulation of Hurst exponent $H = 0.6$ and length $T_1 = 10,000$ points. In this example, the resulting estimator is $H^* = 0.593$ which accurately approximates the Hurst exponent of the simulated process.

![Figure 2: Log-log plot of IMF variance as a function of period for a fBm of $H = 0.6$ and length $T_1 = 10,000$. The blue line represents the least square fit of the dots. The scaling exponent $H^* = 0.593$ can be recovered from half the slope of the least square fit.](image)

**4. Variance scaling in intra-day financial data**

We analysed intra-day prices for 22 different stocks market indices. The data set, extracted from Bloomberg, covers a period of 6 months from May 5th, 2014 to November 5th, 2014. Prices are recorded at 30 second intervals. We excluded weekends and holidays. The number of working days and the number of points for every trading day depend on the opening hours of each stock exchange. The list of analysed stock market indices is reported in Table 1.

We applied EMD to the logarithmic price of each financial time series. For the sake of clarity, in this section we only focus on the decomposition of the S&P 500 index, but a similar analysis has been done for the other stock market indices. For the S&P 500 log-price time series, we extracted 17 IMFs and a residue that describe the local cyclical variability of the original signal and represent it at different time-scales. The original log-price time series and its IMFs are displayed in Figure 3. In this figure, we observe temporary clusters of volatility that characterize some of the components, for example
| Country     | Index    | Length |
|------------|----------|--------|
| Brazil     | BOVESPA  | 105,000|
| China      | SSE      | 60,480 |
| France     | CAC 40   | 136,080|
| Greece     | ASE      | 106,470|
| Hong Kong  | HSI      | 98,154 |
| Hungary    | BUX      | 122,880|
| Italy      | FTSE MIB | 133,056|
| Japan      | NIKKEI 225| 75,600 |
| Malaysia   | KLSE     | 115,320|
| Mexico     | IPC      | 100,620|
| Netherlands| AEX      | 130,680|
| Poland     | WIG      | 64,680 |
| Qatar      | DSM      | 52,080 |
| Russia     | RTSI     | 133,120|
| Singapore  | STI      | 123,840|
| South Africa| JSE   | 117,500|
| Spain      | IBEX     | 135,527|
| Turkey     | XU 100   | 91,760 |
| UAE        | UAED     | 60,000 |
| UK         | FTSE     | 130,560|
| USA        | S&P 500  | 99,840 |
| USA        | NASDAQ   | 100,620|

Table 1: Stock market indices including the length of the time series.

the high volatility at the end of the time series can evidently be seen in components 2,3,4,6 and 7.

The IMF periods, calculated as the total number of data points divided by the total number of zero crossings, are reported in Table 2. These periods are converted into minutes, hours and days. The fastest component has a cycle of 1.6 minutes, contrasting the slowest cycle of 11.6 days. Notice that the first 12 IMFs represent the intra-day activity (6.5 hours of trading), while the remaining IMFs (from 13th to 17th) are associated with the inter-day cycles. The last component is the residue of the EMD.
Figure 3: Top: Log-price time series of the S&P 500 index for the period 05/05/2014 to 05/11/2014. Bottom: The 17 IMFs and the residue obtained through EMD of the log-prices.
| IMF | Period/min | IMF | Period/hr | IMF | Period/days |
|-----|------------|-----|-----------|-----|-------------|
| 1   | 1.6        | 9   | 1.1       | 14  | 1.1         |
| 2   | 2.8        | 10  | 1.9       | 15  | 2.2         |
| 3   | 4.9        | 11  | 3.0       | 16  | 4.3         |
| 4   | 8.4        | 12  | 5.9       | 17  | 11.6        |
| 5   | 13.0       | 13  | 11.7      | 18  | Residue     |
| 6   | 19.3       |     |           |     |             |
| 7   | 28.8       |     |           |     |             |
| 8   | 41.7       |     |           |     |             |

Table 2: Period of the IMFs obtained from the S&P 500 index.

