CHAPTER 4
RESEARCH METHODOLOGY

4.1 Introduction

In the previous chapter we outlined the theoretical framework of the models under study. In this chapter we discuss in detail the methodology used for empirical analysis of the models under test. In the process we also describe the theoretical and mathematical framework of each of the methods used. Most of what follows are summary of some seminal works by notable scientists responsible for the development of these methods. A very detailed study and understanding of the methods are essential in conducting the empirical tests and to achieve meaningful results. The chapter is designed as follows. First we delineate the research methodology that we have adopted wherein we outline what all tests we have performed. Next we devote different sections for different methods. Section[4.2.1] describes the tests for Nonlinearity. Section[4.2.2] elaborates the Cross-correlation method while section[4.2.3] deals with Fourier analysis. Section[4.2.4] and[4.2.5] explain the Empirical Mode Decomposition Method and Recurrence Plot Method.

4.2 Outline of Methodology

Based on the data from Nifty & SENSEX we do a Monte Carlo simulation of first the classical Geometric Brownian Motion. We use the discrete time version of the model:
\[
\frac{\delta S}{S} = \mu \delta t + \sigma \varepsilon \sqrt{\delta t} 
\]

[4.2.1]

Or

\[
\delta S = \mu S \delta t + \sigma S \varepsilon \sqrt{\delta t} 
\]

[4.2.2]

The variable \(\delta S\) is the change in the stock price \(S\) in a small time interval \(\delta t\) (in our case 1 day), and \(\varepsilon\) is a random drawing from a standardized normal distribution. The parameter \(\mu\) is the expected rate of return per annum and \(\sigma\) is the volatility of the stock expressed annually.

Then we do a Monte Carlo simulation of the Borland proposed model.

\[
\delta S = \mu S \delta t + \sigma S \omega \sqrt{\delta t} 
\]

[4.2.3]

where \(\omega\) is a random drawing from Tsallis distribution, all other parameters being same as eqn. [4.2.1] & [4.2.2].

For simulation we used MATLAB platform. We have given our code at the end of chapter 7 in Appendix I. For generating q-Gaussian random numbers we used the Box-Muller generalised function developed by Thristleton and Marsh.
We perform a series of statistical tests to test for nonlinearity in the data sets. We first use the test method developed by Brock, Dechert, and Scheinkman (BDS) and test for nonlinearity in each of the time series. Additionally we also perform the Keenan’s test for nonlinearity. Another popular non-linear test is the Hinich bispectrum test, which involves estimating the bispectrum of the observed time series. We also use this test to find out whether it detects nonlinearity in these time series. To reinforce our findings we also conduct the White’s neural Network tests on the same data set. Another linearity test for time series was introduced based on concepts from the theory of neural networks. Teräsvirta et al. developed its power fully. We use this Terasvirta Neural Network test as a final reinforcement of our findings.

We then use the Cross-Correlation study and Power spectrum analysis to find out the linear relationship between the original and the simulated series respectively. Since the nonlinearity in the data sets are established we then examine the nonlinear dynamic properties of each of the simulated time series using Empirical Mode Decomposition and Recurrence Analysis. We now give the theoretical background and framework of all the methods

---

4.2.1 Tests for Nonlinearity

It is an accepted fact that financial economics has been dominated over the past decade by linear paradigm, which assumes that economic time series conform to linear models or can be well approximated by a linear model. For example, empirical tests of market efficiency, purchasing power parity, tests of stationarity, cointegration, causality and many of the empirical models of asset pricing have implicitly assumed that the underlying dynamics are in linear form or can be made linear by a simple transformation.

However, there is ample empirical evidence against the linear paradigm. Theoretically, there is no reason to believe that economic systems must be intrinsically linear (see, for example, Pesaran and Potter, 1993\textsuperscript{249}; Campbell et al., 2007\textsuperscript{250}; Barnett and Serletis, 2000\textsuperscript{251}. Empirically, there were a great number of studies showing that financial time series exhibit non-linear dependencies (see, for example, Hsieh, 1989\textsuperscript{252}, 1991\textsuperscript{253}; Scheinkman and LeBaron, 1989\textsuperscript{254}, De Grauwe et al., 1993\textsuperscript{255}, Abhyankar et al., 1995\textsuperscript{256}, Steurer,
1995257, Brooks, 1996258, Barkoulas and Travlos, 1998259, Opong, et al., 1999260). With this development, the subject has now moved to a new direction. This new direction is, of course, the study of non-linearity. In the words of Campbell et al. (2007: 467)261, “A natural frontier for financial econometrics is the modelling of non-linear phenomenon”. Testing for non-linearity has become extremely popular in the financial econometrics literature in recent years, though the focus is on financial markets of developed countries. In principle, testing for non-linearity can be viewed as general test of model adequacy for linear models (Hinich and Patterson, 1989)262 and it has been argued that if the underlying generating process for a time series is non-linear in nature, then it would be inappropriate to employ linear methods. For instance, most of the widely applied statistical tests like the unit root or stationary tests, the Granger causality test and the cointegration test are all build on the basis of linear autoregressive model. Taylor and Peel (1997)263 and Sarno (2000)264, amongst others, illustrated that the adoption of linear stationarity tests are inappropriate in detecting mean reversion if the true data generating process is in fact a stationary nonlinear

258 Brooks, C. (1996), op cit
261 Campbell, J.Y., Lo, A.W. and MackInlay, A.C. (2007) op cit, pp 467
process. On the other hand, the Monte Carlo simulation evidence in Bierens (1997)\textsuperscript{265} indicated that the standard linear cointegration framework presents a mis-specification problem when the true nature of the adjustment process is non-linear and the speed of adjustment varies with the magnitude of the disequilibrium. Thus, if the underlying process of a time series is indeed non-linear in nature, one would have to resort to empirical methods like non-parametric cointegration test due to Bierens (1997)\textsuperscript{266}, non-linear stationarity tests (Sarno, 2001\textsuperscript{267}; Chortareas et al., 2002\textsuperscript{268}; Kapetanios et al., 2003\textsuperscript{269}) and non-linear causality test (Baek and Brock, 1992\textsuperscript{270}). To sum, testing for non-linearity is gaining popularity among researchers as a preliminary diagnostic tool to determine the nature of the data generating process before any further empirical analysis.

Over the past few decades, numerous studies have documented the existence of nonlinear dependencies in exchange rates returns series (see, for example, Hsieh, 1989\textsuperscript{271}; De Grauwe et al., 1993\textsuperscript{272}; Steurer, 1995\textsuperscript{273}; Brooks, 1996\textsuperscript{274}).

\textsuperscript{266} Ibid
\textsuperscript{267} Ibid
\textsuperscript{271} Hsieh, D.A. (1989) \textit{op cit}
\textsuperscript{273} Steurer, E. (1995) \textit{op cit}
\textsuperscript{274} Brooks (1996) \textit{op cit}
The stock markets have also attracted the attention of researchers with substantial evidence supporting the presence of non-linearity in stock returns series (see, for example, Scheinkman and LeBaron, 1989\textsuperscript{275}; Hsieh, 1991\textsuperscript{276}; Abhyankar et al.\textsuperscript{1995, 1997}; Barkoulas and Travlos, 1998\textsuperscript{277}; Opong, et al., 1999\textsuperscript{280}). However, much of this evidence has been drawn from the widely traded financial markets of well-developed countries. Not much work has been done to test the linearity in the Indian market in this direction. In this study we wish to bridge this gap.

We take the time series representing the daily close value of the indices as our input value for the tests. We first use the test method developed by Brock, Dechert, and Scheinkman (BDS) (Brock, et al., 1996)\textsuperscript{281} and test for nonlinearity in each of the time series. Additionally we also perform the Keenan’s test for nonlinearity. The BDS test does not provide a direct test for non-linearity because the sampling distribution of the BDS test statistic is not known, either in finite samples or asymptotically, under the null hypothesis of non-linearity. The rejection of the null of independent and identical distribution (IID.) in the BDS test can be due to non-white linear and non-white non-linear dependence in the data.

Thus, the effects of linear serial dependencies have to be filtered out by fitting the

\textsuperscript{275} Scheinkman, J. and LeBaron, B. (1989)
\textsuperscript{276} Hsieh, D.A. (1991) \textit{op cit}
\textsuperscript{277} Abhyankar, A.H., Copeland, L.S. and Wong, W. (1995) \textit{op cit}
\textsuperscript{279} Barkoulas, J. and Travlos, N. (1998) \textit{op cit}
\textsuperscript{280} Opong, K.K., Mulholland, G., Fox, A.F. and Farahmand, K. (1999) \textit{op cit}
\textsuperscript{281} Brock et al (1996) \textit{op cit}
best possible linear model before the BDS test can be applied to detect any non-linear departure from the IID null. However, there is always the concern that the rejection of the null by the BDS test could be due to the possibility of imperfect pre-whitening. This concern is well directed since much of the Monte Carlo research that has been published on the BDS test has emphasized the pre-testing issue and the potential dependence of the properties of the test on the prior linear filter. Some of the test’s sensitivity to non-linearity could be a result of remaining linear dynamics in the data. Another popular non-linear test is the Hinich bispectrum test (Hinich, 1982), which involves estimating the bispectrum of the observed time series (see, for example, De Grauwe et al., 1993; Abhyankar et al., 1995; Brooks, 1996; Vilasuso and Cunningham, 1996). We also use this test to find out whether it detects nonlinearity in these time series. Unlike the BDS test, the Hinich bispectrum test provides a direct test for a non-linear generating mechanism, irrespective of any linear serial dependencies that might be present. Thus, pre-whitening is not necessary in using the Hinich approach. Even if pre-whitening is done anyway, the adequacy of the pre-whitening is irrelevant to the validity of the test. Ashley et al. (1986) presented an equivalence theorem to prove that the Hinich linearity test statistic is invariant to linear filtering of the data, even if the filter is estimated.

