Strong Approximations for Nonlinear Transformations of Integrated Time Series

Joon Y. Park
Department of Economics
Rice University

Abstract

In this paper we establish the strong approximations for the nonlinear transformations of integrated time series. Both the asymptotically homogeneous and integrable transformations are considered, and the explicit rates for the convergence to their limit distributions are obtained under mild regularity conditions that are satisfied by virtually all nonlinear models used in practical applications. The first order asymptotics are also derived under the conditions that are significantly weaker than those required by earlier works.

First Draft: July 25, 2001
This version: January 21, 2003

Key words and phrases: integrated process, nonlinear transformation, strong approximation, invariance principle, Brownian motion, local time

1. Introduction

The asymptotics for the nonlinear transformations of integrated time series were developed earlier by Park and Phillips (1999) (which will be referred to as PP henceafter). Their work laid a ground for many of the subsequent researches on the nonlinear models involving integrated time series. They include Park and Phillips (2001) and Chang, Park and Phillips (2001) for nonlinear regressions, Park and Phillips (2000) for binary choice models, Chang and Park (2001) for index models, and Chang (2002) and Phillips, Park and Chang (2001) for the unit root tests using nonlinear IV tests. Given that many of the important macroeconomic and financial time series exhibit nonstationary characteristics, this line of research would certainly be very helpful in performing inference based on nonlinear econometric models. Recently, de Jong (2002a, 2002b) extended some of the results in PP.

In this paper, we consider the strong approximations for the nonlinear transformations of integrated time series. For the standardized sample moments of the nonlinear transformations of integrated time series, we provide the exact rates of convergence to their limiting distributions. Our results here may thus be seen as the refinements of those in PP. The transformations we investigate in this paper include the asymptotically homogeneous functions and integrable functions, and we assume that the integrated time series are generated by the linear processes with iid innovations. These assumptions are largely comparable to those in PP. However, we impose much less stringent regularity conditions for the transformation functions compared with those used in PP. In particular, our regulaity conditions for the asymptotically homogeneous functions explicitly allow for the logarithms and reciprocals that have poles at the origin. Moreover, our first order asymptotics for the integrable transformations do not require any regularity condition on the transformation function except for Riemann-integrability.

The strong approximations developed in this paper have some important potential applications. First, they make it possible to establish the asymptotic theories for the nonparametric and semiparametric methods used to analyze the models with integrated time series. The methods often require the dimension of the model increase as the sample size, such as in the series estimation like wavelets, and in such cases the strong approximations show us how fast we may expand the model without changing the first order asymptotics. Second, they can be used to establish the theory of bootstrap refinements for the nonlinear models with integrated time series. For the refiement theory of bootstrap, it is necessary to consider the lower order terms and on this regard our strong approximations provide some useful informations. Third, they are useful to statistically analyze the models involving time series with asymptotic unit roots, i.e., roots approaching unity as the sample size increases. The asymptotic analysis of the nonlinear models with such time series require the strong approximations that we develop in this paper.

The rest of the paper is orginized as follows. In Section 2, we introduce the assumptions, and some preliminary notions and results that are necessary to develop and interpret our subsequent developments. The main results on the asymptotics for the nonlinear transformations of integrated time series are given in Section 3. There both the fist order asymptotics and the strong approximations for the asymptotically homogeneous and the integrable transformations are presented. The concluding remark follows in Section 4. All the mathematical proofs of the theorems in the paper are presented in Section 5. A word on notation. The standard notations such as a.s., $\to_p$, $\to_d$, o_p and O_p are used frequently without any specific reference. As usual, $\mathbf{R}$ denotes the real line and $=_d$ signifies the distributional equivalence. Throughout the paper, the integral and integrability of a transformation on $\mathbf{R}$ are interpreted to be the Riemann-integral and Riemann-integrability, respectively.

2. Assumptions and Preliminaries

Consider an integrated time series (x_t) generated by

$$x_t = x_{t-1} + w_t \tag{1}$$

where we set $x_0 = O_p(1)$ and assume that (w_t) satisfies an invariance principle. More specifically, we construct the stochastic process W_n by

$$W_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} w_t \tag{2}$$

and assume

2.1 Assumption $W_n \to_d W$ as $n \to \infty$.

where W is the standard Brownian motion. Throughout the paper, we let the longrun variance of (w_t) be unity, i.e., $\lim_{n\to\infty} \mathbf{E}\left(\sum_{t=1}^n w_t\right)^2/n = 1$ as $n \to \infty$. This is just for simplicity in exposition. The value of the longrun variance of (w_t) only has an unimportant scaling effect in our subsequent analyses. It is well known that the invariance principle in Assumption 2.1 holds for a wide class of weakly dependent and possibly heterogeneous processes.

To obtain the strong approximations for the nonlinear transformations of integrated time series, we need more explicitly data generating processes. For this purpose, we let (w_t) in (1) follow the linear process

$$w_t = \varphi(L)\varepsilon_t = \sum_{k=0}^{\infty} \varphi_k \varepsilon_{t-k} \tag{3}$$

where (ε_t) is a sequence of i.i.d. random variables with mean zero, and $\varphi(1) \neq 0$. For (w_t) generated by (3), the following assumptions will be made frequently in the paper.

2.2 Assumption $\sum_{k=0}^{\infty} k|\varphi_k| < \infty$ and $\mathbf{E}|\varepsilon_t|^p < \infty$ for some $p > 2$.

2.3 Assumption The distribution of (ε_t) is absolutely continuous with respect to the Lebesque measure, and has characteristic function ϕ for which $\lim_{t\to\infty} t^\delta \phi(t) = 0$ for some $\delta > 0$.

