Random Walk or Chaos: A Formal Test on the Lyapunov Exponent[1]

Joon Y. Park[2]
Department of Economics
Indiana University and SKKU

and

Yoon-Jae Whang[3]
Department of Economics
Seoul National University

Abstract

A formal test on the Lyapunov exponent is developed to distinguish a random walk model from a chaotic system. The test is based on the Nadaraya-Watson kernel estimate of the Lyapunov exponent. We show that the estimator is consistent: The estimated Lyapunov exponent converges to zero under the random walk hypothesis, while it converges to a positive constant for the chaotic system. The test is thus expected to have discriminatory powers. We derive the asymptotic distribution of the estimator, and make it possible to formally test for the null hypothesis of random walk against chaos. The proposed test statistic is a simple normalization of the estimated Lyapunov exponent. It is shown that the null distribution of the test statistic is given by the range of standard Brownian motion on the unit interval. We confirm through simulation that our test performs reasonably well in finite samples. We also apply our test to some of the standard macro and financial time series. For most of the series we considered, however, we find no significant empirical evidence of chaos. We also discuss some of the limitations of our empirical findings.

This version: September 30, 2009
Key words and phrases: Lyapunov exponent, chaos, random walk, unit root, kernel regression, Brownian motion, local time, stochastic integrals.

---

[1]We would like to thank the guest co-editors, and two anonymous referees for helpful comments. We are also grateful to Oliver Linton and Mototsugu Shintani for many useful discussions.

[2]Bloomington, IN 47405-7000. E-mail: `joon@indiana.edu`. This research was partially supported by NSF Grant No. SES-0518619.

[3]Seoul 121-746, Korea. Email: `whang@snu.ac.kr`. Web: `http://plaza.snu.ac.kr/~whang`. Thanks to the Korea Science and Engineering Foundation for financial support.

1. Introduction

Since the early work of Brock (1986), there has been much attention directed toward nonlinear chaotic dynamics in the economics literature. That deterministic systems can lead to very complex and essentially unpredictable behavior is certainly a meaningful finding for economics. However, chaos models are important also for the economic modeling in that they suggest a parsimonious representation for seemingly complex systems. The stochastic chaotic models, which introduce stochastic disturbances into the chaotic "skeleton", seem particularly useful in this context. The stochastic disturbances can of course be generated by the presence of measurement errors and/or unobserved quantities. The reader is referred to the special issues of the Journal of the Royal Statistical Society, Series B (1992) and the Journal of the Applied Econometrics (1994) for excellent surveys on the subject of chaos, Cheng and Tong (1992) for more discussions about noisy chaos, and to Rosser (1999) for a more general discussion on complex economic dynamics.

In this paper, we develop a formal test of the random walk hypothesis against the alternative of chaos. To our best knowledge, we are not aware of any existing test for this problem. The test is based on the Lyapunov exponent, which is widely believed to well characterize the chaotic phenomenon (see, e.g., Nychka et. al. (1992)). The reliance on the Lyapunov exponent seems particularly attractive for the test of the random walk model against the chaotic system. One of the essential differences of the random walk model and the chaotic system is their responsiveness to exogenous shocks. The effects of shocks to both models are persistent. The time profiles of their impacts, however, are very different: They remain constant for the random walk model, while they are amplified for the chaotic system. The Lyapunov exponent can be used to effectively discriminate their differing response patterns to shocks, since it measures the rate of divergence (or convergence) of two nearby initial points in a dynamic system.

To construct our test, we use an estimator of the Lyapunov exponent based on the Nadaraya-Watson kernel estimator of the autoregressive regression function. The asymptotic properties of the estimator is well known for the stationary chaotic system. McCaffrey et. al. (1992) established consistency of the estimator and Whang and Linton (1999) obtained its asymptotic distribution. More recently, Shintani and Linton (2003, 2004) extend the latter result to neural network and local polynomial estimators. In this paper, we analyze the asymptotics of the kernel estimator of the Lyapunov exponent when the underlying process is random walk, as is necessary to test for the random walk against the chaotic system. We show that the estimator is consistent, and derive its limiting distribution with appropriate normalization. We note that the existing results such as Whang and Linton (1999) and Shintani and Linton (2003,2004) can also be used to test for the chaos, since they allow us to see whether or not the confidence intervals around the estimated Lyapunov exponents include the positive region. See Gencay (1996), Wolff, Yao and Tong (2004), and Fernández-Rodríguez, Sosvilla-Rivero and Andrada-Félix (2005) and references therein for more existing tests of chaos based on the Lyapunov exponent. However, most of the existing results are valid only for stationary time series and they are not applicable to random walks.

Under the null hypothesis of random walk, our test statistic has limiting distribution given by the range of standard Brownian motion on the unit interval. The density of the limiting distribution (with respect to Lebesgue measure) is derived in the paper. Consequently, the critical values of the test can be easily computed from the integration of the density function. For a stationary dynamic system with positive Lyapunov exponent, the statistic is shown to diverge to infinity. Consequently, the test which rejects the null of the random walk hypothesis when it takes large values is consistent. Though it is not of our primary concern, the statistic can also be used to test for the random walk against the non-chaotic stationary autoregression, since it diverges off to negative infinity in the latter case. From the simulation study, we observe that our test performs quite well in finite samples. In particular, it seems to have desirable discriminatory powers against interesting alternatives.

We apply our test to some of the standard macroeconomic and financial dataset: Standard and Poor composite index 500 and Dow-Jones industrial average, Canada/US and Japan/US bilateral exchange rates, and three-month and one-year US treasury bill rates. For all of the series we considered, we find very little or no evidence of chaos. They seem better fitted to the random walk model or a non-chaotic stationary model. However, we note that our empirical findings are subject to some caveats and should be interpreted more carefully.

The remainder of this paper is organized as follows. Section 2 introduces the model, the nonparametric estimator of the Lyapunov exponent and the test statistic. Section 3 presents the assumptions and preliminary theories, and derives the asymptotic null distribution of the test statistic. The distribution is represented by the range of standard Brownian motion on the unit interval, and its density is obtained. We also establish in this section consistency of the nonparametric estimator of the Lyapunov exponent under the null hypothesis. In Section 4, the test is shown to be consistent against chaotic alternatives. The results from simulation experiments and empirical applications of our testing procedure to some of macro and financial variables are reported in Section 5. Section 6 concludes the paper. The proofs of theorems are given in Section 7.

Finally, a word on notation. For a function $F : \mathbf{R} \to \mathbf{R}$, we define

$$F^{\circ}(x) = \frac{d}{dx} F(x)$$

Moreover, we let

$$\begin{aligned} {}_{\circ}F(x) &= \int_{-\infty}^{x} F(t)\, dt, \\ F_{\circ}(x) &= -1\{x \geq 0\} \int_{x}^{\infty} F(t)\, dt + 1\{x < 0\} \int_{-\infty}^{x} F(t)\, dt. \end{aligned}$$

Note that ${}_{\circ}F \equiv F_{\circ}$ if F is such that $\int_{-\infty}^{\infty} F(x)\, dx = 0$. We also define function ${}_{\iota}F$ by

$${}_{\iota}F(x) = xF(x)$$

The standard order symbols will appear frequently in the paper. We use o and O for the nonstochastic orders, o_p and O_p for the orders in probability and $o_{a.s.}$ and $O_{a.s.}$ for the orders defined in almost sure sense. The usual notations for various notions of the convergences of random sequences, such as $\to_{a.s.}$, $\to_p$ and $\to_d$, are also used throughout the paper. The notation introduced here will be used subsequently without further reference.

2. The Model, Estimator and Test Statistic

We consider the model given by

$$y_t = m(y_{t-1}) + u_t \tag{1}$$

for $t = 1, \ldots, n$, where (u_t) is an iid $(0, \sigma^2)$ sequence of random variables. We assume $y_0 = 0$. This is not critical and just made for simplicity in the subsequent exposition. We may indeed allow for any random y_0 as long as it is of order $O_p(1)$. For the unit root model, we have $m(y) = y$, i.e., m is the identity function. For the chaotic model, m is given generally by a nonlinear function. Of course, purely chaotic series are deterministic, in which case we have $\sigma^2 = 0$. This is allowed in the paper. However, chaotic systems may include stochastic disturbances due to, for instance, measurement errors and unobserved quantities, and this leads to our model (1)[4].

One way to discriminate a random walk model from a chaotic system is to look at the Lyapunov exponent, which is given by

$$\lambda = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \log |m^{\circ}(y_{t-1})|$$

This definition is now standard in the chaos literature, see Nychka et. al. (1992). The quantity λ can be interpreted more generally as a measure of the local stability of the time series (y_t) generated by (1). See, for example, Tong (1990, p309) for more explanations on the interpretation of λ. For stationary linear autoregression $\lambda < 0$, while for the random walk process, we have $\lambda = 0$. For explosive processes, $\lambda > 0$. Chaotic processes have $\lambda > 0$, but they are bounded. Their explosiveness is local in nature and reveals sensitive dependence to initial conditions, see Nychka et. al. (1992). We use λ to test the random walk model against the chaotic alternatives. Specifically, the null hypothesis of interest is that (y_t) is a random walk process with $\lambda = 0$, and tested against the alternative hypothesis of interest that (y_t) is chaotic with $\lambda > 0$.

The Lyapunov exponent can be estimated using the derivative kernel estimator. Let the Nadaraya-Watson kernel estimator m_n of m be given by

$$m_n(y) = \sum_{t=1}^{n} K\left(\frac{y - y_{t-1}}{h_n}\right) y_t \Bigg/ \sum_{t=1}^{n} K\left(\frac{y - y_{t-1}}{h_n}\right)$$

[4]Often the nonlinear dynamic system with the embedding dimension k greater than one, i.e., $y_t = m(y_{t-1}, \ldots, y_{t-k}) + u_t$, is considered in the chaos literature. For simplicity, however, we focus on the system with embedding dimension equal to unity in this paper.

where K is the kernel function and h_n is the bandwidth parameter. The derivative m° of m can be estimated by the derivative m_n° of m_n. We write it as

$$m_n^\circ(y) = P_n(y)\,/Q_n(y)$$

where

$$\begin{aligned} P_n(y) &= \left(\sum_{t=1}^n K^\circ\left(\frac{y-y_{t-1}}{h_n}\right) y_t\right)\left(\sum_{t=1}^n K\left(\frac{y-y_{t-1}}{h_n}\right)\right) \\ &\quad - \left(\sum_{t=1}^n K\left(\frac{y-y_{t-1}}{h_n}\right) y_t\right)\left(\sum_{t=1}^n K^\circ\left(\frac{y-y_{t-1}}{h_n}\right)\right) \\ Q_n(y) &= h_n\left(\sum_{t=1}^n K\left(\frac{y-y_{t-1}}{h_n}\right)\right)^2 \end{aligned}$$

The Lyapunov exponent can then be estimated by

$$\lambda_n = \frac{1}{n}\sum_{t=1}^n \log|m_n^\circ(y_{t-1})| \tag{2}$$

the estimator which will be considered in the paper. As will be shown in the next section, the estimator λ_n of the Lyapunov exponent is consistent and $\lambda_n \to 0$ under the null hypothesis of random walk.

To test for the random walk hypothesis against the alternative of chaos, we consider the statistic

$$T_n = \frac{2n^{1/2}h_n^3}{\kappa^2\sigma_n}\,\lambda_n \tag{3}$$

where

$$\kappa^2 = \int_{-\infty}^{\infty} K^\circ(t)^2 dt$$

and

$$\sigma_n^2 = \sum_{t=1}^n \triangle y_t^2/n \tag{4}$$

with the usual difference operator $\triangle$. The statistic T_n introduced in (3) is just a normalized version of the Lyapunov exponent estimator λ_n. The normalization is done so that the test statistic has a proper limiting distribution. We show in Section 3 that it indeed has an asymptotically nondegenerate distribution represented by the range of standard Brownian motion on the unit interval. The distribution is absolutely continuous with respect to Lebesgue measure, and we also obtain its probability density. It is shown subsequently in Section 4 that the test statistic diverges off to infinity under the alternative of chaos. The test is therefore consistent for rejecting the null of random walk in favor of the alternative of chaos for large values of T_n.

