Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19103-6297

pier@ssc.upenn.edu
http://www.econ.upenn.edu/pier

PIER Working Paper 97-027

"Deciding between I (0) and I (1) FLIL-based bounds"

by

Valentina Corradi

Deciding between $I(0)$ and $I(1)$ via FLIL-based bounds

Valentina Corradi[1]
University of Pennsylvania
Department of Economics
3718 Locust Walk
Philadelphia, PA 19104-6297
e-mail: corradi@econ.sas.upenn.edu
tel. (215) 898-1505

May 1997

[1]I wish to thank a Co-Editor, three anonymous referees, as well as Filippo Altissimo, Frank Diebold, Jin Hahn, Atsushi Inoue, Lutz Kilian, Serena Ng, Shinichi Sakata, Chor-Yiu Sin and the seminar participants at the University of Pennsylvania, University of Florida, Econometric Society 1996 Winter Meeting for very helpful comments and suggestions

Abstract

We construct properly scaled functionals of R^p-valued partial sums of de-meaned data and derive bounds via the functional law of the iterated logarithm (FLIL) for strong-mixing processes. If we obtain a value below or equal to the bound we decide in favor of $I(0)$, otherwise we decide in favor of $I(1)$; this provides a consistent rule for classifying time series as being $I(1)$ or $I(0)$. The nice feature of the procedure lies in the almost sure nature of the bound, guaranteeing a limsup type result. We finally provide conditions for the strong consistency of estimators of the variance in the dependent and heterogeneous case. The finite sample behavior of the procedure is analyzed via Monte Carlo simulations.

1 Introduction

Consistent rules for classifying time series as being integrated of order 1, $I(1)$, or integrated of order zero, $I(0)$, have been recently proposed, see e.g. Phillips and Ploberger (1994, 1996 hereafter PP) and Stock (1994). The main idea underlying these papers is to construct a Bayes factor, or a posterior odd ratio, such that, under the $I(0)$, ($I(1)$), hypothesis the ratio converges in probability to zero and under the $I(1)$, ($I(0)$), hypothesis it diverges to infinity. According to the PP's PIC (Posterior Information Criterion), we choose the $I(1)$ hypothesis, versus the $I(0)$, if the Bayes model likelihood ratio is less than one, and viceversa. On the other hand, Stock (1994) does not study the posterior distribution of the data directly, but he constructs the likelihood ratio of observing some functionals of partial sums of the (de-meaned or de-trended) data under either the $I(0)$ or the $I(1)$ hypothesis; the classification rates are then computed by direct simulation. More recently Sin (1996), extending previous work by Sin and White (1996), proposed a flexible information criterion for selecting strictly nested, linear, possibly non-stationary, models. According to Sin's information criterion, we choose unit root if we obtain a negative value and we choose no unit root if we obtain a strictly positive value; this also delivers a completely consistent procedure.

In this paper we propose a new completely consistent rule for deciding between $I(1)$ and $I(0)$, from a "non-Bayesian" point of view. The main feauture of the procedure we shall consider lies in the almost sure nature of the bound. In what follows we define a purely stochastic process to be $I(d)$ if its d-th differences are strong mixing and satisfy some mild moment conditions. Here we concentrate on the de-meaned case. The definition of an $I(d)$ process as a process whose d-th differences are strong mixing is slightly more restrictive than the more common definition of an $I(d)$ process as a process whose d-th differences have partial sums that, properly scaled, satisfy a functional central limit theorem. Although FLIL are available also for mixingales or for near epoch dependent functions on a mixing basis (e.g. Altissimo and Corradi, 1996), here we confine our attention to the mixing case, this allows us to limit at a minimum the assumptions we need.

Basically we construct a properly scaled functional of the Euclidean norm of R^p-valued partial sums of de-meaned data. Under mild moment conditions, partial sums of R^p-

valued strong mixing processes satisfy an almost sure invariance principle, e.g. Eberlain (1986). It then follows that the set of norm limit points of such partial sums coincides, almost surely, with the set of norm limit points of the p-dimensional Brownian motion; we then obtain a functional law of the iterated logarithm (FLIL) for the Euclidean norm of the partial sums of R^p-valued strong mixing processes. Finally, using a result by Strassen (1964) that plays, in an almost sure context, the same role as the continuous mapping theorem plays in a weak convergence context, we obtain an almost sure bound for functionals of the Euclidean norm of partial sums of strong mixing processes. More precisely, we shall consider the following statistic $b_T^{-1} \sum_{t=1}^{T} \|\hat{\Omega}_T^{-1/2} \sum_{j=1}^{t} (X_j - \bar{X})\|$, where $b_T = T^{3/2}\sqrt{2loglogT}$, log denotes the natural logarithm, X_t is a p-dimensional process and $\hat{\Omega}_T$ is a proper estimator for the variance and $\bar{X}$ denotes the sample mean. We shall show that, if X_t is a strong mixing process, then the limsup of the statistic above is, almost surely, less than or equal to $1/\sqrt{3} + 1/2$ for the case of $p = 1$; while for $p > 1$, the limsup of the statistic above is almost surely less than or equal to $2/3 + 1/2$. In fact, suppose that $E(X_t) = 0$, then in the one-dimensional case $\limsup_{T\to\infty} b_T^{-1} \sum_{t=1}^{T} \|\Omega_0^{-1/2} \sum_{j=1}^{t} X_j\| = \frac{1}{\sqrt{3}}$; instead for the multidimensional case, we can just provide an almost sure upper bound, that is $\limsup_{T\to\infty} b_T^{-1} \sum_{t=1}^{T} \|\Omega_0^{-1/2} \sum_{j=1}^{t} X_j\| \leq \frac{2}{3}$. In both the cases, the additional "1/2" term is due to the fact that we are considering deviations from the sample mean. On the other hand, if X_t is an integrated process, or if at least one component of X_t is $I(1)$, then the statistic above diverges at rate $\frac{\sqrt{T}}{\sqrt{2l_T loglog(T)}}$, where l_T denotes the order of magnitude of the lags used in the estimation of variance. Thus we decide in favor of $I(0)$ whenever we get a value for the statistic smaller or equal to $1/\sqrt{3} + 1/2$, for $p = 1$, or smaller or equal to $2/3 + 1/2$, for $p > 1$, and we decide instead in favor of $I(1)$ whenever we get a larger value.

The statistic proposed above is very similar to the statistic proposed by Kwiatkowski et al. (1992, KPSS hereafter). In principle we could have used their statistic and let the critical values depend on the sample size. While several choices of sequences of critical values assure the zero size property, it is not clear which sequence would also assure the almost sure nature of the bound. Furthermore it is not easy, although very challenging, to find a sequence of critical values, depending on T, such that, as T gets large, the size approaches zero, and in the mean time, for any given T, we have the "highest" possible power. Basically the problem is to find the "optimal" rate of convergence to zero, as $T \to \infty$, for the size, subject to a "power" constraint.

Intuitively in order to preserve the almost sure nature of the bound we need that, in the $I(0)$ case, $\hat{\Omega}_T$ converges almost surely to Ω_0; we then provide conditions under which Newey-West type variance estimators are strongly, rather than weakly, consistent.

As we mentioned above other completely consistent procedures, for choosing between $I(0)$ and $I(1)$ have been already proposed; the advantage of the procedure proposed here lies in the almost sure nature of the bounds, so that we get a limsup type result.

While almost sure bounds are available also for the detrended case, they are too "loose" to be useful in practice; for this reason here we do not consider the detrended case.

The finite sample properties of the procedure are analyzed via some Monte Carlo simulations. For several different DGPs the actual size is very close to zero even for samples as small as 100 observations; the finite sample power properties are instead less satisfactory. As expected, faster rates of growth of the lag truncation parameter l_T have an adverse effect on finite sample power and a beneficial effect on the size; as the former effect is stronger, it is recommendable to let l_T increase with T at a slow rate. Overall if we compare our simulations results with those of PP (Table 2, 1994), our procedure is characterized by lower finite sample type I error and by higher type II error; thus a joint use of the two procedures may be useful in empirical applications.

The paper is organized as follows. Section 2 treats the almost sure bounds for the demeaned case. As a by-product of our analysis we also provide conditions for the strong consistency of Newey-West type estimators of the covariances of strong mixing processes. The behavior of the statistic under the $I(1)$ hypothesis is considered in Section 3. The findings from the Monte Carlo simulations are reported in Section 4. All the proofs are collected in the Appendix.

2 Almost sure bounds for the de-meaned case

In this section we derive almost sure bounds for functionals of the Euclidean norm of partial sums of strong mixing processes, in the de-meaned case. The final outcome is obtained in three steps. First we give a set of conditions, essentially moment conditions, under which partial sums of R^p-valued dependent, heterogeneous processes satisfy a strong invariance principle. In order to have a nuisance parameters free almost sure bound, we need to scale our statistic by the variance term; as in general we do not know it, we need a strong consistent estimator for it. While weak consistency results for estimators of the variance in the heterogeneous and dependent case, have been already obtained, in different frameworks, by several authors, among others White (1984), Newey and West (1987,1994), Gallant and White (1988), Robinson (1988), Andrews (1991), Andrews and Mohanan (1992), Hansen (1992), to the best of our knowledge no result on strong consistency is available so far. We then provide conditions for the almost sure convergence of the estimator of the variance in the dependent and heterogeneous case. Finally, we obtain an almost sure bound for the statistic $1/b_T \sum_{t=1}^{T} \|\hat{\Omega}_T^{-1/2} \sum_{j=1}^{t}(X_j - \bar{X})\|$. Let

$$X_t = \mu + \epsilon_t, \tag{1}$$

we begin by stating the set of assumptions on ϵ_t we shall use below.