As mentioned above, we let $\varphi(1) = 1$ and $\mathbf{E}\,\varepsilon_t^2 = 1$ so that the longrun variance of (w_t) becomes unity.

The stochastic process W_n defined in (2) takes values in $D[0,1]$, the set of cadlag functions on the interval $[0,1]$. Invoking an invariance principle in Assumption 2.1, we have $W_n \to_d W$, which is to be interpreted as the weak convergence in the probability measures on $D[0,1]$. In our context, it is more convenient, and so is assumed, to endow $D[0,1]$ with the uniform topology rather than the usual Skorohod topology [see Billingsley (1968), pp. 150-152]. By virtue of the so-called Skorohod representation theorem [e.g., Pollard(1984), pp. 71-72], it is indeed possible to construct W_n and W on a common probabiity space, up to the distributional equivalence, so that $W_n \to_{a.s.} W$ uniformly on $[0,1]$. Moreover, using strong approximation methods in Csörgő and Horváth (1993, p4) and Akonom (1993), specific rate for the convergence for W_n to W can be obtained under Assumption 2.2.

2.4 Lemma Under Assumption 2.2,

$$\sup_{0\le r\le 1} |W_n(r) - W(r)| = o_p\left(n^{-1/2+1/p}\right)$$

as $n \to \infty$.

For our development of the strong approximations, we assume that the stochastic process W_n in (2) is defined up to the distributional equivalence, rather than directly from the partial sum of (w_t), so that the strong invariance principle in Lemma 2.4 holds.

In this paper, we consider the strong approximations of the sample moments of nonlinear transformations of (x_t). In particular, we let T be a transformation on $\mathbf{R}$ and derive an explicit order for the convergence of

$$\sum_{t=1}^{n} T(x_t)$$

to its limiting distribution, after an appropriate standardization is made. The limiting distributions of such sample moments were obtained earlier by Park and Phillips (1999), which studied the first order asymptotics for the nonlinear transformations of integrated time series. For the transformation function T, we look at integrable and asymptotically homogeneous functions. These are the transformation functions that were investigated by Park and Phillips (1999). Here and elsewhere in the paper, we assume T is real-valued. This is just to simplify the exposition. All our subsequent results hold also for vector-valued functions, only with some trivial notational modifications.

We first define

2.5 Definition Write

$$T(\lambda x) = \kappa(\lambda)S(x) + R(x, \lambda)$$

We say that $T : \mathbf{R} \to \mathbf{R}$ is *asymptotically homogeneous* if
(a) S is locally integrable, and
(b) $|R(x, \lambda)| \le c\,\nu(\lambda)\, Q(x)$ for all λ sufficiently large and for all x over any compact set K, where c is a constant which may depend upon K, $(\kappa^{-1}\nu)(\lambda) \to 0$ as $\lambda \to \infty$, and Q is locally integrable.
We call κ the *asymptotic order* and S the *limit homogeneous function* of T.

2.6 Remark (a) The conditions for the asymptotically homogeneous functions in Definition 2.5 are significantly weaker than those used by Park and Phillips (1999). In particular, the logarithmic function $T(x) = \log|x|$ and the power function $T(x) = |x|^k$ with $-1 < k < 0$, which have poles at the origin, are allowed here. They are satisfied for all the asymptotically homogeneous functions considered in Park and Phillips (2001).

(b) Many of the transformations used in practical nonlinear analyses can be written as the sums of asymptotically homogeneous and integrable transformations. The examples include the shifted integrable or distribution function-like transformations such as $T(x) = 1 + e^{-x^2}$ and $T(x) = e^x/(1 + e^x)$. For such transformations, we may of course apply our results for the asymptotically homogeneous and integrable transformations term by term. For the first order asymptotics, the effects of the integrable transformations become negligible compared to those of the asymptotically homogeneous transformations. The integrable components in the transformations can therefore be ignored to obtain the first order asymptotics. This will be explained in detail later with some concrete examples. See the part (c) of Remark 3.4.

A function is said to be locally integrable if it is integrable over any compact subset of $\mathbf{R}$. To obtain the strong approximations for asymptotically homogeneous transformations, we need to consider the class of locally integrable functions that satisfy a certain set of regularity conditions.

2.7 Definition We say that $T : \mathbf{R}\backslash\{0\} \to \mathbf{R}$ is *regularly locally integrable* if it is locally integrable and, for any $\epsilon > 0$ sufficiently small, it satisfies:
(a) for $|x| < \epsilon$, $T(x)$ is dominated by a constant multiple of $|x|^a$ for some $a > -1$, and
(b) for $|x| \geq \epsilon$, $T(x)$ is locally Lipschitz with Lipschitz constant bounded by a constant multiple of $(1 + \epsilon^b)$ for some b.

2.8 Remark (a) The regularly locally integrable functions are allowed to have pole-type discontinuities at the origin. Of course, we may similarly consider the functions that have the same type of discontinuities elsewhere. We consider the functions having discontinuities at the origin here, since they are the ones that appear frequently in our analysis of the asymptotically homogeneous transformations.

(b) The power functions $T(x) = |x|^k$ are regularly locally integrable as long as $k > -1$. For such functions, we have $a = k$ and $b = k - 1$. For $k > -1$, the power-log functions $T(x) = |x|^k \log |x|$ are also regularly locally integrable with $a = k + \epsilon$ and $b = k - 1 + \epsilon$ for any $\epsilon > 0$.