3. Asymptotic Theory under the Null Hypothesis

3.1 Assumptions and Preliminaries

We first introduce the assumptions for innovation sequence (u_t), the kernel function K and the bandwidth parameter h_n.

3.1 Assumption (a) (u_t) is iid $(0, \sigma^2)$ with $\mathbf{E}\,|u_t|^p < \infty$ for some $p > 4$, and
(b) (u_t) has distribution absolutely continuous with respect to the Lebesgue measure with characteristic function φ such that $\varphi(t) = o(|t|^r)$ for some $r < 0$ as $|t| \to \infty$.

The conditions in Assumption 3.1 are rather stringent. Our asymptotics require the approximations of random walk sample moments as the corresponding continuous functionals of Brownian motion, which are fine enough to have the relevant limit behaviors of Brownian motion appear in the asymptotics. It seems that there are two possibilities to relax our iid assumption. First, it may be possible to extend our subsequent theory to random walks generated by linear processes using the approach used in recent work by Jeganathan (2004), although none of the currently available results in the literature is applicable for the development of our asymptotics. Second, we may consider the discrete samples from continuous time processes such as diffusions, and employ the infill asymptotics. It is clear that our null asymptotics will hold under this setup, as long as the underlying diffusion is close enough to Brownian motion and the sampling interval is sufficiently small.

3.2 Assumption The bandwidth parameter (h_n) satisfies $h_n = c_0 n^{-\nu}$ for some constant $c_0 > 0$ and $0 < \nu < 1/6$.

In this section, we will mostly deal with functions satisfying certain regularity conditions. It is therefore convenient to define: We call a function $F : \mathbf{R} \to \mathbf{R}$ *regular*, if it is bounded and integrable, has integrable Fourier transform and $\int_{-\infty}^{\infty} |t|^{2q} |F(t)| dt < \infty$ for some $q > 1/\nu$, where ν is given in Assumption 3.2.

3.3 Assumption (a) $\int_{-\infty}^{\infty} K(t) = 1$ and K is symmetric, and
(b) $K, {}_{\circ}K$ and ${}_{i}K$ are regular.

Our subsequent theoretical development relies heavily on the probabilistic embedding of the partial sum $\sum_{i=1}^{t} u_i$ of the innovation sequence (u_t) into a Brownian motion in an extended probability space. The basic idea is introduced in e.g., Hall and Heyde (1980). Here we use the results given by Park and Phillips (1999). As they have shown, we may expand the underlying probability space under Assumption 3.1 so that there exist a Brownian motion $(B(t))_{t\in[0,1]}$ and a time change $(\tau_k)_{k\geq 1}$ such that for all $n \geq 1$,

$$B(\tau_k/n) = n^{-1/2} \sum_{i=1}^{k} u_i \tag{5}$$

for $k = 1, \ldots, n$, and

$$\sup_{1 \leq k \leq n} \left| \frac{\tau_k - k}{n^r} \right| \to_{a.s.} 0$$

for any $r > 1/2$.[5] The reader is referred to Park and Phillips (1999) for the proof. If we define

$$B_n(t) = \sum_{i=1}^{n} B(\tau_i/n) 1\{(i-1)/n \leq t \leq i/n\} \tag{6}$$

for $t \in [0,1]$, then we have due to the Hölder continuity of the Brownian sample path

$$\sup_{t \in [0,1]} |B_n(t) - B(t)| \leq \left|\tau_{[nt]}/n - t\right|^{1/2-\varepsilon} = o_{a.s.}(n^{-1/4+\varepsilon})$$

for any $\varepsilon > 0$. In particular, $B_n \to_{a.s.} B$ uniformly on $[0,1]$. For notational brevity, we subsequently write B_t^n and B_t respectively for $B_n(t)$ and $B(t)$. Also, we set $B = \sigma W$ in what follows so that W is the standard Brownian motion with unit variance.

The Brownian local time plays a central role in the development of our theory. The reader is referred to e.g., Chung and Williams (1990) for the introduction of the local time. If we let $L_B(t,s)$ and $L_W(t,s)$ be the local times of B and W respectively, where t and s are time and spatial parameters, then

$$L_B(t,s) = (1/\sigma) L_W(t, s/\sigma)$$

For the expository purpose, it is more convenient to use their scaled version

$$L(t,s) = (1/\sigma^2) L_B(t,s) = (1/\sigma^3) L_W(t, s/\sigma)$$

which is defined as the chronological local time by Phillips and Park (1998). Using the chronological local time L, we may represent an additive functional of Brownian motion B given by any locally integrable function F as

$$\int_0^t F(B_s)\, ds = \int_{-\infty}^{\infty} F(s) L(t,s)\, ds$$

which is called the *occupation times formula*. In the paper, we mostly deal with additive functionals of Brownian motion B over the unit interval $[0,1]$. Therefore, we will abbreviate $L(1,s)$ by $L(s)$, i.e., we will suppress its time parameter whenever it is unity. This convention will be made throughout the paper.

[5] Strictly, the equality in (5) holds only up to the distributional equivalence. We, however, regard it as the usual equality throughout this section. This convention allows us to avoid repetitious embeddings of various sample moments into the probability space where Brownian motion B is defined. Due to this the convention, all the subsequent convergence results of the sample moments with $\to_{a.s.}$, as well as those with $\to_d$, should generally be interpreted as the corresponding ones with $\to_d$. If the convergence is to a nonrandom limit, however, we may as well interpret it as $\to_p$, since $\to_d$ and $\to_p$ are identical in this case.

3.2 Consistency of the Estimated Lyapunov Exponents

We now show that the estimator λ_n of the Lyapunov exponent defined in (2) is consistent, i.e., $\lambda_n \to 0$, under the null hypothesis of random walk. The estimator includes sample moments of kernel function and its derivative, and our consistency proof relies on the approximations of such sample moments by appropriate continuous Brownian functionals. The following lemma allows us to approximate the sample moments such as

$$\begin{aligned} \frac{1}{n}\sum_{t=1}^{n} F\left(\frac{\cdot - y_{t-1}}{h_n}\right) &= \int_0^1 F\left(\frac{\cdot - \sqrt{n}B_t^n}{h_n}\right) dt \\ \frac{1}{\sqrt{n}}\sum_{t=1}^{n} F\left(\frac{\cdot - y_{t-1}}{h_n}\right) u_t &= \int_0^1 F\left(\frac{\cdot - \sqrt{n}B_t^n}{h_n}\right) dB_t \end{aligned}$$

by their corresponding functionals of Brownian motion, up to the desired precision, at fixed y or at the expanding ordinate which is specified as $\sqrt{n}x$. The expanding rate $\sqrt{n}$ is considered here because $\max_{1\le t\le n}|y_t| = O_{a.s.}(\sqrt{n})$. We let $c_n = \sqrt{n}/h_n$ in what follows.

3.4 Lemma (a) Let F be regular. Then

$$\int_0^1 F\left(\frac{y - \sqrt{n}B_t^n}{h_n}\right) dt = \int_0^1 F\left(\frac{y - \sqrt{n}B_t}{h_n}\right) dt + o_{a.s.}(c_n^{-3/2})$$

for all $y \in \mathbf{R}$. Moreover,

$$\int_0^1 F\left(c_n(x - B_t^n)\right) dt = \int_0^1 F\left(c_n(x - B_t)\right) dt + o_{a.s.}(c_n^{-3/2})$$

uniformly in x on any compact interval.

(b) Let F be regular and differentiable with regular derivative. Then

$$\int_0^1 F\left(\frac{y - \sqrt{n}B_t^n}{h_n}\right) dB_t = \int_0^1 F\left(\frac{y - \sqrt{n}B_t}{h_n}\right) dB_t + o_{a.s.}(c_n^{-1/2})$$

for all $y \in \mathbf{R}$. Moreover,

$$\int_0^1 F\left(c_n(x - B_t^n)\right) dB_t = \int_0^1 F\left(c_n(x - B_t)\right) dB_t + o_{a.s.}(c_n^{-1/2})$$

uniformly in x on any compact interval.

Due to the results in Lemma 3.4, it suffices to consider the corresponding functionals of Brownian motion to analyze the asymptotic behaviors of sample moments of various functions of random walks. The asymptotics of the relevant functionals of Brownian motion are given in the following lemma.

3.5 Lemma (a) Let F be bounded and $_iF$ be integrable. Then

$$c_n \int_0^1 F\left(\frac{y-\sqrt{n}B_t}{h_n}\right) dt = L(0) \int_{-\infty}^{\infty} F(t)\, dt + O_{a.s.}(n^{-1/4})$$

for all $y \in \mathbf{R}$. Moreover,

$$c_n \int_0^1 F\left(c_n(x-B_t)\right) dt = L(x) \int_{-\infty}^{\infty} F(t)\, dt + 2\int_0^1 F_\circ\left(c_n(x-B_t)\right) dB_t + O_{a.s.}(c_n^{-1})$$

for all $x \in \mathbf{R}$.

(b) Let F be square integrable and differentiable. Then

$$\int_0^1 F\left(\frac{y-\sqrt{n}B_t}{h_n}\right) dB_t = O_{a.s.}(c_n^{-1/2})$$

for all $y \in \mathbf{R}$. Moreover, for any $\varepsilon > 0$,

$$\int_0^1 F\left(c_n(x-B_t)\right) dB_t = o_{a.s.}(c_n^{-1/2+\varepsilon})$$

uniformly in x on any compact interval.

It follows immediately from the results in Lemmas 3.4 and 3.5 that

3.6 Lemma Suppose Assumptions 3.1 – 3.3 hold. Under the null hypothesis, we have

$$m_n^\circ(y) = 1 + O_{a.s.}(n^{-1/4}h_n^{-3/2})$$

for all $y \in \mathbf{R}$. Moreover, we have

$$\begin{aligned} m_n^\circ(\sqrt{n}x) &= 1 + \frac{h_n^{-2}}{L(x)} \int_0^1 K^\circ(c_n(x-B_t))dB_t + o_{a.s.}(n^{-1/4+\varepsilon}h_n^{-1/2-\varepsilon}) \\ &= 1 + O_{a.s.}(n^{-1/4+\varepsilon}h_n^{-3/2-\varepsilon}) \end{aligned}$$

for arbitrary $\varepsilon > 0$, uniformly in x on any compact interval.

3.7 Remarks (a) For $0 < \nu < 1/6$, we have $n^{-1/4}h_n^{-3/2} = o(1)$. Therefore the derivative estimator m_n° is consistent. This extends the consistency result for m_n obtained by Phillips and Park (1998). In fact, they show that

$$m_n(y) = y + O_p(n^{-1/4}h_n^{-1/2})$$

As expected, the convergence rate for the derivative estimator is slower. More precisely, the rate is reduced by $O_p(h_n^{-1})$. This is the reason that we require the bandwidth contraction parameter ν be smaller than 1/6 instead of 1/2.

(b) The maximum rate of convergence for m_n° is given by $O_p(n^{-1/4})$, which is obtained when we set h_n to be a constant. This is completely analogous to the kernel estimator m_n considered in Phillips and Park (1998). They also allow the bandwidth to grow as the sample size increases, i.e., $h_n \to \infty$. We, however, do not look at such cases because it would make the resulting tests inconsistent.

(c) Lemma 3.6 also gives the convergence results when the ordinate is set to increase along with the sample size. Needless to say, we may let $\varepsilon > 0$ be small enough to make $n^{-1/4+\varepsilon} h_n^{-3/2-\varepsilon} = o(1)$ if $0 < \nu < 1/6$. Therefore, for the changing ordinate $\sqrt{n}x$, the derivative estimator m_n° is uniformly consistent on any compact interval.

We may now easily establish the consistency of λ_n from the results in Lemma 3.6. To see $\lambda_n \to_{a.s.} 0$, note that $\lambda_n = \int_0^1 \log |m_n^\circ(\sqrt{n}B_t^n)|\,dt$ and that $B_n \to_{a.s.} B$ uniformly on $[0,1]$. Of course, Brownian motion on the unit interval takes values in a compact interval a.s., since it has continuous sample paths a.s.

3.8 Theorem Suppose Assumptions 3.1 – 3.3 hold. Under the null hypothesis, we have

$$\lambda_n \to_{a.s.} 0$$

as $n \to \infty$.