A1: $\{\epsilon_t\}_{t=-\infty}^{\infty}$ is a R^p-valued, zero mean, strong mixing process, with mixing coefficients α_k satisfying

$$\sum_{k=0}^{\infty} \alpha_k^{1/q-1/r} = A < \infty$$

where $1 \leq q < r \leq 2+\delta$, $\delta > 0$.

A2: $\sup_t E(|\epsilon_t^j|^{2+\delta}) \leq C_1 < \infty$, $j = 1, 2, \ldots p$, $\delta > 0$, where ϵ_t^j denotes the j-th component of the $p \times 1$ vector ϵ_t.

A3: $\sup_t E(|\epsilon_t^j|^{4+2\delta}) \leq C_2 < \infty$, $\forall j$.

It should be pointed out that the arbitrarily small, strictly positive constant δ, stated in A1-A3, is the same number.

Let $S_T(m) = \sum_{k=m+1}^{m+T} \epsilon_k$ and let $F_m = \sigma(\epsilon_m, \epsilon_{m-1}, \ldots)$, and let $S_T^j(m)$ be the j-th component of $S_T(m)$.

A4: There exists a positive definite $p \times p$ matrix, $\Omega_0 = [\omega_0^{ij}]$,with $tr(\Omega_0) < \infty$, such that

$\forall i, j = 1, \dots p,$

$$|\frac{1}{T}E(S_T^i(m)S_T^j(m)) - \omega_0^{ij}| \leq C_3 T^{-\psi}$$

for some $\psi > 0$, and C_3 independent of m.

A5:

$$E|E(S_T^j(m)S_T^i(m)|F_m) - E(S_T^j(m)S_T^i(m))| \leq C_4 T^{1-\phi}$$

for some $\phi > 0$, and C_4 independent of m.

Let $Z_{t,\tau}^{ij} = \epsilon_t^i \epsilon_{t+\tau}^j - E(\epsilon_t^i \epsilon_{t+\tau}^j)$,

A6: there exists $0 < v_\tau^{ij} \leq v_{ij} < \infty$, such that for $\tau = 1, 2, \dots l_T$ and $j = 1, 2, \dots p$,

$$|\frac{1}{T}E\left(\sum_{t=m+1}^{T+m-\tau} Z_{t,\tau}^{ij}\right)^2 - v_\tau^{ij}| \leq C_5 T^{-\rho}$$

for some $\rho > 0$, C_5 independent of both m and τ

A7

$$E|E\left(\left(\sum_{t=m+1}^{T+m} Z_{t,\tau}^{ij}\right)^2 |F_m\right) - E\left(\sum_{t=m+1}^{T+m} Z_{t,\tau}^{ij}\right)^2 | \leq C_6 T^{1-\theta}$$

$\forall i, j$, for some $\theta > 0$, and C_6 independent of both τ and m.

A1-A3 are standard memory and moment conditions. As we shall see below, A1, A2, A4 and A5 suffice for ϵ_t satisfying a strong invariance principle. By comparing A1 and A2, we note the usual trade-off between the rate of decay to zero of the mixing coefficients and the strength of the moment conditions imposed. Needless to say A3 implies A2. Basically in Lemma 2.2 below, we shall show that the mixing requirement in A1, together with A2, suffices for the satisfaction of the memory condition 1.1 (when $q = 1$) or 1.3 (when $q = 2$) in Eberlain (1986); that is given A1 and A2

$$\|E(S_T(m)|F_m)\|_q \leq C \tag{2}$$

for $0 < C < \infty$, uniformly in m, where $\|.\|_q$ denotes the L^q-norm.

Indeed under few additional conditions, also mixingales processes and near epoch dependent (NED) functions on a mixing basis satisfy the memory condition (2) (see e.g. Altissimo and Corradi, 1996).

A well known example of non-mixing $I(0)$ process, in the sense of having bounded spectral density at zero, is Andrews' (1984) autoregressive process $X_t = \rho X_{t-1} + u_t$, where

$\rho \in (0, 1/2]$ and u_t is a doubly infinite sequence of independent, identically distribuited Bernoulli random variables. Suppose that $P(u_t = 0) = P(u_t = 1) = 1/2, \forall t$, so that $E(u_t) = 1/2$. Clearly X_t cannot satisfy the memory condition (2) as $E(X_t) \neq 0$. Now let us consider the process $\tilde{X}_t = \rho \tilde{X}_t + \tilde{u}_t$, where $\tilde{u}_t = u_t - 1/2$. We have

$$\|E\left(\sum_{t=m+1}^{m+T} \tilde{X}_t|F_m\right)\|_q = \|E(\tilde{X}_{m+1}|F_m) + \ldots E(\tilde{X}_{m+T}|F_m)\|_q$$

$$= \|\sum_{j=1}^{T} \rho^j \tilde{X}_m\|_q = \sum_{j=1}^{T} \rho^j \|\tilde{X}_m\|_q < \tilde{C}$$

uniformly in m, provided $\sup_t E\|X_t\|_q$ is bounded, and recalling that $E(\tilde{u}_{m+j}|F_m) = E(\tilde{u}_{m+j}) = 0, \forall j \geq 1$. Thus also Andrews' (1984) process satisfies the memory condition (2).

Assumption A4 provides a sufficient condition for the linear growth of the variance of the partial sums; for example A4 is implied by the variance growth conditions imposed in McLeish (1975a). Also from Lemma 2.2 below and from Proposition 2 in Eberlain (1986), it follows that covariance stationary processes satisfying A1, with $q = 2$, and A2, also satisfy A4.

Indeed A2,A4 and A5 have been used by Serfling (1968) in order to provide a central limit theorem for dependent and heterogeneous observations, and subsequently have been extensively used in the proofs of weak and strong invariance principles (e.g. Eberlain 1986).

Assumptions A3,A6, and A7 are instead required in order to show the strong consistency of Newey-West type covariance estimators. A6 is the analogous of A4 for the fourth moments; thus it is a sort of linear growth condition for the fourth moments.

Assumptions A5 and A7 require that, on average, the conditional second and fourth moments converge to their unconditional correspectives at an appropriate rate. In the proposition below we shall show that, given A2 and A4 (A3 and A6), A5 (A7) just requires a strengthening of the mixing condition in A1. We have

PROPOSITION 2.1

Let $\{\epsilon_t\}_{t=-\infty}^{\infty}$ be a zero-mean strong mixing process, with size $-\frac{2}{\lambda}, \lambda \in (0,1)$.

(i) If A2 and A4 hold, then A5 holds.

(ii) If A3 and A6 hold, then A7 holds.

□

Basically in the proof of the proposition above, by using a result by Serfling (1968), we show that, under A2 and A4, A5 is equivalent to

$$E|E(S_T^i(m+k)S_T^j(m+k)|F_m) - E(S_T^i(m+k)S_T^j(m+k))| = B(k,T)$$

where $B(T,k) \to 0$, uniformly in m, as $T,k \to \infty$; also we can set $k = o(T)$. The desired result then follows from the mixing inequalities (e.g. Lemma 3.5 in McLeish, 1975a). A similar argument is used to prove part (ii).

We now introduce the three lemmas we need in order to obtain an almost sure bound for the statistic we consider.

LEMMA 2.2

Let A1,A2, A4, A5 hold, then there exists a p-dimensional Brownian motion B with covariance equal to Ω_0*, such that*

$$\frac{1}{\sqrt{2T loglog(T)}} || \sum_{k=1}^{[T\cdot]} \epsilon_k - B(T\cdot)|| = o_{as}(1)$$

and so $\frac{1}{\sqrt{2T loglog(T)}} || \sum_{k=1}^{[Tr]} \epsilon_k - B(Tr)||$ *converges to zero almost surely, continuously on* $r \in [0,1]$.

□.

Basically in Lemma 2.2 we show that strong mixing processes with mixing coefficients decaying to zero sufficiently fast, satisfy the memory conditions required by strong invariance principles for R^p-valued processes. Thus we have, continuously on $r \in [0,1]$

$$\frac{1}{\sqrt{2T loglog(T)}} ||\Omega_0^{-1/2} \sum_{k=1}^{[Tr]} \epsilon_k|| = \frac{1}{\sqrt{2T loglog(T)]}} ||W(Tr)|| + o_{as}(1),$$

where W is a standard Brownian motion with covariance equal to the $p \times p$ identity matrix, i.e. $W = \Omega_0^{-1/2} B$.

It then follows that the set of norm limit points of the scaled partial sums coincides, up to a set of zero probability measure, with the set of norm limit points of the p-dimensional Brownian motion; in this way we obtain a functional law of the iterated logarithm for partial sums of strong mixing processes. From Strassen (1964, theorem 1), we know that the set of norm limit points of the p-dimensional Brownian motion, scaled by $1/\sqrt{2Tloglog(T)}$, coincides, up to a set of measure zero, with K, where K is defined as the set of absolutely continuous functions from $[0,1]$ to R^p, such that $f(0)=0$, and $\int_0^1 \|df(t)/dt\|^2 dt \leq 1$. As $\sup_{t\in[0,1]} \|f(t)\| \leq \left(\int_0^1 \|df(t)/dt\|^2 dt\right)^{1/2}$, it follows that, almost surely

$$\limsup_{T\to\infty} \sup_{r\in[0,1]} \frac{1}{\sqrt{2Tloglog(T)}} \|\Omega_0^{-1/2} \sum_{k=1}^{[Tr]} \epsilon_k\|$$
$$= \limsup_{T\to\infty} \sup_{r\in[0,1]} \frac{1}{\sqrt{2Tloglog(T)}} \|W(Tr)\| = 1$$

Thus for a given r we obtain the ordinary LIL.