---

282 Hinich, M.J. (1982), op cit
284 Abhayankar (1995) op cit
285 Brooks (1996) op cit
286 Vilasuso, J. and S. Cunningham (1996), op cit
287 Ashley, R.A., Patterson, D.M. and Hinich, M.J. (1986), op cit
Thus, the linearity test can be applied to the original returns series, or to the residuals of a linear model with no loss of power.

To reinforce our findings we also conduct the White’s neural Network (T. H. Lee, H. White, and C. W. J. Granger (1993)) tests on the same data set. Lee, White, and Granger (1993) develop a new test, the neural network test for neglected nonlinearity. The neural network test is based on the approximating ability of neural network modelling techniques recently developed by cognitive scientists. This test is a Lagrange multiplier test that statistically determines whether adding ‘hidden units’ to the linear network would be advantageous. This is a powerful test and can be used to determine nonlinearity in a data set quite clearly. This test is further developed and investigated by Teräsvirta and Granger(1993). They are compared by simulation with those of a Lagrange multiplier (LM) type test that they derive from the same single-hidden-layer neural network model. The auxiliary regression of their LM type test is a simple cubic ‘dual’ of the Volterra expansion of the original series, and the power of the test appears superior overall to that of the other test.

We performed all the tests using MATLAB and R software.

4.2.1.1. Theoretical background of tests

We now discuss the background of the tests that we perform on our data sets. We start by trying to understand the econometric definition of nonlinearity. The next few paragraphs under this section give a brief introduction to the theoretical framework and explain the utility of the tests for nonlinearity.

4.2.1.1 Nonlinearity

In the literature, there is no generally agreed definition for 'non-linearity'. From the definition given by De Grauwe et al. (1993: 244)\textsuperscript{290}, a system \( X_t = h(\Omega_t, \alpha) \), is called a non-linear system if it is not possible to regenerate \( X_t \) by one linear model:

\[
X_t = \sum_{i=0}^{\infty} \gamma_i \varepsilon_{t-i} \quad \text{and} \quad \varepsilon \quad \text{is white noise and}
\]

\[
\sum_{i=0}^{\infty} \gamma_i \quad \text{is such that} \quad \sum_{i=0}^{\infty} |\gamma_i| < \infty
\]

According to De Grauwe et al. (1993)\textsuperscript{291}, the definition of non-linearity stems from the negation of linearity. This leaves a lot of other possibilities open for a so-called non-linear system. For example, Hsieh (1989)\textsuperscript{292} divided the realm of non-linear dependencies into three categories. Additive non-linearity, also known as non-linear-in-mean, enters a process through its mean or expected value, so that each element in the sequence can be expressed as the sum of zero-mean random element and a non-linear function of past elements. With multiplicative nonlinearity, or non-linear-in-variance, each element can be expressed as the

\textsuperscript{290} De Grauwe, P., Dewachter, H. and Embrechts, M. (1993), op cit
\textsuperscript{291} Ibid
\textsuperscript{292} Hsieh(1989), op cit
product of a zero-mean random element and a non-linear function of past elements, so that the non-linearity affects the process through its variance. The final category is known as hybrid dependence, in which non-linearity enters through both the mean and the variance.

Tsay(2005)\textsuperscript{293} has considered nonlinearity in the following way.

Let us consider a univariate time series \( x_t \), which, for simplicity, is observed at equally spaced time intervals. We denote the observations by \( \{ x_t \mid t = 1 \ldots T \} \), where \( T \) is the sample size. A purely stochastic time series \( x_t \) is said to be linear if it can be written as

\[
x_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i} \tag{4.2.1.1.1.1}
\]

where \( \mu \) is a constant, \( \psi_i \) are real numbers with \( \psi_0 = 1 \), and \( \{ a_t \} \) is a sequence of independent and identically distributed (IID) random variables with a well defined distribution function. We assume that the distribution of \( a_t \) is continuous and \( E(a_t) = 0 \). In many cases, we further assume that \( \text{Var}(a_t) = \sigma_a^2 \) or, even stronger, that \( a_t \) is Gaussian. If \( \sigma_a^2 \sum_{i=0}^{\infty} \psi_i^2 < \infty \), then \( x_t \) is weakly stationary (i.e., the first two moments of \( x_t \) are time-invariant). Any stochastic process that does not satisfy the condition of Eq. (4.2.1.1.1) is said to be nonlinear. The prior definition of nonlinearity is for purely stochastic time series. One may extend the definition by allowing the mean of \( x_t \) to be a linear function of some exogenous variables, including the time index and some periodic functions.

\textsuperscript{293} Tsay, Ruey(2005), op cit, pp 183-185
Under the null hypothesis of linearity, residuals of a properly specified linear model should be independent. Any violation of independence in the residuals indicates inadequacy of the entertained model, including the linearity assumption. This is the basic idea behind various nonlinearity tests. In particular, some of the nonlinearity tests are designed to check for possible violation in quadratic forms of the underlying time series.

4.2.1.1.2 Brock-Dechert-Schienkman (BDS) Test

Brock, Dechert, and Scheinkman (1996) propose a test statistic, commonly referred to as the BDS test, to detect the IID assumption of a time series. The statistic is, therefore, different from other test statistics discussed because the latter mainly focus on either the second- or third-order properties of \( x_t \). The basic idea of the BDS test is to make use of a “correlation integral” popular in chaotic time series analysis. Given a \( k \)-dimensional time series \( X_t \) and observations \( \{X_t\}_{t=1}^{T_k} \), define the correlation integral as

\[
C_k(\delta) = \lim_{T_k \to \infty} \frac{2}{T_k(T_k - 1)} \sum_{i<j} I_\delta(X_i - X_j) \tag{4.2.1.1.2.1}
\]

where \( I_\delta(u, v) \) is an indicator variable that equals one if \( \|u - v\| < \delta \), and zero otherwise, where \( \|\cdot\| \) is the supnorm. The correlation integral measures the fraction of data pairs of \( \{X_t\} \) that are within a distance of \( \delta \) from each other. Next a time series \( x_t \) is considered. \( k \)-dimensional vectors \( X^k_t = (x_t, x_{t+1}, \ldots, x_{t+k-1}) \) which are called \( k \)-histories are constructed. The idea of the BDS test is as

follows. A k-history is treated as a point in the k-dimensional space. If \([X_t]_{t=1}^{T_k}\) are indeed iid random variables, then the k-histories \([X_t]_{t=1}^{T_k}\) should show no pattern in the k-dimensional space. Consequently, the correlation integrals should satisfy the relation \(C_k(\delta) = [C_1(\delta)]^k\). Any departure from the prior relation suggests that \(X_t\) are not IID. As a simple, but informative example, let us consider a sequence of IID random variables from the uniform distribution over \([0, 1]\). Let \([a, b]\) be a subinterval of \([0, 1]\) and consider the “2-history” \((x_t, x_t + 1)\), which represents a point in the two-dimensional space. Under the IID assumption, the expected number of 2-histories in the subspace \([a, b] \times [a, b]\) should equal the square of the expected number of \(X_t\) in \([a, b]\). This idea can be formally examined by using sample counterparts of correlation integrals. We define

\[
C_l(\delta, T) = \lim_{T_k \to \infty} \frac{2}{T_k(T_k - 1)} \sum_{i<j} I_\delta(X_i^* - X_j^*), \quad l = 1, k \tag{4.2.1.1.2.2}
\]

where \(T_1 = T - 1 + 1\) and \(X_i^* = X_i\) if \(l = 1\) and \(X_i^* = X_i^k\) if \(l = k\). Under the null hypothesis that \([X_t]\) are IID with a nondegenerated distribution function \(F(.)\), Brock, Dechert, and Scheinkman (1996)\(^{295}\) show that

\[
C_k(\delta, T) \to [C_1(\delta)]^k \quad \text{with probability } 1, \quad \text{as } T \to \infty
\]

for any fixed \(k\) and \(\delta\). Furthermore, the statistic \(\sqrt{T} \{C_k(\delta, T) - [C_1(\delta, T)]^k\}\) is asymptotically distributed as normal with mean zero and variance

\[
\sigma_k^2(\delta) = 4 \left( N_k^k + 2 \sum_{j=1}^{k-1} N_k^{k-j} C_j^2 + (k - 1)^2 C_k^k - k^2 NC_k^{k-2} \right) \tag{4.2.1.1.2.3}
\]

\(^{295}\) Brock, W.A., Dechert, W.D., Scheinkman, J.A. and LeBaron, B. (1996) op cit
where \( C = \int \{ F(z + \delta) - F(z - \delta) \} dF(z) \) and \( N = \int \{ F(z + \delta) - F(z - \delta) \}^2 dF(z) \)

It is to be noted that \( C_1(\delta, T) \) is a consistent estimate of \( C \) and \( N \) can be consistently estimated by

\[
N(\delta, T) = \frac{6}{T_k(T_k-1)(T_k-2)} \sum_{1 \leq s < u} I_\delta(x_t, x_s) I_\delta(x_s, x_u)
\]  

[4.2.1.1.2.4]

The BDS test statistic is then defined as

\[
D_k(\delta, T) = \sqrt{T} \{ C_k(\delta, T) - [C_1(\delta, T)]^k \} / \sigma_k(\delta, T),
\]  

[4.2.1.1.2.5]

where \( \sigma_k(\delta, T) \) is obtained from \( \sigma_k(\delta) \) when \( C \) and \( N \) are replaced by \( C_1(\delta, T) \) and \( N(\delta, T) \), respectively. This test statistic has a standard normal limiting distribution.

For further discussion and examples of applying the BDS test, see Hsieh (1989)\(^{296}\) and Brock, Hsieh, and LeBaron (1991)\(^{297}\).

