Uniform approximation to local time with applications in non-linear co-integrating regression

Weidong Liu*, Nigel Chan and Qiying Wang †

May 6, 2014

Abstract

Uniform strong approximation to a local time process is established for a functional of nonstationary time series. The main result is used to investigate uniform convergence for a local linear estimator in a nonlinear cointegrating regression model with non-linear nonstationary heteroskedastic error processes. Sharp convergence rates and optimal range are obtained. Estimates of a heterogeneity generating function (HGF) are also studied. It is shown that, when weighted by the HGF, the uniform convergence rate associated with local linear estimator can be improved in the tail. This feature seems to be new to literature.

Key words and phrases: Strong approximation, Local time, uniform convergence, nonparametric regression, local linear estimate, Kernel estimate, nonstationarity, nonlinearity.
JEL Classification: C14, C22.

*Department of Mathematics and Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai, China (Email: liuweidong99@gmail.com).

†School of Mathematics and Statistics, The University of Sydney, NSW 2006, Australia (Email: qiying@maths.usyd.edu.au).

1 Introduction

Let $x_{k,n}, 1 \le k \le n, n \ge 1$ be a triangular array, constructed from some underlying nonstationary time series and assume that there is a continuous limiting Gaussian process $G(t), 0 \le t \le 1$, to which $x_{[nt],n}$ converges weakly, where $[a]$ denotes the integer part of a. In many applications, we let $x_{k,n} = d_n^{-1} x_k$ where x_k is a nonstationary time series, such as a unit root or long memory process, and d_n is an appropriate standardization factor. A common functional of interest $S_n(x)$ of $x_{k,n}$ is defined by the sample quantity

$$S_n(x) = \frac{c_n}{n} \sum_{k=1}^{n} g[c_n (x_{k,n} + x)], \quad x \in R, \tag{1.1}$$

where c_n is a certain sequence of positive constants and g is a real integrable function on R. These functionals arise in nonparametric estimation and inference problems, particularly, problems involving nonlinear cointegration models. In such situations, the underlying time series x_k is nonstationary, g is a kernel function, and the secondary sequence c_n depends on the bandwidth used in the nonparametric regression. See Park and Phillips (1999, 2001), Karlsen, Myklebust and Tjostheim (2007), Wang and Phillips (2009a, 2009b, 2011, 2012) and the reference therein.

The point-wise limit behavior of $S_n(x)$ in the situation where $\int_{-\infty}^{\infty} g(s)\, ds \neq 0$ was studied in Wang and Phillips (2009a), where it was shown that when $c_n \to \infty$ and $n/c_n \to \infty$,

$$S_n(x) \to_D \int_{-\infty}^{\infty} g(t) dt\, L_G(1, -x), \tag{1.2}$$

where $L_G(t, s)$ is the local time of the process $G(t)$ at the spatial point s, defined in the end of this section. In the related works, Jeganathan (2004) investigated the asymptotic form of similar functionals when $x_{k,n}$ is a the partial sum of linear processes, Borodin and Ibragimov (1995), Akonom (1993) and Phillips and Park (1998) for the particular situation where $c_n x_{k,n}$ is a partial sum of iid random variables. More currently, Wang and Phillips (2011) considered the point-wise asymptotics of the $S_n(x)$ for the so-called zero energy functional, that is, $\int_{-\infty}^{\infty} g(s)\, ds = 0$. Results of the type (1.2) and those appeared in Wang and Phillips (2011) have many statistical applications, especially in nonparametric estimation - see Wang and Phillips (2009a, 2009b, 2011, 2012) and Wang (2014).

The present paper is concerned with developing a uniform approximation of $S_n(x)$ to the local time $L_G(1, -x)$ of the process $G(t)$. Such cases are important in nonlinear cointegrating regression and they appear in the investigation of uniform convergence in relation to non-parametric estimation. In order to investigate the uniform convergence for a local linear estimator, for example, we need to consider the lower bound for $\inf_{|x|\le b_n} S_n(x)$ in the form $g(s) = K(s)$, where b_n is a sequence of positive numbers approaching zero and $K(s)$ is the kernel function used in nonparametric estimation. As a direct consequence of our uniform approximation (Theorem 2.1), Corollary 2.1 provides a uniform lower bound of the $S_n(x)$ under a "optimal" range for the x being held. This result essentially improves the previous those by Chan and Wang (2014). It should be mentioned that similar uniform lower bounds of $S_n(x)$ are required in many other areas, such as the transformation regression and the estimation of the volatility function in a regression model with nonlinear nonstationary heteroskedastic (NNH) error processes. We refer to Wang and Wang (2013), Oliver and Wang (2013) and Section 2.3 for further details.

