Working Paper 97-65
Statistics and Econometrics Series 26
September 1997

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

SEARCHING FOR LINEAR AND NONLINEAR COINTEGRATION: A NEW APPROACH.

F.M. Aparicio and A.Escribano*

Abstract

In this paper we propose several statistics to measure serial dependence that are useful to characterize short and long memory properties. By doing that we can measure nonlinear serial dependence, which allows us to extend the usual linear framework and to discuss short-memory and long-memory properties in a nonlinear context. Bootstrap simulation experiments report promising results for the new indices to detect nonlinearities in the cointegrating relationship, prior to estimating the nonlinear cointegrating function.

Keywords:
Short-memory, long-memory, mean-reversion, linear cointegration, nonlinear cointegration, mutual information.

* Aparicio, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. C/ Madrid, 126 28903 Madrid. Spain. Ph: 34-1-624.98.71, Fax: 34-1-624.98.49, e-mail: aparicio@est-econ.uc3m.es; Escribano, Departamento de Estadística y Econometría. Universidad Carlos III de Madrid. Research partially supported by the Spanish DGCIYT under Grant PB95-0298.

1 Introduction

Many economic time series exhibit important changes in their mean and variance. These series are often said to be *integrated*, since it is possible to simulate the most important features in their patterns with sums of an increasing number of weakly-dependent random variables. Integrated series can be expressed in terms of the *unobserved components model*, where one of the components is a *stochastic trend*. On the other hand, the fact that remote shocks have a permanent influence on the levels of these series is known as the *long-memory* property.

In some cases, the changes in mean behaviour may be correlated accross series. Pairs of series which exhibit a common long-memory component or stochastic trend are said to be *cointegrated*. The concept of *cointegration* was coined by Granger (1981 [14]), and later developed by Engle and Granger, (1987 [7]). In the context of macroeconomics and finance, certain models suggest the presence of economic or social forces preventing two or more series from drifting too far apart from each other. Well-known examples of cointegrating relationships can be found between income and expenditure, prices of a particular good in different markets, interest rates in different parts of a country, etc.

Underlying the idea of cointegration is the existence of an *equilibrium* (i.e. a deterministic relationship that holds on the average) between two integrated variables, x_t, y_t. A strict (linear) equilibrium exists when for some $a \neq 0$, one has $y_t = ax_t$. This unrealistic situation is replaced, in practice, by that of a (linear) cointegrating relationship, in which the equilibrium error $z_t = y_t - ax_t$ is different from zero but fluctuates around this value much more frequently than the individual series, while the size of these fluctuations is much smaller.

However, there is no reason to believe that those equilibrium relationships are linear. In fact, most macroeconomic dynamic models are nonlinear by construction. Similar nonlinear relationships are derived from consumer and production theory. Therefore we need to introduce time series measures of dependence that allow us to formally discuss nonlinear cointegration. This was first pointed by Granger and Hallman (1991)[]. Here we propose an alternative framework which has the advantage of detecting nonlinearities without having to estimate the nonlinear cointegrating relationship.

In summary, in this paper we review the concepts of mean-reversion, short(long)-memory, and cointegration, and introduce a new characterization of the linear cointegration concept, which can be easily generalized to a nonlinear context. This will lead us to proposing some new schemes for exploratory data analysis.

The structure of the paper is as follows. Section 2 introduces several measures of linear dependence in the context of time series that are integrated of order d, $I(d)$, where d can be any real or integer number. In section 3, a new measure of linear cointegration is introduced. We find its relation with the usual concepts for $I(1)$ series and for series with long memory (fractional cointegration). In section 4, the relationship is analyzed in the frequency domain. In section 5, we introduce a measure of departure from linearity in the cointegrating relationship, based on the mutual information. By doing a bootstrap simulation exercise, it is shown that a nonparametric estimator of this nonlinear index seems promising. Finally, section 6 presents some conclusions.

2 Definitions of memory in linear time series

In this section we briefly review the standard characterizations of memory in time series.

An interesting way of representing memory in terms of serial dependence measures. In the linear case, the standard measure is the autocorrelation function (ACF), say $\rho_x(\tau,t)$ for the series x_t, which may not necessarily be wide-sense stationary. We define $\rho_x(\tau,t)$ as

$$\rho_x(\tau,t) = \frac{cov(x_t, x_{t-\tau})}{var(x_{t-\tau})} \tag{1}$$

An early definition of memory in time series has to do with the concept of *mean reversion*:

Definition 1 *A process x_t is said to be* **mean-reverting** *if* $\forall t \; \lim_{\tau\to\infty} \rho_x(\tau,t) = 0$.

Intuitively, the process x_t is mean-reverting if $x_t - E(x_t)$ changes sign with nonzero probability. When the process is not mean-reverting, its memory span is necessarily larger since $\lim_{\tau\to\infty} \rho_x(\tau,t) > 0$, and thus any two infinitely distant variables from the process are still correlated.

However, even for a mean-reverting process, the memory span can be very large in the sense that its decays very slowly as τ grows. This motivates the distinction between *short* and *long memory*:

Definition 2 *A process x_t is said to be* **short-memory**[1] $\forall t \; \exists b_t < \infty$ *such that* $\sum_{\tau>0} |\rho_x(\tau,t)| = b_t$.

Definition 3 *A process x_t is said to be* **long-memory** *if* $\forall t \; \sum_{\tau>0} |\rho_x(\tau,t)| = \infty$.

[1]The concept of short-memory is directly related to that of asymptotic uncorrelation -see White, 1984[28]-. Notice also that mean reversion is a necessary condition for short-memory

Based on these definitions, we could define the concept of integration in the following way:

Definition 4 *A time series of x_t is said to be* **integrated of order d**, *in short $x_t \sim I(d)$, if $\sum_{\tau>0} |\rho_x(\tau,t)| = \infty,\ \forall t$, and d is the smallest real number such that $\sum_{\tau>0} |\rho_z(\tau,t)| < \infty,\ \forall t$ with $z_t = (1-B)^d x_t$.*

Remark:

From the previous definitions it follows that processes that are not mean-reverting must necessarily have long memory. This could be in the form of either stochastic or deterministic trends. On the contrary, processes that are mean-reverting can be either long-memory or short-memory, depending on the rate at which their ACF vanishes with increasing lag.

A particular case of integrated time series are those generated with *ARIMA* models, whose definition goes as follows:

Definition 5 *A time series x_t is said to be $ARIMA(p,d,q),(d \in \Re)$, if after being differenced d times, it has a stationary $ARMA(p,q)$ representation, where p,q are nonnegative integers.*

Thus if $x_t \sim ARIMA(p,d,q)$ there exists polynomials $\Phi(B)$ and $\Theta(B)$ in the delay operator B, of order $p \geq 0$ and $q \geq 0$ respectively, and with all roots outside the unit circle and no factors in common, such that we can write

$$\Phi(B)(1-B)^d x_t = \Theta(B)\epsilon_t, \tag{2}$$

where ϵ_t is generally assumed to be a sequence of zero-mean, independant and identically Normally-distributed errors.