The overall trend of the time series is given by the residue, and each component can be seen as an oscillating trend of the previous component on a shorter time-scale. The effectiveness of EMD as a de-trending and smoothing tool is illustrated in Figure 4. In this figure, the original time series (blue line) is compared with a ‘trend’ (red line), calculated as the sum of the residue plus the last component.

Figure 4: Log-price time series of the S&P 500 index (blue line). The red line represents the “trend” of the data calculated as the sum of the residue plus the last IMF.

In previous section we discussed that for fBm, EMD produces a linear relationship between the logarithmic values of variance of the IMFs and its respective period of oscillation (Equation 4). We tested whether this relationship also holds for financial data. In Figure 5, we show the log-log plot of variance as a function of period for the IMFs obtained from the S&P 500 index (red diamonds). The scaling exponent can be computed as half the slope of the least square fit (red line) and it has a value of $H^* = 0.55$. The goodness of this linear fit was estimated by the coefficient of determination $^3$

$^3$This coefficient of determination is the square of correlation between the dependent
which is $R^2 = 0.992$. We can conclude that this index satisfies the linear relationship of Equation 4 but the scaling exponent, $H^* = 0.55$, is different than for Brownian motion, $H = 0.5$.

We performed the same analysis for the other stock indices, finding both significant deviations from Brownian motion ($H^* \neq 0.5$) and deviations from the scaling law of Equation 4. In Table 3, we report the estimated exponent and the goodness of the linear fit for every stock market index. In this table, we also include the number of IMFs that were obtained for each decomposition. Although the coefficients of determination are all above 0.94, we shall discuss shortly that significant deviations from linearity (fBm behaviour) are observed, especially in less developed markets.

Let us now discuss in more detail the deviations of the scaling laws found in stock markets from the scaling expected in Brownian motion (Bm). With this aim, we generated $N = 100$ paths of Bm with length $T$ equal to the analysed stock market index (see Table 1). We applied EMD to each simulation and obtained its respective intrinsic oscillations denoted as $IMF_k^{Bm}$, $i = 1, 2, \ldots, 100$, $k = 1, 2, \ldots, n_i$. In order to compare the variance of

![Log-log plot of IMF variance as a function of period for the EMD of the S&P 500 index. The red line represents the best least square linear fit. The goodness of the linear fit is $R^2 = 0.992$.](image)

and independent variable. Values of this coefficient range from 0 to 1, with 1 indicating a perfect fit between the data and the linear model, see for example [Rao (1973)](#).
Table 3: Stock market indices including the number of IMFs obtained when applying EMD to the logarithmic price. Indices are reported in descending order of $R^2$, which represents the goodness of the linear fit of Equation 4. Last column reports the estimated exponent $H^*$ of the same equation.

| Index     | No. of IMFs | $R^2$ | $H^*$ |
|-----------|------------|-------|-------|
| S&P 500   | 17         | 0.992 | 0.564 |
| BOVESPA   | 18         | 0.989 | 0.561 |
| FTSE MIB  | 18         | 0.987 | 0.571 |
| XU 100    | 19         | 0.985 | 0.563 |
| RTSI      | 20         | 0.985 | 0.581 |
| CAC 40    | 17         | 0.984 | 0.564 |
| UAED      | 16         | 0.978 | 0.616 |
| FTSE      | 23         | 0.977 | 0.529 |
| ASE       | 15         | 0.974 | 0.587 |
| IBEX      | 18         | 0.973 | 0.531 |
| WIG       | 16         | 0.973 | 0.591 |
| SSE       | 14         | 0.971 | 0.534 |
| DSM       | 18         | 0.971 | 0.618 |
| IPC       | 18         | 0.971 | 0.555 |
| BUX       | 19         | 0.970 | 0.542 |
| HSI       | 19         | 0.969 | 0.554 |
| AEX       | 21         | 0.968 | 0.558 |
| NASDAQ    | 20         | 0.960 | 0.530 |
| NIKKEI 225| 22         | 0.959 | 0.544 |
| JSE       | 19         | 0.956 | 0.518 |
| KLSE      | 22         | 0.943 | 0.540 |
| STI       | 21         | 0.942 | 0.522 |