3.3 Asymptotic Distribution of the Test

The asymptotic distribution of λ_n can also be readily obtained from the result in Lemma 3.6. To derive the limiting distribution of λ_n, we note again that $\lambda_n = \int_0^1 \log |m_n^\circ(\sqrt{n}B_t^n)|\,dt$ and that $\log|1+x| = x - x^2/2 + o(|x|^2)$. Then it follows from the result in Lemma 3.6 that

$$\begin{aligned}\lambda_n \;=\; & h_n^{-2} \int_0^1 \frac{dt}{L(B_t^n)} \int_0^1 ds\, K^\circ(c_n(B_t^n - B_s)) \\ & - \frac{h_n^{-4}}{2} \int_0^1 \frac{dt}{L(B_t^n)^2} \left(\int_0^1 ds\, K^\circ(c_n(B_t^n - B_s)) \right)^2 + o_{a.s.}(n^{-1/2+\varepsilon} h_n^{-2-\varepsilon})\end{aligned}$$

Following the approach taken earlier, we now approximate B_n by B and analyze the resulting functionals of Brownian motion to derive the limiting distribution of λ_n.

The necessary approximations are in fact very straightforward given the results in Lemmas 3.4 and 3.5. For instance, it follows readily from the part (a) of Lemmas 3.4 and 3.5 that

$$\begin{aligned}& \int_0^1 \frac{dt}{L(B_t^n)} \int_0^1 ds\, F\left(c_n(B_t^n - B_s)\right) \\ & \quad = \int_0^1 \frac{dt}{L(B_t)} \int_0^1 ds\, F\left(c_n(B_t - B_s)\right) + o_{a.s.}(c_n^{-3/2}) + o_{a.s.}(c_n^{-1} n^{-1/4+1/2p+\varepsilon})\end{aligned}$$

since, in particular, we have for sufficiently large n

$$\left| \frac{1}{L(B_t^n)} - \frac{1}{L(B_t)} \right| \le c\, \frac{|L(B_t^n) - L(B_t)|}{\inf_{t\in[0,1]} L(B_t)^2} = o_{a.s.}(n^{-1/4+1/2p+\varepsilon})$$

for some constant c and arbitrary $\varepsilon > 0$. This is due to the Hölder continuity of the local time L. Likewise, we have

$$\int_0^1 \frac{dt}{L(B_t^n)} \int_0^1 F\left(c_n(B_t^n - B_s)\right) dB_s = \int_0^1 \frac{dt}{L(B_t)} \int_0^1 F\left(c_n(B_t - B_s)\right) dB_s + o_{a.s.}(c_n^{-1/2})$$

due to the results in part (b) of Lemmas 3.4 and 3.5.

3.9 Lemma (a) Let F be integrable and differentiable. Then we have

$$\int_0^1 \frac{dt}{L(B_t)} \int_0^1 F\left(c_n(B_t - B_s)\right) dB_s = O_{a.s.}(c_n^{-1})$$

where the stochastic integral dB_s is taken in the filtration $\mathcal{F}_s^{(t)} = \sigma\left((B_r)_{r\leq s}, B_t\right)$.

(b) Let F be integrable. Then we have

$$c_n \int_0^1 \frac{dt}{L(B_t)^2} \int_0^1 ds\, F\left(c_n(B_t - B_s)\right) = \left(\int_{-\infty}^{\infty} F(t)\, dt\right)(B_{\max} - B_{\min}) + O_{a.s.}(c_n^{-1/2})$$

where $B_{\max} = \sup_{t\in[0,1]} B_t$ and $B_{\min} = \inf_{t\in[0,1]} B_t$.

(c) Let F and G be square integrable. Then we have

$$\int_0^1 \frac{dt}{L(B_t)^2} \int_0^1 F\left(c_n(B_t - B_s)\right) \int_0^s G\left(c_n(B_t - B_r)\right) dB_r dB_s = O_p(c_n^{-3/2})$$

where the stochastic integrals dB_r and dB_s are taken respectively in the filtrations $\mathcal{F}_r^{(t)} = \sigma\left((B_q)_{q\leq s}, B_t\right)$ and $\mathcal{F}_s^{(t)} = \sigma\left((B_q)_{q\leq s}, B_t\right)$.

The asymptotic null distribution of the appropriately normalized Lyapunov exponent estimator can now be obtained from Lemma 3.9 in a straightforward manner. Let $B_{\max}$ and $B_{\min}$ be defined as above.

3.10 Lemma Suppose Assumptions 3.1 – 3.3 hold. Under the null hypothesis, we have

$$(n^{1/2}h_n^3)\,\lambda_n \to_d -(1/2)\left(\int_{-\infty}^{\infty} K^{\circ}(t)^2 dt\right)(B_{\max} - B_{\min})$$

as $n \to \infty$.

3.11 Remarks (a) The estimator for the Lyapunov exponent, which is the average of the log derivative estimate, converges faster than the derivative estimator. Specifically, the convergence rate is given by $n^{1/2}h_n^3$, while that of the derivative estimator is $n^{1/4}h_n^{3/2}$. It is well known for the stationary models that the average derivative estimation has faster convergence rates than the derivative estimation.

(b) The maximum rate of convergence is achieved when the bandwidth parameter is held fixed, i.e., h_n is set to be a constant and the maximum rate is $\sqrt{n}$.

We now have the asymptotic distribution of our statistic T_n under the null hypothesis. Let $W = (1/\sigma)B$ be standard Brownian motion, and define $W_{\max}$ and $W_{\min}$ similarly as $B_{\max}$ and $B_{\min}$.

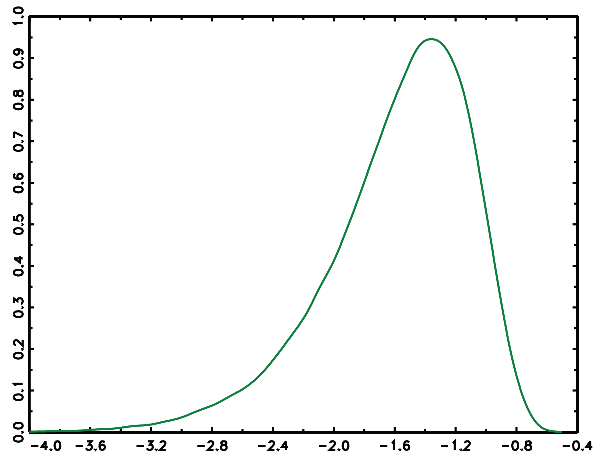

Figure 1: Probability Density of T

3.12 Theorem Suppose Assumptions 3.1 – 3.3 hold. Under the null hypothesis, we have as $n \to \infty$

$$T_n \to_d T,$$

where $T = W_{\min} - W_{\max}$ has probability density

$$f(x) = \sqrt{\frac{2}{\pi}} \sum_{k=-\infty}^{\infty} (-1)^k k(k+1) \left[\exp\left(-\frac{(k+1)^2 x^2}{2} \right) - \exp\left(-\frac{k^2 x^2}{2} \right) \right]$$

on the negative half line $(-\infty, 0]$.

3.13 Remarks (a) The statistic T_n takes asymptotically only negative values under the null hypothesis. For the chaotic system, it is expected to have largely positive values and diverge in probability off to infinity, see Theorem 4.6 below. Therefore, we reject the null hypothesis if the value of T_n is sufficiently large.

(b) The limiting distribution of T_n has mean -1.595 and standard deviation 0.474 approximately. It has median -1.346, and is skewed to the left with skewness -0.940. The asymptotic critical values of T_n are given in Table 1. The limiting probability density f of T is presented in Figure 1.

(c) For the stationary non-chaotic models, our estimator λ_n is known to be consistent for such model (see McCaffrey et al. (1992) and Whang and Linton (1997)), and therefore, $T_n \to_p -\infty$. We may therefore use our statistic T_n to test for the random walk model against the alternative of stationary non-chaotic models with $\lambda < 0$, if we reject the null hypothesis when the statistic takes sufficiently small values. We, however, do not investigate T_n any further along this line in the paper, since our main purpose is to test for random walk against chaos. For a survey of the unit root tests against stationarity, the reader is referred e.g. to Phillips (1997).

(d) Our test is expected to reject the null hypothesis of random walk, when the process is generated by the stationary non-chaotic model with $\lambda = 0$, which has an interesting dynamic behavior and exhibits the 'Lyapunov stability' with a neutral fixed point and orbit. For the stationary non-chaotic models with $\lambda = 0$, we have $\lambda_n = O_p(n^{-1/2})$, and consequently $T_n \to_p 0$. In this case, our test will therefore reject the null hypothesis of random walk with probability one asymptotically.[6]

(e) It is suggested in the definition of our statistic T_n that the unknown variance σ^2 of the innovation sequence (u_t) is estimated under the null hypothesis of the random walk as in (4). It should not affect, at least in any critical way, the asymptotic behavior of T_n under the alternative hypothesis since we normally expect $\sum_{t=1}^n \triangle y_t^2/n = O_p(1)$ for the chaotic models as well as for the random walk models. Clearly, it is also true for many stable stationary models.

Table 1: Distribution of T

$\mathbf{P}\{T \leq x\}$	0.01	0.05	0.10	0.90	0.95	0.99
x	–3.02	–2.50	–2.24	–1.06	–0.97	–0.83

4. Test Consistency

In this section, we show that our test is consistent against the chaotic alternatives. We first introduce the conditions that are assumed to hold under the alternative hypothesis.

4.1 Assumption (a) (y_t) is a sequence of strictly stationary strong mixing random variables with $\mathbf{E}|y_t|^p < \infty$ for some $p > 2$ and the mixing numbers $(\alpha_k)_{k\geq 1}$ satisfying $\sum_{k=1}^\infty \alpha_k^{(p-2)/p} < \infty$,
(b) (y_t) has distribution absolutely continuous with respect to Lebesgue measure with density f, which is strictly positive on the support, and
(c) f and mf are q-times differentiable with bounded derivatives for some $q \geq 3$.

4.2 Assumption (a) $\int_{-\infty}^\infty K(t)dt = 1, \int_{-\infty}^\infty t^k K(t)dt = 0$ for all $1 \leq k \leq q-2$ and $\int_{-\infty}^\infty |t|^{q-1}|K(t)| < \infty$, where q is given in Assumption 4.1,
(b) $K, K^\circ, {}_iK$ and ${}_iK^\circ$ are bounded, and
(c) K and K° are integrable and have Fourier transform satisfying the condition $\int_{-\infty}^\infty (1+|t|) \sup_{\eta\geq 1} |\varphi(\eta t)|dt < \infty$.

4.3 Assumption (a) $(\min_{1\leq t\leq n} |m^\circ(y_t)|)^{-1} = O_p(n^r)$ for some $r \geq 0$, and
(b) for $\varepsilon_n \to 0$, $\mathbf{P}(|m^\circ(y_t)| \leq \varepsilon_n) \leq \varepsilon_n$ for all n sufficiently large.

4.4 Assumption The bandwidth parameter (h_n) satisfies $h_n = c_0 n^{-\nu}$ for some constant $c_0 > 0$ and $1/4(q-1) \leq \nu \leq 1/8$, where q is given by Assumption 4.1.

[6]We are grateful to a referee, who pointed out this to us.

4.5 Remarks (a) Assumption 4.3(a) is quite weak and is often justified by the extreme value theory in a variety of contexts (see, for example Gnedenko(1943)). We expect this assumption holds with $r = 1$ for chaotic processes. (See Whang and Linton (1997) for a verification of this assumption for the Feingenbaum map.)

(b) Assumption 4.3(b) holds if $m^{\circ}(y_t)$ has density bounded in a neighborhood of the origin. For the [zero noise] Feingenbaum map $m(y_t) = 4y_t(1-y_t)$, this assumption is satisfied because the stationary ergodic density of $m^{\circ}(y_t)$ in this case is given by $f(y) = 1/\sqrt{\pi^2 y(1-y)}$. This is bounded at the point $y_t = 1/2$ at which $m^{\circ}(y_t) = 0$.