This paper is organized as follows. In next section, we present our main results. Theorem 2.1 provides a framework for the uniform approximation. It is shown that, under certain conditions and a rich probability space, $S_n(x)$ can be approximated by a local time process over R with certain rate. The rate might be not optimal, but it is enough for many practical applications. Theorem 2.2 gives an important application of Theorem 2.1 to general linear processes. Our result includes the x_k being a partial sum of ARMA processes and fractionally integrated processes, which are most commonly used in practice. Using Theorem 2.2 as a main tool, Theorem 2.3 investigates the uniform asymptotics for the local linear estimators in a non-linear cointegrating regression model with NNH error processes. We also consider the estimates for the heterogeneity generating function in Theorem 2.4. These results improve those in existing literature. All technical proofs are given in Section 3.

Throughout the paper, we denote by $C, C_1, ...$ the constants, which may change at each appearance. The process $\{L_\zeta(t,s), t\ge 0, s\in R\}$ is said to be the local time of a measurable process $\{\zeta(t), t\ge 0\}$ if, for any locally integrable function $T(x)$,

$$\int_0^t T[\zeta(s)]ds = \int_{-\infty}^{\infty} T(s)L_\zeta(t,s)ds, \quad \text{all } t\in R,$$

with probability one.

2 Main results and applications

2.1 Main results. We make use of the following assumptions in the development of main results. Except explicitly mentioned, notation will be the same as in Section 1.

Assumption 2.1. $\sup_x |x|^\rho |g(x)| < \infty$ *for some* $\rho > 1$, $\int_{-\infty}^{\infty} |g(x)|dx < \infty$ *and* $|g(x) - g(y)| \leq C|x - y|$ *whenever* $|x - y|$ *is sufficiently small on* R.

Assumption 2.2. *On a rich probability space, there exist a process* $G(t)$ *that has a local time* $L_G(1, x)$ *satisfying*

$$|L_G(1, x) - L_G(1, y)| \leq C|x - y|^\beta, \quad a.s., \tag{2.3}$$

for some $\beta > 0$ *and a sequence of stochastic processes* $G_n(t)$ *such that* $\{G_n(t); 0 \leq t \leq 1\} =_D \{G(t); 0 \leq t \leq 1\}$ *for each* $n \geq 1$ *and*

$$\sup_{0 \leq t \leq 1} |x_{[nt],n} - G_n(t)| \quad = \quad o_{a.s.}(n^{-\delta}). \tag{2.4}$$

for some $0 < \delta < 1$.

Assumption 2.3. *For all* $0 \leq j < k \leq n, n \geq 1$, *there exist a sequence of* σ*-fields* $\mathcal{F}_{k,n}$ *(define* $\mathcal{F}_{0,n} = \sigma\{\phi, \Omega\}$*, the trivial* σ*-field) such that,*

(i) $x_{j,n}$ *are adapted to* $\mathcal{F}_{j,n}$ *and, conditional on* $\mathcal{F}_{j,n}$, $[n/(k-j)]^d(x_{k,n} - x_{j,n})$ *where* $0 < d < 1$, *has a density* $h_{k,j,n}(x)$ *satisfying that* $h_{k,j,n}(x)$ *is uniformly bounded by a constant* K *and*

(ii) $\sup_{u \in R} \left|h_{k,j,n}(u + t) - h_{k,j,n}(u)\right| \leq C \min\{|t|, 1\}$, *whenever* n *and* $k - j$ *are sufficiently large and* $t \in R$.

Assumption 2.4. *There is a* $\epsilon_0 > 0$ *such that* $c_n \to \infty$ *and* $n^{-1+\epsilon_0} c_n \to 0$.

We remark that Assumption 2.1 is weak and standard for this type of problem, and it is satisfied by many functionals such as $g(x)$ is differentiable and has compact support. Assumption 2.2 is strong approximation version of the result $x_{n,[nt]} \to_D G(t)$ on $D[0, 1]$, and it is widely obtainable for many random sequences. A typical example in statistics and econmetrics is provided in Proposition 2.1 where we establish Assumption 2.2 for general linear processes. Note that, $G_n(x)$ can not be replaced by one single process $G(x)$ which is independent of n. Explanation in this regards can be found in Csörgö and Révész (1981). As for the Lipschitz type condition (2.3), it is satisfied by the classical Gaussian processes, Levy process and many semimartingales.