In order to have an almost sure bound free of nuisance parameters, we need to rescale the partial sums by the variance. As in general we do not know the variance, we need to estimate it. Although for inference purposes, it suffices to have a weak consistent estimator, in this context, in order to preserve the almost sure nature of the bound, we need a strong consistent estimator. Let $\hat{\Omega}_T$ be an estimator for Ω_0 and, for notational simplicity assume that $\hat{\Omega}_T^{-1/2}$ is well defined (we could otherwise use a generalized inverse), we then have

$$\frac{1}{\sqrt{2T\log\log T}} \|\hat{\Omega}_T^{-1/2} \sum_{k=1}^{[Tr]} \epsilon_k\| = \frac{1}{\sqrt{2T\log\log T}} \|\Omega_0^{-1/2} \sum_{k=1}^{[Tr]} \epsilon_k + (\hat{\Omega}_T^{-1/2} - \Omega_0^{-1/2}) \sum_{k=1}^{[Tr]} \epsilon_k\|$$
$$= \frac{1}{\sqrt{2T\log\log T}} \|\Omega_0^{-1/2} \sum_{k=1}^{[Tr]} \epsilon_k\| + o_{as}(1)$$

if $\hat{\Omega}_T - \Omega_0 = o_{as}(1)$.

As mentioned above, to the best of our knowledge, there are no available results for the strong consistency of the estimator of the variance in the dependent and heterogeneous case. Let $S_T = \sum_{k=1}^{T} \epsilon_k$ and

$$\Omega_0 = \lim_{T\to\infty} E(\frac{1}{T} S_T S_T\prime)$$

$$\bar{\Omega}_T = \frac{1}{T} \sum_{t=1}^{T} E(\epsilon_t \epsilon_t\prime) + \frac{1}{T} \sum_{\tau=1}^{T-1} \sum_{t=\tau+1}^{T} E(\epsilon_t \epsilon_{t-\tau}\prime + \epsilon_{t-\tau} \epsilon_t\prime)$$

$$\tilde{\Omega}_T = \frac{1}{T} \sum_{t=1}^{T} E(\epsilon_t \epsilon_t\prime) + \frac{1}{T} \sum_{\tau=1}^{l_T} \sum_{t=\tau+1}^{T} E\left(\epsilon_t \epsilon_{t-\tau}\prime + \epsilon_{t-\tau} \epsilon_t\prime\right)$$

let $w_\tau = (1 - \frac{\tau}{l_T+1})$, and let

$$\ddot{\Omega}_T = \frac{1}{T}\sum_{t=1}^{T} \epsilon_t \epsilon_{t}\prime + \frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T} (\epsilon_t \epsilon_{t-\tau}\prime + \epsilon_{t-\tau}\epsilon_t\prime) \quad .$$

$$\hat{\Omega}_T = \frac{1}{T}\sum_{t=1}^{T}(X_t - \bar{X})(X_t - \bar{X})\prime$$

$$+\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T} ((X_t - \bar{X})(X_{t-\tau} - \bar{X})\prime + (X_{t-\tau} - \bar{X})(X_t - \bar{X})\prime) \tag{3}$$

Basically, in order to prove the strong consistency of $\hat{\Omega}_T$ for Ω_0, we need to show that

$$\hat{\Omega}_T - \ddot{\Omega}_T = o_{as}(1) \tag{4}$$

$$\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T} (\epsilon_t \epsilon_{t-\tau}\prime - E(\epsilon_t \epsilon_{t-\tau}\prime)) = o_{as}(1) \tag{5}$$

It is immediate to see that in the martingale difference (m.d.s.) case, as an estimator for the variance we can just use $T^{-1}\sum_{t=1}^{T}(X_t - \bar{X})(X_t - \bar{X})\prime$; thus the desired result follows straightforwardly from the strong law of large numbers. On the other hand in the non m.d.s. case, we can no longer appeal to the strong law of large numbers. A key fact in the proof of the strong consistency of $\hat{\Omega}_T$ for Ω_0 is provided by the following lemma.

LEMMA 2.3

Let A1,A3, A6, A7 hold. Then

$$\frac{1}{\sqrt{2T loglog(T)}} \sum_{t=\tau+1}^{T} (\epsilon_t \epsilon_{t-\tau}\prime - E(\epsilon_t \epsilon_{t-\tau}\prime)) = O_{as}(1)$$

uniformly in τ*, for* $\tau = o(T)$.

□

We now have

LEMMA 2.4

Let X_t *be generated as (1) and let* ϵ_t *satisfy A1, A3, A6,A7. If, as* $T \to \infty$, $l_T \to \infty$ *and* $\frac{l_T\sqrt{2loglog(T)}}{\sqrt{T}} \to 0$*, then*

$$\hat{\Omega}_T - \Omega_0 = o_{as}(1)$$

□

The price we pay in order to obtain a strong consistent estimator is the uniform boundedness of the $(4+2\delta)$-th moment of ϵ_t, see A3, and A6 and A7 that we discussed above. Under some additional assumptions, in another paper (Altissimo and Corradi 1996), a strong consistency result is shown for the variance of the score of generic m-estimators; in particular the score is allowed to be a near epoch dependent process on a mixing basis. Although we have considered the case of a Bartlett-type kernel, i.e. $w_\tau = 1 - \frac{\tau}{l_T+1}$, from the proofs of Lemma 2.3 and 2.4 we see that the same result would hold for a generic kernel; basically we just need that $w_\tau \to 1$, as $T \to \infty$, and that $\frac{1}{l_T}\sum_{\tau=1}^{l_T} w_\tau$ converges to a well defined positive limit as $T \to \infty$. Thus the statements in the lemmas should apply to the entire class of kernel covariance estimators considered in Newey and West (1987).

We are now ready to state the main result of this section.

THEOREM 2.5

Let X_t be a R^p-valued process generated as in (1) and let ϵ_t satisfy A1, A3-A7. If, as $T \to \infty$, $l_T \to \infty$ and $\frac{l_T\sqrt{2loglog(T)}}{\sqrt{T}} \to 0$, then almost surely

(i) for p=1,

$$\limsup_{T\to\infty} \frac{1}{b_T}\sum_{t=1}^{T} \|\hat{\Omega}_T^{-1/2}\sum_{j=1}^{t}(X_j - \bar{X})\| \leq \frac{1}{\sqrt{3}} + \frac{1}{2}$$

(ii) for $p > 1$,

$$\limsup_{T\to\infty} \frac{1}{b_T}\sum_{t=1}^{T} \|\hat{\Omega}_T^{-1/2}\sum_{j=1}^{t}(X_j - \bar{X})\| \leq \frac{2}{3} + \frac{1}{2}$$

where $b_T = T^{3/2}\sqrt{2loglog(T)}$ and $\hat{\Omega}_T$ is defined as in (3).

□

From the theorem above we see that for the multidimensional case, regardless the specific value of p, we obtain a slightly looser almost sure bound. Now, as we see above, for any finite $p = 1, 2\ldots$, the set of norm limit points of $\frac{1}{\sqrt{2T\log\log T}}\hat{\Omega}_T^{-1/2}\sum_{j=1}^{[T\cdot]}\epsilon_j$ coincides almost surely with the set of norm limit points of $\frac{1}{\sqrt{2T\log\log T}}W_{[T\cdot]}$, where W is a p-dimensional standard Brownian motion. Now the set of norm limit points of

$\frac{1}{\sqrt{2T\log\log T}}W_{[T\cdot]}$ coincides, up to a set of measure zero, with K, the space of absolutely continuous function f, with $f(0)=0$, and such that $\int_0^1 \|\frac{df(t)}{dt}\|^2 \leq 1$. However for $p=1$ we know from Strassen (1964, p.219) that

$$\limsup_{T\to\infty} \frac{1}{T^{3/2}\sqrt{2\log\log T}} \sum_{t=1}^{T} \|\Omega_0^{-1/2} \sum_{j=1}^{t} \epsilon_j\| = \frac{1}{\sqrt{3}}$$

Now the proof of such a result heavily relies on the fact that, when $p=1$, $d\|f(t)\|/dt = \|df(t)/dt\|$; however for the case of $p>1$ in general $d\|f(t)\|/dt \neq \|df(t)/dt\|$. Thus for the case of $p>1$ we just rely on the fact that $\int_0^1 \|df(t)/dt\|^2 dt \leq 1$, and exploiting this we can show that

$$\limsup_{T\to\infty} \frac{1}{T^{3/2}\sqrt{2\log\log T}} \sum_{t=1}^{T} \|\Omega_0^{-1/2} \sum_{j=1}^{t} \epsilon_j\| \leq \frac{2}{3}$$

Thus for $p>1$, regardless the specific value of p, we obtain an almost sure upper bound, rather than an exact almost sure bound. The difference between the two bounds is equal to 0.089. Finally for all finite $p=1,2,\ldots$ the additional "1/2" term appearing in (i) and (ii) of Theorem 2.5 is due to the fact that X_t has an unknown non-zero mean and so we construct the statistic using deviations from the sample mean.
In the proof of the theorem above we make extensive use of results from Strassen (1964). Now Strassen considers the case of iid random variables, while we consider the case of heterogeneous and dependent observations. However this does not create any problem: in fact we use Eberlain (1986) strong invariance principle for dependent and heterogeneous observations, in order to show that the set of the norm limit points of properly scaled partial sums of strong mixing processes coincides almost surely with the set of norm limits points of $\frac{1}{\sqrt{2T\log\log T}}W_{[T\cdot]}$. Once we shown this, then we can rely on the results by Strassen on the set of norm limit points (and on its functionals) of the, properly scaled, p-dimensional Brownian motion.