4.6 Theorem Suppose Assumptions 4.1 – 4.4 hold. If $\lambda > 0$, then we have

$$(n^{1/2}h_n^3)^{-1+\varepsilon}T_n \to_p \infty$$

as $n \to \infty$ for any $\varepsilon > 0$.

5. Simulations and Empirical Applications

We first analyze how our testing procedure works on the random walk specified as $y_t = y_{t-1} + u_t$ with independent standard normal innovations (u_t). For such (y_t), we of course have $\lambda = 0$. We conduct simulations for our testing procedure to examine the finite sample performance. As a benchmark, we consider the augmented Dickey Fullter test ADF_n of the null hypothesis $H_0 : \rho = 1$ for the model $y_t = \mu + \gamma t + \rho y_{t-1} + \sum_{j=1}^{k} \beta_j \Delta y_{t-j} + u_t$.[7] For our test, we use the standard normal kernel $K(t) = (2\pi)^{-1/2}\exp(-t^2/2)$ and the bandwidth given by $h_n = c\,\omega_n$, where $c = .310, \ldots, .400$ and $\omega_n^2 = \sum_{t=1}^{n} y_t^2/n$. A Total of 10,000 replications are used for each experiment. We consider $n = 50, 100, 250$ and 500 and report the rejection probabilities of the test with nominal sizes $\alpha = 0.10, 0.05$ and 0.01. Table 2 shows the size performance of the tests. There appears to be considerable dependence of our test on the bandwidth parameter h_n, but the values $c = .385, .370, .345$ and .325 do reasonably well for the samples sizes $n = 50, 100, 250$ and 500, respectively. On the other hand, the ADF test also has good size performance in general.

[7] A simulation study by Choi and Moh (2007) shows that the ADF test has a reasonable power aganist some class of nonlinear alternatives. Also, Diba and Grossman (1988) show that a standard unit root test can sometimes used to test for a unit root against explosive alternatives.

Table 2: Actual Size

n	c	T_n			k	ADF_n		
		10%	5%	1%		10%	5%	1%
	.370	.157	.102	.039	1	.115	.059	.014
50	.385	.082	.046	.015	2	.132	.069	.018
	.400	.037	.017	.005	3	.140	.079	.023
	.355	.201	.125	.042	1	.111	.059	.012
100	.370	.097	.051	.010	2	.116	.064	.014
	.385	.035	.014	.002	3	.116	.064	.015
	.330	.275	.183	.064	1	.098	.052	.010
250	.345	.122	.058	.010	2	.099	.054	.012
	.360	.032	.009	.000	3	.100	.055	.012
	.310	.285	.201	.079	1	.105	.053	.009
500	.325	.128	.063	.010	2	.108	.055	.011
	.340	.026	.006	.000	3	.107	.055	.010

To examine the power performance of the test, we next consider a chaotic process, the Feingenbaum map with system noise:

$$y_t = 4y_{t-1}(1 - y_{t-1}) + \sigma\varepsilon_t$$

where $(\varepsilon_t/v_t) \sim \text{Uniform}\,(-1, 1)$ independent of y_t, and

$$v_t = \min\{4y_{t-1}(1 - y_{t-1}), 1 - 4y_{t-1}(1 - y_{t-1})\}$$

This particular form of heteroskedasticity ensures that the process y_t is restricted to the unit interval. Several values of the parameter σ were tried to make the signal/noise ratio, as defined in Dechert and Gencay (1992), vary and become .005, .007 and .010. However, the results do not change much with the signal/noise ratio, so we only report the results for $\sigma = .005$ for which $\lambda = .692$. When $\sigma = 0$, we have $\lambda = \ln 2$. Table 3 shows that the power performance of T_n is not very sensitive to h_n and the rejection probability is quite close to one even for the small sample sizes. Interestingly, the ADF test also has non-negligible power against this chaotic alternative, but its finite sample power is substantially dominated by that of T_n, especially when n is small and the number of lagged terms k is large.

Table 3: Finite Sample Power (Feigenbaum map with noise)

n	c	T_n			k	ADF_n		
		10%	5%	1%		10%	5%	1%
	.370	.969	.962	.950	1	.994	.982	.893
50	.385	.957	.948	.932	2	.928	.857	.592
	.400	.945	.936	.919	3	.791	.647	.310
	.355	.993	.991	.989	1	1.00	1.00	1.00
100	.370	.991	.989	.984	2	1.00	1.00	1.00
	.385	.989	.986	.980	3	.999	.995	.953
	.330	1.00	1.00	1.00	1	1.00	1.00	1.00
250	.345	1.00	1.00	1.00	2	1.00	1.00	1.00
	.360	1.00	1.00	1.00	3	1.00	1.00	1.00
	.310	1.00	1.00	1.00	1	1.00	1.00	1.00
500	.325	1.00	1.00	1.00	2	1.00	1.00	1.00
	.340	1.00	1.00	1.00	3	1.00	1.00	1.00

Table 4 presents the rejection probabilities of the test againt an explosive alternative, i.e. $y_t = \rho y_{t-1} + u_t$ for $\rho > 1$. Although an explosive behavior is out of scope of our paper, Table 4 shows that our test has substantial power even for $n = 50$ and is strictly more powerful than the ADF test against mildly explosive explosive alternatives $\rho = 1.01, ..., 1.04$.

Table 4: Finite Sample Power (Explosive AR(1), $n = 50$)

ρ	c	T_n			k	ADF_n		
		10%	5%	1%		10%	5%	1%
	.370	.237	.175	.092	1	.113	.057	.016
1.01	.385	.154	.107	.052	2	.131	.071	.018
	.400	.089	.058	.023	3	.136	.077	.020
	.370	.701	.672	.604	1	.406	.314	.170
1.02	.385	.660	.620	.550	2	.418	.328	.180
	.400	.603	.565	.493	3	.422	.332	.190
	.370	.929	.920	.901	1	.880	.848	.781
1.03	.385	.916	.905	.885	2	.882	.852	.788
	.400	.901	.889	.865	3	.879	.849	.786
	.370	.980	.978	.973	1	.979	.974	.963
1.04	.385	.979	.976	.970	2	.979	.974	.964
	.400	.971	.967	.961	3	.980	.074	.964

We next apply our testing procedure to some of macro and financial time series. The data sets considered are : US treasury bill rates TB3M (3-month) and TB1Y (1-year), stock price indices SP500 (Standard and Poor composite index 500) and DJ

(Dow-Jones industrial average), and two bilateral exchange rates ERJPUS (Japanese Yen/US Dollar) and ERCAUS (Canadian Dollar/US Dollar). For the stock price indices and exchange rates, we took the natural logarithm of the data. The details of the data are described in Table 5.

Table 5: Data Description

Data	Period	Frequency	n	Mean	S.D.
TB3M	1954.1.8-2008.10.10	weekly	2858	5.1107	2.8004
TB1Y	1959.7.17-2001.8.24	weekly	2198	6.1303	2.4152
SP500	1926.1-2008.10	monthly	994	1.9124	0.6865
DJ	1900.1-2008.10	monthly	1302	2.6616	0.6965
ERCAUS	1971.1.4-2008.10.16	daily	9493	0.0888	0.0592
ERJPUS	1979.1.2-2008.10.16	daily	9481	2.2115	0.1730

The empirical results are summarized in Table 6. We find that the values of the test statistic T_n are rather sensitive to the choice of h_n, while those of the Lyapunov exponent estimate are relatively robust to the choice. Table 6 shows that we have stronger evidences for the null hypothesis of random walk against chaotic alternatives as h_n takes larger values. To illustrate sensitivity of our testing results with respect to h_n, we present our testing results with h_n in ranges that are sufficiently wide to include the cases in which we reject and accept at the convention significance level (i.e., 1%, 5%, or 10%) using the asymptotic critical values given in Table 1. To alleviate the dependency of the test on the bandwidth selection, we also perform tests using bootstrap critical values instead of the asymptotic critical values. The bootstrap samples $\{y_t^* : t = 1, ..., n\}$ are generated from the random walk model $y_t^* = y_{t-1}^* + u_t^*$, where the residuals $\{u_t^* : t = 1, ..., n\}$ are drawn (with replacement) from the empirical distribution of the recentered innovations $\{\Delta y_t - \sum_{t=1}^{n} \Delta y_t / n : t = 1, ...n\}$. This step is repeated B times and the bootstrap p-value is then computed by comparing the bootstrap distribution and the original test statistic T_n. To save computational cost, we take $B = 100$ in our applications.

Table 6: Empirical Results

	TB3M				TB1Y		
h_n	λ_n	T_n	pval	h_n	λ_n	T_n	pval
0.10	-0.1284	-0.4852	0.39	0.10	-0.1023	-0.3911	0.24
0.11	-0.1359	-0.6835	0.56	0.11	-0.1082	-0.5509	0.38
0.12	-0.1413	-0.9225	0.61	0.12	-0.1085	-0.7170	0.39
0.13	-0.1485	-1.2325	0.65	0.13	-0.1096	-0.9212	0.41
0.14	-0.1443	-1.4962	0.65	0.14	-0.1084	-1.1381	0.40
0.15	-0.1424	-1.8151	0.65	0.15	-0.1097	-1.4166	0.41
0.16	-0.1429	-2.2110	0.67	0.16	-0.1083	-1.6962	0.39
0.17	-0.1379	-2.5605	0.64	0.17	-0.1096	-2.0592	0.39
0.18	-0.1362	-3.0020	0.60	0.18	-0.1088	-2.4277	0.38
0.19	-0.1365	-3.5363	0.60	0.19	-0.1094	-2.8692	0.37

	SP500				DJ		
h_n	λ_n	T_n	pval	h_n	λ_n	T_n	pval
0.01	-0.2809	-0.0052	0.90	0.01	-0.3059	-0.0066	0.99
0.02	-0.1892	-0.0279	0.79	0.02	-0.2352	-0.0407	0.99
0.03	-0.1894	-0.0943	0.69	0.03	-0.2367	-0.1384	0.96
0.04	-0.2101	-0.2480	0.66	0.04	-0.2500	-0.3464	0.97
0.05	-0.2304	-0.5311	0.63	0.05	-0.2709	-0.7330	0.96
0.06	-0.2537	-1.0106	0.58	0.06	-0.2974	-1.3906	0.93
0.07	-0.2773	-1.7543	0.54	0.07	-0.3269	-2.4273	0.91
0.08	-0.3013	-2.8455	0.55	0.08	-0.3571	-3.9576	0.92
0.09	-0.3257	-4.3797	0.55	0.09	-0.3866	-6.1010	0.91
0.10	-0.3502	-6.4588	0.52	0.10	-0.4147	-8.9771	0.86

	ERJPUS				ERCAUS		
h_n	λ_n	T_n	pval	h_n	λ_n	T_n	pval
0.015	-0.2165	-0.3609	0.89	0.010	-0.3043	-0.2912	0.98
0.016	-0.2194	-0.4440	0.89	0.011	-0.3307	-0.4212	0.96
0.017	-0.2227	-0.5406	0.83	0.012	-0.3546	-0.5863	0.96
0.018	-0.2263	-0.6520	0.76	0.013	-0.3750	-0.7885	0.93
0.019	-0.2299	-0.7789	0.74	0.014	-0.3914	-1.0279	0.94
0.020	-0.2334	-0.9225	0.68	0.015	-0.4034	-1.3029	0.93
0.021	-0.2369	-1.0837	0.70	0.016	-0.4108	-1.6101	0.90
0.022	-0.2403	-1.2641	0.60	0.017	-0.4136	-1.9447	0.87
0.023	-0.2438	-1.4656	0.67	0.018	-0.4123	-2.3011	0.85
0.024	-0.2475	-1.6906	0.68	0.019	-0.4074	-2.6739	0.77

Our results in Table 6 suggest that, for all of the series we considered, we have little empirical evidence to reject the random walk hypothesis in favor of the alternative of chaos. The finding is consistent with some of the existing results in the literature, e.g. Shintani and Linton (2004). However, we should mention that our empirical findings have some limitations and hence should be interpreted more carefully. First, the random walk model we consider is assumed to have iid innovations. If there is conditional heteroskedasticity or other types of dependency in the innovation sequence, we may have different testing results. Second, our test is based on one-dimensional Lyapunov exponent and would not be powerful to detect multi-dimensional chaotic alternatives. Third, our test is justified asymptotically and hence its finite sample performance might be distorted by the small sample bias of the nonparametric estimate of the Lyapunov exponent, see Shintani and Linton (2003, Section 5) for more discussion on this and the other limitations.