To illustrate, let $G(t)$ be a continuous Gaussian process with covariance function satisfying

$$EG(t)G(s) = A\,\{|t|^{2w} + |s|^{2w} - |t-s|^{2w}\}, \tag{2.5}$$

where $0 < w < 1$ and A is a constant. It follows from Theorem 30.4 of Geman and Horowitz (1980) and the remark below it that (2.3) holds for any $0 < \beta < (1-w)/2w$.

Assumption 2.3 is the same as Assumption 2.3 given in Wang and Phillips (2009a), except that the d_{kjn} in cited paper is repalced by $[(k-j)/n]^d$. As explained in Chan and Wang (2014), this additional requirement on $d_{l,k,n}$ is mild, which is only used here for technical convenience. In Assumption 2.4, $c_n \to \infty$ is necessary. If $c_n = 1$, a different limit distribution appears. We refer to Berkes and Horvath (2006) for further details.

We have the following main result.

Theorem 2.1. *Suppose Assumptions 2.1–2.4 hold. On the same probability space as in Assumption 2.2, for any $\beta > 0$, we have*

$$\sup_{x \in R} |S_n(x) - \tau\, L_{G_n}(1, -x)| \;=\; o_P(\log^{-\beta} n) \tag{2.6}$$

where $\tau = \int_{-\infty}^{\infty} g(t)dt \neq 0$. The result (2.6) remains true for $\tau = 0$ and, in this situation, Assumption 2.2 can be removed.

Remark 2.1. *Due to technical difficulties, the convergence rate in (2.6) may not be optimal. We conjecture that the rate should have the form $n^{-\gamma}$, where $\gamma > 0$ is related to $\delta > 0$ given in Assumption 2.2. However, the result (2.6) suffices in many applications. As a direct consequence, we have the following corollary that provides the uniform bounds for $S_n(x)$ under "optimal" range. As stated in Section 1, these uniform bounds are the key to investigate the uniform asymptotics in non-linear regression with non-stationary time series. See Section 2.3 for more details.*

Corollary 2.1. *Suppose Assumptions 2.1–2.4 hold. Then $\sup_{x\in R} |S_n(x)| = O_P(1)$ and, whence $\int_{-\infty}^{\infty} g(t)dt = 0$,*

$$\sup_{x \in R} |S_n(x)| = o_P(\log^{-\beta} n), \quad \textit{for any } \beta > 0. \tag{2.7}$$

Furthermore, if $\int_{-\infty}^{\infty} g(x)dx \neq 0$ and $\lim_{n\to\infty} P(\inf_{x\in\Omega_n} L_G(1,-x) = 0) = 0$ where Ω_n is a subset of $\mathbb{R}$, then

$$\big[\inf_{x \in \Omega_n} |S_n(x)|\big]^{-1} \;=\; O_P(1). \tag{2.8}$$

Remark 2.2. *Corollary 2.1 significantly improves main results of Chan and Wang (2014), where some restrictions are required on the range of x to achieve the optimal convergence rate. To illustrate, let $G(x)$ be a standard Wiener process. In this situation, $P(L_G(1,0)=0)=0$. Hence $\lim_{n\to\infty} P(\inf_{|x|\le r_n} L_G(1,-x)=0)=0$ for any $0<r_n\to 0$, due to the continuity of local time process. Corollary 2.1 yields that*

$$\Big[\inf_{|x|\le r_n} |S_n(x)|\Big]^{-1} = O_P(1), \tag{2.9}$$

for any $0<r_n\to 0$. In comparison, Chan and Wang (2014) only established $\big[\inf_{|x|\le M_0/\log^\gamma n} |S_n(x)|\big]^{-1} = O_P(1)$ for some $\gamma>0$. Furthermore, by noting that $P(L_G(1,x)=0)>0$ for any fixed $x\ne 0$ (see, for instance, Takacs (1995)), the range $|x|\le r_n$ in (2.9) might be optimal. In other words, it can not be improved to $|x|\le b$ for any constant $b>0$.

Remark 2.3. *Due to $\{G_n(t); 0\le t\le 1\} =_D \{G(t); 0\le t\le 1\}$ for each $n\ge 1$, Theorem 2.1 implies that $S_n(x)\Rightarrow \tau L_G(1,x)$ on $C(-\infty,\infty)$, where $C(-\infty,\infty)$ denotes the continuous functional space on $(-\infty,\infty)$ with uniform topology. After the completeness of this manuscript, the authors notice that the latter was established in Duffy (2014a) for partial sum of linear processes in a different method (the authors thank Duffy for his unpublished manuscripts). Using the special structure of linear process, Duffy (2014b) constructed a refine uniform estimates of $S_n(x)$ in zero energy situation, which has better convergence rate than that of (2.7). However, duo to the generality of Assumption 2.3, Duffy's methodology cannot be extended to this paper. Since (2.7) is enough for many applications (see Section 2.3), we leave the investigation for sharp convergence rate (in zero energy situation) under Assumption 2.3 for future work.*