4.2.1.1.3 Hinich Bi-Spectrum Test

Hinich (1982) laid out a statistical test for determining whether an observed stationary time series \( (y_t) \) is linear. It is possible that \( y_t \) is linear without being Gaussian, but all of the stationary Gaussian time series are linear. The Hinich bispectrum test involves estimating the bispectrum of the observed time series, which is the double Fourier transform of the third order cumulant function.

\(^{296}\) Hsieh (1989) \textit{op cit}
\(^{297}\) Brock, W.A., Hsieh, D.A. and LeBaron, B. (1991) \textit{op cit}
In this section, we present a brief description of the testing procedures presented by Hinich (1982). Let $y_t$ denote a third order stationary time series, where the time unit, $t$, is an integer.

The third-order cumulant function of $y_t$ is defined to be

$$C_{yy}(m,n) = E[(y_{t+m}y_{t+n}y_t)]$$

for each $(m,n)$ when $E[y_t] = 0$, in which $n < m$ and $m = 0, 1, 2, \ldots$

Since third-order cumulants are difficult to interpret, and their estimates are even difficult to fathom, the double Fourier transform of the third-order cumulant function (called the bispectrum) is calculated.

The bispectrum at frequency pair $(f_1, f_2)$ is the double Fourier transform of $C_{yy}(m,n)$:

$$B_y(f_1, f_2) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} C_{yy}(m,n) \exp[-i2\pi(f_1m + f_2n)]$$

assuming that $|C_{yy}(m,n)|$ is summable. The symmetries of $C_{yy}(m,n)$ translate into symmetries of $B_y(f_1, f_2)$ that yield a principal domain for the bispectrum, which is the triangular set $\Omega = \left\{ 0 < f_1 < \frac{1}{2}, f_1 < f_2, 2f_1 + f_2 < 1 \right\}$.

Since the spectrum of $y_t$ is $S_y(f) = \sigma_y^2 |A(f)|^2$

$$\psi^2(f_1, f_2) = \frac{|B_y(f_1, f_2)|^2}{S_y(f_1)S_y(f_2)S_y(f_1 + f_2)} = \frac{\sigma_y^2}{\sigma_u^2}$$

298 Hinich (1982) op cit
for all \( f_1 \) and \( f_2 \) in \( \Omega \), where \( A(f) = \sum_{n=0}^{\infty} a(n) \exp(-i2\pi f_n) \)

The left hand side of Equation (4) defines the square of the skewness function of \( \{ y_t, \psi(f_1, f_2) \} \). Linearity and Gaussianity of \( y_t \) are tested through the null hypotheses that \( \psi(f_1, f_2) \) is constant over all frequencies and that \( \psi(f_1, f_2) \) is zero over all frequencies respectively using the estimated bispectrum.

The test statistics for both hypotheses are reduced to:

\[
\hat{S} = 2 \sum_m \sum_n |\hat{y}_{m,n}|^2
\]  \[\text{[4.2.1.1.3.3]}\]

at the frequency pair \((m, n)\) where:

\[
\hat{y}_{m,n} = \frac{B_y(m,n)}{[N/M^2]^{1/2}[S_y(g_m)\hat{S}_y(g_{mn})\hat{S}_y(g_{m+n})]^{1/2}}
\]  \[\text{[4.2.1.1.3.4]}\]

Under the null hypothesis of Gaussianity, \( \hat{S} \) is distributed chi-squared with \( 2P \) degree of freedom, with \( P \) being the number of squares whose centres are in the principal domain. Hinich (1982)\(^{299}\) showed that, asymptotically, the transformation of \( \hat{S} \) is well approximated by a normal distribution with zero mean and unit variance. Thus, the significance of the test statistics is readily determined from standard normal tables.

\(^{299}\) Hinich(1982), op cit
On the other hand, if \( y_t \) is linear but not Gaussian, the sample dispersion of, \( 2|\hat{y}_{m,n}|^2 \) should not differ significantly from the population dispersion of \( \chi^2(2, \lambda) \), where \( \lambda = \{S/P\} - 2 \). Linearity test statistics examine whether the sample dispersion is significantly different from that of \( \chi^2(2, \lambda) \). The distribution of the standard normal is used to produce a one-sided test, in which the null is rejected if the test statistic is greater than the critical value at the chosen level of significance.

4.2.1.1.4 Keenan Test

Ramsey (1969)\textsuperscript{300} proposes a specification test for linear least squares regression analysis. The test is referred to as a RESET test and is readily applicable to linear AR models. Consider the linear AR(p) model

\[
x_t = X_{t-1}^\prime \phi + a_t
\]  \hspace{1cm} [4.2.1.1.4.1]

where \( X_{t-1} = (1, x_{t-1}, \ldots, x_{t-p}) \) and \( \phi = (\varphi_0, \varphi_1, \ldots, \varphi_p)' \). The first step of the RESET test is to obtain the least squares estimate \( \hat{\phi} \) of Eq. (1) and compute the fit \( \hat{x}_t = \hat{X}_{t-1} \hat{\phi} \), the residual \( a_t = x_t - \hat{x}_t \), and the sum of squared residuals \( SSR_0 = \sum_{t=p+1}^T \hat{a}_t^2 \), where \( T \) is the sample size. In the second step, consider the linear regression

\[
\hat{a}_t = X_{t-1}^\prime \alpha_1 + M_{t-1}^\prime \alpha_2 + \upsilon_t
\]  \hspace{1cm} [4.2.1.1.4.2]

Where \( M_{t-1} = (\hat{x}_t^2, \ldots, \hat{x}_t^{s+1})' \) for some \( s \geq 1 \), and compute the least squares residuals

\[ v_t = \hat{a}_t - X'_{t-1} \alpha_1 - M'_{t-1} \alpha_2 \]  \[4.2.1.1.4.3\]

and the sum of squared residuals \( \text{SSR}_1 = \sum_{t=p+1}^{T} \hat{\delta}_t \) of the regression. The basic idea of the RESET test is that if the linear AR(p) model in Eq. [4.2.1.1.4.1] is adequate, then \( \alpha_1 \) and \( \alpha_2 \) of Eq. [4.2.1.1.4.3] should be zero. This can be tested by the usual F statistic of Eq. [4.2.1.1.4.2] given by

\[
F = \frac{(\text{SSR}_0 - \text{SSR}_1)/g}{\text{SSR}_1/(T-p-g)} \quad \text{with } g = s + p + 1 \tag{4.2.1.1.4.4}
\]

which, under the linearity and normality assumption, has an F distribution with degrees of freedom \( g \) and \( T - p - g \).

Keenan (1985)\(^\text{301}\) proposes a nonlinearity test for time series that uses \( \hat{x}_t^2 \) only and modifies the second step of the RESET test to avoid multicollinearity between \( \hat{x}_t^2 \) and \( X_{t-1} \). Specifically, the linear regression (Eq. [4.2.1.1.4.2]) is divided into two steps. In step 2(a), one removes linear dependence of \( \hat{x}_t^2 \) on \( X_{t-1} \) by fitting the regression

\[
\hat{x}_t^2 = X'_{t-1} \beta + u_t \tag{4.2.1.1.4.5}
\]

and obtaining the residual \( u_t = \hat{x}_t^2 - X'_{t-1} \beta \). In step 2(b), we consider the linear regression

\[
a_t = u_t \alpha + v_t \tag{4.2.1.1.4.6}
\]

and obtain the sum of squared residuals to test the null hypothesis \( \alpha = 0 \).

4.2.1.1.5 Tests based on Neural Networks and Taylor series approximations

Let us consider a nonlinear autoregression involving the last p lags of the variable \( u_t \) as follows

\[
 u_t = F(u_{t-1}, \ldots, u_{t-p}) + \varepsilon_t \tag{4.2.1.1.5.1}
\]

Implementation of the ANN testing framework specifies that the nonlinear part of \( F(\cdot) \) in [4.2.1.1.5.1] is given by

\[
 \sum_{j=1}^{q} \beta_j \times \varphi\left(\sum_{i=1}^{p} \gamma_{ij} \hat{u}_{t-i}\right),
\]

where \( \varphi(\lambda) \) is the logistic function, given by \([1 + \exp(-\lambda)]^{-1}\). As noted by Lee et al. (1993), this functional form can approximate any continuous function arbitrarily well. The coefficients \( \gamma_{ij} \) are randomly generated from a uniform distribution over \([\gamma_1, \gamma_h]\). It should be noted that using random \( \gamma_{ij} \) has two purposes. First, it bypasses the need for computationally expensive estimation techniques; second, and most importantly, it solves the identification problem for \( \gamma_{ij} \) because these parameters are not identified under the null hypothesis of linearity. For a given \( q \), the constructed regressors \( \varphi(\sum_{i=1}^{p} \gamma_{ij} \hat{u}_{t-i}), j = 1, \ldots, q \), may suffer from multicollinearity.