6. Conclusion

In this paper, we have developed a test of the random walk hypothesis against the chaotic alternative. The test is based on the Nadaraya-Watson kernel estimate of the Lyapunov exponent, which is expected to capture the differing sensitivities of the random walk and chaos models to their dependencies on the initial condition. Our test performs reasonably well in finite samples. In particular, it seems to have quite good discriminatory powers against interesting alternatives. The test is applied to several important macroeconomic and financial time series. There appears to be little empirical evidence in favor of chaos for the interested rates, stock price indices, exchange rates data we considered. They seem to be better modelled as random walks or non-chaotic stationary processes. However, we note that our empirical findings have some limitations and should be taken more carefully in practice.

7. Mathematical Proofs

7.1 Proof of Lemma 3.4 We first prove part (a). Write

$$\frac{1}{n}\sum_{t=1}^{n} F\left(\frac{y-y_{t-1}}{h_n}\right) = \int_0^1 F\left(\frac{y-\sqrt{n}B_t^n}{h_n}\right) dt$$

It follows from Phillips and Park (1998, Proof of Theorem 3.1) that

$$\begin{aligned}\int_0^1 F\left(\frac{y-\sqrt{n}B_t^n}{h_n}\right) dt &= \int_0^1 F\left(\frac{y-\sqrt{n}B_t}{h_n}\right) dt + o_{a.s.}\left(\frac{n^{-3/4+\varepsilon}}{c_n^{r/2}}\left(1+\frac{\kappa_n}{c_n^{2r}}\right)^{1/2}\right) \\ &\quad + O_{a.s.}\left(\frac{n^{-1/2+1/p}}{\kappa_n}\right) + o_{a.s.}\left(\frac{1}{c_n^{2q(1-r)}}\right) \end{aligned} \tag{7}$$

for any $\varepsilon > 0$, where the constants p and q are given in Assumptions 3.1 and 3.3, and (κ_n) is a sequence of numbers such that

$$1 \leq \kappa_n/c_n^r \leq n^{1/2}\log n \tag{8}$$

with $r \in (0,1)$.

We may choose r so that

$$\frac{1}{1+2\nu} < r < 1 - \frac{2}{(1+2\nu)q} \tag{9}$$

which is possible since $q > 1/\nu$ as we assume in Assumption 3.3. With the choice of r in (9), we have

$$\frac{n^{-3/4+\varepsilon}}{c_n^{r/2}},\ \frac{1}{c_n^{2q(1-r)}} = o(n^{-1}) \tag{10}$$

Furthermore, if we let $\kappa_n = n^\kappa$ with κ satisfying

$$\max\left(\frac{r(1+2\nu)}{2}, \frac{p+2}{2}\right) < \kappa < \frac{1+r(1+2\nu)}{2}$$

then the inequalities in (8) hold, $\kappa_n/c_n^{2r} = o(1)$ and

$$\frac{n^{-1/2+1/p}}{\kappa_n} = o(n^{-1}) \tag{11}$$

It now follows from (10) and (11) that the residual term in (7) is

$$o_{a.s.}(n^{-1}) = o_{a.s.}(n^{-3/4}h_n^{3/2})$$

due to the condition $0 < \nu < 1/6$ in Assumption 3.2 on the bandwidth parameter. It is easy to see that the approximation results (7) in Park and Phillips (1999) hold uniformly in x, when we specify $y = \sqrt{n}x$. also The proof for part (a) is therefore complete.

For part (b), we define

$$\begin{aligned} R_n &= \int_0^1 F\left(\frac{y-\sqrt{n}B_t}{h_n}\right)dB_t - \int_0^1 F\left(\frac{y-\sqrt{n}B_t^n}{h_n}\right)dB_t \\ &= \sum_{i=1}^n \int_{\tau_{i-1}/n}^{\tau_i/n}\left[F\left(\frac{y-\sqrt{n}B_t}{h_n}\right) - F\left(\frac{y-\sqrt{n}B_t^n}{h_n}\right)\right]dB_t \qquad (12) \end{aligned}$$

and show that $R_n = o_{a.s.}(c_n^{-1/2})$, which we now set out to do. Write $B_i^n = B(\tau_i/n)$ for notational brevity, and set $\sigma^2 = 1$. For the first term in (12), we apply the Ito formula to obtain

$$\begin{aligned} \int_{\tau_{i-1}/n}^{\tau_i/n} F\left(\frac{y-\sqrt{n}B_t}{h_n}\right)dB_t &= -c_n^{-1}\left[{}_\circ F\left(\frac{y-\sqrt{n}B_i^n}{h_n}\right) - {}_\circ F\left(\frac{y-\sqrt{n}B_{i-1}^n}{h_n}\right)\right] \\ &\quad + \frac{c_n}{2}\int_{\tau_{i-1}/n}^{\tau_i/n} F^\circ\left(\frac{y-\sqrt{n}B_t}{h_n}\right)dt \qquad (13) \end{aligned}$$

On the other hand, we may write the second term in (12) as

$$\begin{aligned} F\left(\frac{y-\sqrt{n}B_{i-1}^n}{h_n}\right)\left(B_i^n - B_{i-1}^n\right) &= -c_n^{-1}\left[{}_\circ F\left(\frac{y-\sqrt{n}B_i^n}{h_n}\right) - {}_\circ F\left(\frac{y-\sqrt{n}B_{i-1}^n}{h_n}\right)\right] \\ &\quad + \frac{c_n}{2}F^\circ\left(\frac{y-\sqrt{n}\overline{B}_{i-1}^n}{h_n}\right)\left(B_i^n - B_{i-1}^n\right)^2 \qquad (14) \end{aligned}$$

using the second order Taylor expansion, where $\overline{B}_{i-1}^n$ is a random variable taking values between B_{i-1}^n and B_i^n.

Since the Brownian motion B has continuous sample path, we may define a time change $(\overline{\tau}_k)_{k\geq 1}$ such that $\overline{\tau}_i$ is the infimum of $t \geq \tau_i$ for which $\overline{B}_i^n = B(\overline{\tau}_i/n)$ for $i = 1, \ldots, n$. Then we have

$$\tau_k \leq \overline{\tau}_k \leq \tau_{k+1}$$

for all $k \geq 1$. Furthermore, if we define $(\overline{u}_t)$ by

$$B\left(\overline{\tau}_k/n\right) = n^{-1/2}\sum_{i=1}^k \overline{u}_i$$

analogously as in (5), then $(\overline{u}_t)$ becomes a new innovation sequence, which is iid and satisfies Assumption 2.1. This is due to the strong Markov property of Brownian motion (see e.g. Hida (1980, Theorem 2.8, p80). We therefore have from the result in part (a)

$$\int_0^1 F^\circ\left(\frac{y-\sqrt{n}\overline{B}_t^n}{h_n}\right)dt = \int_0^1 F^\circ\left(\frac{y-\sqrt{n}B_t}{h_n}\right)dt + o_{a.s.}(c_n^{-3/2}) \tag{15}$$

where $\left(\overline{B}_t^n\right)_{t\in[0,1]}$ is constructed from $(B(\overline{\tau}_k/n))_{1\le k\le n}$, similarly as $(B_t^n)_{t\in[0,1]}$ defined from $(B(\tau_k/n))_{1\le k\le n}$ in (6).

Now we show that

$$\begin{aligned} S_n &= \sum_{i=1}^n F^\circ\left(\frac{y-\sqrt{n}\overline{B}_{i-1}^n}{h_n}\right)\left(B_i^n - B_{i-1}^n\right)^2 - \frac{1}{n}\sum_{i=1}^n F^\circ\left(\frac{y-\sqrt{n}\overline{B}_{i-1}^n}{h_n}\right) \\ &= o_{a.s.}(c_n^{-3/2}) \end{aligned} \tag{16}$$

from which and (12)-(15) the stated result follows. To deduce (16), we define for each n

$$A_i^n = \left((B_i^n)^2 - (i/n)\right) - 2\sum_{k=1}^i B_{k-1}^n\left(B_k^n - B_{k-1}^n\right)$$

with $A_0^n = 0$, $i = 1, \dots, n$, so that

$$A_i^n - A_{i-1}^n = \left(B_i^n - B_{i-1}^n\right)^2 - (1/n)$$

We therefore have

$$S_n = \sum_{i=1}^n F^\circ\left(\frac{y-\sqrt{n}\overline{B}_{i-1}^n}{h_n}\right)\left(A_i^n - A_{i-1}^n\right) \tag{17}$$

Introduce time changes $(\tau_k^u)_{k\ge1}$ and $(\tau_k^v)_{k\ge1}$ such that, if we let $U_i^n = B(\tau_i^u/n)$ and $U_i^n = B(\tau_i^u/n)$ for $i = 1, \dots, n$, then

$$U_i^n - V_i^n = A_{i+1}^n - A_i^n$$

and

$$\begin{aligned} &F\left(\frac{y-\sqrt{n}\left(\overline{B}_{i-1}^n + U_{i-1}^n\right)}{h_n}\right) - F\left(\frac{y-\sqrt{n}\left(\overline{B}_{i-1}^n + V_{i-1}^n\right)}{h_n}\right) \\ &= c_n F^\circ\left(\frac{y-\sqrt{n}\overline{B}_{i-1}^n}{h_n}\right)\left(A_i^n - A_{i-1}^n\right) \end{aligned} \tag{18}$$

for $i = 1, \dots, n$. Clearly, we may choose $(\tau_k^u)_{k\ge1}$ and $(\tau_k^v)_{k\ge1}$ such that

$$\tau_k \le \tau_k^v \le \overline{\tau}_k \le \tau_k^u \le \tau_{k+1}$$

for all $k \ge 1$.

It follows from (17) and (18) that

$$S_n = c_n^{-1}\left[\sum_{i=1}^n F\left(\frac{y-\sqrt{n}\left(\overline{B}_{i-1}^n + U_{i-1}^n\right)}{h_n}\right) - \sum_{i=1}^n F\left(\frac{y-\sqrt{n}\left(\overline{B}_{i-1}^n + V_{i-1}^n\right)}{h_n}\right)\right]$$

However, we have

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n F\left(\frac{y-\sqrt{n}\left(\overline{B}_{i-1}^n + U_{i-1}^n\right)}{h_n}\right) &= \int_0^1 F\left(\frac{y-\sqrt{n}B_t}{h_n}\right)dt + o_{a.s.}(c_n^{-3/2})\\ \frac{1}{n}\sum_{i=1}^n F\left(\frac{y-\sqrt{n}\left(\overline{B}_{i-1}^n + V_{i-1}^n\right)}{h_n}\right) &= \int_0^1 F\left(\frac{y-\sqrt{n}B_t}{h_n}\right)dt + o_{a.s.}(c_n^{-3/2})\end{aligned}$$

due to the result in part (a). Consequently,

$$S_n = o_{a.s.}(nc_n^{-5/2}) = o_{a.s.}(c_n^{-1/2}h_n^2)$$

as was to be shown. It is straightforward to see that all of the above results hold uniformly in x on any compact interval if we let $y = \sqrt{n}x$. ■

7.2 Proof of Lemma 3.5 The first part of (a) follows readily from the result in Phillips and Park (1998, Proof of Theorem 3.1) and Park and Phillips (1999, Theorem 5.1). For the second part, we deduce by the successive applications of the occupation times formula and the change of variables that

$$\begin{aligned}c_n\int_0^1 F\left(c_n(x-B_t)\right)dt &= c_n\int_{-\infty}^{\infty} F\left(c_n(x-t)\right)L(t)\,dt\\ &= \int_{-\infty}^{\infty} F(t)L(x-t/c_n)\,dt\end{aligned}$$

which we write as

$$L(x)\int_{-\infty}^{\infty} F(t)\,dt + D_n(x,F)$$

where

$$D_n(x,F) = \int_{-\infty}^{\infty} F(s)\left[L(x-s/c_n) - L(x)\right]ds \tag{19}$$

Now it suffices to show that

$$D_n(x,F) = 2\int_0^1 F_\circ\left(c_n(x-B_t)\right)dB_t + O_{a.s.}(c_n^{-1}) \tag{20}$$

to obtain the stated result.