For the rescaled $\hat{\text{var}}(\text{IMF}_{k}^{B_{mi}})$, we estimated the intercept constant $c_{o_i}$ of Equation 4, fixing $H = 0.5$. In Figure 6, we present all these 100 linear fits as light blue lines. In the same figure, we plotted the variance of the
IMFs extracted from the S&P 500 index, same as reported in Figure 5. We observe that the Brownian motion linear fits (blue lines) and the linear fit of the S&P 500 index (red line) are close to each other, suggesting an efficient behaviour in this market.

Figure 6: Log-log plot of variance as a function of period for the S&P 500 IMFs (red diamonds) compared with 100 rescaled Bm linear fits of slope $H^* = 0.5$ (blue lines).

The goodness of the linear fit between the financial data points (red diamonds) and each of the Brownian motion linear fit (blue lines) was calculated as follow:

$$ R_{Bm_i}^2 = 1 - \frac{\sum_{k=1}^{n} [\log (\text{var} (IMF^X_k)) - \log (c_i c_0 \tau^X_k)]^2}{\sum_{k=1}^{n} [\log (\text{var} (IMF^X_k)) - \langle \log (\text{var} (IMF^X_k)) \rangle]^2}. $$

(7)

The number of IMFs extracted from the stock index $X$ is denoted by $n$ and $\langle \cdot \rangle$ indicates the mean over these $n$ IMF variances.

To measure the deviation from Bm, we calculated the mean over the goodness of the linear fits, i.e. we calculated $\langle R_{Bm_i}^2 \rangle = \frac{1}{n} \sum_{i=1}^{100} R_{Bm_i}^2$. For the S&P 500 index, we obtained a coefficient equal to $\langle R_{Bm}^2 \rangle = 0.979$, demonstrating the similarity between the scaling properties of this financial index and Brownian motion.

4.1. Scaling properties of stock market indices

In previous section, we introduced two measures that quantify the deviations from the scaling behaviour of fractional Brownian motion and from
Brownian motion. These measures are given by:

1. $R^2$, coefficient of determination (square of correlation) between the logarithm of period and logarithm of variance of IMFs obtained from stock market indices.

2. $\langle R^2_{Bm} \rangle$, mean of the relative squared residuals between the IMF variances obtained from financial data and each of the linear fits for Brownian motion simulations.

In Table 4 we report the values of $\langle R^2_{Bm} \rangle$ for the 22 stock market indices. For comparison purposes, we repeated the $R^2$ values. The last column in this table indicates the ordering of the markets if $R^2$ were used as the ranking measure.

| Country  | Index     | $\langle R^2_{Bm} \rangle$ | Rank $\langle R^2_{Bm} \rangle$ | $R^2$ | Rank $R^2$ |
|----------|-----------|-----------------------------|----------------------------------|-------|------------|
| USA      | S&P 500   | 0.979                       | 1                                | 0.992 | 1          |
| Brazil   | BOVESPA   | 0.977                       | 2                                | 0.989 | 2          |
| UK       | FTSE      | 0.973                       | 3                                | 0.977 | 8          |
| Turkey   | XU 100    | 0.972                       | 4                                | 0.985 | 4          |
| Italy    | FTSE MIB  | 0.971                       | 5                                | 0.987 | 3          |
| France   | CAC 40    | 0.970                       | 6                                | 0.984 | 6          |
| Spain    | IBEX      | 0.969                       | 7                                | 0.973 | 10         |
| China    | SSE       | 0.967                       | 8                                | 0.971 | 12         |
| Russia   | RTSI      | 0.964                       | 9                                | 0.985 | 5          |
| Hungary  | BUX       | 0.963                       | 10                               | 0.970 | 15         |
| Mexico   | IPC       | 0.960                       | 11                               | 0.971 | 14         |
| Hong Kong| HSI       | 0.958                       | 12                               | 0.969 | 16         |
| USA      | NASDAQ    | 0.954                       | 13                               | 0.960 | 18         |
| Netherlands | AEX    | 0.953                       | 14                               | 0.968 | 17         |
| South Africa | JSE | 0.952                       | 15                               | 0.956 | 20         |
| Japan    | NIKKEI 225| 0.949                       | 16                               | 0.959 | 19         |
| Greece   | ASE       | 0.948                       | 17                               | 0.974 | 9          |
| Poland   | WIG       | 0.947                       | 18                               | 0.973 | 11         |
| UAE      | UAED      | 0.939                       | 19                               | 0.978 | 7          |
| Singapore| STI       | 0.934                       | 20                               | 0.942 | 22         |
| Malaysia | KLSE      | 0.933                       | 21                               | 0.943 | 21         |
| Qatar    | DSM       | 0.928                       | 22                               | 0.971 | 13         |