If we define Δ as the first difference operator, that is, $\Delta x_t = (1-B)x_t = x_t - x_{t-1}$, it turns out that when d is not an integer we can write (e.g. Hosking, 1981 [22])

$$\begin{aligned} \Delta^d x_t &= \sum_{k=0}^{\infty} \frac{d!}{k!(d-k)!}(-B)^k x_t \\ &= 1 - dB - \frac{1}{2}d(1-d)B^2 - \frac{1}{6}d(1-d)(2-d)B^3 - \cdots, \end{aligned} \tag{3}$$

When the parameter in this model is fractional, it is sometimes referred to as the *long-memory parameter*, and it determines the rate of decay of the ACF of x_t with increasing lag. Only when $d < 1$, x_t is mean-reverting. If $d > 0$ the process has long memory, while it has short-memory when

$d = 0$. Moreover, if $d < \frac{1}{2}$ then x_t is stationary, while it is nonstationary for $d \geq \frac{1}{2}$ (see Granger and Joyeux, 1980 [20]).

It is known that if x_t is Gaussian and short-memory then its ACF, $\rho_x(\tau, t)$, will converge at an *exponentially* rate to zero as τ grows to infinity (Box and Jenkins, 1970 [5]), while this rate would only be *hyperbolical* for $d > 0$.

3 Linear cointegration in the time domain

Let x_t, y_t be two zero-mean integrated time series of orders $d_x, d_y \in \Re^+$, respectively. In short, $x_t \sim I(d_x)$, $y_t \sim I(d_y)$, which means that $(1-B)^{d_x} x_t = \epsilon_t$, $(1-B)^{d_y} y_t = \xi_t$, where ϵ_t, ξ_t are *short-memory* series. And let $z_t = y_t - a x_t$, for some generic nonzero real number, a.

Definition 6 *(Granger, 1981 [14]) Two $I(d)$ time series x_t, y_t, with $d > 0$, are said to be (linearly [2]) **cointegrated** if $\exists\, a \in \Re - 0$ such that the series $z_t = y_t - a x_t$ is $I(d_z)$ with $d_z < d$.*

Remarks:

- That x_t, y_t are cointegrated means that these series tend to move jointly in the long-run, even though their short-run movements may not be "aligned".
- From the macroeconomic point of view, a most important case is when $d = 1$, $d_z = 0$, since this situation can be clearly interpreted as the existence of a (linear) equilibrium for the series. However, the previous definition of cointegration does not imply the existence of an equilibrium between the two $I(d)$ series ($d > 0$), since for the latter we need that their cointegration residuals z_t be $I(d_z)$ with $d_z < \min(1, d)$. That is, an observable equilibrium also requires that z_t be mean-reverting ($d_z < 1$).

Figure 1 illustrates a simulation example of linear cointegration between a pair of correlated random walks and for $a = 0.72$. The scatter plot clearly shows the linearity of the relationship between x_t and y_t.

The Engle and Granger (1987) [7] standard tests for cointegration can be decomposed in two stages:

[2] In Granger (1981) [14], there is no explicit mention to the term *linear*, although it is implicit.

- A test for *long-memory* in the variables, say $x_t \sim I(d_x), y_t \sim I(d_y)$, and estimation of the long-memory parameters d_x, d_y. If we cannot reject the hypothesis that $d_x = d_y$ then we make $d = d_x = d_y$ and go to the next step. Otherwise, the series cannot be cointegrated.

- A test for long-memory in the cointegrating residuals $z_t = y_t - \hat{a}x_t$, and estimation of its long-memory parameter, d_z. Then a test of significance for the stochastic difference $\nu = \hat{d} - \hat{d}_z$. Values of ν significantly larger than 0, can be taken as evidence of the existence of a cointegrating relationship between x_t and y_t.

A most investigated case corresponds to where the long-memory features of the variables are generated by unit roots (i.e. $d_x = d_y = 1$). This simplifies the testing procedure since there is no need to estimate the long-memory parameter. Only a *test for unit roots* is needed to confirm this hypothesis (see Fuller, 1976 [12]; Phillips, 1987 [26]; Mackinnon, 1990 [25]).

In the sequel, we propose an alternative characterization of linear cointegration. We will restrict our discussion to the non-trivial case where all series are mutually dependent. Let x_t, y_t denote two $I(d)$ time series, and let $\rho_{y,x}(\tau, t)$ represent the *cross-correlation function* (CCF) of x_t, y_t, which we define as

$$\rho_{y,x}(\tau, t) = \frac{cov(y_t, x_{t-\tau})}{var(x_{t-\tau})}. \tag{4}$$

Once again, we consider a possible time dependence of $\rho_{y,x}(\tau, t)$ so as to allow for some degree of heterogeneity in the series.

Proposition 1 *A pair of $I(d)$ time series x_t, y_t are said to be* **linearly cointegrated** *if and only if*

$$\lim_{\tau \to \infty} \frac{\rho_{y,x}(\tau, t)}{\rho_x(\tau, t)} = b, \ \forall t. \tag{5}$$

with b equal to a nonzero and finite real number.

PROOF:

Using definition 5, it is easy to see that when linear cointegration holds then there exists a nonzero finite real number, b, such that

$$\lim_{\tau \to \infty} \rho_{y,x}(\tau, t) = b \lim_{\tau \to \infty} \rho_x(\tau, t) + \lim_{\tau \to \infty} \rho_{z,x}(\tau, t) \tag{6}$$

Thus

$$\lim_{\tau \to \infty} \frac{\rho_{y,x}(\tau, t)}{\rho_x(\tau, t)} = b + \lim_{\tau \to \infty} \frac{\rho_{z,x}(\tau, t)}{\rho_x(\tau, t)}. \tag{7}$$

Now since z_t must be an $I(d')$ series, with $d' < d$, and since x_t is $I(d)$ then:

$$\Delta^d x_t = u_t \tag{8}$$

$$\Delta^{d'} z_t = v_t, \tag{9}$$

with u_t, v_t representing zero-mean $I(0)$ series. By inversion of the differencing operator, we obtain the following binomial expansions:

$$z_t = \sum_{k=0}^{\infty} \theta_k v_{t-k} \tag{10}$$

$$x_{t-\tau} = \sum_{k'=0}^{\infty} \phi_{k'} u_{t-\tau-k'}. \tag{11}$$

Thus:

$$E(z_t x_{t-\tau}) = \sum_k \sum_{k'} \phi_{k'} \theta_k cov(u_{t-\tau-k'}, v_{t-k}). \tag{12}$$

Now, recalling that the cross-spectrum of z_t and x_t (see Granger and Hatanaka, 1964 [17]) is given by

$$S_{z,x}(\lambda) = \sum_\tau E(z_t x_{t-\tau}) exp(-j\lambda\tau), \tag{13}$$

with $j^2 = -1$, we obtain:

$$\begin{aligned} S_{z,x}(\lambda) &= \left(\sum_k \sum_{k'} \phi_{k'} \theta_k exp[-j\lambda(k-k')] \right) S_{u,v}(\lambda) \\ &= \Phi^*(\lambda)\Theta(\lambda)S_{u,v}(\lambda), \end{aligned} \tag{14}$$

where

$$\Phi^*(\lambda) = \sum_k \phi_k exp(j\lambda k) = [1 - exp(j\lambda)]^{-d} \tag{15}$$

$$\Theta(\lambda) = \sum_k \theta_k exp(-j\lambda k) = [1 - exp(-j\lambda)]^{-d'}. \tag{16}$$

Since u_t, v_t are $I(0)$ processes, as λ approaches 0 we have:

$$S_{z,x}(\lambda) \propto \Phi^*(\lambda)\Theta(\lambda), \tag{17}$$

and thus

$$S_{z,x}(\lambda) \propto \lambda^{-d-d'}. \tag{18}$$

Finally, using the inversion formula

$$cov(z_t, x_{t-\tau}) = \frac{1}{2\pi} \int_{-\pi}^{pi} exp(j\lambda\tau) S_{z,x}(\lambda) d\lambda, \tag{19}$$

and recalling that the Fourier transform of $|x|^{-p}$ is equal to $|\lambda|^{1-p}$ up to a scale factor, we have that

$$cov(z_t, x_{t-\tau}) \propto \tau^{d+d'-1}, \tag{20}$$

for τ sufficiently large. Therefore

$$\frac{cov(z_t, x_{t-\tau})}{cov(x_t, x_{t-\tau})} \propto \tau^{d'-d}, \tag{21}$$

for τ large enough. And this converges to zero, since cointegration implies that $d' < d$.