Following a suggestion of Lee et al. (1993)\textsuperscript{302}, the \( \bar{q} \) in this study is taken to be the largest principle components of the constructed regressors excluding the largest one used as regressors in

\[
 \hat{u}_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \hat{u}_{t-i} + \sum_{i=1}^{\bar{q}} \beta_j \bar{\varphi}_{j,t} + \varepsilon_t, \tag{4.2.1.1.5.2}
\]

where \( \bar{\varphi}_{j,t} \) denotes the \((j + 1)\)th principal component. A standard LM test is then be performed; Lee et al. (1993)\textsuperscript{303} suggested constructing the test statistic as \( \hat{\tau}^2 \),

\textsuperscript{302} Lee, T. H., White, H., and Granger, C. W. J. (1993) \textit{op cit}

\textsuperscript{303} \textit{Ibid}
where $R^2$ is the uncentered squared multiple correlation coefficient of a
regression of $\hat{\varepsilon}_t$ on a constant; $\hat{u}_{t-i}, i = 1, \ldots, p$; $\phi_{i,j}, j = 1, \ldots, q$, where $\hat{\varepsilon}_t$ is the
residual of the regression of $\hat{u}_t$ on a constant; and $\hat{u}_{t-i}, i = 1, \ldots, p$. Under the
null hypothesis, this test statistic has an asymptotic $\chi^2_q$ distribution. Under the
alternative hypothesis, this test is consistent, as discussed by Stinchcombe and
White (1998).\textsuperscript{304}

An alternative two-step approach is to apply the logistic neural network test
proposed by Teräsvirta, Lin, and Granger (1993)\textsuperscript{305} to the fractionally filtered
series. This test approximates the logistic neural network by a Taylor series
expansion and tests for the significance of the additional terms when they are
subsequently substituted into the model for $\hat{u}_t$. (See Blake and Kapetanios
2003\textsuperscript{306} for an alternative interpretation of the logistic neural network test without
long memory.) The appropriate order of terms will typically depend on the degree
of nonlinearity in the data. The second-order expansion is

\[
\hat{u}_t = \beta_0 + \sum_{i=1}^{p} \beta_i \hat{u}_{t-i} + \sum_{i=1}^{p} \gamma_{0,i,2} \hat{u}_{t-i}^2 + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{1,i,j} \hat{u}_{t-i} \hat{u}_{t-j} + \varepsilon_t. [4.2.1.1.5.3]
\]

The third-order expansion, which was recommended by Teräsvirta et al. (1993),
is of the form

\textsuperscript{304} Stinchcombe, M. B., and H. White. "Consistent Specification Testing With Nuisance Parameters Present
\textsuperscript{305} Teräsvirta, T., Fu-Lin C, and Granger, CWJ. (1993) op cit
\textsuperscript{306} Blake, A. P., and G. Kapetanios. "A Radial Basis Function Artificial Neural Network Test for Neglected
\[ \hat{u}_t = \beta_0 + \sum_{i=1}^{p} \beta_i \hat{u}_{t-i} + \sum_{j=2}^{3} \sum_{l=1}^{p} \gamma_{0,1,2}^{l} \hat{u}_{t-i}^{l} + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \gamma_{1,i,j} \hat{u}_{t-i} \hat{u}_{t-j} + \sum_{s=0}^{1} \sum_{l=1}^{p-1} \sum_{j=l+1}^{p} \gamma_{2,s,i,j}^{l} \hat{u}_{t-i}^{l} \hat{u}_{t-j}^{s} + \varepsilon_t, \]

[4.2.1.1.5.4]

Clearly, these are all very general approximations with a considerable number of terms and interactions. For the purpose of this study, we decided to restrict the number of parameters in the third- and fourth-order Taylor series expansions by considering only cross-products and powers of up to two lags. Given this restriction, the null hypothesis corresponding to the absence of nonlinearity is equivalent to the \( \gamma \) coefficients being zero.

4.2.2 Cross Correlation function between two data series

For further investigation we have studied the nonlinear correlation analysis for the two sets. The autocorrelation is basically a mathematical tool for finding repeating patterns, such as the presence of a periodic signal, which has been buried under noise, or identifying the missing fundamental frequency in a signal implied by its harmonic frequencies. It is used frequently in signal processing for analysing functions or series of values, such as time domain signals. In other words it is a mathematical representation of the degree of similarity between a given time series and a lagged version of itself over successive time intervals. It
is the same as calculating the correlation between two different time series, except that the same time series is used twice - once in its original form and once lagged one or more time periods. The discrete version of the autocorrelation function can be written as

\[
R_{xx}(j) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} x_n x_{n-j}
\]

[4.2.2.1]

Where N is the total sample and \(x_n, \bar{x}_i\) are the values from two data series. When computed, the resulting number can range from +1 to -1. An autocorrelation of +1 represents perfect positive correlation (i.e. an increase seen in one time series will lead to a proportionate increase in the other time series), while a value of -1 represents perfect negative correlation (i.e. an increase seen in one time series results in a proportionate decrease in the other time series).

The cross-correlation is a statistical measure timing the movements and proximity of alignment between two different information sets of a series of information. Cross correlation is generally used when measuring information between two different time series. The range of the data is -1 to 1 such that the closer the cross-correlation value is to 1; the more closely related the information sets are.
In signal processing, cross-correlation is a measure of similarity of two waveforms as a function of a time lag applied to one of them. This is also known as a sliding dot product or inner product. It is commonly used to search a long duration signal for a shorter, known feature. It also has applications in pattern recognition and cryptanalysis.

For continuous functions $f$ and $g$ the cross-correlation is defined as:

$$\left(f \ast g\right)(t) = \int f^*(\tau)g(t + \tau)d\tau$$

[4.2.2.2]

The cross-correlation is similar in nature to the convolution of two functions. Whereas convolution involves reversing a signal, then shifting it and multiplying by another signal, correlation only involves shifting it and multiplying (no reversing). In autocorrelation, which is the cross-correlation of a signal with itself, there will always be a peak at a lag of zero. If $X$ and $Y$ are two independent random variables with probability distributions $f$ and $g$, respectively, then the probability distribution of the difference $X - Y$ is given by the cross-correlation figure. In contrast, the convolution $f \ast g$ gives the probability distribution of the sum $X + Y$.

4.2.3 Power Spectrum Analysis

One of the simplest and useful tools to investigate nonlinear dynamics is the power spectrum analysis. For a given data series, the power spectrum gives a plot of the portion of a signal’s power (energy per unit time) falling within given
frequency bins. The most common and effective way of generating a power spectrum is by using a discrete Fourier transform, but there are some other techniques such as the maximum entropy method that can also be used.

The continuous Fourier transform \( \overline{f}(\nu) \) of a function \( f(t) \) is defined as

\[
\overline{f}(\nu) = F[f(t)](\nu) = \int_{-\infty}^{\infty} f(t) \exp(-2\pi i \nu t) dt
\]  

[4.2.3.1]

For the discrete function \( f(t) \rightarrow f(t_k) \) we can modify our definition as

\[
F_n = \sum_{k=0}^{N-1} f_k \exp(2\pi i k / N)
\]  

[4.2.3.2]

The corresponding inverse transform can be written as

\[
f_k = \sum_{k=0}^{N-1} F_n \exp(2\pi i k n / N)
\]  

[4.2.3.3]

Where \( f(t_k) = f_k \) for \( k=0,1,2... \ N-1 \).

Very often, the most efficient way of estimating \( f_k \) from some data is via the Fourier transform, using the Convolution Theorem (the convolution of two infinite sequences is equal to the inverse Fourier transform of the product of the Fourier transforms of the individual sequences). The Power spectrum analysis is an effective tool to quantify chaos. Any data series, which is in chaotic state shows random behaviour and lots of peaks, arise corresponding to the spectrum
analysis. Also a data set in either steady or periodic state will converge to its corresponding number of peaks in Power spectrum diagram.

4.2.4 Empirical Mode Decomposition

Data analysis is at the heart of all empirical research work, for data is the only link we have with the reality. Consequently, data analysis serves two purposes:

Firstly, it acts as a validation of scientific theories or models. Secondly, it unearths the clues related to the underlying mechanisms which act as a basis for innovations, invention or improvements of the theories and models. Whatever the case may be, the data contains information we are seeking; the goal of data analysis is to find the information in the data. As one does not have a complete knowledge base of the underlying mechanisms for most of the physical problems we face today, one should inject as little subjective specifications as possible in the process of data analysis, so that we do not prejudice the results. A truly objective data analysis method should be adaptive to the data and let the data set speaks for itself. Traditional time-frequency analysis methods, however, all follows the well established mathematical rules: the methods all start with a definition of a basis, and convolve the signal with the basis to get amplitude and frequency either for distributions or for filtering. Such an approach has the great advantage of having a solid mathematical foundation. Once the algorithm is established, data analysis can go forward mechanically. Unfortunately, within the comfortable fold of solid mathematic foundation, the methods cannot be adaptive
at all. Furthermore, this well trodden path also restricts the methods developed under this paradigm to linear and stationary assumptions.

As data can come from all sources ranging from relatively well established physical sciences, to complicated biologic processes and social-economic phenomena., most of the driving mechanisms are so complicadly intertwined and interacting that the data one obtains are also highly variable, not only from one case to another but also from time to time even limited to one single case. In other words, one has to face data from nonlinear and non-stationary processes. This requirement is known for a long time, but remedy is slow to come. To accommodate for data from non-stationary processes, one has met many success. Methods (see for example, Flandrin, 1999) such as spectrogram, Wigner-Ville distribution, Wavelet analysis are all examples. To accommodate for data from non-linear processes, however, progress has been very slow. The available methods (see, for example, Tong, 1990, Krantz and Schreiber, 1997 and Diks, 1998) are limited to handle data from deterministic low dimensional chaotic systems.

Even for data from non-stationary processes, the available methods are also limited to linear systems, for the methods were mostly based on the well

---

established a priori basis approach, where all the analysis is based on convolution of the data with the established basis. This approach, unfortunately, has drained all the physical meaning out of the analysed results, for any a priori basis could not possibly fit all the variety of data from drastically different underlying driving mechanisms. Any misfit will automatically be assigned to the various orders of harmonics with respect to the selected basis. Though results so obtained satisfy the mathematical requirements, they lack physical meaning. Furthermore, the convolution processes involve integration, which make the results suffering the limitation imposed by the uncertainty principle, and preventing us from examine the details of the data and their underlying mechanisms.