As we shall show in Section 3 below, under the $I(1)$ hypothesis the statistic given above diverges at rate $\sqrt{\frac{T}{2l_T\log\log(T)}}$; thus we can decide in favor of the $I(0)$ hypothesis, if we compute a value smaller or equal to $1/\sqrt{3}+0.5$, for $p=1$, and smaller or equal to $2/3+0.5$ for $p>1$, and we can decide in favor of the $I(1)$ hypothesis if we get a larger value.

3 Asymptotic behavior under the $I(1)$ hypothesis

We now consider the asymptotic behavior of the statistic proposed in the previous section when (i) all variables are $I(1)$, (ii) at least one component of X_t is I(1). Consider the following DGPs

$$X_t = \mu + \sum_{j=1}^{t} \epsilon_j \tag{6}$$

and

$$X_{1t} = \mu_1 + \sum_{j=1}^{t} \epsilon_{1j} \tag{7}$$

$$X_{2t} = \mu_2 + \epsilon_{2t} \tag{8}$$

where X_{1t} and X_{2t} are respectively R^h-valued and R^{p-h}-valued processes, with $1 \leq h < p$. Thus in (6) all the components of X_t are $I(1)$, while in (7)-(8) a subset of h random variables is $I(1)$, while the remaining $p-h$ variables are $I(0)$.

THEOREM 3.1

Let X_t be as in (6) or as in (7)-(8) and let ϵ_t satisfy A1-A2 and A4-A5. If as $T \to \infty$, $l_T \to \infty$ and $\frac{l_T\sqrt{2\log\log(T)}}{\sqrt{T}} \to 0$, then

$$\lim_{T\to\infty} P\left[\frac{1}{b_T}\sum_{t=1}^{T} \|\hat{\Omega}_T^{-1/2}\sum_{j=1}^{t}(X_j - \bar{X})\| > C_T\right] = 1$$

where $C_T \times \left(\frac{2l_T\log\log(T)}{T}\right)^{1/2-\eta} \to \infty$, as $T \to \infty$, for $\eta > 0$ arbitrarily small, and $\hat{\Omega}_T$ is defined as in (3).

□

For the case in which all the variables are $I(1)$, the proof of the result above comes straightforwardly from a generalization of the results in KPSS (1992, section 4). In fact we can show that $\frac{1}{\sqrt{Tl_T}}\hat{\Omega}_T^{1/2}$ weakly converges to a non degenerate weak limit and $\frac{1}{T^{3/2}}\sum_{j=1}^{[T\cdot]}(X_j - \bar{X})$ also weakly converges to a non degenerate limit. The desiderd result then follows from the continuous mapping theorem. In the case in which some variables are $I(1)$ and other are $I(0)$, then certain components or blocks of $\frac{1}{\sqrt{Tl_T}}\hat{\Omega}_T^{1/2}$ weakly converge to a non-degenerate limit, while the others converge in probability to zero, and similarly certain elements of $\frac{1}{T^{3/2}}\sum_{j=1}^{[T\cdot]}(X_j - \bar{X})$ weakly converge to a non degenerate limit, while the others converge to zero. It follows that at least one component of $\hat{\Omega}_T^{-1/2}\frac{1}{\sqrt{T}}\sum_{j=1}^{t}(X_j - \bar{X})$ diverges at rate $\sqrt{\frac{T}{2l_T \log\log(T)}}$, and so the Euclidean norm (and so our statistic) diverges at the same rate. Concluding if we obtain a number larger than $1/\sqrt{3} + 1/2$, in the one-dimensional case, or larger than $2/3 + 1/2$, in the multidimensional case, then we decide in favor of the $I(1)$ hypothesis, or in favor of the fact that at least one variable is $I(1)$.

4 Monte Carlo simulations

Tables 1-4 display some simulations results, about the finite sample size and power of the statistic we considered in the previous sessions. We consider two sample sizes $n = 100$ and $n = 500$; for all experiments we have run 1,000 replications. Two different choices of the lag truncation parameter have been used l_{T0} and l_{T4}, where $l_{T0} = 0$ and $l_{T4} = 3T^{1/4}$. The bound used in Tables 1 and 3 is $1/\sqrt{3} + 1/2$, while the bound used in Tables 2 and 4 is $2/3 + 1/2$. In Table 1 we consider the DGP in (1) with $\mu = 0$ and $\epsilon_t = a\epsilon_{t-1} + \eta_t$, $\eta_t \sim N(0,1)$, and $a = -0.3, 0.2, 0.7, 0.9$. When we use l_{T4} then the size is virtually zero for all the cases considered; as expected for $a = 0.9$ and $l_T = l_{T0}$ the size increases, instead of declining, with the sample size, in fact the estimator of the variance constructed using $l_T = 0$ is clearly inconsistent for the true variance. We then consider (Table 2) the two-dimensional DGP $X_t = \epsilon_t$ and $\epsilon_{it} = a_i\epsilon_{i,t-1} + u_{it}$, $i = 1, 2$, $u_t \sim N(0, \Sigma_u)$, $E(u_{1t}u_{2t}) \neq 0$, for $a_1 = 0.2, a_2 = 0.3$ and $a_1 = 0.9, a_2 = 0.8$. When the roots are large, $0.9, 0.8$ and $l_T = 0$, the size is pretty high, above 0.20, while, in all the other cases is virtually zero. Table 3 show the actual power for the one-dimensional case, $X_t = \sum_{j=1}^{t}\epsilon_j$, with ϵ_j being iid, following a first order autoregressive process and following a first order moving average process, the actual power is clearly negatively affected by the choice of $l_T = 3T^{1/4}$. The power in the two-dimensional case is considered in Table 4, for the case of one variable being $I(0)$ and the other $I(1)$ and for the case of both variables being partial sums of iid, of first order autoregressive and and of first order moving average processes. Although in the multidimensional case we use a slightly looser bound, by comparing Table 3 and 4 it seems that the finite sample power improves with the dimensionality of the system, in particular when both variables are $I(1)$, in fact Models B,C,D for $l_T = 0$ have power very close to one, even for $n = 100$.

The fact that a fast rate of growth of the lag truncation parameter negatively affects the finite sample power is clear from Theorem 3.1 and has been already noted by KPPS (1992); in our context as the negative effect on the finite sample power is definetely stronger than the beneficial effect on the size, it seems recommendable to let l_T increase very slowly. Overall if we compare our simulation excercise with the PIC performance (PP 1994, Table 2), we note that our procedure is characterized by a lower finite sample Type I error and by a higher Type II error. Thus in empirical applications a joint use of PIC and of our procedure is highly recommendable.

5 APPENDIX

PROOF OF PROPOSITION 2.1

To avoid notational confusion, note that below our m is Serfling's (1968) a, and our k is Serfling's m.

(i) Serfling (1968), in his Lemma 2.2, has shown that, given A2 and A4,

$$E|E(S_T^i(m)S_T^j(m)|F_m) - E(S_T^i(m)S_T^j(m))|$$

$$\leq E|E(S_T^i(m+k)S_T^j(m+k)|F_m) - E(S_T^i(m+k)S_T^j(m+k))| + O(k^{1/2}T^{-1/2})$$

where the last term on the right hand side (RHS) holds uniformly in m. Without loss of generality, we can pick $k = T^\lambda$, $\lambda \in (0,1)$, so that, as $T \to \infty$ the second term on the RHS approaches zero, uniformly in m. Now it suffices to show that the first term on the RHS approaches zero uniformly in m. We have

$$E|E(S_T^i(m+k)S_T^j(m+k)|F_m) - E(S_T^i(m+k)S_T^j(m+k))|$$

$$\leq E|E(\epsilon_{m+k+1}^i\epsilon_{m+k+1}^j|F_m) - E(\epsilon_{m+1+k}^i\epsilon_{m+k+1}^j)|$$

$$+\ldots E|E(\epsilon_{m+T+k}^i\epsilon_{m+T+k}^j)|F_m) - E(\epsilon_{m+T+k}^i\epsilon_{m+T+k}^j)|$$

$$+\sum_{h=2}^{T} E|E(\epsilon_{m+k+1}^i\epsilon_{m+k+h}^j|F_m) - E(\epsilon_{m+k+1}^i\epsilon_{m+k+h}^j)|$$

$$+\sum_{h=2}^{T} E|E(\epsilon_{m+k+h}^i\epsilon_{m+k+1}^j|F_m) - E(\epsilon_{m+k+h}^i\epsilon_{m+k+1}^j)|$$

$$+\ldots\sum_{h=l+1}^{T} E|E(\epsilon_{m+k+l}^i\epsilon_{m+k+h}^j|F_m) - E(\epsilon_{m+k+l}^i\epsilon_{m+k+h}^j)|$$

$$+\sum_{h=l+1}^{T} E|E(\epsilon^i_{m+k+h}\epsilon^j_{m+k+l}|F_m) - E(\epsilon^i_{m+k+h}\epsilon^j_{m+k+l})|$$

$$+\dots E|E(\epsilon^i_{m+k+T-1}\epsilon^j_{m+k+T}|F_m) - E(\epsilon^i_{m+k+T-1}\epsilon^j_{m+k+T})|$$

$$+E|E(\epsilon^i_{m+k+T}\epsilon^j_{m+k+T-1}|F_m) - E(\epsilon^i_{m+k+T}\epsilon^j_{m+k+T-1})|$$

$$\leq \tilde{C}\left(\sup_t |\epsilon^i_t|^{2(1+\delta)} \sup_t |\epsilon^j_t|^{2(1+\delta)}\right)^{1/2} T^2\alpha(k)$$

for some $0 < \tilde{C} < \infty$, where the first inequality comes Minkowski inequality, while the last inequality follows from the mixing moment inequalities (e.g. McLeish Lemma 3.5, 1975a), Cauchy-Schwartz inequality, A2 and the fact that $\alpha(k+\tau) \leq \alpha(k), \forall \tau > 0$. Now set $k = T^{\lambda}$, $\lambda \in (0,1)$. As ϵ_t is α-mixing with size $-\frac{2}{\lambda}$ and $\epsilon_t\epsilon_j$ is mixing of the same size for any given j, it follows that $\alpha(T^{\lambda})T^2 \to 0$, as $T \to \infty$; in fact $\alpha(T^{\lambda}) = O(T^{-\lambda\beta})$, with $\beta > \frac{2}{\lambda}$.