2.2 An application to linear processes. In what follows we consider an application of Theorem 2.1 to general linear processes. Let $\{\xi_j, j\ge 1\}$ be linear processes defined by

$$\xi_j = \sum_{k=0}^{\infty} \phi_k\, \epsilon_{j-k}, \tag{2.10}$$

where $\{\epsilon_j, -\infty<j<\infty\}$ is a sequence of i.i.d. random variables with $E\epsilon_0=0$, $E\epsilon_0^2=1$, $\mathbb{E}|\epsilon_0|^r<\infty$ for some $r>2$ and the characteristic function $\varphi(t)$ of ϵ_0 satisfies $\int_{-\infty}^{\infty}|\varphi(t)|dt<\infty$. Throughout the section, the coefficients ϕ_k, $k\ge 0$ are assumed to satisfy one of the following conditions:

C1. $\phi_k \sim k^{-\mu}\,\rho(k)$, where $1/2 < \mu < 1$ and $\rho(k)$ is a function slowly varying at ∞, satisfying $|\rho(m+n)/\rho(n) - 1| \le C_0\, m/n$ for $1 \le m \le n$, where C_0 is a positive constant.

C2. $\sum_{k=0}^{\infty} k|\phi_k| < \infty$ and $\phi \equiv \sum_{k=0}^{\infty} \phi_k \neq 0$.

We remark that the requirement on $\rho(k)$ under condition **C1** is weak, which is satisfied by a large class of slowly varying functions such as $\log^{\alpha} x$, $\log\log^{\alpha} x$ and $\exp(\log^{\beta} x)$, where $\alpha \in R$ and $0 < \beta < 1$. See, e.g., Wang et al. (2003). Put $x_k = \sum_{j=1}^{k} \xi_j$ and $d_n^2 = Ex_n^2$. It is well-known that

$$d_n^2 = Ex_n^2 \sim \begin{cases} c_\mu\, n^{3-2\mu}\,\rho^2(n), & \text{under } \mathbf{C1}, \\ \phi^2\, n, & \text{under } \mathbf{C2}, \end{cases} \tag{2.11}$$

where $c_\mu = \frac{1}{(1-\mu)(3-2\mu)} \int_0^\infty x^{-\mu}(x+1)^{-\mu} dx$. See, e.g., Wang et al. (2003) for instance.

We consider the uniform limit behavior of sample functions of the form:

$$S_{1n}(x) \;=\; \frac{d_n}{nh} \sum_{k=1}^{n} g\big[h^{-1}\,(x_k + x\,d_n)\big], \tag{2.12}$$

when $h \to 0$. Let $W_d(t)$ be a fractional Brownian motion with Hurst parameter $-1/2 < d < 1/2$ on $D[0,1]$, defined by

$$W_d(t) = \frac{1}{A(d)}\,\Big[\int_{-\infty}^{0} \Big[(t-s)^d - (-s)^d\Big] dW^*(-s) + \int_0^t (t-s)^d dW(s)\Big],$$

where

$$A(d) = \Big(\frac{1}{2d+1} + \int_0^\infty \Big[(1+s)^d - s^d\Big]^2 ds\Big)^{1/2},$$

$W(s), 0 \le s < \infty$ is a standard Brownian motion, and $W^*(u), 0 \le u < \infty$ is an independent copy of $W(s), 0 \le s < \infty$. It is readily seen that $W_0(t) = W(t)$ and $W_d(t)$ have a continuous local time $L_{W_d}(t,s)$ with regard to (t,s) in $[0,\infty) \times R$, satisfying (2.3). See, e.g., Theorems 22.1 and 30.4 of Geman and Horowitz (1980).

The following results are direct consequences of Theorem 2.1.

Theorem 2.2. *Suppose Assumptions 2.1 holds, $h \to 0$ and $n^{1-\epsilon_0}h/d_n \to \infty$ for some $\epsilon_0 > 0$. Then, on a rich probability space, there exists a fractional Brownian motion*

$$Y_n(t) \;=\; \begin{cases} c_\mu^{-1/2}\, n^{-(3/2-\mu)}\,\rho^{-1}(n)\, W_{1-\mu}(nt), & \textit{under } \boldsymbol{C1}, \\ \phi^{-1} n^{-1/2}\, W(nt), & \textit{under } \boldsymbol{C2}, \end{cases} \tag{2.13}$$

such that, for all $\beta > 0$

$$\sup_{x\in R} |S_{1n}(x) - \tau\, L_{Y_n}(1, -x)| \quad = \quad o_P(\log^{-\beta} n), \tag{2.14}$$

where $\tau = \int_{-\infty}^{\infty} g(x)dx$.