To show (20), we first let $s<0$ fixed and apply the result in Exercise 1.28 (p.226) of Revuz and Yor (1994) with $f = 1(x, x-s/c_n]$ and $S=1$. Note that

$$\int_{-\infty}^{\infty}|f(t)|\,dt = s/c_n$$

It can therefore be deduced that

$$\begin{aligned} L\left(x-s/c_n\right)-L(x) &= 2\int_0^1 1\left\{x\le B_t\le x-s/c_n\right\}dB_t+O_{a.s.}(c_n^{-1}) \\ &= 2\int_0^1 1\left\{s\le c_n(x-B_t)\le 0\right\}dB_t+O_{a.s.}(c_n^{-1}) \qquad (21)\end{aligned}$$

uniformly for all $s\in\mathbf{R}$. Moreover, we have

$$\begin{aligned} &\int_{-\infty}^{\infty}F(s)\left(\int_0^1 1\left\{s\le c_n(x-B_t)\le 0\right\}dB_t\right)ds \\ &= \int_0^1\left(\int_{-\infty}^{\infty}F(s)1\left\{s\le c_n(x-B_t)\le 0\right\}ds\right)dB_t \qquad (22)\end{aligned}$$

by the Fubini's theorem for stochastic integrals (see, e.g., Karatzas and Shreve (1988), Problem 6.12, p 209). Finally, note that

$$\int_{-\infty}^{\infty}F(s)1\left\{s\le x\le 0\right\}ds=F_{\circ}(x) \qquad (23)$$

for $x<0$. Now the result in (20) follows easily from (19), due to (21)-(23). The proof for the case of $s>0$ is entirely analogous. In this case, we write

$$\begin{aligned} L\left(x-s/c_n\right)-L(x) &= -2\int_0^1 1\left\{x-s/c_n\le B_t\le x\right\}dB_t+O_{a.s.}(c_n^{-1}) \\ &= -2\int_0^1 1\left\{0\le c_n(x-B_t)\le s\right\}dB_t+O_{a.s.}(c_n^{-1})\end{aligned}$$

and note that

$$-\int_{-\infty}^{\infty}F(s)1\left\{0\le x\le s\right\}ds=F_{\circ}(x)$$

for $x\ge 0$.

We now prove the result in (b). The first part of (b) is shown in Phillips and Park (1998). For the second part, we apply the Ito formula to obtain

$$\int_0^1 F\left(c_n(x-B_t)\right)dB_t=c_n\int_0^1 F^{\circ}\left(c_n(x-B_t)\right)dt+O_{a.s.}(c_n^{-1}) \qquad (24)$$

and notice that

$$\begin{aligned} c_n\int_0^1 F^{\circ}\left(c_n(x-B_t)\right)dt &= c_n\int_{-\infty}^{\infty}F^{\circ}\left(c_n(x-t)\right)L(t)\,dt \\ &= \int_{-\infty}^{\infty}F^{\circ}(t)L(x-t/c_n)\,dt \qquad (25)\end{aligned}$$

due to the occupation times formula and the change of variables. The stated result now follows easily from (24) and (25), because $\int_{-\infty}^{\infty} F^\circ(t)dt = 0$ and

$$|L(x - t/c_n) - L(x)| \le c\,|t/c_n|^{1/2-\varepsilon}$$

for some constant c and arbitrary $\varepsilon > 0$, uniformly in $x \in \mathbf{R}$. Recall that we assume $\int_{-\infty}^{\infty} |t|\,|F^\circ(t)|\,dt < \infty$. The proof is therefore complete. ■

7.3 Proof of Lemma 3.6 Write

$$\begin{aligned}
\sum_{t=1}^{n} K\left(\frac{y - y_{t-1}}{h_n}\right) y_{t-1} &= y \sum_{t=1}^{n} K\left(\frac{y - y_{t-1}}{h_n}\right) - h_n \sum_{t=1}^{n} {}_iK\left(\frac{y - y_{t-1}}{h_n}\right) \\
\sum_{t=1}^{n} K^\circ\left(\frac{y - y_{t-1}}{h_n}\right) y_{t-1} &= y \sum_{t=1}^{n} K^\circ\left(\frac{y - y_{t-1}}{h_n}\right) - h_n \sum_{t=1}^{n} {}_iK^\circ\left(\frac{y - y_{t-1}}{h_n}\right)
\end{aligned}$$

and partition $P_n(y)$ as

$$P_n(y) = P_{1n}(y) + P_{2n}(y) + P_{3n}(y) + P_{4n}(y) \tag{26}$$

where

$$\begin{aligned}
P_{1n}(y) &= -h_n \sum_{t=1}^{n} {}_iK^\circ\left(\frac{y - y_{t-1}}{h_n}\right) \sum_{t=1}^{n} K\left(\frac{y - y_{t-1}}{h_n}\right) \\
P_{2n}(y) &= \sum_{t=1}^{n} K^\circ\left(\frac{y - y_{t-1}}{h_n}\right) u_t \sum_{t=1}^{n} K\left(\frac{y - y_{t-1}}{h_n}\right) \\
P_{3n}(y) &= -\sum_{t=1}^{n} K\left(\frac{y - y_{t-1}}{h_n}\right) u_t \sum_{t=1}^{n} K^\circ\left(\frac{y - y_{t-1}}{h_n}\right) \\
P_{4n}(y) &= h_n \sum_{t=1}^{n} {}_iK\left(\frac{y - y_{t-1}}{h_n}\right) \sum_{t=1}^{n} K^\circ\left(\frac{y - y_{t-1}}{h_n}\right)
\end{aligned}$$

Note that

$$\int_{-\infty}^{\infty} K(t)dt = 1, \int_{-\infty}^{\infty} iK^\circ(t)dt = -1, \int_{-\infty}^{\infty} K^\circ(t)dt = \int_{-\infty}^{\infty} {}_iK(t)dt = 0 \tag{27}$$

by Assumption 3.3. It follows from Lemmas 3.4(a) and 3.5(a) that

$$\begin{aligned}
\sum_{t=1}^{n} K\left(\frac{y - y_{t-1}}{h_n}\right) &= n^{1/2} h_n L(0) + O_{a.s.}(n^{1/4} h_n) \\
\sum_{t=1}^{n} {}_iK^\circ\left(\frac{y - y_{t-1}}{h_n}\right) &= -n^{1/2} h_n L(0) + O_{a.s.}(n^{1/4} h_n)
\end{aligned}$$

and that

$$\sum_{t=1}^{n} K^\circ\left(\frac{y - y_{t-1}}{h_n}\right), \sum_{t=1}^{n} {}_iK\left(\frac{y - y_{t-1}}{h_n}\right) = O_{a.s.}(n^{1/4} h_n)$$

Moreover, we have from Lemmas 3.4(b) and 3.5(b) that

$$\sum_{t=1}^{n} K\left(\frac{y-y_{t-1}}{h_n}\right)u_t,\ \sum_{t=1}^{n} K^{\circ}\left(\frac{y-y_{t-1}}{h_n}\right)u_t = O_{a.s.}(n^{1/4}h_n^{1/2})$$

We therefore have

$$\begin{aligned} P_{1n}(y) &= nh_n^3 L(0) + O_{a.s.}(n^{3/4}h_n^3) \\ P_{2n}(y) &= O_{a.s.}(n^{3/4}h_n^{3/2}) \\ P_{3n}(y) &= O_{a.s.}(n^{1/2}h_n^{3/2}) \\ P_{4n}(y) &= O_{a.s.}(n^{1/2}h_n^3) \end{aligned}$$

so that

$$P_n(y) = nh_n^3 L(0) + O_{a.s.}(n^{3/4}h_n^{3/2}) \tag{28}$$

Furthermore,

$$Q_n(y) = nh_n^3 L(0) + O_{a.s.}(n^{3/4}h_n^3) \tag{29}$$

The stated result for fixed y can then be easily deduced from (28) and (29).

Now consider the case with $y = \sqrt{n}x$. Define $y_t = \sqrt{n}x_t^n$ and

$$M_n(x, F) = \int_0^1 F\left(c_n(x - B_t)\right) dB_t$$

Then we have from Lemmas 3.4(a) and 3.5(a) that

$$\begin{aligned} \sum_{t=1}^{n} K\left(c_n(x - x_{t-1}^n)\right) &= n^{1/2}h_n\left[L(x) + 2M_n(x, K_{\circ})\right] + o_{a.s.}(n^{1/4}h_n^{3/2}) \\ \sum_{t=1}^{n} {}_iK^{\circ}\left(c_n(x - x_{t-1}^n)\right) &= n^{1/2}h_n\left[-L(x) + 2M_n(x, ({}_iK^{\circ})_{\circ})\right] + o_{a.s.}(n^{1/4}h_n^{3/2}) \\ \sum_{t=1}^{n} K^{\circ}\left(c_n(x - x_{t-1}^n)\right) &= 2n^{1/2}h_n M_n(x, K) + o_{a.s.}(n^{1/4}h_n^{3/2}) \\ \sum_{t=1}^{n} {}_iK\left(c_n(x - x_{t-1}^n)\right) &= 2n^{1/2}h_n M_n(x, ({}_iK)_{\circ}) + o_{a.s.}(n^{1/4}h_n^{3/2}) \end{aligned}$$

due to (27). Moreover, we have from Lemma 3.4(b) that

$$\begin{aligned} \sum_{t=1}^{n} K\left(c_n(x - x_{t-1}^n)\right)u_t &= n^{1/2}M_n(x, K) + o_{a.s.}(n^{1/4}h_n^{1/2}) \\ \sum_{t=1}^{n} K^{\circ}\left(c_n(x - x_{t-1}^n)\right)u_t &= n^{1/2}M_n(x, K^{\circ}) + o_{a.s.}(n^{1/4}h_n^{1/2}) \end{aligned}$$

If we partition $P_n(\sqrt{n}x)$ similarly as in (26), it follows from Lemma 3.5(b) that

$$\begin{aligned} P_{1n}(\sqrt{n}x) &= nh_n^3 L(x)^2 + o_{a.s.}(n^{3/4+\varepsilon}h_n^{5/2-\varepsilon}) \\ P_{2n}(\sqrt{n}x) &= nh_n L(x)M_n(x,K^\circ) + o_{a.s.}(n^{1/2+\varepsilon}h_n^{2-\varepsilon}) \\ P_{3n}(\sqrt{n}x) &= o_{a.s.}(n^{1/2+\varepsilon}h_n^{2-\varepsilon}) \\ P_{4n}(\sqrt{n}x) &= o_{a.s.}(n^{1/2+\varepsilon}h_n^{4-\varepsilon}) \end{aligned}$$

uniformly in x on any compact interval with arbitrary $\varepsilon > 0$. Therefore, we have

$$P_n(\sqrt{n}x) = nh_n^3 L(x)^2 + nh_n L(x)M_n(x,K^\circ) + o_{a.s.}(n^{3/4+\varepsilon}h_n^{5/2-\varepsilon}) \tag{30}$$

for $\varepsilon > 0$ arbitrary, uniformly in x on any compact interval. We also have

$$\begin{aligned} Q_n(\sqrt{n}x) &= nh_n^3 L(x)^2 + o_{a.s.}(n^{3/4+\varepsilon}h_n^{5/2-\varepsilon}) \\ &= nh_n^3 \left[L(x)^2 + o_{a.s.}(n^{-1/4+\varepsilon}h_n^{-1/2-\varepsilon}) \right] \end{aligned} \tag{31}$$

for $\varepsilon > 0$ arbitrary, uniformly in x on any compact interval. Consequently, it follows from (30) and (31) that

$$\begin{aligned} m_n^\circ(\sqrt{n}x) &= \frac{P_n(\sqrt{n}x)}{Q_n(\sqrt{n}x)} \\ &= 1 + h_n^{-2}\frac{M_n(x,K^\circ)}{L(x)} + o_{a.s.}(n^{-1/4+\varepsilon}h_n^{-1/2-\varepsilon}) \\ &= 1 + O_{a.s.}(n^{-1/4+\varepsilon}h_n^{-3/2-\varepsilon}) \end{aligned} \tag{32}$$

as was to be shown. ■

7.4 Proof of Theorem 3.8 The stated result follows immediately from the result in Lemma 3.6. The proof is therefore omitted. ■