(ii) By a similar argument as in (i), by noting that if ϵ_t is α-mixing with size $-\frac{2}{\lambda}, \lambda \in (0,1)$, then $\epsilon_t\epsilon_j\epsilon_h\epsilon_i$ is also mixing of the same size, for any given i, j, h, and by using A3 and A6, instead of A2 and A4.

PROOF OF LEMMA 2.2

By Lemma (3.5) in McLeish (1975a) and by Minkowski inequality,

$$\|E(S_T(m)|F_m)\|_q \leq \|E(\epsilon_{m+1}|F_m)\|_q + \|E(\epsilon_{m+2}|F_m)\|_q \dots + \|E(\epsilon_{m+T}|F_m)\|_q$$

$$\leq 6\alpha_1^{1/q-1/r}\|\epsilon_{m+1}\|_r + 6\alpha_2^{1/q-1/r}\|\epsilon_{m+2}\|_r \dots + 6\alpha_T^{1/q-1/r}\|\epsilon_{m+T}\|_r$$

$$\leq 6\alpha_1^{1/q-1/r}\|\epsilon_{m+1}\|_{2+\delta} + 6\alpha_2^{1/q-1/r}\|\epsilon_{m+2}\|_{2+\delta} + \dots 6\alpha_T^{1/q-1/r}\|\epsilon_{m+T}\|_{2+\delta}$$

$$\leq 6\sum_{j=1}^{\infty}\alpha_j^{1/q-1/r}\sup_{k\geq 1}\|\epsilon_k\|_{2+\delta} \leq 6ApC_1^{1/(2+\delta)}$$

because of A1 and A2, as $\|\epsilon_k\|_{2+\delta} = \sum_{i=1}^{p}\left(E|\epsilon^i_k|^{2+\delta}\right)^{\frac{1}{2+\delta}}$. For $q = 1$, condition (1.1) (and for $q = 2$ condition (1.3)) in Eberlain (1986) is satisfied. The result then follows from theorem 1 in Eberlain (1986), by noting that $A4$ and $A5$ correspond to condition 1.5 and 1.6 in Eberlain's.

PROOF OF LEMMA 2.3

Let $Z_{t,\tau}^{i,j} = \epsilon_t^i \epsilon_{t+\tau}^j - E(\epsilon_t^i \epsilon_{t+\tau}^j)$.

Given A6, and recalling the LIL for the Brownian motion, it suffices to show that

$$\limsup_{T\to\infty} \frac{1}{\sqrt{2T\log\log(T)}} \left| \sum_{t=1}^{T-\tau} Z_{t,\tau}^{i,j} - B_{T-\tau} \right| = o_{as}(1) \tag{9}$$

uniformly in τ, where B is a Brownian motion with variance equal to v_τ^{ij}.

We begin by showing that the properly scaled partial sums of $Z_{t,\tau}^{ij}$ satisfies a strong invariance principle, then we shall show that the implied constant does not depend on τ. As ϵ_t is α-mixing, for any given τ, $\epsilon_t\epsilon_\tau$ is also mixing of the same size. By Minkowski inequality and by Lemma 1 in Hansen (1992), recalling the definition of $Z_{t,\tau}^{ij}$, we have

$$\|E\left(\sum_{t=m+1}^{T+m-\tau} Z_{t,\tau}^{ij} \mid F_m\right)\|_q \leq \|E(Z_{m+1,\tau}^{ij}|F_m)\|_q$$

$$+ \ldots + \|E(Z_{m+k+1,\tau}^{ij}|F_m)\|_q + \ldots + \|E(Z_{m+T-\tau,\tau}^{ij}|F_m)\|_q$$

$$\leq 12\alpha_1^{1/q-1/r}\|\epsilon_{m+1}^i\|_{2r}\|\epsilon_{m+1+\tau}^j\|_{2r} + \ldots 12\alpha_{T-\tau}^{1/q-1/r}\|\epsilon_{T+m-\tau}^i\|_{2r}\|\epsilon_{T+m}^j\|_{2r} \leq 12AC_2^2 \tag{10}$$

because A1 and A3. Given A6 and A7 we know that $\frac{1}{\sqrt{2T\log\log(T)}}\sum_{t=1}^{T-\tau} Z_{t,\tau}^{ij}$ satisfies Eberlain's strong invariance principle; now we need to show that the implied constant holds uniformly in τ. Below we shall need the following,

$$E\left|\sum_{t=1}^{T-\tau} Z_{t,\tau}^{ij}\right|^{2+\nu} \leq 2^{1+\nu}\left(E|Z_{1,\tau}^{ij}|^{2+\nu} + \ldots + E|Z_{T-\tau,\tau}^{ij}|^{2+\nu}\right)$$

$$\leq 2^{1+\nu}T\left(\sup_h E|\epsilon_h^i|^{2(2+\nu)} \sup_h E|\epsilon_h^j|^{2(2+\nu)}\right)^{1/2} \leq 2^{1+\nu}TC_2 \tag{11}$$

for $0 < \nu < \delta$, with δ and C_2 as defined in A3. Note that the first inequality above comes from the C_r-inequality (e.g. White, 1984, p.33), the second inequality from Cauchy-Schwartz inequality and the last one is a straightforward consequence of A3.

Now we shall go through the main steps of the proof of Theorem 1 in Eberlain (1986) and show that given A3,A6,A7 and (11), the constant term implied in (9) does not depend on τ. From eqs.(3.3)-(3.5) in Eberlain, define $s_k = \sum_{i\leq k-1} n(i)$, with

$$n(i) = [i^{(1+\gamma^{-1})\lambda}][i^{(1+\gamma^{-1})(1-\lambda)}]$$

with $\lambda \in (1/2, 1)$ and $0 \leq \gamma \leq 1/120$; let

$$W_k^\tau = n(k)^{-1/2} \sum_{t=1+s_k}^{s_{k+1}} Z_{t,\tau}^{ij}$$

and let

$$A_{k-1}^\tau = \sigma(W_1^\tau, \ldots W_{k-1}^\tau)$$

and finally define λ_k as any sequence that satisfies

$$E\left[|E\left[\exp(i < u, W_k^\tau >)|A_{k-1}^\tau\right] - \exp(- < u, v_\tau^{ij} u > /2)| \right] \leq \lambda_k$$

with $< \cdot, \cdot >$ denoting the inner product for all u such that $< u, u > \leq n(k)^a$ with a depending only on δ, ρ, θ as defined in A3, A6 and A7. Thus λ_k is the rate at which the characteristic function of the scaled partial sum approaches the characteristic function of independent identically distribuited, $N(0, v_\tau^{ij})$ random variables. From p.266-268 in Eberlain we see that $\lambda_k = M_0 n(k)^{\gamma^*}$ with M_0 depending only on the constants in A3,A6,A7 and in (11) and γ^* depending only on $\lambda, \rho, \theta, \delta$; thus the convergence of the characteristic function of the scaled partial sums to the characteristic function of $N(0, v_\tau^{ij})$ random variables is uniform in τ. Eberlain then invokes Theorem 1 in Berkes and Philipp (1979), according to which

$$P\left[|W_k^\tau - Y_k^\tau| \geq \alpha_k\right] \leq \alpha_k$$

where Y_k^τ are iid $N(0, v_\tau^{ij})$, and

$$\begin{aligned} \alpha_k &= 16 n(k)^{-a} \log(n(k)^a) + 4\lambda_k^{1/2} n(k)^a + P\left[|N(0, v_\tau^{ij})| > n(k)^a/4\right] \\ &\leq 16 n(k)^{-a} \log(n(k)^a) + 4\lambda_k^{1/2} n(k)^a + P\left[|N(0, v^{ij})| > n(k)^a/4\right] \end{aligned}$$

where the second inequality follows from A6. As shown in Eberlain the right hand side of the inequality is of order $k^{-(1+\gamma)}$; also by the definition of $n(k)$ and λ_k, the implied constant does not depend on τ. Thus, as $\gamma > 0$, by the Borel-Cantelli lemma we have that uniformly in τ,

$$|W_k^\tau - Y_k^\tau| = O_{as}(k^{-(1+\gamma)}) \tag{12}$$

We can assume, without loss of generality, that there exists a real-valued Brownian motion $B_t, t \geq 0$, with mean zero and variance v_τ^{ij}, such that almost surely, for $k \geq 1$

$$(s_{k+1} - s_k)^{-1/2}(B_{s_{k+1}} - B_{s_k}) = Y_k^\tau$$

By (12) and eq.(3.4) in Eberlain

$$|\sum_{\nu=s_k+1}^{s_{k+1}} Z_{\nu,\tau}^{ij} - (B_{s_{k+1}} - B_{s_k})| = n(k)^{1/2}|W_k^{\tau} - Y_k^{\tau}| = O_{as}\left(k^{(1+\gamma)(1/(2\gamma)-1)}\right)$$

uniformly in τ.