Corollary 2.2. *Under the condition of Theorem 2.2, we have*

$$\sup_{x\in R} |\sum_{k=1}^{n} g\big[h^{-1}(x_k + x)\big]| = O_P(nh/d_n). \tag{2.15}$$

If in addition $\int_{-\infty}^{\infty} g(x)dx = 0$, then

$$\sup_{x\in R} |\sum_{k=1}^{n} g\big[h^{-1}(x_k + x)\big]| = O_P\big[(nh/d_n)\log^{-\beta} n\big], \tag{2.16}$$

for any $\beta > 0$. If in addition $\int_{-\infty}^{\infty} g(x)dx \neq 0$, then

$$\Big[\inf_{|x|\le r_n d_n} |\sum_{k=1}^{n} g\big[h^{-1}(x_k + x)\big]|\Big]^{-1} = O_P(d_n/(nh)), \tag{2.17}$$

for any $0 < r_n \to 0$.

Remark 2.4. *The key to prove Theorem 2.2 is to verify that $x_{k,n} := x_k/d_n$ satisfies Assumptions 2.2 and 2.3. The verification of Assumption 2.3 is given in Chan and Wang (2014). Recalling that a fraction Brownian motion is a Gaussian process satisfying (2.3) for some $0 < w < 1$, Assumption 2.2 can be easily verified by using the following proposition. We omit the details.*

Proposition 2.1. *On a rich probability space, there exists a fraction Brownian motion $\{W_{1-\mu}(t), 0 \le t < \infty\}$ such that*

$$\sup_{0\le t\le 1} \Big| c_\mu^{-1/2} \sum_{k=1}^{[nt]} \xi_k - \rho(n)\, W_{1-\mu}(nt) \Big| = o_{a.s.}[n^{(r+1)/r-\mu}\, \rho(n)] \tag{2.18}$$

*provided the condition **C1** holds.*

*Similarly, under the condition **C2**, on a rich probability space, there exists a Brownian motion $\{W(t), 0 \le t < \infty\}$ such that*

$$\sup_{0\le t\le 1} \Big| \phi^{-1} \sum_{k=1}^{[nt]} \xi_k - W(nt) \Big| \quad = \quad o_{a.s.}(n^{1/r}). \tag{2.19}$$

The proof of (2.18) is given in Wang et al. (2003) with minor improvement. The proof of (2.19) is given in Csörgö and Horvath (1993, Page 18).

2.3 Uniform convergence in non-linear cointegrating regression with NNH errors. As stated in Introduction, Theorems 2.1-2.2 and their corollaries play a key role in the investigation of non-stationary cointegration regression. Using these results, this section investigates the uniform convergence in the following nonlinear cointegrating regression model with NNH errors:

$$y_t = m(x_t) + \sigma(x_t)u_t, \quad t = 1, 2, ..., n, \tag{2.20}$$

where m is an unknown function to be estimated with the observed data $\{x_t, y_t\}_{t=1}^n$, $\sigma(x)$ is a heterogeneity generating function (HGF) and for a filtration $\mathcal{F}_t$ to which x_{t+1} is adapted, $\{u_t,\ \mathcal{F}_t\}$ forms a martingale difference.

When there are no data x_t incorporating in error process, namely $\sigma(x) = 1$, the issue on uniform convergence for the conventional Nadaraya-Watson estimator in model (2.20) has been currently considered in Chan and Wang (2014), Gao et al. (2014) and Wang and Chan (2014). In Chan and Wang (2014), the authors also investigated the uniform convergence for the local linear non-parametric estimator $\widehat{m}_n(x)$ of $m(x)$, defined by

$$\widehat{m}_n(x) = \sum_{i=1}^n w_i(x)y_i / \sum_{i=1}^n w_i(x), \tag{2.21}$$

where $K(x)$ be a non-negative real function, the bandwidth $h \equiv h_n \to 0$, $K_j(x) = x^j K(x)$, $V_{n,j}(x) = \sum_{i=1}^n K_j[(X_i - x)/h]$ and

$$w_i(x) = K[(x_i - x)/h]V_{n,2}(x) - K_1[(x_i - x)/h]V_{n,1}(x).$$

They further proved that, unlike the point-wise situation where the local linear estimate has no advantages in bias deduction (up to the second order), the performance of local linear estimator $\widehat{m}_n(x)$ is superior to that of the conventional Nadaraya-Watson estimator in uniform asymptotics. See, also, Duffy (2014a, b) for some similar arguments.