The converse follows using the same argument. As an example, we show here below how the statement in the proposition is not satisfied by a pair of non-cointegrated time series.

Example 1:

Consider the following pair of non-cointegrated series:

$$\begin{aligned} x_t &= w_t + \xi_t & (22) \\ y_t &= q_t + v_t & (23) \end{aligned}$$

where w_t, q_t are independent $I(d)$ series, and where ξ_t, v_t are series from $ARMA(p,q)$ processes with possibly different AR and MA orders, and independent of w_t and q_t, respectively. For any a we can write $y_t = ax_t + z_t$ with $z_t = q_t - aw_t + v_t - a\xi_t$. We also have

$$\begin{aligned} cov(y_t, x_{t-\tau}) &= cov(ax_t + z_t, x_{t-\tau}) & (24) \\ &= a\, cov(w_t, w_{t-\tau}) + a\, cov(\xi_t, \xi_{t-\tau}) + cov(q_t - aw_t + v_t - a\xi_t, w_{t-\tau} + \xi_{t-\tau}) \\ &= cov(v_t, \xi_{t-\tau}). \end{aligned}$$

The covariance $cov(v_t, \xi_{t-\tau})$ will tail off exponentially as τ grows to infinity. On the contrary, $cov(x_t, x_{t-\tau})$ will either be a constant (case where $d = 1$), or decay hyperbolically with growing τ. Thus our condition in the proposition follows with $b = 0$.

Example 2:

Consider the following linear *common factor model*:

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} a \\ 1 \end{pmatrix} w_t + \begin{pmatrix} v_t \\ \xi_t \end{pmatrix} \tag{25}$$

with $a \neq 0$ and where $w_t = w_{t-1} + \epsilon_t$ and (v_t, ξ_t, ϵ_t) are independent sequences of independent and identically Normally distributed $r.v.$'s with zero mean and joint covariance matrix equal to the identity matrix. Let $\beta'_\perp = (a, 1)$, where $\beta'_\perp$ is the transpose of $\beta_\perp$. Thus the orthogonal complement of $\beta'_\perp$ is $\beta' = (1, -a)$. The cointegrating relationship is therefore obtained as

$$z_t = \beta' \begin{pmatrix} y_t \\ x_t \end{pmatrix} = y_t - ax_t \tag{26}$$

where $z_t = v_t - a\xi_t$ and is obviously $I(0)$.

If we now define the ACF of x_t as

$$\rho_x(\tau, t) = \frac{cov(x_t, x_{t-\tau})}{\sigma^2_{x_{t-\tau}}}, \tag{27}$$

with $\sigma_{x_{t-\tau}} = \sqrt{var(x_{t-\tau})}$, we obtain after some algebra

$$\rho_x(\tau, t) = \frac{(t-\tau)\sigma^2_\epsilon}{(t-\tau)\sigma^2_\epsilon + \sigma^2_\xi}, \tag{28}$$

which clearly converges to 1 as $\tau \to \infty$, for any τ. Similarly, it is easy to show that $\rho_{y,x}(\tau, t) \to a$ as $\tau \to \infty$. Thus, since $a \neq 0$ by assumption, we may conclude that the series y_t, x_t are linearly cointegrated.

Example 3:

Consider the following pair of non-cointegrated series:

$$x_t = w_t + \xi_t \tag{29}$$

$$y_t = q_t + v_t \tag{30}$$

with

$$w_t = w_{t-1} + \epsilon_t \tag{31}$$

$$q_t = q_{t-1} + \eta_t \tag{32}$$

where $v_t, \xi_t, \epsilon_t, \eta_t$ are independent sequences of *i.i.d. r.v.*'s. For any a we can write $y_t = ax_t + z_t$ with $z_t = q_t - aw_t + v_t - a\xi_t$. We also have

$$\sigma^2_{x_t} = t\sigma^2_\epsilon + \sigma^2_\xi \tag{33}$$

$$\sigma^2_{y_t} = t\sigma^2_\eta + \sigma^2_v \tag{34}$$

$$cov(x_t x_{t-\tau}) = (t-\tau)\sigma^2_\epsilon \tag{35}$$

with $a \neq 0$ and where $w_t = w_{t-1} + \epsilon_t$ and (v_t, ξ_t, ϵ_t) are independent sequences of independent and identically Normally distributed *r.v.*'s with zero mean and joint covariance matrix equal to the identity matrix. Let $\beta'_\perp = (a, 1)$, where $\beta'_\perp$ is the transpose of $\beta_\perp$. Thus the orthogonal complement of $\beta'_\perp$ is $\beta' = (1, -a)$. The cointegrating relationship is therefore obtained as

$$z_t = \beta' \begin{pmatrix} y_t \\ x_t \end{pmatrix} = y_t - a x_t \tag{26}$$

where $z_t = v_t - a\xi_t$ and is obviously $I(0)$.

If we now define the ACF of x_t as

$$\rho_x(\tau, t) = \frac{cov(x_t, x_{t-\tau})}{\sigma^2_{x_{t-\tau}}}, \tag{27}$$

with $\sigma_{x_{t-\tau}} = \sqrt{var(x_{t-\tau})}$, we obtain after some algebra

$$\rho_x(\tau, t) = \frac{(t-\tau)\sigma^2_\epsilon}{(t-\tau)\sigma^2_\epsilon + \sigma^2_\xi}, \tag{28}$$

which clearly converges to 1 as $\tau \to \infty$, for any τ. Similarly, it is easy to show that $\rho_{y,x}(\tau, t) \to a$ as $\tau \to \infty$. Thus, since $a \neq 0$ by assumption, we may conclude that the series y_t, x_t are linearly cointegrated.