7.5 Proof of Lemma 3.9 We write $a = B_{\max}$ and $b = B_{\min}$ to simply the notation. To prove the result stated in part (a), we first obtain

$$\int_0^1 F\left(c_n(B_t - B_s)\right) dB_s = c_n \int_0^1 F^\circ\left(c_n(B_t - B_s)\right) ds + O_{a.s.}(c_n^{-1}) \tag{33}$$

using the Ito formula. Now, we have

$$
\begin{aligned}
&\int_0^1 \frac{dt}{L(B_t)} \int_0^1 ds\, F^\circ\left(c_n(B_t - B_s)\right) \\
&= \int_{-\infty}^{\infty} dt\, 1\{a \le t \le b\} \int_{-\infty}^{\infty} ds\, F^\circ\left(c_n(t-s)\right) L(s) \\
&= c_n^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 1\{c_n(a-s) \le t \le c_n(b-s)\} F^\circ(t) L(s)\, ds dt \\
&= c_n^{-1} \int_{-\infty}^{\infty} ds\, L(s) \int_{-\infty}^{\infty} dt\, 1\{c_n(a-s) \le t \le c_n(b-s)\} F^\circ(t) \\
&= c_n^{-1} \int_{-\infty}^{\infty} L(s) \left[F\left(c_n(b-s)\right) - F\left(c_n(b-s)\right)\right] ds \\
&= c_n^{-2} \int_{-\infty}^{\infty} \left[L(b - s/c_n) - L(a - s/c_n)\right] F(s) ds \\
&= c_n^{-2} \int_{-\infty}^{\infty} \left[L(b) - L(a)\right] F(s) ds + o_{a.s.}(c_n^{-5/2+\varepsilon}) \qquad (34)
\end{aligned}
$$

by the successive applications of the occupation times formula and the changes of variables. Part (a) follows immediately from (33) and (34).

The result in part (b) can also be easily deduced, since

$$
\begin{aligned}
&c_n \int_0^1 \frac{dt}{L(B_t)^2} \int_0^1 ds\, F\left(c_n(B_t - B_s)\right) \\
&= c_n \int_{-\infty}^{\infty} \frac{dt}{L(t)} 1\{a \le t \le b\} \int_{-\infty}^{\infty} ds\, F\left(c_n(t-s)\right) L(s) \\
&= c_n \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{L(s)}{L(t)} 1\{a \le t \le b\} F\left(c_n(t-s)\right) dsdt \\
&= \int_{-\infty}^{\infty} dt\, 1\{a \le t \le b\} \int_{-\infty}^{\infty} ds\, \frac{L(t - s/c_n)}{L(t)} F(s) \\
&= \left(\int_{-\infty}^{\infty} F(s)\, ds\right) \left(\int_{-\infty}^{\infty} 1\{a \le t \le b\}\, dt\right) + O_{a.s.}(c_n^{-1/2}) \\
&= \left(\int_{-\infty}^{\infty} F(s)\, ds\right) (b-a) + O_{a.s.}(c_n^{-1/2})
\end{aligned}
$$

as was to be shown.

We now prove the result in (c). By the Fubini's theorem for stochastic integrals (see, e.g., Revuz and Yor (1994, Exercise 5.17, pp 167-168)), we may consider

$$
\int_0^1 \int_0^s \left[\int_0^1 \frac{dt}{L(B_t)^2} F\left(c_n(B_t - B_s)\right) G\left(c_n(B_t - B_r)\right)\right] dB_r dB_s \qquad (35)
$$

We have

$$\begin{aligned}
& c_n \int_0^1 \frac{dt}{L(B_t)^2} F\left(c_n(B_t - B_s)\right) G\left(c_n(B_t - B_r)\right) \\
&= c_n \int_{-\infty}^{\infty} \frac{dt}{L(t)} 1\left\{a \le t \le b\right\} F\left(c_n(t - B_s)\right) G\left(c_n(t - B_r)\right) \\
&= \int_{-\infty}^{\infty} \frac{dt}{L(B_s + t/c_n)} F(t) G\left(t + c_n(B_s - B_r)\right) \\
&= \frac{1}{L(B_s)} H\left(c_n(B_s - B_r)\right) + R_n \qquad (36)
\end{aligned}$$

where

$$H(x) = \int_{-\infty}^{\infty} F(t)G(t + x)dt$$

Note that H is in particular bounded and integrable. Since we have for large n

$$\left| \frac{1}{L(B_s + t/c_n)} - \frac{1}{L(B_s)} \right| \le c \, \frac{|t/c_n|^{1/2-\varepsilon}}{\inf_{s\in[0,1]} L(B_s)^2}$$

for some constant c and arbitrary $\varepsilon > 0$, the effect of R_n becomes negligible in the subsequent computation.

Now we consider

$$M_n = c_n^{1/2} \int_0^1 \frac{1}{L(B_s)} \int_0^s H\left(c_n(B_s - B_r)\right) dB_r dB_s \qquad (37)$$

Two stochastic integrals are involved in the definition of M_n in (37). The first stochastic integral with respect to dB_r is in the filtration $\mathcal{F}_r^{(s)} = \sigma\left((B_q)_{q\le r}, B_s\right)$. It needs a little more elaboration to define the second integral with respect to dB_s, since the integrand includes $L(B_s) \equiv L(1, B_s)$. It requires the knowledge of (B_t) not just up to time s, but up to time 1. To properly interpret the stochastic integral with respect to dB_s in (37), we introduce the following lemma.

Lemma A1 We define an extended σ-field $\mathcal{G}_t$ by

$$\mathcal{G}_t = \sigma\left((B_s, L(B_s)_{s\le t}\right)$$

Then $(B_t, \mathcal{G}_t)$ is a martingale. The process (B_t) is indeed $(\mathcal{G}_t)$-Brownian motion.

Proof of Lemma A1 Fix $t \in [0, 1]$, and write

$$L(1, B_s) = L(t, B_t) + [L(1, B_s) - L(t, B_s)]$$

for $s \in [0, t]$. Note that $L(t, B_s)_{s\le t}$ is $\mathcal{F}_t$-measurable, where $\mathcal{F}_t = \sigma\left((B_s)_{s\le t}\right)$ is the natural filtration of (B_t). Moreover, conditional on $\mathcal{F}_t$,

$$\sigma\left(L(r, B_s)_{r\ge t}\right) = \sigma\left(|B_r - B_t|_{r\ge t}\right)$$

for all $s \in [0, t]$. It therefore follows that

$$\mathcal{G}_t = \sigma\Big((B_s)_{s\le t}, |B_r - B_t|_{r\ge t}\Big)$$

See, e.g., Revuz and Yor (1994, Corollary 2.2, p222).

Now we may easily see that $(B_t, \mathcal{G}_t)$ is a martingale. For any s and t such that $s < t$, we have

$$\mathbf{E}(B_t|\mathcal{G}_s) = B_s + \mathbf{E}(B_t - B_s|\mathcal{G}_s) = B_s$$

Observe that, conditional on $\mathcal{G}_s$, $B_t - B_s$ is symmetrically distributed with point probability mass 1/2 at each of $\pm|B_t - B_s|$. Since $(B_t, \mathcal{G}_t)$ is a continuous martingale with quadratic variation (t), (B_t) is trivially $(\mathcal{G}_t)$-Brownian motion, due to the Levy's characterization theorem. ■

Now we go back to the proof of Lemma 3.9. Due to Lemma A1, the stochastic integral with respect to dB_s in (37) can be well defined in the filtration $(\mathcal{G}_s)$. It therefore follows that

$$\mathbf{E}M_n^2 = \mathbf{E}\left(c_n \int_0^1 \frac{ds}{L(B_s)^2} \int_0^s dr\, H\left(c_n(B_s - B_r)\right)^2\right) \tag{38}$$

However, we have

$$\begin{aligned} & c_n \int_0^1 \frac{ds}{L(B_s)^2} \int_0^s dr\, H\left(c_n(B_s - B_r)\right)^2 \\ &= \int_{-\infty}^{\infty} \frac{ds}{L(s)} 1\{a \le s \le b\} \int_{-\infty}^{\infty} dr\, L(s,r) H\left(c_n(s-r)\right)^2 \\ &= \int_{-\infty}^{\infty} \frac{ds}{L(s)} 1\{a \le s \le b\} \int_{-\infty}^{\infty} dr\, L(s, s - r/c_n) H(r)^2 \\ &= \left(\int_a^b \frac{L(s,s)}{L(s)} ds\right)\left(\int_{-\infty}^{\infty} H(r)^2 dr\right) + o_{a.s.}(c_n^{-1/2+\varepsilon}) \end{aligned} \tag{39}$$

The second equality, in particular, follows from the change of variables $(r, s) \mapsto (c_n(s - r), s)$. The third equality is due to the uniform Holder continuity of L, i.e.,

$$|L(s, s - r/c_n) - L(s,s)| \le c\,|r/c_n|^{1/2-\varepsilon}$$

uniformly in $s \in [0, 1]$, for some constant c and arbitrary $\varepsilon > 0$. We now have from (38) and (39) that $M_n = O_p(1)$, and the stated result in (c) follows from (35)-(37).■

7.6 Proof of Lemma 3.10 We have from (32) in the proof of Lemma 3.6 that

$$\log\left|m_n^\circ(\sqrt{n}x)\right| = h_n^{-2}\frac{M_n(x, K^\circ)}{L(x)} - \frac{h_n^{-4}}{2}\left(\frac{M_n(x, K^\circ)}{L(x)}\right)^2 + o_{a.s.}(n^{-1/2+\varepsilon} h_n^{-2-\varepsilon})$$

for arbitrary $\varepsilon > 0$, uniformly in x on any compact interval. Therefore, we have

$$\begin{aligned}
\lambda_n &= \frac{1}{n}\sum_{t=1}^{n} \log\left|m_n^\circ(\sqrt{n}x_{t-1}^n)\right| \\
&= h_n^{-2}\frac{1}{n}\sum_{t=1}^{n}\frac{M_n(x_{t-1}^n, K^\circ)}{L(x_{t-1}^n)} - \frac{h_n^{-4}}{2}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{M_n(x_{t-1}^n, K^\circ)}{L(x_{t-1}^n)}\right)^2 + o_{a.s.}(n^{-1/2+\varepsilon}h_n^{-2-\varepsilon}) \\
&= h_n^{-2}\int_0^1 \frac{M_n(B_t^n, K^\circ)}{L(B_t^n)}dt - \frac{h_n^{-4}}{2}\int_0^1\left(\frac{M_n(B_t^n, K^\circ)}{L(B_t^n)}\right)^2 dt + o_{a.s.}(n^{-1/2+\varepsilon}h_n^{-2-\varepsilon}) \\
&= h_n^{-2}\int_0^1 \frac{M_n(B_t, K^\circ)}{L(B_t)}dt - \frac{h_n^{-4}}{2}\int_0^1\left(\frac{M_n(B_t, K^\circ)}{L(B_t)}\right)^2 dt + o_{a.s.}(n^{-1/2}h_n^{-3/2})
\end{aligned}$$

However, by Lemma 3.9(a), we have

$$\int_0^1 \frac{M_n(B_t, K^\circ)}{L(B_t)}dt = O_{a.s.}(n^{-1/2}h_n)$$

Moreover,

$$\begin{aligned}
&\int_0^1\left(\frac{M_n(B_t, K^\circ)}{L(B_t)}\right)^2 dt \\
&\quad= \int_0^1 \frac{dt}{L(B_t)^2}\int_0^1 ds\, K^\circ\left(c_n(B_t - B_s)\right)^2 \\
&\qquad + 2\int_0^1 \frac{dt}{L(B_t)^2}\int_0^1 K^\circ\left(c_n(B_t - B_s)\right)\int_0^s K^\circ\left(c_n(B_t - B_r)\right) dB_r dB_s
\end{aligned}$$

and the stated result follows directly from Lemma 3.9(b) and (c). ■

7.7 Proof of Theorem 3.12 We have $\sigma_n^2 \to_{a.s.} \sigma^2$, and therefore,

$$T_n = \frac{2n^{1/2}h_n^3}{\kappa^2\sigma_n}\lambda_n \to_d W_{\min} - W_{\max}$$