For our developments of the strong approximations of integrable transformations, we need to define

2.9 Definition We say that $T : \mathbf{R} \to \mathbf{R}$ is *regularly integrable* if
(a) T is piecewise Lipschitz, and
(b) $\int_{-\infty}^{\infty} |x|^q |T(x)| dx < \infty$ for some $q \geq (p-2)/6$ for p introduced in Assumption 2.1.

For an integrable transformation, we require that the tail decrease at a faster rate as $p \to \infty$, i.e., as higher moments exist for the innovations. Our approximation becomes more precise as $p \to \infty$, and, for such an improved approximaion, the tail should vanish accordingly at a faster rate.

Our subsequent asymptotics relies on the local time of the limit Brownian motion W. The local time L of W is defined as

$$L(t, s) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t 1\{|W(r) - s| < \epsilon\} \, dr$$

The reader is referred to Chung and Williams (1990) for an introduction to the local time. Roughly, it denotes the time spent by the process in the vicinity of a spatial point s over the time interval $[0,t]$. It is well known that the Brownian local time L is continuous a.s. with respect to both parameters t and s. The local time yields the occupation time formula

$$\int_0^t T(W(r))\,dr = \int_{-\infty}^{\infty} T(s)L(t,s)\,ds$$

for any locally integrable $T:\mathbf{R}\to\mathbf{R}$. The formula allows us to represent the integral of any locally integrable transformation of Brownian motion as the integral of the function itself spatially weighted by the local time.

3. Main Results

3.1 Asymptotics for Asymptotically Homogeneous Transformations

To obtain the strong approximations for the asymptotically homogeneous transformations, we first note that

3.1 Lemma Let Assumption 2.1 hold, and let $T:\mathbf{R}\to\mathbf{R}$ be locally integrable. Then

$$\frac{1}{n}\sum_{t=1}^{n} T\left(\frac{x_t}{\sqrt{n}}\right) \to_d \int_0^1 T(W(r))\,dr$$

as $n\to\infty$.

Moreover, if we let

$$\Delta_n = \sup_{0\le r\le 1} |W_n(r) - W(r)| \tag{4}$$

so that $\Delta_n = o_p(n^{-1/2+1/p})$, due to Lemma 2.3, then we have

3.2 Theorem Let Assumption 2.2 hold, and let $T:\mathbf{R}\to\mathbf{R}$ be strongly locally integrable. Then

$$\int_0^1 T(W_n(r))\,dr = \int_0^1 T(W(r))\,dr + O_p\left(c_n^{1+a}\right) + O_p\Big((1+c_n^b)\Delta_n\Big)$$

as $n\to\infty$, for any sequence (c_n) such that $c_n \ge \Delta_n$ a.s. and $c_n \to_p 0$.

The actual order of magnitude for the approximation error involved in the result of Theorem 3.2 depends upon the values of a and b. For $a\ge 0$ and $b\ge 0$, we may choose $c_n=\Delta_n$ to show that the error is of order $O_p(\Delta_n)$. If $-1<a<0$ and $a\le b$, then the same choice of c_n gives the optimal rate $O_p(\Delta_n^{1+a})$. Finally, when $a>-1$, $a>b$ and $b<0$, the optimal choice of c_n reduces to $c_n=\Delta_n^{1/(1+a-b)}$, which yields the error of order $O_p(\Delta_n^{(1+a)/(1+a-b)})$.

3.3 Remark (a) For $T(x) = |x|^k$ with $k > -1$, we have $a = k$ and $b = k - 1$ as noted earlier. If $k \geq 1$, we may choose $c_n = \Delta_n$ to deduce that the approximation error is of order $o_p(n^{-1/2+1/p})$. For the case of $-1 < k < 1$, the optimal choice of c_n becomes $c_n = \Delta_n^{1/2}$, in which case the order of the approximation error reduces to $o_p(n^{-(p-2)(k+1)/4p})$. Note that it approaches to $o_p(1)$ as $k \to -1$.

(b) For $T(x) = \log|x|$, a can be any positive number arbitrarily small and $b = -1$. The optimal choice of c_n in this case is given by $c_n = \Delta_n^{1/2-\epsilon}$ for $\epsilon > 0$ arbitrarily small. The resulting approximation error becomes $o_p(n^{-1/2+1/p+\epsilon})$ for an abitrarily small $\epsilon > 0$.

(c) For $T(x) = 1\{x \geq 0\}$, we have $a = 0$ and $b = \infty$. In this case, we may choose $c_n = \Delta_n$ so that the approximation error becomes of order $o_p(n^{-1/2+1/p})$.

The asymptotics for the asymptotically homogeneous transformations now follow readily from Lemma 3.1 and Theorem 3.2. To obtain the first order asymptotics, we simply observe that

$$\frac{1}{n\kappa(\sqrt{n})}\sum_{t=1}^{n} T(x_t) = \frac{1}{n}\sum_{t=1}^{n} S\left(\frac{x_t}{\sqrt{n}}\right) + O_p\left((\kappa^{-1}\nu)(\sqrt{n})\right)$$

holds for any asymptotically homogeneous function T introduced in Definition 2.5. For any asymptotically homogeneous function T, we therefore have

$$\frac{1}{n\kappa(\sqrt{n})}\sum_{t=1}^{n} T(x_t) = \frac{1}{n}\sum_{t=1}^{n} S\left(\frac{x_t}{\sqrt{n}}\right) + o_p(1) \to_d \int_0^1 S(W(r))\,dr$$

due to Lemma 3.1.