In this section, we will consider the uniform convergence of $\widehat{m}_n(x)$ with the model that nonlinear nonstationary heterogeneity is incorporated into error process. We also investigate the uniform convergence for an estimator of the HGF $\sigma(x)$ in model (2.20). In this regard, some initial results was established in Wang and Wang (2013), requiring some strong restrictions on the range of x. An application of Theorem 2.2 essentially improves these existing results, in particular, our results provide an "optimal" range for the values of x to be held in the establishment of uniform convergence for the $\hat{m}_n(x)$ and a sharp (may be optimal) convergence rate. Furthermore it is shown that, if weighted by the HGF $\sigma(x)$, the unform convergence rate for the local linear estimator $\hat{m}_n(x)$ can be improved in the tail of the range for x. This feature seems to be new to literature.

We make use of the following assumptions in the development of our unform asymptotics.

Assumption 2.5. $x_t = \sum_{j=1}^t \xi_j$, *where* ξ_j *is defined as in (2.10) with* ϕ_k *satisfying* ***C1*** *or* ***C2****.*

Assumption 2.6. $\{u_t, \mathcal{F}_t\}_{t\geq 1}$ *is a martingale difference, where* $\mathcal{F}_t = \sigma(x_1, ..., x_{t+1}, u_1, ..., u_t)$*, satisfying* $\sup_{t\geq 1} E(|u_t|^{2p} \mid \mathcal{F}_{t-1}) < \infty$*, where* $p > 1/\epsilon_0$ *for some* $0 < \epsilon_0 \leq 1/2$*.*

Assumption 2.7. K *has a compact support,* $\int_{-\infty}^{\infty} xK(x)dx = 0$ *and* $|K(x) - K(y)| \leq C|x-y|$ *for all* $x, y \in R$*.*

Assumption 2.8. *The first derivative of* $m(x)$ *exists and there exist a* $0 < \tau \leq 1$ *and a real positive function* $m_0(x)$ *such that*

$$|m'(y) - m'(x)| \;\leq\; C\,|y-x|^{\tau}\, m_0(x), \tag{2.22}$$

uniformly for $x \in R$ *and* $|y-x|$ *sufficiently small.*

Assumption 2.9. $\inf_{x\in R}\sigma(x) > 0$ *and for any* $|y|$ *sufficiently small,*

$$\sup_{x\in R} \frac{|\sigma(x+y) - \sigma(x)|}{\sigma(x)} \;\leq\; C\,|y|, \tag{2.23}$$

where C *is a positive constant.*

We remark that Assumptions 2.5-2.9 all are quite natural and easy to verify. Assumption 2.5 is not necessary, which can be replaced by Assumption 2.3 with some corresponding modifications in the following theorem. We use Assumption 2.5 in this paper to avoid the complexity of notation. When $\sigma(x)$ is a positive constant, Assumption 2.9 is trivially satisfied. In this situation, the model (2.20) and the conditions imposed are reduced to those of Chan and Wang (2014).

We have the following uniform asymptotic result.

Theorem 2.3. *Suppose Assumptions 2.5–2.9 hold. Let $\epsilon_0 > 0$ be given as in Assumption 2.6, d_n be defined as in (2.11) and $0 < r_n \to 0$. Then, for any h satisfying $h \to 0$ and $n^{1-\epsilon_0} h/d_n \to \infty$, we have*

$$\sup_{|x|\le b_n} \frac{|\widehat{m_n}(x) - m(x)|}{\sigma(x)} = O_P\left\{ \big(nh/d_n\big)^{-1/2} \log^{1/2} n + h^{1+\tau}\, \delta_n \right\}, \tag{2.24}$$

where $\delta_n = \sup_{|x|\le b_n} \big[m_0(x)/\sigma(x)\big]$ and $b_n \le r_n\, d_n$.

Remark 2.5. *The convergence rate in (2.24) is sharp and probably optimal. In the situation where x_t is stationary regressor, the sharp rate of convergence is $O_P[(nh)^{-1/2} \log^{1/2} n]$ (see, e.g., Hansen (2008)). The reason behind the difference is due to the fact that, the integrated series wanders over the entire real line but spent only $O(d_n)$ amount of sample time around any specific point, while stationary time series spent $O(n)$.*

Similarly to the explanation in Remark 2.2, the range $|x| \le r_n\, d_n$ is optimal in the situation that b_n in (2.24) can not be extended to $b_n/d_n \to C > 0$. This essentially improves Theorem 4.1 of Chan and Wang (2014), where the result is established under $\sigma(x) = 1$ and the range $|x| \le M_0 d_n/\log^{1+\gamma} n$ for some $\gamma > 0$.