The density of

$$T = W_{\min} - W_{\max}$$

can easily be obtained from the well known result by Levy (see e.g. Hida (1980, Proposition 2.10, pp85-86)) as follows. As shown in Billingsley (1968, p79), we have

$$\mathbf{P}\left\{a < W_{\min} \le W_{\max} < b\right\} = \sum_{k=-\infty}^{\infty}(-1)^k\left[{}_\circ\varphi(b + k(b-a)) - {}_\circ\varphi(a + k(b-a))\right]$$

where φ is the standard normal density function. The joint density $f(a,b)$ of $(W_{\min}, W_{\max})$ can now be easily obtained, since

$$\begin{aligned}
f(a,b) &= -\frac{\partial^2}{\partial a\partial b}\mathbf{P}\left\{a < W_{\min} \le W_{\max} < b\right\} \\
&= \sum_{k=-\infty}^{\infty}(-1)^k k\left[(k+1)\varphi^\circ(b + k(b-a)) - (k-1)\varphi^\circ(a + k(b-a))\right]
\end{aligned}$$

over the support $\{(a,b)|a \leq 0, b \geq 0\}$, where the interchange of differentiation and infinite summation is justified due to the absolute convergence of the series functions. To obtain the given marginal density of $T = W_{\min} - W_{\max}$, make the change of variables $x = a-b$ and $y = b$ and get the joint density $f(x,y)$ of $(W_{\min}-W_{\max}, W_{\max})$, and then integrate it out with respect to y. ■

7.8 Proof of Theorem 4.6 Since we have $\sigma_n^2 \to_p \sigma^2 = \mathbf{E}\Delta y_t^2 < \infty$ by the weak law of large numbers for the strong mixing random variables and $n^{1/2}h_n^3 \to \infty$ by Assumption 4.4, it suffices to establish the following consistency result

$$\lambda_n \to_p \lambda > 0$$

We need the following lemma to establish consistency of our test:

Lemma A2 Suppose Assumptions 4.1-4.2 and 4.4 hold. Then, we have

$$\sup_{y\in\mathbf{R}} |m_n^\circ(y) - m^\circ(y)| = O_p(n^{-1/2}h_n^{-2}) + O_p(h_n^{q-1}).$$

Proof of Lemma A2 The result of Lemma A2 follows directly from Theorem 1 of Andrews (1995) because Assumptions NP1-NP5 of the latter paper are implied by our Assumptions 4.1-4.2 and 4.4 with $\eta = \infty$, $|\omega| = q$, $|\lambda| = 1, \sigma_{1T} = \sigma_{2T} = n^{-\nu}$ and (Y_{Tt}, X_{Tt}), f_{Tt}, g_T, and $\hat{\sigma}_T$ given by (y_t, y_{t-1}), f, m, and h_n respectively. ■

To prove the consistency of our test, we first note that

$$\sup_{y\in\mathbf{R}} |m_n^\circ(y) - m^\circ(y)| = O_p(n^{-1/4}) \tag{40}$$

which holds due to Assumption 4.4 and Lemma A2. Consequently, it follows from (40) and Assumption 4.3(a) that

$$\left(\min_{1\leq t\leq n} |m_n^\circ(y_t)|\right)^{-1} = O_p(n^{r+1/4}) \tag{41}$$

Moreover, if we let (ε_n) be a sequence of the truncation numbers given by

$$\varepsilon_n = n^{-\delta} \tag{42}$$

with $0 < \delta < 1/4$, then we have by Assumption 4.3(b) and (40)

$$\mathbf{P}\left\{|m_n^\circ(y_t)| \leq \varepsilon_n\right\} \leq \mathbf{P}\left\{|m^\circ(y_t)| \leq \varepsilon_n + n^{-1/4}\right\} \leq \varepsilon_n + n^{-1/4} \tag{43}$$

for n sufficiently large.

We now let

$$R_n = \left|\frac{1}{n}\sum_{t=1}^{n} \log|m_n^\circ(y_t)| - \frac{1}{n}\sum_{t=1}^{n} \log|m^\circ(y_t)|\right|$$

and show that $R_n = o_p(1)$. Notice that

$$R_n \leq A_n + B_n + C_n$$

where

$$\begin{aligned} A_n &= \left| \frac{1}{n} \sum_{t=1}^{n} \log |m_n^\circ(y_t)| 1\{|m_n^\circ(y_t)| > \varepsilon_n\} \right. \\ &\quad \left. - \frac{1}{n} \sum_{t=1}^{n} \log |m^\circ(y_t)| 1\{|m^\circ(y_t)| > \varepsilon_n\} \right| \\ B_n &= \left| \frac{1}{n} \sum_{t=1}^{n} \log |m^\circ(y_t)| 1\{|m^\circ(y_t)| \leq \varepsilon_n\} \right| \\ C_n &= \left| \frac{1}{n} \sum_{t=1}^{n} \log |m_n^\circ(y_t)| 1\{|m_n^\circ(y_t)| \leq \varepsilon_n\} \right| \end{aligned}$$

We first consider B_n. We have

$$\begin{aligned} |B_n| &\leq O_p(\log n^r) \frac{1}{n} \sum_{t=1}^{n} 1\{|m^\circ(y_t)| \leq \varepsilon_n\} \\ &= O_p(\log n) O_p(\varepsilon_n) \to_p 0 \end{aligned} \quad (44)$$

where the inequality holds by Assumption 4.3(a), the equality holds by Assumption 4.3(b) and the last convergence to zero holds by the definition of ε_n in (42). Similarly, C_n is $o_p(1)$ because we have

$$\begin{aligned} |C_n| &\leq O_p(\log n^{r+1/4}) \frac{1}{n} \sum_{t=1}^{n} 1\{|m_n^\circ(y_t)| \leq \varepsilon_n\} \\ &= O_p(\log n) O_p(\varepsilon_n + n^{-1/4}) \to_p 0 \end{aligned} \quad (45)$$

using (41) – (43).

It now remains to show that

$$A_n = o_p(1) \quad (46)$$

which we set out to do. We first note that

$$A_n \leq A_{1n} + A_{2n} + A_{3n}$$

where

$$\begin{aligned} A_{1n} &= \frac{1}{n} \sum_{t=1}^{n} \Big| \log |m_n^\circ(y_t)| - \log |m^\circ(y_t)| \Big| 1\{|m^\circ(y_t)| > \varepsilon_n\} 1\{|m_n^\circ(y_t)| > \varepsilon_n\} \\ A_{2n} &= \frac{1}{n} \sum_{t=1}^{n} \Big| \log |m_n^\circ(y_t)| \Big| 1\{|m_n^\circ(y_t)| > \varepsilon_n\} 1\{|m^\circ(y_t)| \leq \varepsilon_n\} \\ A_{3n} &= \frac{1}{n} \sum_{t=1}^{n} \Big| \log |m^\circ(y_t)| \Big| 1\{|m^\circ(y_t)| > \varepsilon_n\} 1\{|m_n^\circ(y_t)| \leq \varepsilon_n\} \end{aligned}$$

However, due to (40), we have

$$A_{1n} \leq \frac{1}{\varepsilon_n} \sup_{y \in \mathbf{R}} |m_n^\circ(y) - m^\circ(y)| = O_p(n^{-1/4}\varepsilon_n^{-1}) \to_p 0$$

Also, similarly in (44) and (45), we have

$$\begin{aligned} A_{2n} \; &\leq \; O_p(\log n^{r+1/4}) \frac{1}{n} \sum_{t=1}^{n} 1\{|m^\circ(y_t)| \leq \varepsilon_n\} = O_p(\varepsilon_n \log n) \to_p 0 \\ A_{3n} \; &\leq \; O_p(\log n^{r}) \frac{1}{n} \sum_{t=1}^{n} 1\{|m_n^\circ(y_t)| \leq \varepsilon_n\} = O_p((\varepsilon_n + n^{-1/4}) \log n) \to_p 0 \end{aligned}$$

and hence the required result (46) holds. ■

References

Andrews, D. W. K. (1995). "Nonparametric kernel estimation for semiparametric models," *Econometric Theory*, 11, 560-596.

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

Brock, W. A. (1986). "Distinguishing random and deterministic systems," *Journal of Economic Theory*, 40, 168-195.

Cheng, B. and H. Tong (1992). "On consistent nonparametric order determination and chaos," *Journal of the Royal Statistical Society, Series B*, 54, 427-449.

Choi, C.-Y. and Y.-K. Moh (2007). "How useful are tests for unit-root in distinguishing unit-root processes from stationary but nonlinear processes?," *Econometrics Journal*, 10, 82-112.

Chung, K.L. and R.J. Williams (1990). *Introduction to Stochastic Integration*, 2nd ed. Birkhäuser, Boston.

Dechert, W. D. and R. Gencay (1992). "Lyapunov exponent as a nonparametric diagnostic for stability analysis," *Journal of Applied Econometrics,* 7, S41-S60.

Diva, B.T. and H.I. Grossman (1988). "Explosive rational bubbles in stock prices?," *American Economic Review*, 78, 520-530.

Fernández-Rodríguez, Sosvilla-Rivero and Andrada-Félix (2005). "Testing chaotic dynamics via Lyapunov exponents," *Journal of Applied Econometrics*, 20, 911-930.

Gencay (1996). "A statistical framework for testing chaotic dynamics via Lyapunov exponents," *Physica D*, 89, 261-266.

Gnedenko, B. V. (1943), "Sur la distribution limite du term maximum d'une série aléatoire," *Annals of Mathematical Statistics*, 44, 423-453.

Hall, P. and C.C. Heyde (1980). *Martingale Limit Theory and Its Application.* Academic Press, New York.

Hida, T. (1980). *Brownian Motion.* Springer-Verlag, New York.

Jeganathan, P. (2004). "Convergence of functionals of sums of r.v.'s to local times of fractional stable motions," *Annals of Probability*, 32, 1771-1795.

Karatzas, I. and S.E. Shreve (1988). *Brownian Motion and Stochastic Calculus.* Springer–Verlag, New York.

McCaffrey, D., S. Ellner, A. R. Gallant, and D. Nychka (1992). "Estimating the Lyapunov exponent of a chaotic system with nonparametric regression," *Journal of the American Statistical Association*, 87, 682-695.

Nychka, D., S. Ellner, A. R. Gallant, and D. McCaffrey (1992). "Finding chaos in noisy systems," *Journal of the Royal Statistical Society, Series B*, 54, 399-426.

Park, J. Y. and P. C. B. Phillips. (1999). "Asymptotics for nonlinear transformations of integrated time series," *Econometric Theory*, 15, 269-298.

Phillips, P.C.B. and J.Y. Park (1998). "Nonstationary density estimation and kernel autoregression," *Cowles Foundation Discussion Paper*, No. 1181, Yale University.

Phillips, P. C. B. (1997). "Unit root tests." In *Encyclopedia of Statistical Sciences*, Vol. 12, Wiley: New York.

Revuz, D. and M. Yor (1994). *Continuous Martingale and Brownian Motion*, 2nd ed. Springer–Verlag, New York.

Rosser, J. B., Jr. (1999). "On the complexity of complex economic dynamics," *Journal of Economic Perspectives*, 13, 169-192.

Shintani, M. and O. Linton (2003). "Is there chaos in the world economy? A nonparametric test using consistent standard errors," *International Economic Review*, 44, 331-358.

Shintani, M. and O. Linton (2004). "Nonparametric neural network estimation of Lyapunov exponents and a direct test for chaos," *Journal of Econometrics*, 120, 1-33.

Tong, H. (1990). *Nonlinear Time Series: A Dynamical Systems Approach*. Clarendon Press, Oxford.

Whang, Y.-J. and O. Linton (1999). "The asymptotic distribution of nonparametric estimates of the Lyapunov exponent for stochastic time series," *Journal of Econometrics*, 91, 1-42.

Wolff, R.C., Q. Yao and H. Tong (2004). "Statistical tests for Lyapunov exponents of deterministic systems," *Studies in Nonlinear Dynamics and Econometrics* 8, Article 10.