Let us take a simple example to examine the characteristics of data from a nonlinear system. Consider the Duffing equation without damping given as

\[
\frac{d^2 x}{dt^2} + x + \varepsilon x^3 = \gamma \cos \omega t \tag{4.2.4.1}
\]

where \( \varepsilon \) is a parameter, not necessarily small; \( \gamma \) is the magnitude of the driving force. We can easily rewrite this equation slightly as follows:

\[
\frac{d^2 x}{dt^2} + x(1 + \varepsilon x^2) = \gamma \cos \omega t \tag{4.2.4.2}
\]
If one treats the quantity in the parenthesis as a single number designated as $L$:

$$L = 1 + \varepsilon x^2 \quad \text{[4.2.4.3]}$$

then the quantity $L$ can be treated as the pendulum length or the spring constant. Either way, $L$ changes with position; therefore, the frequency of the system should also change with position even within one oscillation period. Such intra-wave frequency modulation is the special characteristics of a nonlinear system; and it requires a detailed frequency representation that is unattainable from a priori basis approach. For example, following the classic perturbation analysis by imposing a linear structure on a nonlinear system, one would find the solution consisted of endless harmonics. The effect of the harmonics is to find enough sinusoidal components to fit the deformed final waveform, commonly known as harmonic distortions. It is well known that each term in this perturbation solution does not have physical meaning, only the summation of all the terms represents the physics. But using any a priori basis analysis, one would inevitably obtain a collection of the harmonics of one form or the other depending on the basis function selected; thus rendered the interpretation of spectral analysis problematical. The harmonics representation here is a poor substitute of the detailed instantaneous frequency description of the intra-wave frequency modulation. But such a detailed description will call for a drastic new approach. In fact to describe intra-wave frequency modulation, one cannot use a priori basis approach. An easy alternative is to use the Hilbert Transform, which is defined as
\[ y(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} \, d\tau \quad [4.2.4.4] \]

in which \( x(t) \) is the given function of \( L_p \) class, \( y(t) \) is the Hilbert transform, which is the complex conjugate of \( x(t) \), and \( P \) indicates the principal value of the singular integral. As \( y(t) \) is the complex conjugate, one has

\[ z(t) = x(t) + jy(t) = a(t)e^{j\theta(t)}, \quad [4.2.4.5] \]

where

\[ a(t) = (x^2 + y^2)^{1/2} ; \quad \theta(t) = \tan^{-1} \frac{y}{x} \quad [4.2.4.6] \]

Here \( a \) is the instantaneous amplitude, and \( \theta \) is the phase function; thus the instantaneous frequency, with the stationary phase approximation, is simply

\[ \omega = \frac{d\theta}{dt} \quad [4.2.4.7] \]

This definition also coincides with the classical wave theory. This definition of instantaneous frequency appears to be local, for it is defined through differentiation rather than integration, and hence, the resulting instantaneous frequency may be able to describe the intra-wave frequency modulation. This approach has been recommended by Hahn (1996)\textsuperscript{311} for applications signal processing. Unfortunately, this straightforward and simple-minded approach does not work well. Although the Hilbert transform is valid under a very general

\textsuperscript{311} Hahn, S. L. \textit{Hilbert Transforms in Signal Processing}. Artech House, 1996,
condition, for the instantaneous frequency derived from the above approach to make physical sense, the function has to be ‘mono-component’ as discussed by Cohen (1995) and Huang et al. (1998, 1999). This has been illustrated by Huang et al (1998) with a simple function as

\[ x(t) = a + \cos \alpha t \]  \hspace{1cm} [4.2.4.8]

with \( a \) as an arbitrary constant. Its Hilbert transform is simply

\[ y(t) = \sin \alpha t \]  \hspace{1cm} [4.2.4.9]

therefore, the instantaneous frequency according to Equation (4.2.4.7) is

\[ \omega = \frac{\alpha (1 + a \sin \alpha t)}{1 + 2a \cos \alpha t + a^2} \]  \hspace{1cm} [4.2.4.10]

Equation (35) can give any value for the instantaneous frequency, depending on the value of \( a \). In order to recover the frequency of the input sinusoidal signal, the constant has to be zero. This simple example illustrates a crucial condition for the Hilbert Transform approach to work here: the function will have to be zero mean locally. This seemingly trivial condition has created great misunderstanding, which has prompted Cohen (1995) to list a number of ‘paradoxes’ concern instantaneous frequency. Some of the paradoxes concerning negative frequency are direct consequence of this condition.

---

315 Ibid
316 Cohen, L., 1995, op cit
Another obvious consequence of this condition is the difficult experience by all previous attempts to use the Hilbert transform: how to reduce or decompose an arbitrary function to a ‘mono-component’ one with local zero mean? And more fundamentally, if the function is non-stationary, how can one find the local mean? These difficulties have forced the past applications of Hilbert transform to extract a narrow band component with a band-pass filter on the original data (Melville, 1983)\textsuperscript{317}. As the band-pass filter is a linear operator, any signal passing through it will lost all its ‘harmonics’, and suffer deformation of the fundamental wave shape. Such approach certainly satisfies the condition demanded by the instantaneous frequency computation through Hilbert transform. However, it has unwittingly drained some interesting information from the data, the nonlinear characteristics associated with the signal.

With all these difficulties, the real applications of Hilbert transform will have to wait for the development of the Empirical Mode Decomposition (EMD) (Huang et al. 1998\textsuperscript{318}, 1999\textsuperscript{319}, 2003\textsuperscript{320}). Together with the Hilbert Spectral Analysis (HSA), the combination established a new adaptive time-frequency analysis method.


\textsuperscript{319} Huang, N. E., Z. Shen, R. S. Long, 1999, \textit{op cit}

The details of both Empirical Mode Decomposition (EMD) and the Hilbert Spectral Analysis (HSA) are given in Huang et al. 1996\textsuperscript{321}, 1998\textsuperscript{322} and 1999\textsuperscript{323}). The following summary is based on a simplified version given in Huang (2005)\textsuperscript{324}. The EMD method is necessary to reduce any data from non-stationary and nonlinear processes into simple oscillatory function that will yield meaningful instantaneous frequency through the Hilbert transform. Contrary to almost all the previous decomposing methods, EMD is empirical, intuitive, direct, and adaptive, with the a posteriori defined basis derived from the data. The decomposition is designed to seek the different simple intrinsic modes of oscillations in any data based on the principle of scale separation.

The data, depending on it complexity, may have many different coexisting modes of oscillation at the same time. Each of these oscillatory modes is represented by an Intrinsic Mode Function (IMF) with the following definitions:

(a) in the whole data set, the number of extrema and the number of zero-crossings must either equal or differ at most by one, and

(b) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The IMF is a counterpart to the simple harmonic function, but it is much more general: instead of constant amplitude and frequency, IMF can have both

\textsuperscript{323} Huang, N. E., Z. Shen, R. S. Long, 1999, \textit{op cit}
\textsuperscript{324} Huang, N. E. and N. Attoh-Okine, Ed. \textit{Hilbert-Huang Transform in Engineering}. CRC Press, 2005
variable amplitude and frequency as functions of time. This definition is inspired by the simple example of constant plus sinusoidal function given above. The total number of the IMF components is limited to $ln2N$, where $N$ is the total number of data points. It satisfies all the requirements for a meaningful instantaneous frequency through Hilbert transform.

Pursuant to the above definition for IMF, one can implement the needed decomposition of any function, known as sifting, as follows: take the test data; identify all the local extrema; divide the extrema into two sets: the maxima and the minima. Then connect all the local maxima by a cubic spline line to form an upper envelope. Repeat the procedure for the local minima to form a lower envelope. The upper and lower envelopes should encompass all the data between them. Their mean is designated as $m_1$, and the difference between the data and $m_1$ is designated as, $h_1$, a proto-IMF:

$$X(t) - m_1 = h_1 \quad [4.2.4.11]$$

Ideally, $h_1$ should satisfy the definition of an IMF by construction of $h_1$ described above, which should have made it symmetric and having all maxima positive and all minima negative. Yet, in changing the local zero from a rectangular to a curvilinear coordinate system some inflection points could become additional extrema. New extrema generated this way actually reveal the hidden modes missed in the initial treatment. The sifting process sometimes can recover signals representing low amplitude riding waves with repeated siftings. The sifting
process serves two purposes: to eliminate riding waves and to make the wave profiles more symmetric. While the first condition is absolute necessary for Hilbert transform to give a meaningful instantaneous frequency, the second condition is also necessary in case the neighbouring wave amplitudes having too large a disparity. As a result, the sifting process has to be repeated many times to reduce the extracted signal an IMF. In the subsequent sifting process, \( h_1 \) is treated as the data for the next round of sifting; therefore,

\[
h_1 - m_{11} = h_{11}
\]  \[4.2.4.12\]

After repeated sifting, up to \( k \) times, \( h_{1k} \):

\[
h_{1(k-1)} - m_{1k} = h_{1k}
\]  \[4.2.4.13\]

If \( h_{1k} \) becomes an IMF, it is designated as \( c_1 \):

\[
c_1 = h_{1k}
\]  \[4.2.4.14\]

the first IMF component from the data. Here one has a critical decision to make: when to stop. Too many rounds of sifting will reduce the IMF to FM page criterion; too few rounds of sifting will not have a valid IMF. In the past, different criteria have been used, including Cauchy type criterion (Huang et al. 1998)\(^{325}\), S-number criterion (Huang et al. 2003)\(^{326}\), fixed-number criterion (Wu and Huang 2004)\(^{327}\), and etc. With any stoppage criterion, the, \( c_1 \) should contain the finest

\(^{326}\) Huang, N. E. and Long, S. R. 2003, op cit
scale or the shortest period component of the signal. One can, then, remove $c_1$ from the rest of the data by

$$X(t) - c_1 = r_1$$ \[4.2.4.15\]

Since the residue, $r_1$, contains all longer period variations in the data, it is treated as the new data and subjected to the same sifting process as described above. This procedure can be repeated to all the subsequent $r_j$'s, and the result is

$$r_1 - c_2 = r_2,$$

$$\ldots$$

$$r_{n-1} - c_n = r_n$$ \[4.2.4.16\]

The sifting process should stop when the residue, $r_n$, becomes a constant, a monotonic function, or a function contains only a single extrema, from which no more IMF can be extracted. By summing up Equations (4.2.4.16) and (4.2.4.17), we finally obtain

$$X(t) = \sum_{j=1}^{n} c_j + r_n$$ \[4.2.4.17\]

Thus, sifting process produces a decomposition of the data into $n$-intrinsic modes, and a residue, $r_n$. When apply the EMD method, a mean or zero
reference is not required; EMD needs only the locations of the local extrema. The sifting process generates the zero reference for each component. Without the need of the zero reference, EMD avoids the troublesome step of removing the mean values for the large non-zero mean. Two special notes here deserve our attention. First, the sifting process offered a way to circumvent the difficulty of defining the local mean in a nonstationary time series, where no length scale exists for one to implement the traditional mean operation. The envelope mean employed here does not involve time scale; however, it is local. Second, the sifting process is a Reynolds-type decomposition: separating variations from the mean, except that the mean is a local instantaneous mean, so that the different modes are almost orthogonal to each other, except for the nonlinearity in the data.