Remark 2.6. *Under less restrictions on the kernel $K(x)$ and the regression function $m(x)$, a similar result can be established for the conventional kernel estimator $\widetilde{m}_n(x)$ defined by*

$$\widetilde{m}_n(x) = \frac{\sum_{k=1}^n K\big[(x_k - x)/h\big] y_k}{\sum_{k=1}^n K\big[(x_k - x)/h\big]}.$$

For details, we refer to Chan and Wang (2014) and Duffy (2014a, b). The latter papers investigated the uniform asymptotics under the random optimal range for the x to be held.

Remark 2.7. *Due to the definition of δ_n, when it is weighted by the HGF $\sigma(x)$, the uniform convergence rate of $\widehat{m}_n(x) - m(x)$ is improved in the tail. Note that model (2.20) can be restated as*

$$y_t/\sigma(x_t) = m(x_t)/\sigma(x_t) + u_t.$$

The limit behavior of $[\widehat{m}_n(x) - m(x)]/\sigma(x)$ is improved in the tail is not strange. It is the first, however, that this feature is noticed in literature.

As in Wang and Wang (2013), the HGF $\sigma(x)$ can be estimated by

$$\widehat{\sigma}^2(x) = \frac{\sum_{t=1}^n [y_t - \widehat{m}_n(x_t)]^2 K[(x_t - x)/h]}{\sum_{t=1}^n K[(x_t - x)/h]}.$$

The following result provides the limit behavior of $\widehat{\sigma}_n^2(x)$.

Theorem 2.4. *Suppose Assumptions 2.5, 2.7–2.9 hold, and in addition to Assumption 2.6,* $E(u_t^2 \mid \mathcal{F}_{t-1}) \to 1, a.s.$ *and* $\sup_{t\geq 1} E(|u_t|^{4p} \mid \mathcal{F}_{t-1}) < \infty$*, where* $p \geq 1 + 1/\epsilon_0$ *for some* $\epsilon_0 > 0$*. Let* d_n *be defined as in (2.11) and* $0 < r_n \to 0$*. Then, for any* h *satisfying* $h \to 0$ *and* $n^{1-\epsilon_0} h/d_n \to \infty$*, we have*

$$\sup_{|x|\leq b_n} \frac{|\widehat{\sigma}^2(x) - \sigma^2(x)|}{\sigma^2(x)} \;=\; O_P\Big\{h + \big(nh/d_n\big)^{-1/2} \log^{1/2} n + h^{1+\tau}\,\delta_n\Big\}, \quad (2.25)$$

where $\delta_n = \sup_{|x|\leq b_n} \big[m_0(x)/\sigma(x)\big]$ *and* $b_n \leq r_n\, d_n$*. Consequently,*

$$\sup_{|x|\leq b_n} \frac{|\hat{m_n}(x) - m(x)|}{\widehat{\sigma}(x)} = O_P\Big\{\big(nh/d_n\big)^{-1/2} \log^{1/2} n + h^{1+\tau}\,\delta_n\Big\} \quad (2.26)$$

Remark 2.8. *A similar result to (2.25) was established in Theorem 3.1 of Wang and Wang (2013) under* $|x| \leq A$*, where* A *is a constant. Theorem 2.4 provides both optimal convergence rate and optimal range for the values of* x *to be held.*

3 Proofs of main results

This section provides proofs of the main results. We start with several preliminary lemmas in Section 3.1. These lemmas, in particular Lemmas 3.5 and 3.6, are of interests in their own rights. The proofs of Theorem 2.1, 2.3 and 2.4 are given in Sections 3.2-3.4, respectively.

3.1 Preliminary Lemmas. Throughout this section, we set $f_{t,s}(x) = g(c_n x + t) - g(c_n x + s)$ where $g(x)$ satisfies Assumption 2.1, and assume that $x_{k,n}$ is defined as in Assumption 2.3.