Example 3:

Consider the following pair of non-cointegrated series:

$$x_t = w_t + \xi_t \tag{29}$$

$$y_t = q_t + v_t \tag{30}$$

with

$$w_t = w_{t-1} + \epsilon_t \tag{31}$$

$$q_t = q_{t-1} + \eta_t \tag{32}$$

where $v_t, \xi_t, \epsilon_t, \eta_t$ are independent sequences of *i.i.d. r.v.*'s. For any a we can write $y_t = a x_t + z_t$ with $z_t = q_t - a w_t + v_t - a\xi_t$. We also have

$$\sigma^2_{x_t} = t\sigma^2_\epsilon + \sigma^2_\xi \tag{33}$$

$$\sigma^2_{y_t} = t\sigma^2_\eta + \sigma^2_v \tag{34}$$

$$cov(x_t x_{t-\tau}) = (t-\tau)\sigma^2_\epsilon \tag{35}$$

Finally, we have

$$
\begin{aligned}
cov(y_t, x_{t-\tau}) &= cov(ax_t + z_t, x_{t-\tau}) \\
&= a\, cov(w_t, w_{t-\tau}) + cov(q_t - aw_t + v_t - a\xi_t, w_{t-\tau} + \xi_{t-\tau}) \\
&= 0.
\end{aligned}
\tag{36}
$$

Thus $\rho_{y,x}(\tau, t) = 0\ \forall \tau, t$, and the series are not linearly cointegrated according to our criterion.

Remarks:

- The condition in the proposition need not be checked in the limit for most practical cases where we look for cointegration. For example, suppose $y_t = ax_t + z_t$, where $a \neq 0$, x_t, y_t are $I(1)$ and z_t is a sequence of *i.i.d. r.v.*'s and independent of x_t. In this case, $\rho_{y,x}(\tau, t)/\rho_x(\tau, t) = a\ \forall \tau$. If we now allow for some memory in z_t then the constancy of this ratio will only take place for τ's beyond some value.
- Intuitively, the proposition states that, under linear cointegration, the remote past of y_t should be as useful as the remote past of x_t in long-term forecasting x_t.
- The proposition implies that the rates of convergence of $\rho_{y,x}(\tau, t)$ and of $\rho_x(\tau, t)$ as τ increases without bound, should be the same. For example, suppose $\rho_{y,x}(\tau, t) \sim b\, \tau^{-\beta}$ and that $\rho_x(\tau, t) \sim \tau^{-\alpha}$ for large τ. In general, we expect $\alpha \leq \beta$, but equality should hold under linear cointegration. In a preliminary analysis, a plot of $\rho_{y,x}(\tau, t)/\rho_x(\tau, t)$ versus τ should help in identifying the existence of a linear cointegrating relationship.
- Obviously, this approach to linear cointegration using the cross-correlation or any other measure of mutual dependence is not useful for the analysis of non-cointegration with independent long-memory variables.
- Notice that if the series are short-memory then the ratio of the CCF to the ACF could also eventually converge to a nonzero value, and thus we need to impose that the individual series be long-memory.

We propose now an alternative condition for linear cointegration that implicitely constraints the individual series to be long-memory. Let $s_n^{(y,x)} = \sum_{\tau=1}^{n} \rho_{y,x}(\tau, t)$.

Proposition 2 *The series y_t, x_t are* **long-memory** *and* **linearly cointegrated** *if and only if as $n \to \infty$.*

1. *The sequence of partial sums $s_n^{(y,x)}$ diverges.*

2. *The ratio of sequences $s_n^{(y,x)}/s_n^{(x,x)}$ converges to a nonzero and finite real number, b.*

PROOF:

First, let us prove that our conditions are necessary if the series are long-memory and linearly cointegrated. Under the latter assumption, from proposition 1, there exists a nonzero and finite real number b such that

$$\rho_{y,x}(\tau, t) = b\, \rho_x(\tau, t) + o\left(\rho_x(\tau, t)\right), \tag{37}$$

where $o\left(\rho_x(\tau, t)\right)$ denotes a function of τ and t, which converges towards 0 faster than $\rho_x(\tau, t)$. On the other hand, since x_t is long-memory, $s_n^{(x,x)}$ diverges. Now, taking partial sums in the previous equation, it is inmediate that $s_n^{(y,x)}$ must also diverge, whereas the ratio $s_n^{(y,x)}/s_n^{(x,x)}$ will converge to b as n grows to infinity.

Now, for the converse, we prove that our two conditions in the proposition are sufficient to ensure that the individual series are both long-memory and linearly cointegrated. First, let us prove that if condition 1 holds then the individual series are necessarily long-memory. To see this, remark that we have $\rho_{y,x}(\tau, t) \leq C_x\, \rho_x(\tau, t)$ and $\rho_{y,x}(\tau, t) \leq C_y\, \rho_y(\tau, t)$, for some positive constants C_x, C_y and for sufficiently large τ's. Thus there exists finite constants C_1, C_2 such that

$$s_n^{(y,x)} \leq C_1 + C_x\, s_n^{(x)} \tag{38}$$

$$s_n^{(y,x)} \leq C_2 + C_y\, s_n^{(y)} \tag{39}$$

And since $s_n^{(y,x)}$ diverges, both $s_n^{(x)}$ and $s_n^{(y)}$ must also diverge.

Now, let us prove that if condition 2 in the proposition holds then the individual series are linearly cointegrated. To see it, remark that condition 2 can be rewritten as

$$s_n^{(y,x)} = b\, s_n^{(x,x)} + o\left(s_n^{(x,x)}\right), \tag{40}$$

for some nonzero and finite b. Or equivalently, that

$$\rho_{y,x}(\tau, t) = b\, \rho_x(\tau, t) + o\left(\rho_x(\tau, t)\right) \tag{41}$$

and thus $\lim_{\tau \to \infty} \frac{\rho_{y,x}(\tau,t)}{\rho_x(\tau,t)} = b$, which by proposition 1, establishes the linear cointegration of the long-memory series x_t, y_t.

Corollary 1 *If y_t, x_t are $I(d)$, respectively, with d restricted to the interval $[0,1]$, then a necessary and sufficient for the series to be cointegrated with $d = 1$, or equivalently,* **not jointly linearly mean-reverting** *is that $\lim_{\tau\to\infty} \rho_{y,x}(\tau,t) = b$, with $b \neq 0$.*

PROOF:

That the condition is necessary follows from the fact that $\rho_x(\tau,t) \to 1$ if $x_t \sim I(1)$, and that $\lim_{\tau\to\infty} \rho_{y,x}(\tau,t)/\rho_x(\tau,t) = b$ for a nonzero finite b, when the individual series x_t, y_t are linearly cointegrated.

To prove that the condition is also sufficient we remark that, $\mid \rho_{y,x}(\tau,t) \mid \leq C_x \mid \rho_x(\tau,t) \mid$, for a positive constant C_x. Thus the condition implies that $\mid b \mid \leq C_x \lim_{\tau\to\infty} \mid \rho_x(\tau,t) \mid$, and therefore that $\lim_{\tau\to\infty} \mid \rho_x(\tau,t) \mid \geq 0$. It follows that x_t is not mean-reverting. And since $x_t \sim I(d)$ with $d \in [0,1]$, the only possibility is that $d = 1$. Since $\mid \rho_{y,x}(\tau,t) \mid \leq C_y \mid \rho_y(\tau,t) \mid$ for some positive constant C_y, using the same argument we can prove that $y_t \sim I(1)$. Finally, that $\lim_{\tau\to\infty} (\rho_{y,x}(\tau,t)/\rho_x(\tau,t)) = b$ follows automatically, since $\rho_x(\tau,t) = 1 \; \forall\tau$, and the proof is complete.

Remarks:

- The name of joint mean-reversion conveys the idea that the mean of one series reverts around the mean of the other. Therefore under linear cointegration there cannot be joint mean-reversion, and viceversa.
- There is no loss of generality by restricting d_x, d_y to lie within the unit interval, since by proper differencing of the series we can determine the integers closest to d_x and d_y.