Recent studies by Flandrin et al. (2004)\textsuperscript{328} and Wu and Huang (2004)\textsuperscript{329} established that the EMD is equivalent to a dyadic filter bank, and it is also equivalent to an adaptive wavelet. Being adaptive, they have avoided the shortcomings of using any a priori-defined wavelet basis, and also avoided the spurious harmonics that would have resulted. The components of the EMD are usually physically meaningful, for the characteristic scales are defined by the physical data. Having established the decomposition, we can also identify a new use of the IMF components as filtering. Traditionally, filtering is carried out in

\textsuperscript{329} Wu, Z. and Huang, N. E. 2004, op cit
frequency space only. But there is a great difficult in applying the frequency filtering when the data is either nonlinear or non-stationary or both, for both nonlinear and nonstationary data generate harmonics of all ranges. Therefore, any filtering will eliminate some of the harmonics, which will cause deformation of the data filtered. Using IMF, however, we can devise a time space filtering. For example, a low pass filtered results of a signal having n-IMF components can be simply expressed as

\[ X_{lk}(t) = \sum_{k}^{n} c_j + r_n \]  

[4.2.4.18]

a high pass results can be expressed as

\[ X_{hk}(t) = \sum_{1}^{k} c_j \]  

[4.2.4.19]

and a band pass result can be expressed as

\[ X_{bk}(t) = \sum_{b}^{n} c_j \]  

[4.2.4.20]

The advantage of this time space filtering is that the results preserve the full nonlinearity and nonstationarity in the physical space. Having obtained the Intrinsic Mode Function components, one can compute the instantaneous frequency for each IMF component as the derivative of the phase function. And one can also designate the instantaneous amplitude from the Hilbert transform to
each IMF component. Finally, the original data can be expressed as the real part, $\text{RP}$, of the sum of the data in terms of time, frequency and energy as:

$$X(t) = \text{RP} \sum_{j=1}^{n} a_j(t)e^{i \int \omega_j(t)dt} \quad [4.2.4.21]$$

Equation [4.2.4.21] gives both amplitude and frequency of each component as a function of time. The same data, if expanded in a Fourier representation, would have a constant amplitude and frequency for each component. The contrast between EMD and Fourier decomposition is clear: The IMF represents a generalized Fourier expansion with a time varying function for amplitude and frequency. This frequency-time distribution of the amplitude is designated as the Hilbert Amplitude Spectrum, $H(\omega, t)$, or simply the Hilbert spectrum. From the Hilbert spectrum, we can also define the marginal spectrum, $h(\omega)$, as

$$h(\omega) = \int_{0}^{T} H(\omega, t)dt \quad [4.2.4.22]$$

The marginal spectrum offers a measure of total amplitude (or energy) contribution from each frequency value. It represents the cumulated amplitude over the entire data span in a probabilistic sense. The combination of the Empirical Mode Decomposition and the Hilbert Spectral Analysis is designated by NASA as the Hilbert-Huang Transform (HHT) for short. Recent studies by various investigators indicate that HHT is a super tool for time-frequency analysis...
of nonlinear and nonstationary data (Huang and Attoh-Okine, 2005\textsuperscript{330}, Huang and Shen, 2005\textsuperscript{331}). It is based on an adaptive basis, and the frequency is defined through the Hilbert transform. Consequently, there is no need for the spurious harmonics to represent nonlinear waveform deformations as in any of the a priori basis methods, and there is no uncertainty principle limitation on time or frequency resolution from the convolution pairs based also on a priori bases.

4.2.5 Recurrence Analysis\textsuperscript{332}

4.2.5.1 Recurrence Plot

Definition: A recurrence plot (RP) is a graph that shows all those times at which a state of the dynamical system recurs. In other words, the RP reveals all the times when the phase space trajectory visits roughly the same area in the phase space. [A phase space, introduced by Willard Gibbs in 1901, is a space in which all possible states of a system are represented, with each possible state of the system corresponding to one unique point in the phase space. In a phase space, every degree of freedom or parameter of the system is represented as an axis of a multidimensional space. For every possible state of the system, or allowed combination of values of the system's parameters, a point is plotted in the multidimensional space. Often this succession of plotted points is analogous to the system's state evolving over time. In the end, the phase diagram represents all that the system can be, and its shape can easily elucidate qualities of the

\textsuperscript{330} N. E. Huang and N. Attoh-Okine(2005), \textit{op cit}
\textsuperscript{331} Huang N. E., S. R. Long, and Z. Shen(1996), \textit{op cit}
\textsuperscript{332} For the details of Recurrence Analysis we have followed Dr Norbert Marwan's PhD thesis (2003).
system that might not be obvious otherwise. A phase space may contain very many dimensions.

Natural processes can have a distinct recurrent behaviour, e.g. periodicities (as seasonal or Milankovich cycles), but also irregular cyclicities (as El Niño Southern Oscillation). Moreover, the recurrence of states, in the meaning that states are arbitrary close after some time, is a fundamental property of deterministic dynamical systems and is typical for nonlinear or chaotic systems. The recurrence of states in nature has been known for a long time and has also been discussed in early publications (e.g. recurrence phenomena in cosmic-ray intensity, Monk, 1939\textsuperscript{333}). Eckmann et al. (1987)\textsuperscript{334} have introduced a tool which can visualize the recurrence of states $x_i$ in a phase space. Usually, a phase space does not have a dimension (two or three) which allows it to be pictured. Higher dimensional phase spaces can only be visualized by projection into the two or three dimensional sub-spaces. However, Eckmann’s tool enables us to investigate the m-dimensional phase space trajectory through a two-dimensional representation of its recurrences. Such recurrence of a state at time $i$ at a different time $j$ is marked within a two-dimensional squared matrix with ones and zeros dots (black and white dots in the plot), where both axes are time axes. This representation is called recurrence plot (RP).

$$R_{i,j} = \Theta(\varepsilon_i - \|x_i - x_j\|), x_i \in \mathbb{R}^n, i, j = 1, \ldots, N$$

\[4.2.5.1.1\]


\textsuperscript{334} Eckmann J.P., S.O. Kamphorst, D. Ruelle (1987), op cit
Where $R_{i,j}$ is the recurrence plot, $N$ is the number of considered states $x_i$, is a threshold distance, $||\cdot||$ a norm and $\Theta(\cdot)$ the Heaviside function. The Heaviside step function is given by:

$$\Theta(x) = 0 \text{ if } x < 0$$

$$\Theta(x) = 1 \text{ if } x \geq 0$$

[4.2.5.1.2]

---

**Fig 4.1(A)** Segment of the phase space trajectory of the Lorenz system (for standard parameters $r=28$, $\sigma=10$, $b=8/3$; Lorenz, 1963$^{335}$) by using its three components and (B) its corresponding recurrence plot. A point of the trajectory at $j$ which falls into the neighbourhood (gray circle in (A)) of a given point at $i$ is considered as a recurrence point (black point on the trajectory in (A)). This is marked with a black point in the RP at the location $(i, j)$. A point outside the neighbourhood (small circle in (A)) causes a white point in the RP. The radius of the neighbourhood for the RP is $\varepsilon=5$.

---

4.2.5.1.1 Embedding Parameters

The most natural question pertains to how to choose an appropriate value for the time delay \( d \) and the embedding dimension \( m \). Several methods have been developed to best guess \( m \) and \( d \). The most often used methods are the Average Mutual Information Function (AMI) for the time delay, as introduced by Fraser and Swinney in 1986\(^{336}\) and the False Nearest Neighbours (FNN) method for the embedding dimension developed by Kennel (1992)\(^{337}\).

**Time Delay**

First of all, the time delay has to be estimated, since most logically the method to find the embedding dimension needs an estimation of \( d \). There are two main methods. In the first one, the value for which the autocorrelation function first passes through zero is searched, which gives \( d \). The function is given by,

\[
C(d) = \frac{1}{N - d} \sum_{i=1}^{N-d} (X_i - \overline{X})(X_{i+d} - \overline{X})
\]

[4.2.5.1.3]

Where \( N \) is the length of the sample and \( X_i \) are the random sample.

In the second method, one chooses the first minimum location of the average mutual information function, where the mutual information function is defined as follows. Let us start with partitioning the real numbers. Let \( p_i \) be the probability to find a time series value in the \( i \)-th interval of the partition, let \( p_{ij}(d) \) be the joint

\(^{336}\) Fraser, Andrew M. and Swinney, Harry L. (1986), *op cit*

\(^{337}\) Kennel, Matthew B., Brown, Reggie and Abarbanel, Henry D. I (1992), *op cit*
probability to find a time series value in the i-th interval and a time series value in the j-th interval after a time d, i.e. the probability of transition in d time from the i-th to the j-th interval. The average mutual information function is

$$S(d) = - \sum_{ij} p_{ij}(d) \ln \frac{p_{ij}(d)}{p_i p_j} \quad [4.2.5.1.4]$$

The value d that firstly minimizes the quantity S(d) is the method choice for finding a reasonable time delay.