Summing up over k and using the fact that $k^{2+\gamma^{-1}} \leq s_{k+1}$, we note that (see Eberlain p.269)

$$|\sum_{\nu=1}^{s_{k+1}} Z_{\nu,\tau}^{ij} - B_{s_{k+1}}| = O_{as}(s_{k+1}^{1/2-\gamma/2})$$

uniformly in τ. Now we can find a k, such that $s_k \leq s - \tau \leq s_{k+1}$ and

$$|\sum_{\nu=1}^{s-\tau} Z_{\nu,\tau}^{ij} - B_{s-\tau}| \leq |\sum_{\nu=1}^{s_k} Z_{\nu,\tau}^{ij} - B_{s_k}| + |\sum_{\nu=s_k+1}^{s-\tau} Z_{\nu,\tau}^{ij}| + |B_{s-\tau} - B_{s_k}| \tag{13}$$

We know that the first term on the RHS above is bounded by an $O_{as}(s_k^{1/2-\gamma/2})$ term, uniformly in τ; by the same argument used in Proposition 2.2 by Kuelbs and Philipp (1979) and by A3, for any given τ

$$\sup_{s_k \leq s-\tau \leq s_{k+1}} |\sum_{\nu=s_k+1}^{s-\tau} Z_{\nu,\tau}^{ij}| \leq C_7(s_k)^{1/2-\zeta} \leq C_7 s^{1/2-\zeta}$$

with C_7 depending on the constant in A3 and so independent of τ and for some $\zeta > 0$. Finally for any given τ, because of the modulus of continuity of the Brownian motion, we have almost surely

$$\sup_{s_k \leq s-\tau \leq s_{k+1}} |B_{s-\tau} - B_{s_k}| \leq \sup_{\tau} v_{\tau}^{ij} \sup_{s_k \leq s-\tau \leq s_{k+1}} |W_{s-\tau} - W_{s_k}|$$

$$\leq v^{ij} s_k^{1/2-\zeta} \leq v^{ij} s^{1/2-\zeta}$$

where W_s denotes a standard BM.

Thus the left hand side of (13) is of almost sure order $s^{1/2-\eta}$, with $\eta = \min(\zeta, \gamma/2)$, uniformly in τ, consequently, by putting $s = T$, we see that the constant implied in (9) does not depend on τ. The desired outcome then follows.

PROOF OF LEMMA 2.4

We first show that $\ddot{\Omega}_T - \Omega_0 = o_{as}(1)$. By lemma 6.17 in White (1984), if, as $T \to \infty$, $l_T \to \infty$, then $\tilde{\Omega}_T - \bar{\Omega}_T \to 0$. As $\bar{\Omega}_T \to \Omega_0$, we need to show that $\ddot{\Omega}_T - \tilde{\Omega}_T = o_{as}(1)$.

As almost sure convergence element by element implies almost sure convergence of the matrix, it suffices to show that $\ddot{\omega}_{ij} - \tilde{\omega}_{ij} = o_{as}(1)$, $\forall i, j = 1, 2, \ldots p$. We can write

$$\ddot{\omega}_{ij} - \tilde{\omega}_{ij} = \frac{1}{T}\sum_{t=1}^{T}(\epsilon_t^i \epsilon_t^j - E(\epsilon_t^i \epsilon_t^j)) \tag{14}$$

$$+\frac{1}{T}\sum_{\tau}^{l_T} w_\tau \sum_{t=1}^{T-\tau}\left((\epsilon_t^i \epsilon_{t+\tau}^j - E(\epsilon_t^i \epsilon_{t+\tau}^j)) + (\epsilon_t^j \epsilon_{t+\tau}^i - E(\epsilon_t^j \epsilon_{t+\tau}^i))\right)$$

$$+\frac{1}{T}\sum_{\tau=1}^{l_T}(w_\tau - 1)\left(\sum_{t=1}^{T-\tau}\left(E(\epsilon_t^i \epsilon_{t+\tau}^j) + E(\epsilon_t^j \epsilon_{t+\tau}^i)\right)\right)$$

The last term converges to zero by the same argument used in the proof of Theorem 2 in Newey and West (1987). Given A1 and A3, the first term on the RHS of (14) is $o_{as}(1)$ by the strong law of large numbers (e.g. McLeish, 1975b, Theorem 2.10). To complete the proof, we need to show that the second term is also $o_{as}(1)$. Let

$$Z_{t,\tau}^{ij} = \epsilon_t^i \epsilon_{t+\tau}^j - E(\epsilon_t^i \epsilon_{t+\tau}^j)$$

We can write absolute value of the "first half" of the second term on the RHS of (14), as

$$\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \left[\sum_{t=1}^{T-\tau} Z_{t,\tau}^{ij} - B_{T-\tau} + B_{T-\tau}\right]$$

$$\leq \frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \mid \sum_{t=1}^{T-\tau} Z_{t,\tau}^{ij} - B_{T-\tau} \mid + \frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \mid B_{T-\tau}|$$

where $B_{T-\tau}$ is a one-dimensional Brownian motion with variance equal to v_τ^{ij}. We can rewrite the last term on the RHS above, as

$$\frac{1}{l_T}\sum_{\tau=1}^{l_T} w_\tau \frac{l_T\sqrt{2\log\log(T)}}{\sqrt{T}} \frac{1}{\sqrt{2T\log\log(T)}}|B_{T-\tau}|$$

Because of the LIL for the standard BM, e.g. Karatzas and Shreve (1990, p.112), almost surely

$$\limsup_{T\to\infty} \frac{|B_{T-\tau}|}{\sqrt{2T\log\log(T)}} = \sqrt{v_{ij}^\tau} \leq \sqrt{v_{ij}}$$

and so

$$\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \mid B_{T-\tau} \mid = o_{as}(1)$$

uniformly in τ, provided $l_T = o\left(\frac{\sqrt{T}}{\sqrt{2T loglog(T)}}\right)$.
Now because of Lemma 2.3 and because $l_T = o\left(\sqrt{\frac{T}{2\log\log(T)}}\right)$, we have that

$$\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \left[\sum_{t=1}^{T-\tau} Z_{t,\tau}^{ij} - B_{T-\tau}\right] = o_{as}(1), \tag{15}$$

uniformly in τ.
Now it remains to show that $\hat{\Omega}_T - \ddot{\Omega}_T = o_{as}(1)$,

$$\hat{\Omega}_T = \frac{1}{T}\sum_{t=1}^{T}((X_t-\mu)-(\bar{X}-\mu))((X_t-\mu)-(\bar{X}-\mu))'$$
$$+\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=1}^{T-\tau}[((X_t-\mu)-(\bar{X}-\mu))((X_{t+\tau}-\mu)-(\bar{X}-\mu))'$$
$$+((X_{t+\tau}-\mu)-(\bar{X}-\mu))((X_t-\mu)-(\bar{X}-\mu))']$$
$$= \ddot{\Omega}_T + (\bar{X}-\mu)(\bar{X}-\mu)' - (\bar{X}-\mu)\frac{1}{T}\sum_{t=1}^{T}(X_t-\mu)' + \frac{1}{T}\sum_{t=1}^{T}(X_t-\mu)(\bar{X}-\mu)'$$
$$-(\bar{X}-\mu)\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=1}^{T-\tau}((X_{t+\tau}-\mu)' + (X_t-\mu)')$$
$$-\frac{1}{T}\sum_{\tau=1}^{l_T} w_\tau \sum_{t=\tau+1}^{T-\tau}((X_t-\mu)+(X_{t+\tau}-\mu))(\bar{X}-\mu)'$$
$$+2(\bar{X}-\mu)(\bar{X}-\mu)' l_T \frac{T-\tau}{T} = \ddot{\Omega}_T + o_{as}(1)$$

by the same argument as in the proof of Lemma 2.2 and because, given A1-A2, A4-A5, by the same argument as above $\frac{1}{\sqrt{2T\log\log T}}\sum_{t=1}^{T-\tau}(X_t-\mu) = O_{as}(1)$, uniformly in τ. The result then follows.

PROOF OF THEOREM 2.5

Most of this proof is the same for $p = 1$ and $p > 1$. We shall just stress the specific points in which the one-dimensional and the multidimensional case require a different treatment.
Let $b_T = T^{3/2}\sqrt{2loglog(T)}$.

$$\frac{1}{b_T}\sum_{t=1}^{T}\|\hat{\Omega}_T^{-1/2}\sum_{j=1}^{t}(X_j-\bar{X})\| = \frac{1}{b_T}\sum_{t=1}^{T}\|\Omega_0^{-1/2}\sum_{j=1}^{t}(X_j-\mu) - \Omega_0^{-1/2}t(\bar{X}-\mu)$$

$$+(\hat{\Omega}_T^{-1/2}-\Omega_0^{-1/2})\sum_{j=1}^{t}(X_j-\mu)-(\hat{\Omega}_T^{-1/2}-\Omega_0^{-1/2})t(\bar{X}-\mu)\|$$

$$\leq \frac{1}{b_T}\sum_{t=1}^{T}\|\Omega_0^{-1/2}\sum_{j=1}^{t}\epsilon_j\|+\frac{1}{Tb_T}\sum_{t=1}^{T}t\|\Omega_0^{-1/2}\sum_{s=1}^{T}\epsilon_s\|$$

$$+\frac{1}{b_T}\sum_{t=1}^{T}\|(\hat{\Omega}_T^{-1/2}-\Omega_0^{-1/2})\sum_{j=1}^{t}\epsilon_j\|$$

$$+\frac{1}{Tb_T}\sum_{t=1}^{T}t\|(\hat{\Omega}_T^{-1/2}-\Omega_0^{-1/2})\sum_{s=1}^{T}\epsilon_s\| \quad (16)$$

We shall organize the proof in three different steps.