Corollary 2 *If y_t, x_t are $I(d_y), I(d_x)$, respectively, with d_x, d_y restricted to the interval $[0,1]$, then a necessary and sufficient condition for the series to be* **fractionally linearly cointegrated**, *or that they are cointegrated with $0 < d_x = d_y < 1$ is that*

1. $\lim_{\tau\to\infty} \rho_{y,x}(\tau,t) = 0$.
2. *The sequence $s_n^{(y,x)}$ diverges as n grows to infinity.*
3. *The ratio of sequences $s_n^{(y,x)}/s_n^{(x,x)}$ converges to a nonzero and finite real number, b, as n grows to infinity.*

PROOF:

That the conditions are necessary follows from the fact that if the individual series are (linearly) mean-reverting then $\lim_{\tau\to\infty}\rho_z(\tau,t) = 0\ \forall t$ and for $z = x, y$. Therefore if, in addition, the series are linearly cointegrated, then by propositions 1 and 2, conditions 1 to 3 above follow. For the converse, conditions 2 and 3 above are sufficient to ensure that the individual series are long-memory and linearly cointegrated. If condition 3 also holds then the individual series must be linearly mean-reverting, that is $x_t, y_t \sim I(d)$ with $0 < d < 1$.

Remark:

Once we know that $0 < d < 1$, we can inquire as to whether $d < 1/2$ (in which case the series are stationary) or not. To answer the question we only need to test whether the variance of x_t diverges or not.

Based on the new characterization above, we propose the following exploratory approach to linear cointegration testing. Again, we assume that the series have been differenced to such extent that their long-memory parameters lie within the interval $[0, 1]$:

1. Estimate by OLS the regression parameter a in the linear regression equation $y_t = ax_{t-\tau} + z_t$ for an increasingly large number of values of τ. Plot these estimates, say $\hat{a}_\tau$ as a function of τ. If the sequence of estimates $\hat{a}_\tau$ converges to a nonzero value then there is room to believe that the series are not jointly linearly mean-reverting. If not proceed to the next step.

2. Check the convergence of the ratio of partial sums $s_n^{(y,x)}/s_n^{(x,x)}$ as n grows without bound. If this sequence appears to diverge then there are good chances that the series are fractionally linearly cointegrated. If not, we may suspect that the series are not linearly cointegrated.

4 Linear cointegration in the frequency domain

Consider again $x_t \sim I(d_x)$, $y_t \sim I(d_y)$, and the series z_t formed as $z_t = y_t - ax_t$. To illustrate the meaning of linear cointegration in the frequency domain, let us also assume that we can define the spectral and cross-spectral densities for the different series (see Granger and Hatanaka, 1964[17], and Granger, 1983[15]). From the definition of z_t, it is easy to see that

$$S_z(\lambda) = a^2 S_x(\lambda) + S_y(\lambda) - a\left(S_{y,x}(\lambda) + S^*_{y,x}(\lambda)\right), \tag{42}$$

where $S_u(\lambda)$ and $S_{y,x}(\lambda)$ represent the spectrum of u_t ($u = x, y, z$) and the cross-spectrum of the pair x_t, y_t, respectively, and $S^*_{y,x}(\lambda)$ denotes the complex conjugate of $S_{y,x}(\lambda)$. Since $|S_{y,x}(\lambda)|^2 \leq S_x(\lambda)S_y(\lambda)$, and $S_x(\lambda) \sim A_x\lambda^{-2d_x}$, $S_y(\lambda) \sim A_y\lambda^{-2d_y}$ as $\lambda \to 0$, it is clear that the term $\lambda^{-2\max(d_x,d_y)}$ will dominate at low frequencies, and thus $z_t \sim I(\max[d_x, d_y])$ in general.

However, the previous algebraic rule breaks down under cointegration, that is, a situation in which $x_t, y_t \sim I(d)$, $d > 0$ and $\exists a \in \Re - \{0\}$ such that $z_t = y_t - ax_t \sim I(d_z)$, $d_z < d$. Defining $S_u(0) = \lim_{\lambda\to\infty} S_u(\lambda)$ (u=x,y,z), it is also straightforward to see that under cointegration, there exists a nonzero finite real number c such that $S_y(0)/S_x(0) = c$, and on the other hand, $S_z(0)/S_x(0) = 0$ (Granger, 1980[14]). Therefore, $(a^2+c)S_x(0) - aS_{y,x}(0) = 0$, and since both a and c are nonzero, the ratio $S_{y,x}(0)/S_x(0)$ must be nonzero and finite, as implied by proposition 2 in the previous section. Intuitively, this means that x_t, y_t have the long-wave component in common (*comovement*).

It is also possible to re-write the statement of proposition 1 in the frequency domain. For this, consider the inverse discrete Fourier transforms of the sequences $\rho_{y,x}(\tau)$ and $\rho_x(\tau)$. That is, the representations:

$$\rho_x(\tau) = S_x(0) + \sum_{\lambda>0} S_x(\lambda)exp(j\tau 2\pi\lambda) \tag{43}$$

$$\rho_{y,x}(\tau) = S_{y,x}(0) + \sum_{\lambda>0} S_{y,x}(\lambda)exp(j\tau 2\pi\lambda), \tag{44}$$

where $j^2 = -1$. Now, if x_t, y_t are long-memory, the long-wave component (i.e. the spectral component at frequency $\lambda = 0$) will tend dominate the spectrum as the lag, τ, goes to infinity. A heuristic explanation for this could be that the terms in the sum on the left of the previous equation tend to cancel each other because of the rapidly oscillating exponentials (as $\tau \to \infty$). Thus we can write:

$$\lim_{\tau\to\infty} \frac{\rho_{y,x}(\tau)}{\rho_x(\tau)} = \frac{S_{y,x}(0)}{S_x(0)}, \tag{45}$$

which by propositions 1 and 2, is known to be nonzero and finite. A more formal way of showing the previous equivalence is the following. Recalling that

$$\frac{S_{y,x}(0)}{S_x(0)} = \frac{s_n^{(y,x)}}{s_n^{x,x}} \tag{46}$$

$$= \frac{\sum_{\tau>0}\rho_{y,x}(\tau)}{\sum_{\tau>0}\rho_x(\tau)}. \tag{47}$$

Now let us take an arbitrarily large positive real number, T, and split the correlation sums in both the denominator and numerator, so that we can write

$$\frac{\sum_{\tau>0}\rho_{y,x}(\tau)}{\sum_{\tau>0}\rho_x(\tau)} = \frac{\sum_{0<\tau<T}\rho_{y,x}(\tau)+\sum_{\tau\geq T}\rho_{y,x}(\tau)}{\sum_{0<\tau<T}\rho_x(\tau)+\sum_{\tau\geq T}\rho_x(\tau)}. \tag{48}$$

Finally, divide every term in the righthand side ratio by $\sum_{\tau\geq T}\rho_x(\tau)$, which is unbounded for any finite T. Thus we obtain

$$\frac{S_{y,x}(0)}{S_x(0)} = \frac{\sum_{\tau\geq T}\rho_{y,x}(\tau)}{\sum_{\tau\geq T}\rho_x(\tau)}. \tag{49}$$

Now, letting T tend to infinity, we obtain our result:

$$\frac{S_{y,x}(0)}{S_x(0)} = \lim_{\tau\to\infty}\frac{\rho_{y,x}(\tau)}{\rho_x(\tau)}. \tag{50}$$

Notice that both propositions 1 and 2 point to the same spectral result, namely that under linear cointegration, the ratio $S_{y,x}(0)/S_x(0)$ is nonzero and finite. The later result also suggests using the distance between the time domain estimator of $\frac{\rho_{y,x}(\tau)}{\rho_x(\tau)}$, as τ approaches infinity, and the spectral estimator of $\frac{S_{y,x}(0)}{S_x(0)}$, as a way of testing the hypothesis of linear cointegration. In one such test, if the hypothesis is rejected then it could be because either the series are not long-memory or because there is no comovement at the zero frequency.