We may also easily obtain from Theorem 3.2 the strong approximations for the asymptotically homogeneous transformations if the transformation function has a regularly locally integrable limit homogeneous function. If an asymptotically homogeneous function T has the regularly locally integrable limit homogeneous function S, then we have

$$\begin{aligned}
\frac{1}{n\kappa(\sqrt{n})}\sum_{t=1}^{n} T(x_t) &= \frac{1}{n}\sum_{t=1}^{n} S\left(\frac{x_t}{\sqrt{n}}\right) + O_p\big((\kappa^{-1}\nu)(\sqrt{n})\big) \\
&=_d \int_0^1 S(W_n(r))\,dr + O_p\big((\kappa^{-1}\nu)(\sqrt{n})\big) \\
&= \int_0^1 S(W(r))\,dr + O_p\big(c_n^{1+a}\big) + O_p\Big((1+c_n^b)\Delta_n\Big) + O_p\big((\kappa^{-1}\nu)(\sqrt{n})\big)
\end{aligned}$$

for all n sufficiently large.

3.4 Remark (a) For the logarithmic function $T(x) = \log|x|$, we may apply Theorem 3.2 and the part (b) of Remark 3.3 to get

$$\frac{1}{n\log n}\sum_{t=1}^{n} \log|x_t| = \frac{1}{2} + O_p((\log n)^{-1})$$

for all n sufficiently large.

(b) Our results here can also be used to derive the asymptotics for the transformations that can be written as the sums of the asymptotically homogeneous and integrable transformations. Consider, for an example, the logistic function given by $T(x) = e^x/(1+e^x)$, which can be written as $T = S+R$, where $S(x) = 1\{x \geq 0\}$ and $R(x) = e^x/(1+e^x)-1\{x \geq 0\}$. Let Assumptions 2.2 and 2.3 hold. Clearly, R is integrable, and we have $\sum_{t=1}^n R(x_t) = O_p(\sqrt{n})$, as will be shown later in Corollary 3.7. Therefore, it follows from Theorem 3.2 and the part (c) of Remark 3.3 that

$$\frac{1}{n}\sum_{t=1}^n \frac{\exp(x_t)}{1+\exp(x_t)} =_d \int_0^1 1\{W(r) \geq 0\}\, dr + o_p(n^{-1/2+1/p})$$

for all large n.

3.2 Asymptotics for Integrable Transformations

Now we consider the asymptotics for the integrable transformations. First, we show that

3.5 Theorem Let Assumptions 2.2 and 2.3 hold, and let $T : \mathbf{R} \to \mathbf{R}$ be strongly integrable. Then we have as $n \to \infty$

$$\sqrt{n}\int_0^1 T(\sqrt{n}W_n(r))dr = \sqrt{n}\int_0^1 T(\sqrt{n}W(r))dr + o_p\left(n^{-1/6+1/3p+\epsilon}\right)$$

for any $\epsilon > 0$.

For the strongly integrable function T, it therefore follows that

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{t=1}^n T(x_t) &=_d \sqrt{n}\int_0^1 T(\sqrt{n}W_n(r))dr \\ &= \sqrt{n}\int_0^1 T(\sqrt{n}W(r))dr + o_p\left(n^{-1/6+1/3p+\epsilon}\right)\end{aligned}$$

However, we have

$$\begin{aligned}\sqrt{n}\int_0^1 T(\sqrt{n}W(r))dr &= \sqrt{n}\int_{-\infty}^{\infty} T(\sqrt{n}s)L(1,s)ds &(5)\\ &= \int_{-\infty}^{\infty} T(s)L\left(1,\frac{s}{\sqrt{n}}\right)ds \\ &= L(1,0)\int_{-\infty}^{\infty} T(x)dx + O(n^{-1/4}) \text{ a.s.} &(6)\end{aligned}$$

by the the successive applications of the occupation times formula and change of variables for integrals, and using the fact

$$L(1,c_n) = L(1,0) + O(c_n^{1/2}) \text{ a.s.}$$

as $c_n \to 0$. Consequently, we have

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} T(x_t) =_d L(1,0)\int_{-\infty}^{\infty} T(x)dx + o_p\left(n^{-1/6+1/3p+\epsilon}\right) \tag{7}$$

for all large n, which provides the strong approximations for the integrable transformations.

3.6 Remark (a) Our strong approximation for the integrable transformation in (7) may be viewed as an extension of Borodin (1989, Theorem 3.1, pp. 40-41)'s result on pure random walks to general integrated processes driven by linear processes. For random walks with iid innovations, he showed that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} T(x_t) =_d L(1,0)\int_{-\infty}^{\infty} T(x)dx + O_p(n^{-1/4})$$

He allows for the innovations to have a discrete distribution, as long as it has the characteristic function φ such that $|\varphi(t)| = 1$ if and only if t is a multiple of 2π. His result requires that the innovations have finite third moment. Moreover, T should be bounded, square integrable and $\int_{-\infty}^{\infty} |x|^{1/2+\epsilon}|T(x)|\, dx < \infty$ for the innovations having a continuous distribution, and $\sum |x|^{1/2+\epsilon}|T(x)| < \infty$ for the innovations having a discrete distribution, for some $\epsilon > 0$.

(b) The order of approximation given in (7) is precisely the same as the one obtained by Akonom (1993) for T being an indicator on a compact interval.

The first order asymptotics for the integrable transformations do not require any regularity conditions other than integrability. We have

3.7 Corollary Let Assumptions 2.2 and 2.3 hold, and let $T : \mathbf{R} \to \mathbf{R}$ be integrable. Then we have

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} T(x_t) \to_d L(1,0)\int_{-\infty}^{\infty} T(x)\, dx$$

as $n \to \infty$.

3.8 Remark (a) The first order asymptotics in Corollary 3.7 significantly improves upon Park and Phillips (1999). For the moment condition on the innovation process, their results require $p > 8$, while ours hold as long as $p > 2$. Moreover, they assume the Lipschitz condtion, while we do not require anything other than integrability, on the transformation function T.