STEP 1

We show that, for $p=1$, almost surely

$$(a)\limsup_{T\to\infty}\frac{1}{T^{3/2}\sqrt{2\log\log(T)}}\sum_{t=1}^{T}\|\Omega_0^{-1/2}\sum_{j=1}^{t}\epsilon_j\|=\frac{1}{\sqrt{3}}$$

and for $p>1$, almost surely,

$$(b)\limsup_{T\to\infty}\frac{1}{T^{3/2}\sqrt{2\log\log(T)}}\sum_{t=1}^{T}\|\Omega_0^{-1/2}\sum_{j=1}^{t}\epsilon_j\|\leq\frac{2}{3}$$

Let $t=[Tr]$, and let $\sum_{j=1}^{[Tr]}\Omega_0^{-1/2}\epsilon_j=S_{[Tr]}=S_T(r)$, and also let

$$\eta_T(r)=([Tr]+1-Tr)S_T(r)+(Tr-[Tr])S_{T+1}(r)$$

Let $d_T=\sqrt{2T\log\log(T)}$, from Lemma 2.2, for any finite p,

$$\limsup_{T\to\infty}\frac{1}{d_T}\|S_{[T\cdot]}-W(T\cdot)\|=o_{as}(1)$$

and, as

$$\limsup_{T\to\infty}\frac{1}{d_T}\|\eta_{[T\cdot]}-S_{[T\cdot]}\|=o_{as}(1),$$

it follows that $\{d_T^{-1}S_T\}$ and $\{d_T^{-1}\eta_T\}$, for $T\geq 3$, are relatively norm compact and the set of their accumulation points coincides with the set of norm limit points of $d_T^{-1}W_T$ with probability 1. From Theorem 1 in Strassen (1964), it follows that $\{d_T^{-1}W_T, T\geq 3\}$ is relatively norm compact and the set of its limit points coincides, almost surely, with K, where $K=[f\in[0,1]\to R^p; f(0)=0$, f absolutely continuous,$\int_0^1\|df(r)/dr\|^2dr\leq 1]$

We now make use of the following result, due to Strassen (1964, p.218),

LEMMA A

If ϕ is a continuous map from C, the space of continuous functions on $[0,1]$, closed with respect to the sup norm, to some Hausdorf space, e.g. an Euclidean space as in our case, then, with probability one, $\{\phi(d_T^{-1}\eta_T), T \geq 3\}$, where η_T is defined as above, is relatively norm compact and the set of norm limit points coincides, almost surely, with $\phi(K)$, with K defined as above. □.

Now,

$$P\left[\limsup_{T\to\infty} \frac{1}{b_T}\sum_{t=1}^{T} \|\Omega_0^{-1/2}\sum_{j=1}^{t}\epsilon_j\| = \limsup_{T\to\infty} d_T^{-1}\int_0^1 \|\eta_T(t)\|dt\right] = 1$$

From Lemma A, for $p = 1$, we have,

$$P\left[\limsup_{T\to\infty} d_T^{-1}\int_0^1 \|\eta_T(t)\|dt = \sup\phi(K)\right] = 1$$

Given the definition of K, following Strassen (1964, p.218-219),

$$\sup\phi(K) = \sup_{u\in K}\int_0^1 \|u_t\|dt = \sup_{u\in K}\int_0^1 \|\dot{u}_t\|(1-t)dt$$

$$= \sup\left[\int_0^1 (1-t)\|y(t)\|dt : \int_0^1 \|y(t)\|^2 dt \leq 1\right] = \left(\int_0^1 (1-t)^2 dt\right)^{1/2}$$

where the second equality on the RHS comes from integration by parts and the last equality follows from the evaluation of the supremum of a linear functional on the unit sphere of a Hilbert space. As,

$$\left(\int_0^1 (1-t)^2 dt\right)^{1/2} = \frac{1}{\sqrt{3}}$$

the statement of Step 1(a) follows.

The same argument does not apply when $p > 1$, as $d\|u_t\|/dt \neq \|du_t/dt\|$. Now for arbitrary finite p,

$$\sup\phi(K) = \sup_{u\in K}\int_0^1 \|u_t\|dt$$

Now

$$\sup_{u\in K}\int_0^1 \|u_t\|dt = \sup_{u\in K}\int_0^1 \|\int_0^t \dot{u}_s ds\|dt \leq \int_0^1 \left(\sup_{u\in K}\left(\int_0^t ds\int_0^t \dot{u}_s\prime\dot{u}_s ds\right)^{1/2}\right)dt \quad (17)$$

By the definition of K,

$$\int_0^1 \dot{u}_s\prime\dot{u}_s ds \leq 1,$$

as the bound does not depend on the specific u we pick, we also have that

$$\sup_{u\in K}\int_0^1 \dot{u}_s\prime\dot{u}_s ds \leq 1$$

Thus the right hand side in (17) is majorized by $\int_0^1 t^{1/2}dt = \frac{2}{3}$. The statement in Step 1(b) then follows.

STEP 2

For arbitrary finite p,

$$\limsup_{T\to\infty}\frac{1}{b_T}\sum_{t=1}^{T} t\|\Omega_0^{-1/2}(\bar{X}-\mu)\|$$

$$=\limsup_{T\to\infty}\frac{T(T+1)}{2T^2}\frac{1}{\sqrt{2T loglog(T)}}\|\Omega_0^{-1/2}\sum_{t=1}^{T}\epsilon_t\| = \frac{1}{2}$$

recalling that

$$\limsup_{T\to\infty}\frac{1}{\sqrt{2T loglog(T)}}\|\Omega_0^{-1/2}\sum_{t=1}^{T}\epsilon_t\| = 1$$

STEP 3

It remains to show that the last two terms on the RHS of (16) are $o_{as}(1)$. Let $\hat{\Omega}_T^{-1/2} = [\hat{u}_{ij}]$ and $\Omega_0^{-1/2} = [u_{ij}]$, $\forall i,j = 1,2,\ldots p$.

$$\frac{1}{b_T}\sum_{t=1}^{T}\|(\hat{\Omega}_T^{-1/2}-\Omega_0^{-1/2})\sum_{j=1}^{t}\epsilon_j\|$$

$$\leq \frac{1}{b_T}\sum_{t=1}^{T}|\sum_{i=1}^{p}(\hat{u}_{i1}-u_{i1})\sum_{j=1}^{t}\epsilon_j^1|$$

$$+\ldots\frac{1}{b_T}\sum_{t=1}^{T}|\sum_{i=1}^{p}(\hat{u}_{ip}-u_{ip})\sum_{j=1}^{t}\epsilon_j^p| = o_{as}(1),$$

because of Lemmas 2.2-2.4.

As for the last term on the RHS of (16), it can be majorized by

$$\frac{1}{T^2}\frac{T(T+1)}{2}\frac{1}{\sqrt{2Tloglog(T)}} \mid \sum_{i=1}^{p}(\hat{u}_{i1}-u_{i1})\sum_{t=1}^{T}\epsilon_t^1 \mid$$

$$+\ldots+\frac{T(T+1)}{2T^2}\frac{1}{\sqrt{2Tloglog(T)}} \mid \sum_{i=1}^{p}(\hat{u}_{ip}-u_{ip})\sum_{t=1}^{T}\epsilon_t^p \mid = o_{as}(1)$$

as $\hat{u}_{ij}-u_{ij}=o_{as}(1)$, $\forall ij$, and because $\forall j=1,2\ldots p$,

$$\limsup_{T\to\infty}\frac{1}{\sqrt{2Tloglog(T)}} \mid \sum_{t=1}^{T}\epsilon_t^j \mid = O_{as}(1)$$

The desired result then follows by putting together Step 1, Step 2 and Step 3 and by recalling that the limsup of the sum is less than or equal to the sum of the limsups.

PROOF OF THEOREM 3.1

We begin by considering the case in which all the variables are $I(1)$. Basically we generalize the argument in KPSS (1992, section 4) to the multidimensional case. As for $\hat{\Omega}_T$,

$$\frac{1}{Tl_T}\hat{\Omega}_T=\frac{1}{l_T}\sum_{\tau=1}^{l_T}w_\tau\frac{1}{T}\sum_{t=\tau+1}^{T}\left(\frac{1}{\sqrt{T}}\sum_{j=1}^{t}(\epsilon_j-T^{-1}\sum_{t=1}^{T}\epsilon_t)\right)\left(\frac{1}{\sqrt{T}}\sum_{j=1}^{t-\tau}(\epsilon_j-T^{-1}\sum_{t=1}^{T}\epsilon_t)\right)'+o_p(1)$$

$$=\frac{1}{l_T}\sum_{\tau=1}^{l_T}w_\tau\frac{1}{T}\sum_{t=\tau+1}\left(\frac{1}{\sqrt{T}}\sum_{j=1}^{t-\tau}(\epsilon_j-T^{-1}\sum_{t=1}^{T}\epsilon_t)\right)\left(\frac{1}{\sqrt{T}}\sum_{j=1}^{t-\tau}(\epsilon_j-T^{-1}\sum_{t=1}^{T}\epsilon_t)\right)'+o_p(1)$$

as $l_T=o\left(\sqrt{\frac{T}{2\log\log(T)}}\right)$.