5 Detecting nonlinearities in cointegration

All the measures of memory and serial dependence discussed so far are linear. In order to extent those indices to a nonlinear context we have to introduce new concepts that take into account eventual nonlinear dependencies. Only few works consider the simultaneous treatment of long-memory and nonlinearity. See for example Escribano (1986,1987)[8, 9], Granger and Hallman (1991)[18, 19], Granger and Terasvirta (1993)[21], Granger (1995)[16], Aparicio(1995)[2, 4, 3], Escribano and Mira (1997)[10, 11], and Aparicio and Escribano (1997)[1]. Several papers have analyzed the concept of $I(0)$ versus $I(1)$, or of long-memory in a general contex, by considering that a series is $I(0)$ if it satisfies a *functional central limit theorem*; see for example Lo (1991)[24]. Kwiatkowski et al. (1992)[23], and Escribano and Mira (1997) [10]. Those papers are based on the concept of *mixing* as a measure of dependence. Alternatively, we could have considered the concept of *near epoch dependence* (NED) (see Gallant and White, 1988[13]) instead of that of mixing (see Escribano and Mira, 1997 [11] for the use of the NED concept in *error correction models*). Finally, in Aparicio (1995)[2, 4, 3] and in Aparicio and Escribano (1997)[1] the concept of mutual information is used as a measure of serial dependence and of cross-dependence.

Henceforth we consider the problem of discriminating between linear and nonlinear cointegration. One way to detect the presence of nonlinearity in a relationship, which bypasses the previous difficulties, is by comparing a measure of cross-dependence for the given pair of series and for pairs of series constructed in such a way as to preserve the linear part of the relationship while obliterating any higher-order dependence. A candidate test statistic for this detection problem could be the following one:

$$R_{y,x} = \lim_{\tau \to \infty} \left(\frac{i_{y,x}(\tau)}{i_x(\tau)} - \frac{S_{y,x}(0)}{S_x(0)} \right), \tag{51}$$

where $i_x(\tau)$ and $i_{y,x}(\tau)$ represent measures of dependence and cross-dependence which generalize the auto- and cross-correlation function, respectively. The concepts of *mixing* and that of *mutual information* (see Aparicio and Escribano, 1997 [1]) offer some possibilities to implement these measures.

Suppose we construct a pair of series (x'_t, y'_t) from the given pair (x_t, y_t), satisfying the following constraints:

$$S_{x'}(\lambda) = S_x(\lambda) \tag{52}$$

$$S_{y'}(\lambda) = S_y(\lambda) \tag{53}$$

$$|S_{y',x'}(\lambda)| = |S_{y,x}(\lambda)| \tag{54}$$

To detect the presence of neglected nonlinearity in a cointegration relationship using this technique, the linear features of the relationship between x_t and y_t must be preserved in the linear replicas, x'_t and y'_t. Suppose that $x_t = cy_t + f(y_t) + \epsilon_t$, where $f(.)$ is a nonlinear function and ϵ_t is a short-memory disturbance. Therefore, we may obtain an estimate of the Fourier transfer function $H_{y,x}(\lambda) = \mathcal{F}_x(\lambda)/\mathcal{F}_y(\lambda)$, where $\mathcal{F}_u(\lambda)$ denotes the discrete Fourier transform of the sequence u_t. Remark that the linear part of the relationship between x_t and y_t "lives" only in $|H_{y,x}(\lambda)|$. Now let $H'_{y,x}(\lambda_j) = |H_{y,x}(\lambda_k)|\, exp[j\phi(\lambda_k)]$, where $\phi(\lambda_k)$ is an *i.i.d.* sequence (indexed by k) of *r.v.*'s. Obtain y'_t as the inverse discrete Fourier transform of $\mathcal{F}_x(\lambda)/H'_{y,x}(\lambda)$. Notice that each *i.i.d.* phase sequence $\phi(\lambda_k)$ serves to generate one linear replica. The relationship between the series x'_t, y'_t preserves the linear features of the one between x_t and y_t, but should be devoid of the nonlinearity in the latter. This difference will translate into a value of the statistic $R_{y,x}$ significantly larger than zero.

Bootstrap replicas could also be obtained in the time domain. To do so, we first bootstrap the linear residuals of one of the series, say x_t, fitted with a sufficiently long autoregressive model. In this way, we generate the x'_t replica. The second step is to generate y'_t as $y'_t = \hat{a}x'_t + z'_t$ where $\hat{a}$ is the OLS estimate of the parameter of the regression of y_t on x_t (this is the simplest of all possible linear cointegration models, but we may also consider models including several lags of x_t, or allowing for delays of one series with respect to the other), and z'_t represent the bootstrap linear regression residuals.

Clearly, the problem with a parametric (time-domain) bootstrapping approach is that there is a high risk of uncorrectly specifying the model for x_t (i.e. misspecification will certainly occur if x_t is fractionally integrated, since the fitted AR model order will necessarily be finite). In the sequel, any pair of bootstrapped series (x'_t, y'_t) which preserve the linear structure of the relationship between x_t and y_t, will be referred to as *linear bootstrap replicas* of the pair (x_t, y_t). Our next step could be to define a *nonlinear coherence* measure.

$$\eta(x, y, x', y'; \tau) = \begin{cases} \mid i_{y,x}(\tau)/i_x(\tau) - i_{y',x'}(\tau)/i_{x'}(\tau) \mid, & \text{if } i_x(\tau), i_{x'}(\tau) \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

Since the pair (x'_t, y'_t) preserves only the dependencies in (x_t, y_t) captured by their *amplitude cross-spectrum*, $\eta(x, y, x', y'; \tau)$ will provide a measure of the higher-order dependencies living at any lag, τ. To reduce the variability of $\eta(x, y, x', y'; \tau)$ we could average this statistic or a transformation of it on a number M of linear replicas (x'_t, y'_t). For instance, we may consider the nonlinear average

$$R_{y,x}(M, \tau) = 1 - M^{-1} \sum_{m=1}^{M} r(x, y, x'^{(m)}, y'^{(m)}; \tau) \tag{55}$$

with

$$r(x, y, x'^{(m)}, y'^{(m)}; \tau) = exp\left(\frac{\eta(x, y, x'^{(m)}, y'^{(m)}; \tau)}{\sigma_\eta}\right) \tag{56}$$

where σ^2_η represents a bootstrap estimate of the variance in the sequence $\left\{\eta(x, y, x'^{(m)}, y'^{(m)}; \tau)\right\}_m$. The statistic $R_{y,x}(M, \tau)$ provides a measure of nonlinearity in the cointegrating relationship, for sufficiently large τ. This measure is confined to the interval $(0, 1]$ and behaves in a simple way. Indeed, as the incidence of nonlinearity is higher, the statistic will approach 1, while it will tend to concentrate around 0 as m grows, when the relationship is truly linear. In fact, we expect that for a linear relationship, assuming sufficient moment conditions are satisfied by the sequence $\eta(x, y, x'^{(m)}, y'^{(m)}; \tau)$, indexed by m, it may be possible to find a scaling law M^a with $a > 0$ so that the standardized sum $M^a R_{y,x}(M, \tau)$ has a well-defined limiting distribution. Obviously, the standardized sum will diverge towards infinity under the alternative of nonlinearity in the relationship.