(b) Our result in Corollary 3.7 is comparable to Theorems 2.1 in Borodin and Ibragimov (1995). For the random walks driven by iid innovations, they establish the same result. Their result requires that T is integrable with $|T(x)| < c/(1 + |x|^{1+\epsilon})$ for some constant c and $\epsilon > 0$. They allow for the innovations having a discrete distribution if their characteristic function φ is such that $|\varphi(t)| = 1$ if and only if t is a multiple of 2π.

4. Concluding Remark

In this paper, we develop the strong approximations for the nonlinear transformations of integrated time series. As noted earlier, they are potentially useful in many different contexts. In particular, they can be used directly for: the nonparametric and semiparametric estimation of the models involving integrated time series, the theory of the bootstrap refinements for nonlinear models with integrated time series, and the statistical analysis of models with the asymptotic unit roots. These researches are currently under way, and will be reported later.

5. Mathematical Proofs

5.1 Proof of Lemma 3.1 The stated result follows from Theorem 1.1, pp. 80-81, of Borodin and Ibragimov (1995). ∎

5.2 Proof of Theorem 3.2 Let (c_n) be a sequence satisfying the conditions in Theorem 3.2, and write

$$|T(W_n(r)) - T(W(r))| \leq A_n(r) + B_n(r) + C_n(r) + D_n(r)$$

where

$$\begin{aligned}
A_n(r) &= |T(W_n(r)) - T(W(r))| 1\{|W_n(r)| \geq c_n\} 1\{|W(r)| \geq c_n\} \\
B_n(r) &= |T(W_n(r)) - T(W(r))| 1\{|W_n(r)| < c_n\} 1\{|W(r)| \geq c_n\} \\
C_n(r) &= |T(W_n(r)) - T(W(r))| 1\{|W_n(r)| \geq c_n\} 1\{|W(r)| < c_n\} \\
D_n(r) &= |T(W_n(r)) - T(W(r))| 1\{|W_n(r)| < c_n\} 1\{|W(r)| < c_n\}
\end{aligned}$$

for all $r \in [0,1]$.

Define

$$c_m = 1 + \max_{0 \leq r \leq 1} |W(r)|$$

and let M be the Lipschitz constant such that

$$|T(x) - T(y)| \leq M|x - y|$$

for all $x, y \in [c_n, c_m]$. We may now deduce that

$$M \leq (1 + c_n^b) Z$$

for some random variable Z such that $Z < \infty$ a.s., due to the condition in the part (b) of Assumption 2.4, and therefore,

$$A_n(r) \leq (1 + c_n^b) Z |W_n(r) - W(r)|$$

for all large n. Consequently,

$$\int_0^1 A_n(r)\, dr = O_p\left((1 + c_n^b)\Delta_n\right) \tag{8}$$

for all large n.

We may similarly show that for all n sufficiently large

$$\begin{aligned} B_n(r) &\le |T(W_n(r)) - T(W(r))| 1\{|W_n(r)| \ge c_n - \Delta_n\} 1\{|W(r)| \ge c_n\} \\ &\le \left(1 + (c_n - \Delta_n)^b\right) Z |W_n(r) - W(r)| \end{aligned}$$

with some random variable Z such that $Z < \infty$ a.s., by considering the interval $[c_n - \Delta_n, c_m]$ in place of $[c_n, c_m]$. By the same token, we also have for all n sufficiently large

$$\begin{aligned} C_n(r) &\le |T(W_n(r)) - T(W(r))| 1\{|W(r)| \ge c_n - \Delta_n\} 1\{|W_n(r)| \ge c_n\} \\ &\le \left(1 + (c_n - \Delta_n)^b\right) Z |W_n(r) - W(r)| \end{aligned}$$

with some random variable Z such that $Z < \infty$ a.s. It therefore follows that

$$\int_0^1 B_n(r)\, dr, \ \int_0^1 C_n(r)\, dr = O_p\left((1 + c_n^b)\Delta_n\right) \tag{9}$$

for all large n.

Finally, we have

$$\begin{aligned} \int_0^1 D_n(r)\, dr &\le \int_0^1 |T(W_n(r))| 1\{|W_n(r)| < c_n\}\, dr + \int_0^1 |T(W(r))| 1\{|W(r)| < c_n\}\, dr \\ &\le 2(1 + o_p(1)) \int_0^1 |T(W(r))| 1\{|W(r)| < c_n\}\, dr \end{aligned}$$

for large n, since, in particular, we may show that

$$\int_0^1 |T(W_n(r))| 1\{|W_n(r)| < c_n\}\, dr = (1 + o_p(1)) \int_0^1 |T(W(r))| 1\{|W(r)| < c_n\}\, dr$$

for large n, following the proof of Theorem 1.1, pages 81-82, of Borodin and Ibragimov (1995). Moreover, due to the condition in the part (a) of Assumption 2.4,

$$|T(W(r))| 1\{|W(r)| < c_n\} \le |W(r)|^a 1\{|W(r)| < c_n\}$$

for all $r \in [0, 1]$ and for all n sufficiently large, and we have

$$\begin{aligned} \int_0^1 |W(r)|^a 1\{|W(r)| < c_n\}\, dr &= \int_{-\infty}^{\infty} |x|^a 1\{|x| < c_n\} L(1, x)\, dx \\ &\le \left(\max_{x \in \mathbf{R}} L(1, x)\right) \int_{-\infty}^{\infty} |x|^a 1\{|x| < c_n\}\, dx \\ &= \left(\max_{x \in \mathbf{R}} L(1, x)\right) \frac{c_n^{1+a}}{1 + a} \end{aligned}$$