The difference between these two methods resides in the fact that while the first looks for linear independence, the second measures a general dependence of two variables. For this reason the second method seems to be preferred in non linear time series analysis.

**Embedding Dimension**

The method used to find the embedding dimension is based on the concept of false neighbour. A false neighbour is a point in the data set that looks like a neighbour to another because the orbit is seen in a too small embedding space. For example two points on a circle can appear close to each other, even though they are not, if e.g. the circle is seen sideways (as a projection), thus is appearing like a line segment, whence increasing by one the dimension m of the reconstructed space often permits to differentiate between the points of the
orbit, i.e those which are true neighbours and those which are not.

Let \( y(i) \) be a point of the reconstructed space. Note as \( y(i)r \) the \( r \)-th nearest neighbour and compute the Euclidean distance \( L^2 \) between them

\[
R^2_m \left( y(i), y(i)^r \right) = \sum_{k=1}^{m-1} \left[ y(i + kd) - y^r(i + kd) \right]^2
\]

[4.2.5.1.5]

Next increase \( m \) to \( m + 1 \) and compute the new distance, i.e. \( R^2 \)

\[
R^2_{m+1} \left( y(i), y^r(i) \right) = R^2_m \left( y(i), y^r(i) \right) + \left[ y(i + kd) - y^r(i + kd) \right]^2
\]

[4.2.5.1.6]

The point \( y^r(i) \) is said a false nearest neighbour if

\[
\left[ \frac{R^2_{m+1} \left( y(i), y^r(i) \right) - R^2_m \left( y(i), y^r(i) \right)}{R^2_m \left( y(i), y^r(i) \right)} \right] > R_{tol}
\]

[4.2.5.1.7]

where \( R_{tol} \) is a predetermined threshold. Note that the number of false nearest neighbours depends on \( R_{tol} \). The sensitivity of the criterion to \( R_{tol} \) is not discussed here. Kennel (1992)\textsuperscript{338} found that for \( R_{tol} = 10 \) the false nearest neighbour is clearly identified and we stick by this value below, but for a more profound discussion we suggest to read Kennel (1992)\textsuperscript{339}

In practice, the percentage of false nearest neighbours (FNN) is computed for each \( m \) of a set of values; the embedding dimension is said to be found for the

\textsuperscript{338} Kennel, Matthew B., Brown, Reggie and Abarbanel, Henry D. I (1992), op cit

\textsuperscript{339} Ibid
first percentage of FNN dropping to zero. Notice that when the signal is noisy this percentage never reaches a true zero value.

4.2.5.1.2 Structures in Recurrence Plots

The initial purpose of RPs is the visual inspection of higher dimensional phase space trajectories. The view on RPs gives hints about the time evolution of these trajectories. The advantage of RPs is that they can also be applied to rather short non-stationary data.

The RPs exhibit characteristic large scale and small scale patterns. The first patterns were denoted by Eckmann et al. (1987)\textsuperscript{340} as typology and the latter as texture. The typology offers a global impression which can be characterized as homogeneous, periodic, drift and disrupted.

Homogeneous RPs are typical of stationary and autonomous systems in which relaxation times are short in comparison with the time spanned by the RP. An example of such an RP is that of a random time series.

Oscillating systems have RPs with diagonal oriented, periodic recurrent structures (diagonal lines, checkerboard structures). For quasi-periodic systems, the distances between the diagonal lines are different. However, even for those

\textsuperscript{340} Eckmann et al. (1987), \textit{op cit}
oscillating systems whose oscillations are not easily recognizable, the RPs can be used in order to find their oscillations.

The drift is caused by systems with slowly varying parameters. Such slow (adiabatic) change brightens the RP’s upper-left and lower-right corners.

Abrupt changes in the dynamics as well as extreme events cause white areas or bands in the RP. RPs offer an easy possibility to find and to assess extreme and rare events by using the frequency of their recurrences.

Fig 4.2 Characteristic typology of recurrence plots: (A) homogeneous (uniformly distributed noise), (B) periodic (super-positioned harmonic oscillations), (C) drift (logistic map corrupted with a linearly increasing term) and (D) disrupted (Brownian motion). These examples illustrate how different RPs can be. The used data have the length 400 (A, B, D) and 150 (C), respectively; no embeddings are used; the thresholds are $\epsilon = 0.2$ (A, C, D) and $\epsilon = 0.4$ (B).

The closer inspection of the RPs reveals small scale structures (the texture) which are single dots, diagonal lines as well as vertical and horizontal lines (the combination of vertical and horizontal lines obviously forms rectangular clusters of recurrence points).
Single, isolated recurrence points can occur if states are rare, if they do not persist for any time or if they fluctuate heavily. However, they are not a unique sign of chance or noise (for example in maps).

A diagonal line $R_{i+k, j+k} = 1$ (for $k=1\ldots l$, where $l$ is the length of the diagonal line) occurs when a segment of the trajectory runs parallel to another segment, i.e. the trajectory visits the same region of the phase space at different times. The length of this diagonal line is determined by the duration of such similar local evolution of the trajectory segments. The direction of these diagonal structures can differ. Diagonal lines parallel to the Line of identity (LOI) (angle $\pi$) represent the parallel running of trajectories for the same time evolution. The diagonal structures perpendicular to the LOI represent the parallel running with contrary times (mirrored segments; this is often a hint for an inappropriate embedding). Since the definition of the Lyapunov exponent uses the time of the parallel running of trajectories, the relationship between the diagonal lines and the Lyapunov exponent is obvious.

A vertical (horizontal) line $R_{i+k} = 1$ (for $k=1\ldots v$, where $v$ is the length of the vertical line) marks a time length in which a state does not change or changes very slowly. It seems, that the state is trapped for some time. This is a typical behaviour of laminar states (intermittency).

These small scale structures are the base of a quantitative analysis of the RPs.
Summarizing the last mentioned points we can establish the following list of observations and give the corresponding qualitative interpretation:

<table>
<thead>
<tr>
<th>Observation</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogeneity</td>
<td>the process is obviously stationary</td>
</tr>
<tr>
<td>Fading to the upper left and lower right corners</td>
<td>nonstationarity; the process contains a trend or drift</td>
</tr>
<tr>
<td>Disruptions (white bands) occur</td>
<td>nonstationarity; some states are rare or far from the normal; transitions may have occurred</td>
</tr>
<tr>
<td>Periodic/ quasi-periodic patterns</td>
<td>cyclicities in the process; the time distance between periodic patterns (e.g. lines) corresponds to the period; long diagonal lines with different distances to each other reveal a quasi-periodic process</td>
</tr>
<tr>
<td>Single isolated points</td>
<td>heavy fluctuation in the process; if only single isolated points occur, the process may be an uncorrelated random or even anti-correlated process</td>
</tr>
<tr>
<td>Diagonal lines (parallel to the)</td>
<td>the evolution of states is similar at different times; the process could be deterministic; if these diagonal lines</td>
</tr>
<tr>
<td>Observation</td>
<td>Interpretation</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
</tr>
<tr>
<td>LOI)</td>
<td>occur beside single isolated points, the process could be chaotic (if these diagonal lines are periodic, unstable periodic orbits can be retrieved)</td>
</tr>
<tr>
<td>Diagonal lines (orthogonal to the LOI)</td>
<td>the evolution of states is similar at different times but with reverse time; sometimes this is a sign for an insufficient embedding</td>
</tr>
<tr>
<td>Vertical and horizontal lines/clusters</td>
<td>some states do not change or change slowly for some time; indication for laminar states</td>
</tr>
<tr>
<td>Long bowed line structures</td>
<td>the evolution of states is similar at different epochs but with different velocity; the dynamics of the system could be changing (but note: this is not fully valid for short bowed line structures)</td>
</tr>
</tbody>
</table>

Recurrence Plot is essentially done to reveal the dynamic characteristics of the time series under investigation. For an economic time series, the evolving patterns tell us whether the series is disrupted, non-stationary or nonlinear in nature. By comparing the RPs of two economic time series we can make out whether the dynamic systems governing the time series are similar, or not.
The visual interpretation of RPs requires some experience. The study of RPs from paradigmatic systems gives a good introduction into characteristic typology and texture. However, their quantification offers a more objective way for the investigation of the considered system. With this quantification, the RPs have become more and more popular within a growing group of scientists who use RPs and their quantification techniques for data analysis.

4.2.5.2 Quantification of Recurrence Plots (Recurrence Quantification Analysis)

**Definition:** The recurrence quantification analysis (RQA) is a method of nonlinear data analysis which quantifies the number and duration of recurrences of a dynamical system presented by its phase space trajectory.

A quantification of recurrence plots was developed by Zbilut and Webber Jr. (Zbilut and Webber Jr., 1992\textsuperscript{341}; Webber Jr. and Zbilut, 1994\textsuperscript{342}) and extended with new measures of complexity by Marwan et al. (2002)\textsuperscript{343}. Measures which base on diagonal structures are able to find chaos-order transitions (Trulla et al., 1996\textsuperscript{344}), measures based on vertical (horizontal) structures are able to find chaos-chaos transitions (laminar phases, Marwan et al., 2002\textsuperscript{345}).