Finally, if the series x_t, y_t are not cointegrated, $M^a R_{y,x}(M,\tau)$ will diverge but with wandering sign. In practice, we can test the linearity of the long-run relationship without knowing the limiting distribution of our statistic under any hypothesis. In this case, for a given M and sufficiently large τ, we can estimate the empirical critical value b_α for a one-sided test, such that $P\left(R_{y,x}(M,\tau) > b_\alpha\right) = \alpha$ under linearity, for a given significance level, α. The null hypothesis will be rejected at this level whenever $R_{y,x}(M,\tau) > b_\alpha$.

In the sequel we present some simulation results obtained by using for $i_x(\tau)$ and $i_{y,x}(\tau)$ in $\eta(.)$, the *mutual information function* introduced in Aparicio and Escribano (1997 [1]). That is, here we take $i_x(\tau) = E_t I(X_t, X_{t-\tau})$, and $i_{y,x}(\tau) = E_t I(Y_t, X_{t-\tau})$ where E_t is a time-averaging operator, and $I(V,W)$ represents the mutual information of the *r.v.*'s V, W, which can be defined as

$$I(V,W) = E\left[log \frac{f_{v,w}(V,W)}{f_v(V) f_w(W)}\right], \tag{57}$$

with $f_{y,x}(Y,X)$ and $f_x(X)$ representing joint and univariate probability density functions, respectively. The function $i_x(\tau)$ was evaluated using the following estimator by Robinson (1991) [27], where N is the sample size,

$$\begin{aligned} \hat{i}^{(N)}_{y,x}(\tau) &= N^{-1} \sum_{t=1}^{N} \hat{i}_{y,x}(\tau, t) \\ &\approx N_\gamma^{-1} \sum_{t \in \mathcal{S}} c_t(\gamma) log \left(\frac{\hat{f}_{y,x}(Y_t, X_{t-\tau})}{\hat{f}_y(Y_t) \hat{f}_x(X_{t-\tau})} \right), \end{aligned} \tag{58}$$

with

$$c_t(\gamma) = \begin{cases} 1+\gamma, & \text{for } t \text{ odd} \\ 1-\gamma, & \text{for } t \text{ even} \end{cases}$$

where $\gamma \geq 0$, $N_\gamma = N$ for N even, and $N_\gamma = N + \gamma$, for N odd. Here, $\hat{f}_{y,x}(.,.)$ and $\hat{f}_x(.)$ are estimators of the corresponding bivariate and univariate *pdf*'s (which may be time-varying), and the set $\mathcal{S}$ is introduced to make explicit the exclusion of certain inocuous summands, which can occur, for example, when $\hat{f}_{y,x}(.,.) \leq 0$ or $\hat{f}_x(.) \leq 0$, or when logarithms cannot be taken. The densities can be estimated using *kernel smoothers* (Breiman et al., 1977 [6]).

We simulated $M = 1, 5, 10, 50$, and 100 replications of cointegrated and non-cointegrated series, with varying sample sizes, $N = 500, 1000$. The linear cointegrated series were generated as in figure 1. The nonlinearly cointegrated ones were computed applying third polynomial transformations to a common random walk. The coefficients of the polynomials were chosen at random. Finally, the non-cointegrated series were either pairs of independent random walks ($H_{2,1}$) or mutually dependent short-memory series ($H_{2,2}$). In the latter case, the series were generated according to the model $y_t = x_t + \epsilon_t$, $z_t = a_0 + a_1 x_t + a_2 x_t^2 + a_3 x_t^3 + \epsilon_t'$, where $x_t = a_4 e_{t-2} e_{t-1} + e_t$, $\epsilon_t, \epsilon_t', e_t$ are *i.i.d.* sequences, and the a_i were chosen at random. For the experiment, we took a sample size of $N = 1000$ and a lag of $\tau = 0$. The linear bootstrap replicas were generated in the time-domain, by resampling the linear OLS residuals of one of the series, and the linear regression residuals $\hat{z}_t$ from the simple regression model $y_t = ax_t + z_t$.

We found that the values of $R_{y,x}(M, 0)$ were comparatively low under linear cointegration, but exhibited a small negative bias, probably caused by the asymmetry in the distribution of the statistic under H_0. In spite of this small negative bias in $M^{1/2}R_{y,x}(M, 0)$, there was a remarkable difference between the values of this statistic under linear cointegration, and those under nonlinear cointegration. In the former case, these values were small in magnitude and negative, while in the latter, they were generally positive and comparatively large (by at least one order magnitude for $M = 1$). Under non-cointegration, most of the times, $M^{1/2}R_{y,x}(M, 0)$ took values comparable in magnitude to those under nonlinear cointegration (sometimes much larger), but with varying sign.

For the sample sizes $N = 500, 1000$, $M^{1/2}R_{y,x}(M, 0)$ was positive in approximately 90% of the replications under nonlinearity, while it was only positive in 60% of the replications, for $N = 100$. In table 2, we show the sample mean, absolute mean, and standard deviation (between brackets) of $R_{y,x}(1, 0)$ for $N = 500, 1000$, from an experiment involving 100 replications of linear, nonlinear and non-cointegrated (independent random walks) series. As N was increased, $R_{y,x}(1, 0)$ tended to take systematically positive values in the inteval $(0, 1)$, under H_1 This was so to such extent that it seemed possible testing for linearity in cointegration from just the sign fluctuations of this statistic in moderate to large samples.

The frequency of rejection of the linearity hypothesis (assuming cointegration) for a one-sided test based on $R_{y,x}(1, 0)$ and applied to 100 replications of nonlinearly cointegrated series, was approximately 90% for both sample sizes, using the bootstrap critical values estimated under H_0 of linear cointegration, from 1000 bootstrap replications. Moreover, the inconsistency of the sample mean of $R_{y,x}(1, 0)$ under non-cointegration, and the positivity of this statistic under H_1 of nonlinearity in moderate to large samples, suggest a way of testing the hypothesis H_2 of non-cointegration against the joint hypothesis of cointegration (H_1, H_0). For instance, one may generate a sufficiently large number of linear replicas of the pair (y, x), preserving only their linear relationship, and estimate

$R_{y,x}(1,0)$ for each one. Each pair of linear replicas could be considered as an independent realization from the linear part of the model which produced (y, x). Under non-cointegration, the values of the statistic will exhibit wandering signs, but large absolute values, contrary to its expected behaviour under either H_1 or H_0 (as we show under H_1, it takes very often comparatively large positive values, while it has a small absolute value under H_0).