using the occupation times formula and the boundedness of $L(1, \cdot)$. Recall that we assume $a > -1$. We may now readily establish that

$$\int_0^1 D_n(r)\, dr = O_p(c_n^{1+a}) \tag{10}$$

The stated result now follows from (8) - (10). ∎

5.3 Proof of Theorem 3.5 We assume that $x_0 = 0$, and that the support of T is included on the positive half of the real line. These assumptions will simplify the exposition and cause no loss in generality. We also assume, by taking piece by piece if necessary, that T satisfies the conditions in Definition 2.9 over its entire support. We let (κ_n) and (δ_n) be sequences of numbers that are given by

$$\kappa_n = n^{1/6+5/3p+\epsilon}, \quad \delta_n = n^{-1/6-2/3p}$$

for some small $\epsilon > 0$. It follows that

$$\kappa_n \delta_n = n^{1/p+\epsilon} \to \infty \tag{11}$$

Note that $\kappa_n \leq n$ and $\delta_n \geq n^{-1/2}$ for sufficiently small $\epsilon > 0$, since we assume $p > 2$. These are necessary to use the results in Akonom (1993) for our subsequent proof. Moreover, we let (Δ_n) be defined as in (4). Note that $\Delta_n = o_p(n^{-1/2+1/p})$, and therefore,

$$\sqrt{n}\Delta_n = o_p(\kappa_n \delta_n)$$

and

$$\kappa_n \delta_n \pm 2\sqrt{n}\Delta_n \geq \kappa_n \delta_n (1 + o_p(1)) \tag{12}$$

due to (11).

Define

$$\begin{aligned} T_n(x) &= T(x)1\{0 \leq x < \kappa_n \delta_n\} \\ T_{nn}(x) &= \sum_{k=1}^{\kappa_n} T(k\delta_n)1\{(k-1)\delta_n \leq x < k\delta_n\} \end{aligned}$$

The function T_n is a truncated version of T, and the function T_{nn} is a simple function approximating T_n. These two functions play important roles in what follows. It follows from the part (b) of Definition 2.9 and (11) that

$$\begin{aligned} \int_{-\infty}^{\infty} |(T - T_n)(x)| dx &\leq c \int_{-\infty}^{\infty} x^{-q-1} 1\{x \geq \kappa_n \delta_n\}\, dx \\ &= n^{-q/p-\epsilon} \\ &= o\left(n^{-1/6+1/3p-\epsilon}\right) \end{aligned} \tag{13}$$

for some constant c and for some small $\epsilon > 0$. Moreover,

$$\sup_{-\infty < x < \infty} |T_n(x) - T_{nn}(x)| \leq c\, \delta_n \tag{14}$$

for some constant c.

First, we show that

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} T(x_t) - \frac{1}{\sqrt{n}} \sum_{t=1}^{n} T_n(x_t) &=_d \sqrt{n} \int_0^1 (T - T_n)(\sqrt{n} W_n(r)) dr \\ &= o_p\left(n^{-1/6+1/3p+\epsilon}\right) \end{aligned} \tag{15}$$

for any $\epsilon > 0$. By taking n sufficiently large, we may assume that $T - T_n$ is monotone decreasing on its support. This can be done without loss of generality, since we may always bound $T - T_n$ by such a function satisfying the same condition as $T - T_n$. We then have

$$(T - T_n)(\sqrt{n}W_n(r)) \leq T(\sqrt{n}(W(r) - \Delta_n))1\{\sqrt{n}(W(r) + \Delta_n) > \kappa_n\delta_n\}$$

and it follows that

$$\begin{aligned}
&\sqrt{n}\int_0^1 (T - T_n)(\sqrt{n}W_n(r))dr \\
&\leq \sqrt{n}\int_0^1 T(\sqrt{n}(W(r) - \Delta_n))1\{\sqrt{n}(W(r) + \Delta_n) > \kappa_n\delta_n\}dr \\
&= \sqrt{n}\int_{-\infty}^{\infty} T(\sqrt{n}(s - \Delta_n))1\{\sqrt{n}(s + \Delta_n) > \kappa_n\delta_n\}L(1, s)ds \\
&= \int_{-\infty}^{\infty} T(s)1\{s > \kappa_n\delta_n - 2\sqrt{n}\Delta_n\}L\left(1, \frac{s}{\sqrt{n}} + \Delta_n\right)ds \\
&\leq \left(\max_{x\in\mathbf{R}} L(1, x)\right)\int_{-\infty}^{\infty} T(s)1\{s > \kappa_n\delta_n - 2\sqrt{n}\Delta_n\}ds \\
&= o_p\left(n^{-1/6+1/3p-\epsilon}\right)
\end{aligned}$$

for some small $\epsilon > 0$. The second and third equalities are due respectively to the occupation times formula and a simple change of variables for integrals, and the fourth inequality follows from the boundedness of $L(1, \cdot)$. The last equality can be deduced from (12) and (13).