Thus

$$\frac{1}{Tl_T}\hat{\Omega}_T\Rightarrow\int_0^1\left(B_s-\int_0^1 B_r dr\right)\left(B_s-\int_0^1 B_r dr\right)' ds$$

where B_s is a p-dimensional Brownian motion with covariance matrix equal to Ω_0. Now

$$\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{T}\sum_{j=1}^{t}\left(\frac{1}{\sqrt{T}}\sum_{i=1}^{j}(\epsilon_i-T^{-1}\sum_{t=1}^{T}\epsilon_t)\right)\right)\Rightarrow\int_0^1\left(\int_0^r\left(B_s-\int_0^1 B_t dt\right)ds\right)dr$$

Thus the statistic diverges at rate $\sqrt{\frac{T}{2l_T\log\log(T)}}$.

For the case in which only a subset of variables is $I(1)$, while the remaining subset is $I(0)$, we know that certain elements, or blocks, of $\sqrt{\frac{1}{Tl_T}}\hat{\Omega}_T^{1/2}$ weakly converge to a non degenerate limit, while the remaining converge to zero in probability. Similarly some elements of $\frac{1}{T^{5/2}}\sum_{t=1}^{T}\left(\sum_{j=1}^{t}(X_j-\bar{X})\right)$ weakly converge to a non degenerate limit, while the others converge to zero.

Thus, provided at least one variable is $I(1)$, at least one element of $\hat{\Omega}_T^{-1/2}\frac{1}{\sqrt{T}}\sum_{j=1}^{t}(X_j-\bar{X})$ diverges at rate $\sqrt{\frac{T}{l_T}}$. It follows that $\frac{1}{T^{3/2}\sqrt{2\log\log(T)}}\sum_{t=1}^{T}\|\hat{\Omega}_T^{-1/2}\sum_{j=1}^{t}(X_j-\bar{X})\|$ diverges at rate $\sqrt{\frac{T}{2l_T\log\log(T)}}$.

6 References

Altissimo, F. & V. Corradi (1996), A LIL for m-estimators and applications to hypothesis testing with nuisance parameters, manuscript, University of Pennsylvania.

Andrews D.W.K. (1984) Non-strong mixing autoregressive processes, *Journal of Applied Probability* , 21, 930-934

Andrews, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* 59, 817-858.

Andrews, D.W.K. & J.C. Monahan (1992) An improved heteroskedasticity and autocorrelation consistent variance estimator *Econometrica* 60, 953-966.

Berkes, I & W. Philipp (1979) Approximation theorems for independent and weakly dependent random vectors, *Annals of Probability* 7, 29-54.

Eberlain E. (1986) On strong invariance principles under dependence assumptions *Annals of Probability* 14, 260-270.

Gallant A.R & H. White (1988) *A Unified Theory for Estimation and Inference for Nonlinear Econometric Models* Basil Blackwell, New York.

Hansen B.E. (1992) Consistent covariance matrix estimation for dependent heterogeneous processes *Econometrica* 60, 967-972.

Karatzas, I. & S.E. Shreve (1990) *Brownian Motion and Stochastic Calculus* (2nd edn). New York: Springer and Verlag.

Kuelbs J. & W. Philipp (1980) Almost sure invariance principles for partial sums of mixing B-valued random variables. *Annals of Probability* 8, 1003-1036.

Kwiatkowski D., P.C.B. Phillips, P. Schmidt & Y. Shin (1992) Testing the null hypothesis of stationarity against the alternative of a unit root *Journal of Econometrics* 54, 159-178.

McLeish D.L. (1975a) Invariance principles for dependent variables, *Z. Wahrsheinlichkeitstheorie verw. Gebiete* 32, 165-178.

McLeish D.L. (1975b) A maximal inequality and dependent strong laws *Annals of Probability* , 3, 829-839.

Newey W.K. & K.D. West (1987) A simple, positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix *Econometrica* 55, 703-708.

Newey W.K. & K.D. West (1994) Automatic lag selection in covariance matrix estimation *Review of Economic Studies* 61, 631-653.

Phillips P.C.B. & W. Ploberger (1994) Posterior odds testing for a unit root with data-based model selection *Econometric Theory* 10, 774-808.

Phillips P.C.B. & W. Ploberger (1996) An asymptotic theory of Bayesian inference for time series *Econometrica* 64, 381-412

Robinson P.M. (1991) Automatic frequency domain inference on semiparametric and nonparametric models, *Econometrica* 1329-1363.

Serfling R.J. (1968) Contributions to central limit theory for dependent variables, *Annals of Mathematical Statistics* , 39, 1158-1175.

Sin C.Y. & H. White (1996) Information criteria for selecting possibly mispecified parametric models. *Journal of Econometrics* 71, 207-225.

Sin C.Y. (1996) On selecting strictly nested linear regression models with information criteria, City University of Honk Kong, mimeo.

Stock J.H. (1994) Deciding between I(1) and I(0) *Journal of Econometrics* 63, 105-131

Strassen V. (1964) An Invariance principle for the law of the iterated logarithm *Z. Wahrscheinlichkeitstheorie,* 3, 211-226.

White H. (1984) *Asymptotic Theory for Econometrician* Academic Press, New York.

Table 1: Actual size of T1(T)

	lt0		lt4	
	n=100	n=500	n=100	n=500
a=-0.3	0.00	0.00	0.00	0.00
a=0.2	0.00	0.00	0.00	0.00
a=0.7	0.04	0.05	0.00	0.00
a=0.9	0.07	0.35	0.00	0.00

Notes:

$$X_t = \epsilon_t$$

$$\epsilon_t = a\epsilon_{t-1} + \eta_t$$

$$\eta_t \sim N(0,1)$$

$$T1(T) = b_T^{-1} \sum_{t=1}^{T} \|\hat{\Omega}_T^{-1/2} \sum_{j=1}^{t} (X_j - \bar{X})\|$$

Table 2: Actual size of T1(T)

	lt0		lt4	
	n=100	n=500	n=100	n=500
Model A	0.04	0.01	0.00	0.00
Model B	0.23	0.24	0.04	0.00

Notes:

$$X_t = (X_{1t}, X_{2t})$$
$$X_t = \epsilon_t$$

Model A is :

$$\epsilon_{1t} = 0.2\epsilon_{1,t-1} + u_{1t}$$
$$\epsilon_{2t} = 0.3\epsilon_{2,t-1} + u_{2t}$$

Model B is:

$$\epsilon_{1t} = 0.9\epsilon_{1,t-1} + u_{1t}$$
$$\epsilon_{2t} = 0.8\epsilon_{2,t-1} + u_{2t}$$

and for both Model and B, $u_t \sim iidN(0, \Sigma)$, $\Sigma = [\sigma_{ij}]$, $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{12} = \sigma_{21} = 0.4$. And

$$T1(T) = b_T^{-1} \sum_{t=1}^{T} \|\hat{\Omega}_T^{-1/2} \sum_{j=1}^{t} (X_j - \bar{X})\|$$

Table 3: Actual Power of T1(T)

	lt0		lt4	
	n=100	n=500	n=100	n=500
Model A	0.56	0.95	0.42	0.54
Model B	0.57	0.93	0.17	0.56
Model C	0.58	0.94	0.15	0.55

Notes: For Model A-C:

$$X_t = \sum_{j=1}^{t} \epsilon_j$$

Model A: $\epsilon_t \sim iidN(0,1)$
Model B: $\epsilon_t = 0.2\eta_{t-1} + \eta_t$, $\eta_t \sim iidN(0,1)$
Model C: $\epsilon_t = 0.2\epsilon_{t-1} + \eta_t$, $\eta_t \sim iidN(0,1)$

$$T1(T) = b_T^{-1} \sum_{t=1}^{T} \|\hat{\Omega}_T^{-1/2} \sum_{j=1}^{t} (X_j - \bar{X})\|$$

Table 4: Actual power of T1(T)

	lt0		lt4	
	n=100	n=500	n=100	n=500
Model A	0.65	0.95	0.34	0.64
Model B	0.92	1.00	0.32	0.54
Model C	0.93	1.00	0.39	0.60
Model D	0.93	1.00	0.22	0.54

Notes: $X_t = (X_{1t}, X_{2t})$.

In all Models A-D, $\Sigma = [\sigma_{ij}]$, $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{12} = \sigma_{21} = 0.4$.

Model A is : $X_{1t} = \epsilon_{1t}$, $X_{2t} = \sum_{j=1}^{t} \epsilon_{2j}$, with $\epsilon_{1t} = 0.8\epsilon_{t-1} + u_{1t}$, and $\epsilon_{2t} = u_{2t}$, $u_t \sim iidN(0, \Sigma)$.

In Models B-D, $X_{1t} = \sum_{j=1}^{t} \epsilon_{1j}$, $X_{2t} = \sum_{j=1}^{t} \epsilon_{2j}$.

Model B: $\epsilon_t \sim iidN(0, \Sigma)$.

Model C: $\epsilon_{1t} = 0.1u_{1,t-1} + u_{1t}$, $\epsilon_{2t} = 0.2u_{2,t-1} + u_{2t}$, $u_t \sim iidN(0, \Sigma)$.

Model D: $\epsilon_{1t} = 0.3\epsilon_{1,t-1} + u_{1t}$, $\epsilon_{2t} = 0.2\epsilon_{2t} + u_{2t}$, $u_t \sim iidN(0, \Sigma)$.

$$T1(T) = b_T^{-1} \sum_{t=1}^{T} \|\hat{\Omega}_T^{-1/2} \sum_{j=1}^{t} (X_j - \bar{X})\|$$