Table 4: Stock Market indices ranked in descending order of $\langle R^2_{Bm} \rangle$. The last column indicates the ordering of the markets with respect to $R^2$. 

14
We observe that the S&P 500 index (USA) is ranked the highest in both scales. Developed markets tend to be at the top of the table with some exceptions that may arise from the specific characteristics of the analysed period of time, May 5th, 2014 to November 5th, 2014.

In Figure 7, we plot the financial market ranking. The horizontal bars represent the 5th and 95th percentiles of the $R^{2}_{Bm}$ distribution. The blue dot inside each bar indicates the mean value $\langle R^{2}_{Bm} \rangle$ as reported in Table 4. We observe a clear tendency for developed markets to present larger values of $\langle R^{2}_{Bm} \rangle$ with narrower distributions.

![Figure 7: Percentiles 5th and 95th of the $R_{Bm}$ distribution for the analysed stock market indices. The blue dot inside each bar indicates the value of $\langle R^{2}_{Bm} \rangle$ used for the financial market ranking.](image)

In order to visualize the anomalous scaling in some stock markets and to understand the origin of the differences in the results, we present the cases of the NASDAQ (USA), BOVESPA (Brazil), NIKKEI 225 (Japan) and DSM
(Qatar) indices in more detail.

For the NASDAQ index (USA), we obtained 20 IMFs and a residue. In Figure 8(a) we present the log-price time series (blue line) and the “trend” consisting of the residue plus the last IMF (red line). In Figure 8(b) we observe that the deviation from the linear relationship of Equation 4 is significant. Thus, the log-log relationship between period and variance is not completely satisfied. The resultant coefficient of determination is $R^2 = 0.960$, ranking this index at the 18th position. Moreover, when compared with Bm, we identify that most of the components deviate from the Bm linear fits (blue lines). We also note that the total number of component (21) is considerable larger than what would be expected from a process with uniform scales, i.e., $\log_2(100620) = 16.6$. The presence of these extra oscillations with reduced variance suggests a more complex structure than Bm. The deviations from Bm model, quantified by the coefficient $\langle R^2_{Bm} \rangle = 0.958$, rank this index at the 13th position.

The variance scaling properties of the BOVESPA index (Brazil) are presented in Figure 9. For this stock index, the EMD identifies long period cycles with larger variance than what would be expected from Bm, see Figure 9(b). However, the linear fit between logarithmic value of IMF variances and periods is in general good with $R^2 = 0.989$. The goodness of the linear fit between the Bm simulations is $\langle R^2_{Bm} \rangle = 0.977$, placing this index at the
second position. Such a good ranking for this market may be unexpected, but we must stress that it only reflects the six-month period of observations. From Figure 9, we can see that this was a rather random but calm period.

Figure 9: EMD analysis for the BOVESPA index. Captions for figures (a) and (b) are the same as figures (4) and (6) respectively.