Second, we have from the Lipschitz condition for T in the part (a) of Definition 2.9 that

$$\begin{aligned}
\left|\frac{1}{\sqrt{n}}\sum_{t=1}^n T_n(x_t) - \frac{1}{\sqrt{n}}\sum_{t=1}^n T_{nn}(x_t)\right| &\leq c\kappa_n\delta_n^2\frac{1}{\kappa_n\delta_n\sqrt{n}}\sum_{t=1}^n 1\{0 \leq x_t < \kappa_n\delta_n\} \\
&= O_p(\kappa_n\delta_n^2) \\
&= O_p\left(n^{-1/6+1/3p+\epsilon}\right)
\end{aligned} \tag{16}$$

where c is the constant introduced in (14). Note that

$$\begin{aligned}
\frac{1}{\kappa_n\delta_n\sqrt{n}}\sum_{t=1}^n 1\{0 \leq x_t < \kappa_n\delta_n\} &=_d \frac{\sqrt{n}}{\kappa_n\delta_n}\int_0^1 1\{0 \leq \sqrt{n}W_n(r) < \kappa_n\delta_n\}dr \\
&\leq \frac{\sqrt{n}}{\kappa_n\delta_n}\int_0^1 1\{0 \leq \sqrt{n}W(r) < \kappa_n\delta_n + \sqrt{n}\Delta_n\}dr \\
&= \frac{\sqrt{n}}{\kappa_n\delta_n}\int_{-\infty}^{\infty} 1\{0 \leq \sqrt{n}s < \kappa_n\delta_n + \sqrt{n}\Delta_n\}L(1, s)ds \\
&= \int_{-\infty}^{\infty} 1\{0 \leq s < 1 + o_p(1)\}L\left(1, \frac{\kappa_n\delta_n s}{\sqrt{n}}\right)ds \\
&= O_p(1)
\end{aligned}$$

by applying occupation times formula and change of variables for integrals, and using the boundedness of $L(1,\cdot)$ and the result in (11).

Third, we show that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} T_{nn}(x_t) = \left(\int_{-\infty}^{\infty} T(x)dx\right)\frac{1}{\delta_n\sqrt{n}}\sum_{t=1}^{n} 1\{0 \le x_t < \delta_n\} + o_p\left(n^{-1/6+1/3p+\epsilon}\right) \tag{17}$$

for any $\epsilon > 0$. We have

$$\mathbf{E}\left(\sum_{t=1}^{n} 1\{0 \le x_t < \delta_n\} - \sum_{t=1}^{n} 1\{(k-1)\delta_n \le x_t < k\delta_n\}\right)^2 \le c\,n^{1/2}\delta_n \tag{18}$$

for $k = 1,\ldots,\kappa_n$, where c is some number, which is dependent only upon the distribution of (ε_t) and bounded by some absolute constant. This follows from Lemma 6 of Akonom (1993) for random walks driven by iid innovations, and can be shown to hold for general integrated processes as in his proof of Lemma 13 in Akonom (1993). Since $p > 2$, we have

$$\kappa_n\delta_n^2 \log n = n^{-1/6+1/3p+\epsilon}\log n \le 1$$

for sufficiently small $\epsilon > 0$, and therefore, the term $(1 + \kappa_n\delta_n^2 \log n)$ included in his result becomes redundant and is not presented in (18).

It follows from (18) that

$$\sum_{t=1}^{n} 1\{0 \le x_t < \delta_n\} = \sum_{t=1}^{n} 1\{(k-1)\delta_n \le x_t < k\delta_n\} + O_p\left(n^{1/4}\delta_n^{1/2}\right) \tag{19}$$

Since the error term in (19) can be bounded uniformly in $k = 1,\ldots,\kappa_n$ as explained below (18), we may multiply both sides of (19) by $T(k\delta_n)$ for $k = 1,\ldots,\kappa_n$ and divide their sum by $\sqrt{n}$ for normalization to get

$$\left(\int_{-\infty}^{\infty} T_{nn}(x)dx\right)\frac{1}{\delta_n\sqrt{n}}\sum_{t=1}^{n} 1\{0 \le x_t < \delta_n\} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} T_{nn}(x_t) + O_p\left(n^{-1/4}\delta_n^{-1/2}\right)$$

However, we have $n^{-1/4}\delta_n^{-1/2} = n^{-1/6+1/3p}$. Moreover, it follows from (13) and (14) that

$$\begin{aligned}\int_{-\infty}^{\infty} T_{nn}(x)dx &= \int_{-\infty}^{\infty} T_n(x)dx + O(\kappa_n\delta_n^2)\\ \int_{-\infty}^{\infty} T_n(x)dx &= \int_{-\infty}^{\infty} T(x)dx + o\left(n^{-1/6+1/3p-\epsilon}\right)\end{aligned}$$

and therefore

$$\int_{-\infty}^{\infty} T_{nn}(x)dx = \int_{-\infty}^{\infty} T(x)dx + o\left(n^{-1/6+1/3p+\epsilon}\right)$$

for any $\epsilon > 0$. The result in (17) is now immediate.

Fourth, we may easily deduce from Theorem 14 of Akonom (1993)

$$\frac{1}{\delta_n\sqrt{n}}\sum_{t=1}^{n}1\{0\le x_t<\delta_n\}=\frac{\sqrt{n}}{\delta_n}\int_0^1 1\{0\le\sqrt{n}W(r)<\delta_n\}dr+o_p\left(n^{-1/6+1/3p+\epsilon}\right)\tag{20}$$

with the missing δ_n term corrected to make it comparable as Theorem 8. Finally, we have

$$\begin{aligned}\frac{\sqrt{n}}{\delta_n}\int_0^1 1\{0\le\sqrt{n}W(r)<\delta_n\}dr &= \frac{\sqrt{n}}{\delta_n}\int_{-\infty}^{\infty}1\{0\le\sqrt{n}s<\delta_n\}L(1,s)ds\\ &= \int_{-\infty}^{\infty}1\{0\le s<1\}L\left(1,\frac{\delta_n s}{\sqrt{n}}\right)ds\\ &= L(1,0)+O\left(n^{-1/4}\delta_n^{1/2}\right)\text{ a.s.}\end{aligned}\tag{21}$$

and, due to the result in (6),

$$\sqrt{n}\int_0^1 T(\sqrt{n}W(r))dr=L(1,0)\int_{-\infty}^{\infty}T(x)dx+O(n^{-1/4})\text{ a.s.}\tag{22}$$