---

\textsuperscript{341} Zbilut J. P, Webber C. L (1992.), *op cit*
\textsuperscript{343} Marwan, Norbert and Kurths , Jurgen (2002), *op cit*
\textsuperscript{345} *Ibid*
These measures can be computed in windows along the main diagonal. This allows us to study their time dependence and can be used for the detection of transitions (Trulla et al., 1996\textsuperscript{346}). Another possibility is to define these measures for each diagonal parallel to the main diagonal separately (Marwan and Kurths, 2002\textsuperscript{347}). This approach enables the study of time delays, unstable periodic orbits (UPOs; Lathrop and Kostelich, 1989\textsuperscript{348}, Gilmore, 1998\textsuperscript{349}), and by applying to cross recurrence plots, the assessment of similarities between processes (Marwan and Kurths, 2002\textsuperscript{350}).

The fraction of recurrence points forming diagonal lines is measured by the \textit{determinism} \(DET\),

\[
DET = \frac{\sum_{l=l_{\text{min}}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)} \quad [4.2.5.2.1]
\]

where \(P(l)\) is the histogram of diagonal lines of exactly length \(l\), and \(l_{\text{min}}\) is a minimal length a diagonal structure should have to be counted as a line. Processes with uncorrelated or weakly correlated, stochastic or irregular chaotic behaviour cause none or very short diagonals, hence, small \(DET\). In contrast, regular deterministic processes lead to longer diagonals and less isolated recurrence points, resulting in higher values of \(DET\). This measure can also be interpreted as characterizing the predictability of the system.

\textsuperscript{346} Ibid
\textsuperscript{347} Ibid
\textsuperscript{349} Gilmore, R. (1998), \textit{op cit}
\textsuperscript{350} Ibid
The average diagonal line length \( L \),

\[
L = \frac{\sum_{l=l_{\text{min}}}^{N} lP(l)}{\sum_{l=l_{\text{min}}}^{N} P(l)}
\]  

[4.2.5.2.2]

gives the average time that two segments of the trajectory are close to each other, and can be interpreted as the mean prediction time.

A vertical (or, equivalently, horizontal) line \((R_{i,j+k} \equiv 1 \mid_{k=0}^{v-1})\) of length \( v \) allows identifying a time interval during which a state does not change or changes very slowly (hence, the state is trapped for some time). This is a typical characteristic of laminar states (i.e., intermittency) [Marwan et al. (2007)]

Analogously to the definition of the determinism in Eq. [4.2.5.2.1], the fraction of recurrence points forming vertical structures in the RP is defined as

\[
LAM = \frac{\sum_{v=v_{\text{min}}}^{N} \nu P(\nu)}{\sum_{\nu=1}^{N} \nu P(\nu)}
\]  

[4.2.5.2.3]

and is called laminarity. The computation of LAM is realized for those \( \nu \) that exceed a minimum length \( \nu_{\text{min}} \) in order to decrease the influence of the tangential motion (time-continuous systems that are discretized with sufficiently high sampling rate and an appropriately large threshold \( \epsilon \) result in a large amount of recurrences coming from succeeding states \( \vec{x}_i, \vec{x}_{i+1}, \vec{x}_{i+2}, \ldots \)). LAM represents the occurrence of laminar states in the system without describing the length of

---

\[351 \text{Marwan, N. Romano, M. C., Thiel, M., Kurths, J. (2007), op cit}\]
these laminar phases. In particular, LAM decreases if the RP consists of more isolated recurrence points than vertical structures.

The average length of vertical structures is given by

\[
TT = \frac{\sum_{v=v_{min}}^{N} v P(v)}{\sum_{v=v_{min}}^{N} P(v)}
\]

and is called trapping time. As in the case of LAM, the computation of TT requires the consideration of a minimal length \( v_{min} \) as well. The trapping time estimates the mean time that the system will abide at a specific state, i.e., how long the state will be trapped.

Both LAM and TT have been proven to be useful for describing the dynamics of discrete systems and studying chaos-chaos transitions. RQA as the whole is a very powerful technique for quantifying differences in the dynamics of complex systems and has meanwhile found numerous applications

**Confidence bounds of RQA measures**

The RQA measures have been quite useful for the analysis of a variety of data. Yet, in order to not only detect qualitative changes in a system's dynamics but to be able to judge their significance or to compare two univariate time series, it is necessary to derive a quantitative judgment such as a confidence interval. However, Schienkel et al. (2009)\textsuperscript{352} have recently shown that for recurrence-based complexity measures those intervals can be estimated using a resampling paradigm.

\textsuperscript{352} Schinkel, S., Marwan, N., Dimigen, O., Kurths, J., 2009, op cit
Statistical techniques based on resampling were among the first methods ever thought of. Sir R.A. Fisher(1935)\textsuperscript{353} himself introduced this idea when pondering over Gosset’s(1908)\textsuperscript{354} t-distribution. Due to lacking computational power, these ideas were not feasible at that time. With the advent of powerful, low-cost computers these methods have gained a broad interest and have been proven to be very reliable and powerful. In this Letter we focus on one particular resampling method – the bootstrap [Efron 1993]. The bootstrap is a nonparametric method for estimating the variance of a statistic of interest. It relies on resampling of a given distribution with replacement and does not require any specific probability distribution.

The bootstrap procedure works as follows:

Given a random sample $x_i$ ($i = 1, 2, \ldots, n$) of size $n$, from an unspecified probability distribution we compute a statistic of interest, say, the mean $\langle x \rangle$. In order to estimate the variance of that statistic we draw at random and with replacement the same number ($n$) of elements from $x_i$ to obtain the resampled distribution $x_i^*$. From $x_i^*$we again compute the statistic of interest. With replacement means that one can draw the individual elements in $x_i$ more than once. Doing this a larger number of times\textsuperscript{1} we obtain the empirical distribution of the statistic of interest, $\hat{P}_{\langle x \rangle}$. From the empirical distribution one can compute the

\textsuperscript{354} Gosset, W. S. "Probable error of a correlation coefficient." \textit{Biometrika} 6, no. (2/3) (1908): 302–310.
percentiles $\alpha/2$ and $1 - \alpha/2$ and define the $(100 - \alpha)\%$ confidence interval (CI) as the range between those two percentiles.

The empirical distribution could also be used to perform hypothesis testing. We opt for the estimation of confidence intervals only. The interpretation of hypothesis tests, especially p-values, the chosen indicator of significance, is currently under discussion and not agreed upon by the frequentist and Bayesian schools. Therefore the authors follow the suggestions of Hubbard and Lindsay (2008) and only estimate the confidence intervals of the RQA measures. This allows them to not only detect transitions in the dynamics of one system or to differences between the dynamics of two systems but to provide a judgment whether those differences are statistically significant.

Since the bootstrap relies on resampling with replacement one cannot simply bootstrap the RP matrix as such for two reasons.

First of all, one could draw one of the black points more than once. As the RP is a binary matrix by definition this is not possible.

---

Secondly, randomly resampling an RP would necessarily result in a loss of most of the small-scale structures in it (i.e. diagonal and vertical lines). A loss of structures would result in an RP corresponding to noise. This is not desirable because we want to compare different systems against each other and not test against randomness/noise.

As stated above, RQA measures like DET or LAM rely on the distribution of line structures $P(l)$ and $P(v)$. Therefore a method is devised that ensures that the structural elements are preserved during resampling. Only the distributions of diagonal and vertical lines, $P(l)$ and $P(v)$ are resampled. It is important to note that one needs to resample all lines in $P(l)$ and $P(v)$, even those of only length 1, thereby obtaining $P^*(l)$ and $P^*(v)$, respectively. The value of determinism is then computed for each bootstrapped sample. The computation for a bootstrapped sample of $\langle L \rangle^*$, LAM* and TT* is done accordingly. Repeating this procedure $n_{bs}$ times we obtain the empirical distributions $\hat{P}_{DET}$, $\hat{P}_{(L)}$, $\hat{P}_{LAM}$ and $\hat{P}_{TT}$. From the empirical distributions one can calculate the percentiles $\alpha/2$ and $1 - \alpha/2$. The two-sided $(100 - \alpha)$ % confidence interval is then defined as the range between those two percentiles. The value $\alpha$ determines the spread of the interval, the smaller $\alpha$, the broader the interval. As the structures in the RP are left intact, this procedure is referred to as structure preserving resampling.
For obvious reasons this approach is restricted to DET, $\langle L \rangle$, LAM and TT. $V_{\text{max}}$ and $L_{\text{max}}$ already represent maxima in the distribution and are therefore very unlikely to show variation and the upper bound cannot vary at all. Furthermore, one can apply this procedure either to the whole RP or to moving windows in order to obtain CIs over time. For this purpose the resampling and CI estimation is applied to each window separately.

One can then plot the CIs as coloured bands ranging from the lower to the upper bound of the estimated CI.

We do not use the values of DET, LAM, L, TT as an absolute index of dynamic state (viz., chaotic, random, laminar etc.). Instead we try to find out for our analysis how the respective values are changing in the two systems. By comparing their movement, we try to detect whether they move concurrently or absolutely independent of each other.

For Recurrence Plot a Cross Recurrence Plot Toolbox, Version 5.15 (R28.4) 21-Jul-2009 is used (Copyright (c) 1998-2008 Norbert Marwan, Potsdam University, Germany; http://www.agnld.uni-potsdam.de.) For computation of RQA measures with confidence intervals we have used the software available for download at http://tocsy.agnld.uni-potsdam.de. It is freely available on the internet.
4.3 Summary of research methodology used

To summarise our work, at first we have used the nonlinearity tests as discussed in section[4.2.1] above. Next we proceed with the simple cross correlation and power spectrum analysis. These tests can establish a prima facie similarity between time series. But since they are based on linearity assumption, to establish the similarity between two nonlinear time series we need other methods for which we used the EMD and RP analysis. Our discussion above clearly shows how these tools by design can take care of nonlinearity in time series and can bring out comparative graphical and statistical features of the same. In the next two chapters we use the two major analytical tools from nonlinear dynamics viz., EMD and Recurrence Analysis for empirical analysis of data from stock markets in India and abroad.