For the NIKKEI 225 index (Japan), we obtained 22 IMFs and a residue. Similar as the NASDAQ index, the number of components is considerably larger than what would be expected from $B_m$, i.e., $\log_2(75600)=16.2$. These many oscillations, specially the high frequency components, generate a non-linear behaviour that deviates from $B_m$. Given this anomalous scaling, we obtained $\langle R^2_{B_m} \rangle = 0.949$, ranking this index at the 16th position.
Finally, the DSM stock index (Qatar) is displayed in Figure 11. The log-price time series and its respective “trend” are displayed in Figure 11(a). In Figure 11(b) we observe the poor liner fit of Equation 4 that is characterized by a considerable steep slope. We obtained $R^2 = 0.971$, ranking this index at the lowest position. Furthermore, if we compare its IMF variances against the Bm linear fits, we observe that most of the variances values (red diamonds) follow outside the band expected from Bm. The large variance of the low frequency components suggests the presence of important long period cycles. Given its deviations from Bm behaviour, this index is also ranked the lowest with respect to the measure $\langle R^2_{Bm} \rangle = 0.928$. 

Figure 10: EMD analysis for the NIKKEI 225 index. Captions for figures (a) and (b) are the same as figures (4) and (6) respectively.
5. Conclusions

We explored the scaling properties of Empirical Mode Decomposition (EMD), an algorithm that de-trends and separates time series into a set of oscillating components (IMFs) associated with specific time-scales. We empirically showed that fractional Brownian motion (fBm) obeys a scaling law that relates linearly the logarithm of the variance and the logarithm of the period of IMFs. For fBm, the extracted coefficient of proportionality equals the Hurst exponent multiplied by two. When applied to stock market indices, EMD reveals instead different scaling laws that can deviate significantly from both Brownian motion and fractional Brownian motion behaviour. In particular, we noted that the EMD of high frequency financial data results in a larger number of IMFs than what would be expected from Brownian motion. Moreover, sometimes the linearity in the log-log relation between IMF variance and period is not satisfied. This is a direct indication of anomalous scaling that reveals a more complex structure in financial data than in self-similar processes.

In this study, we applied EMD to 22 different stock indices and observed that developed markets (European and North American markets) tend to have a behaviour closer to Brownian motion scaling properties. Conversely, larger deviations from uni-scaling laws are observed in some emerging markets such as Malaysian and Qatari. These findings are in agreement with
the discernible characteristics of developed and emerging markets, the former type being more likely to exhibit an efficient behaviour, see for example Di Matteo et al. (2005, 2003). Compared to previous approaches, the EMD method has the advantage to directly quantify the cyclical components with strong deviations, giving a further instrument to understand the origin of market inefficiencies.

Acknowledgements

The authors wish to thank Bloomberg for providing the data. NN would like to acknowledge the financial support from Conacyt-Mexico. TDM wishes to thank the COST Action TD1210 for partially supporting this work. TA acknowledges support of the UK Economic and Social Research Council (ESRC) in funding the Systemic Risk Centre (ES/K002309/1).

References

Balocchi, R., Menicucci, D., Santarcangelo, E., Sebastiani, L., Gemignani, A., Ghelarducci, B., Varanini, M., 2004. Deriving the respiratory sinus arrhythmia from the heartbeat time series using empirical mode decomposition. Chaos, Solitons & Fractals 20 (1), 171 – 177, towards a new vision of complexity.

Bartolozzi, M., Mellen, C., Di Matteo, T., Aste, T., 2007. Multi-scale correlations in different futures markets. The European Physical Journal B 58 (2), 207–220.

Barunik, J., Aste, T., Di Matteo, T., Liu, R., 2012. Understanding the source of multifractality in financial markets. Physica A: Statistical Mechanics and its Applications 391 (17), 4234 – 4251.

Calvet, L., Fisher, A., 2002. Multifractality in asset returns: Theory and evidence. The Review of Economics and Statistics 84 (3), 381–406.

Cont, R., 2001. Empirical properties of asset returns: stylized facts and statistical issues. Quantitative Finance 1, 223–236.