6 Conclusions

In this paper we have proposed an alternative characterization of long-memory and of linear cointegration, both integer and fractional, in the univariate case. It is based on simple statistics constructed from a defintion of the autocorrelation and cross-correlation functions of the series, that allows for heterogeneity in the latter. We formulated this characterization in both the time and the frequency domains, and showed the equivalence between them. Finally, we showed how it can be used to test for linear cointegration and for nonlinearities in a cointegrating relationship. The results obtained with some simulation experiments seemed support the validity of the nonlinearity testing device.

References

[1] F.M. Aparicio and A. Escribano. Information-theoretic analysis of serial dependence and cointegration. Working paper of the Depart. of Statistics of Universidad Carlos III de Madrid, 1997.

[2] F.M. Aparicio Acosta. *Nonlinear modelling and analysis under long-range dependence with an application to positive time series.* PhD thesis, Swiss Federal Institute of Technology (EPFL) -Signal Processing Laboratory-, Lausanne, Switzerland, 1995.

[3] F.M. Aparicio Acosta and C.W.J. Granger. Information-theoretic schemes for linearity testing under long-range dependence and under cointegration. Working Paper (Depart. of Economics of the University of California at San Diego), March 1995.

[4] F.M. Aparicio Acosta and C.W.J. Granger. Nonlinear cointegration and some new tests for comovements. Working Paper (Depart. of Economics of the University of California at San Diego), March 1995.

[5] G.E.P. Box and G.M. Jenkins. *Time series analysis: forecasting and control.* Holden-Day, San Francisco, 1970.

[6] L. Breiman, W.S. Meisel, and E. Purcell. Variable kernel estimates of multivariate densities and their calibration. *Technometrics*, 19:135–144, 1977.

[7] R.F. Engle and C.W.J. Granger. Co-integration and error correction: representation, estimation and testing. *Econometrica*, 55:251–276, 1987.

[8] A. Escribano. *Nonlinear error-correction. The case of Money dmand in the U.K. (1878-1970). Chapter IV.* PhD thesis, Department of Economics of the University of California at San Diego, La Jolla, USA, 1986.

[9] A. Escribano. Error-correcting systems: nonlinear adjustment to linear long-run relationships. CORE Discussion paper 8730, C.O.R.E., 1987.

[10] A. Escribano and S. Mira. Nonlinear cointegration with mixing errors. Working paper 97-25 of the Depart. of Statistics of Universidad Carlos III de Madrid, 1997.

[11] A. Escribano and S. Mira. Nonlinear error correction models. Working paper 97-26 of the Depart. of Statistics of Universidad Carlos III de Madrid, 1997.

[12] W. A. Fuller. *Introduction to statistical time series.* John Wiley and Sons, N.Y., 1976.

[13] R. Gallant and H. White. *A unified theory of estimation and inference for nonlinear dynamic models.* Basil Blackwell, N.Y., 1988.

[14] C.W.J. Granger. Some properties of time series data and their use in econometric model specification. *Journal of econometrics*, 16:121–130, 1981.

[15] C.W.J. Granger. Co-integrated variables and error-correcting models. Discussion paper 83-13 (Depart. of Economics of the University of California at San Diego), 1983.

[16] C.W.J. Granger. Modelling nonlinear relationships between extended-memory variables. *Econometrica*, 63(2):265–279, 1995.

[17] C.W.J. Granger and M. Hatanaka. *Spectral analysis of economic time series.* Princeton University Press, New Jersey, 1964.

[18] C.W.J. Granger and Hallman J.J. Long-memory series with attractors. *Oxford Bulletin of Economics and Statistics*, 53(1):11–26, 1991.

[19] C.W.J. Granger and Hallman J.J. Nonlinear transformations of integrated time series. *Journal of Time Series Analysis*, 12(3):207–224, 1991.

[20] C.W.J. Granger and R. Joyeux. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–29, 1980.

[21] C.W.J. Granger and T. Terasvirta. *Modeling nonlinear economic relationships.* Oxford University Press, Oxford, 1993.

[22] J.R.M. Hosking. Fractional differencing. *Biometrika*, 68:165–176, 1981.

[23] D. Kwiatkowski, P.D.B. Phillips, P. Schmidt, and Shin. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of econometrics*, 54:159–178, 1992.

[24] A.W. Lo. Long-term memory in stock market prices. *Econometrica*, 59(4):1279–1313, 1991.

[25] J. G. MacKinnon. Critical values for cointegration. Working Paper (Depart. of Economics of the University of California at San Diego), January 1990.

[26] P.C.B. Phillips. Time series regression with a unit root. *Econometrica*, 55:277–301, 1987.

[27] P.M. Robinson. Consistent nonparametric entropy-based testing. *Review of Economic Studies*, 58:437–453, 1991.

[28] H. White. *Asymptotic theory for econometricians.* Academic Press, San Diego, 1984.

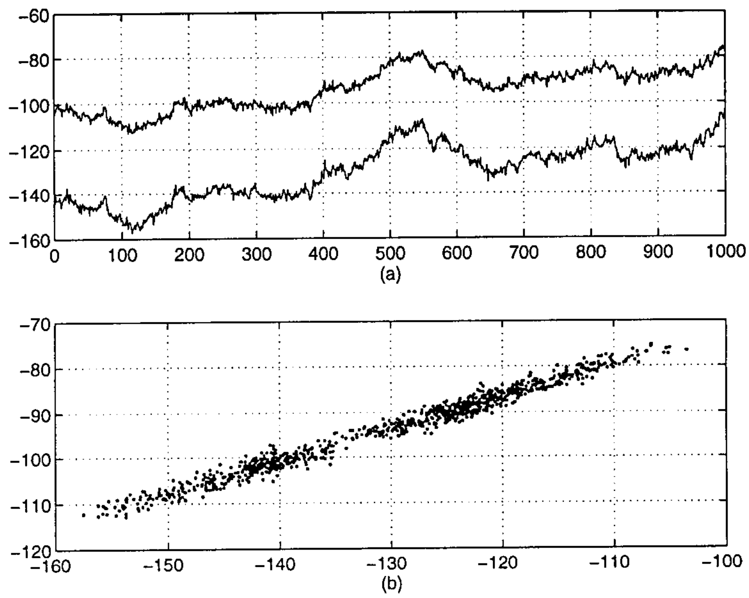

Figure 1: Two simulated linearly cointegrated random walks (a) and their scatter plot (b). The series, x_t, x'_t were generated with the model: $x_t = aw_t + \epsilon_t$, $x'_t = w_t + \epsilon'_t$, $w_t = w_{t-1} + \xi_t$, where $\epsilon_t, \epsilon'_t, \xi_t$ are independent sequences of *i.i.d.* Gaussian *r.v.*'s.

Test statistic	*linear cointeg.*	*nonlin. cointeg.*	*non-cointeg.*
$E(R_{y,x}(1,0)$, *($N = 500$)*	-0.052 (0.213)	0.221 (0.91)	0.231 (1.349)
$E(\|R_{y,x}(1,0)\|)$, *($N = 500$)*	0.087	0.56	0.83
$E(R_{y,x}(1,0)$, *($N = 1000$)*	-0.05 (0.157)	0.425 (0.735)	0.048 (1.304)
$E(\|R_{y,x}(1,0)\|)$, *($N = 1000$)*	0.086	0.464	0.822

Table 1: Means, absolute means and standard deviations of a mutual-information-based statistic for testing linearity in cointegration.