Note that $n^{-1/4}\delta_n^{-1/2}=n^{-1/6+1/3p}$. The stated result now follows from (15), (16), (17), (20), (21) and (22). ∎

5.4 Proof of Corollary 3.7 Let (κ_n) and (δ_n) be defined as in the proof of Theorem 3.5, and let (Δ_n) be given as in (4). Moreover, we define T_n as in the proof of Theorem 3.5. For any integrable T, we have

$$\sqrt{n}\int_0^1(T-T_n)(\sqrt{n}W_n(r))\,dr\le\left(\max_{x\in\mathbf{R}}L(1,\cdot)\right)\int_{-\infty}^{\infty}T(s)1\{s>\kappa_n\delta_n-2\sqrt{n}\Delta_n\}\,ds=o_p(1)$$

by (11) and the dominated convergence. Here we assume without loss of generality that $T-T_n$ is monotone decreasing on its support, as in the proof of Theorem 3.5.

We now let

$$I_k(\delta_n)=\{(k-1)\delta_n\le x<k\delta_n\}$$

and define

$$\begin{aligned}\overline{T}_n(x) &= \sum_{k=1}^{\kappa_n}\left(\sup_{x\in I_k(\delta_n)}T(x)\right)1(I_k(\delta_n))\\ \underline{T}_n(x) &= \sum_{k=1}^{\kappa_n}\left(\inf_{x\in I_k(\delta_n)}T(x)\right)1(I_k(\delta_n))\end{aligned}$$

Then we may deduce as in the proof of Theorem 3.5

$$\begin{aligned}\left(\int_{-\infty}^{\infty}\overline{T}_n(x)\,dx\right)\frac{1}{\delta_n\sqrt{n}}\sum_{t=1}^{n}1\{0\le x_t<\delta_n\} &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\overline{T}_n(x_t)+O_p(n^{-1/4}\delta_n^{-1/2})\\ \left(\int_{-\infty}^{\infty}\underline{T}_n(x)\,dx\right)\frac{1}{\delta_n\sqrt{n}}\sum_{t=1}^{n}1\{0\le x_t<\delta_n\} &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\underline{T}_n(x_t)+O_p(n^{-1/4}\delta_n^{-1/2})\end{aligned}$$

and consequently,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\overline{T}_n(x_t) = L(1,0)\left(\int_{-\infty}^{\infty}T(x)\,dx\right)+o_p(1)$$
$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\underline{T}_n(x_t) = L(1,0)\left(\int_{-\infty}^{\infty}T(x)\,dx\right)+o_p(1)$$

Note that

$$\int_{-\infty}^{\infty}\overline{T}_n(x)\,dx,\ \int_{-\infty}^{\infty}\underline{T}_n(x)\,dx = \int_{-\infty}^{\infty}T_n(x)\,dx + o(1)$$

since T is Riemann-integrable, and

$$\int_{-\infty}^{\infty}|(T-T_n)(x)|\,dx = \int_{-\infty}^{\infty}|T(x)|1\{x\geq\kappa_n\delta_n\}\,dx = o(1)$$

by (11) and the dominated convergence. The stated result now follows immediately, since

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\underline{T}_n(x_t)\leq\frac{1}{\sqrt{n}}\sum_{t=1}^{n}T_n(x_t)\leq\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\overline{T}_n(x_t)$$

and the proof is complete. ∎

References

Akonom, J. (1993). Comportement asymptotique du temps d'occupation du processus des sommes partielles. *Annales de l'Institut Henri Poincaré*, 29, 57-81.

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

Borodin, A.N. (1989). "Brownian local time," *Russian Mathematical Surveys*, 44, 1-51.

Borodin, A.N. and I.A. Ibragimov (1995). *Limit Theorems for Functionals of Random Walks*, Proceedings of the Steklov Institute of Mathematics, Vol. 195. American Mathematical Soceity.

Chang, Y. (2002). "Nonlinear IV unit root tests in panels with cross-sectional dependency," *Journal of Econometrics*, 110, 261-292.

Chang, Y. and J.Y. Park (2001). "Index models with integrated time series," forthcoming in *Journal of Econometrics.*

Chang, Y., J.Y. Park and P.C.B. Phillips (2001). "Nonlinear econometric models with cointegrated and deterministically trending regressors," *Econometrics Journal*, 4, 1-36.

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

Csörgő, M. and L. Horváth (1993). *Weighted Approximations in Probability and Statistics.* Wiley: New York.

de Jong, R.M. (2002a). "Appendum to "Asymptotics for nonlinear transformations of integrated time series"," Mimeograph, Department of Economics, Michigan State University.

de Jong, R.M. (2002b). "Asymptotics for periodic transformations of integrated time series"," Mimeograph, Department of Economics, Michigan State University.

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

Park, J. Y. and P. C. B. Phillips (2001). "Nonlinear regressions with integrated time series", *Econometrica*, 69 1452-1498.

Park, J.Y. and P.C.B. Phillips (2000). "Nonstationary binary choice," *Econometrica*, 68, 1249-1280.

Phillips, P.C.B., J.Y. Park and Y. Chang (2001). "Nonlinear instrumental variable estimation of an autoregression," forthcoming in *Journal of Econometrics.*

Pollard, D. (1984). *Convergence of Stochastic Processes.* Springer-Verlag, New York.