Dacorogna, M. M., Gençay, R., Müller, U., Olsen, R. B., Pictet, O. V., 2001. An introduction to high frequency finance. Academic Press, San Diego.
Di Matteo, T., Aste, T., Dacorogna, M., 2003. Scaling behaviors in differently
developed markets. Physica A: Statistical Mechanics and its Applications
324 (1), 183–188.

Di Matteo, T., Aste, T., Dacorogna, M., 2005. Long-term memories of de-
veloped and emerging markets: Using the scaling analysis to characterize
their stage of development. Journal of Banking & Finance 29 (4), 827 –
851.

Fama, E. F., 1965. The behavior of stock-market prices. The Journal of
Business 38 (1), 34–105.

Feller, W., 1966. An introduction to probability theory and its applications.
Vol. II, 2nd Edition. John Wiley & Sons Inc., New York.

Flandrin, P., Gonçalves, P., 2004. Empirical mode decompositions as data-
driven wavelet-like expansions. Int. J. of Wavelets, Multiresolution and
Information Processing 2 (4), 477–496.

Flandrin, P., Rilling, G., Goncalves, P., Feb 2004. Empirical mode decompo-
sition as a filter bank. Signal Processing Letters, IEEE 11 (2), 112–114.

Glattfelder, J. B., Dupuis, A., Olsen, R. B., 2011. Patterns in high-frequency
fx data: discovery of 12 empirical scaling laws. Quantitative Finance 11 (4),
599–614.

Guhathakurta, K., Mukherjee, I., Chowdhury, A. R., 2008. Empirical mode
decomposition analysis of two different financial time series and their com-
parison. Chaos, Solitons & Fractals 37 (4), 1214 – 1227.

Guo, Z., Zhao, W., Lu, H., Wang, J., 2012. Multi-step forecasting for wind
speed using a modified emd-based artificial neural network model. Renewable
Energy 37 (1), 241 – 249.

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., 1998. The empirical mode decomposition
and the Hilbert spectrum for nonlinear and non-stationary time series anal-
ysis. Proceedings of the Royal Society of London. Series A: Mathematical,
Physical and Engineering Sciences 454 (1971), 903–995.
Huang, N. E., Wu, M.-L., Qu, W., Long, S. R., Shen, S. S. P., 2003. Applications of Hilbert-Huang transform to non-stationary financial time series analysis. Appl. Stochastic Models Bus. Ind. 19 (3), 245–268.

Huang, Y., Schmitt, F. G., Lu, Z., Liu, Y., 2009. Analysis of daily river flow fluctuations using empirical mode decomposition and arbitrary order hilbert spectral analysis. Journal of Hydrology 373 (12), 103 – 111.

Lillo, F., Farmer, J. D., 2004. The long memory of the efficient market. Studies in Nonlinear Dynamics & Econometrics 8 (3), 1–35.

Mandelbrot, B., 1963. The Variation of Certain Speculative Prices. The Journal of Business 36, 394.

Mandelbrot, B., Fisher, A., Calvet, L., 1997. A multifractal model of asset returns (1164).

Mandelbrot, B., Taylor, H. M., 1967. On the distribution of stock price differences. Operations Research 15 (6), 1057–1062.

Mandelbrot, B. B., 1982. The fractal geometry of nature. Freeman.

Muller, U. A., Dacorogna, M. M., Olsen, R. B., Pictet, O. V., Schwarz, M., Morgenegg, C., December 1990. Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis. Journal of Banking & Finance 14 (6), 1189–1208.

Nava, N., Di Matteo, T., Aste, T., 2014. Time-dependent scaling patterns in high frequency financial data, submitted.

Peng, Z., Tse, P. W., Chu, F., 2005. A comparison study of improved Hilbert-Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing. Mechanical Systems and Signal Processing 19 (5), 974 – 988.

Peters, E. E., 1994. Fractal Market Analysis: Applying Chaos Theory to Investment and Economics. Wiley, New York.

Rao, C. R., 1973. Linear statistical inference and its applications. Wiley.