Lemma 3.1. *We have, for any* $k > j$*,*

$$|E[f_{t,s}(x_{k,n})|\mathcal{F}_{j,n}]| \leq C\, n^d c_n^{-1}(k-j)^{-d} \min\{|t-s| n^d c_n^{-1}(k-j)^{-d}, 1\},$$

$$E[|f_{t,s}(x_{k,n})||\mathcal{F}_{j,n}] \leq C\, n^d c_n^{-1}(k-j)^{-d},$$

$$E[f_{t,s}^2(x_{k,n})|\mathcal{F}_{j,n}] \leq C\, n^d c_n^{-1}(k-j)^{-d}, \quad (3.27)$$

where C *is a uniformly bounded constant on* t, s, k *and* j*.*

Proof. Let $d_{k,j,n} = [(k-j)/n]^d$. Due to Assumption 2.3(i), we have

$$\begin{aligned} &\mathbb{E}(f_{t,s}(x_{n,k}) \mid \mathcal{F}_{n,j}) \\ =\;& \int_{-\infty}^{\infty} \big[g(c_n x_{j,n} + c_n d_{k,j,n} y + t) - g(c_n x_{j,n} + c_n d_{k,j,n} y + s)\big] h_{k,j,n}(y)\, dy \\ =\;& c_n^{-1} d_{k,j,n}^{-1} \int_{-\infty}^{\infty} g(y) \Big[h_{k,j,n}\big(\frac{y - t - c_n x_{j,n}}{c_n d_{k,j,n}}\big) - h_{j,k,n}\big(\frac{y - s - c_n x_{j,n}}{c_n d_{k,j,n}}\big)\Big]\, dy \end{aligned}$$

Now, Assumption 2.3 (ii) and $\int_{-\infty}^{\infty}|g(x)|dx < \infty$ yield that, for any $k > j$,

$$\begin{aligned}|\mathbb{E}(f_{t,s}(x_{n,k}) \mid \mathcal{F}_{n,j})| &\leq C\, c_n^{-1}\, d_{k,j,n}^{-1} \min\{|t-s|c_n^{-1}d_{k,j,n}^{-1}, 1\} \\ &\leq C\, n^d c_n^{-1}\, (k-j)^{-d} \min\{|t-s|n^d c_n^{-1}(k-j)^{-d}, 1\}.\end{aligned}$$

Similarly, using Assumptions 2.1 and 2.3, it follows that

$$\begin{aligned}&\mathbb{E}(f_{t,s}^2(x_{n,k}) \mid \mathcal{F}_{n,j}) \leq C\, \mathbb{E}(|f_{t,s}(x_{n,k})| \mid \mathcal{F}_{n,j}) \\ &= C \int_{-\infty}^{\infty} \big|g(c_n x_{j,n} + c_n d_{k,j,n} y + t) - g(c_n x_{j,n} + c_n d_{k,j,n} y + s)\big|\, h_{k,j,n}(y)\, dy \\ &= C\, c_n^{-1} d_{k,j,n}^{-1} \int_{-\infty}^{\infty} \big|g(y+t) - g(y+s)\big|\, h_{k,j,n}(\frac{y - c_n x_{j,n}}{c_n d_{k,j,n}})\, dy \\ &\leq C_1\, n^d c_n^{-1} / (k-j)^d \int_{-\infty}^{\infty} \big[g(y+t) - g(y+s)\big]^2\, dy \\ &\leq C_2 n^d c_n^{-1}/(k-j)^d.\end{aligned}$$

The proof of Lemma 3.1 is complete. $\square$

Lemma 3.2. *There exist constants H_0 (not depending on t_1, t_2, t_3) and m such that*

$$\begin{aligned}&\sup_{t,s} E\big(|\sum_{k=t_2}^{t_3} f_{t,s}(x_{k,n})|^m \mid \mathcal{F}_{n,t_1}\big) \\ &\leq H_0^m\, (m+1)!\, n^d\, c_n^{-1}(t_3 - t_1)^{1-d}\big[1 + \big\{(t_3 - t_2)^{1-d} n^d\, c_n^{-1}\big\}^{m-1}\big].\end{aligned} \tag{3.28}$$

for all $0 \leq t_1 < t_2 < t_3 \leq n$ and integer $m \geq 1$. In particular, by letting $t_1 = 0, t_2 = 1$ and $t_3 = n$, we have

$$\sup_{t,s} E|\sum_{k=1}^{n} f_{t,s}(x_{k,n})|^m \leq H_0^m\, (m+1)!\, (n/c_n)^m. \tag{3.29}$$

Proof. See Lemma 4.1 of Chan and Wang (2014) with minor improvements. $\square$

Lemma 3.3. *We have*

$$\sup_{t,s} \big|\sum_{k=1}^{b_n} f_{t,s}(x_{k,n})\big| = O_{a.s.}\big[(b_n/c_n)\log n\big] \tag{3.30}$$

for any $b_n, c_n \to \infty$ and $c_n/n \to 0$.

Proof. By virtue of Lemma 3.2, the proof follows from the similar arguments as in the proof of Theorem 2.1 of Chan and Wang (2014). We omit the details. $\square$

Lemma 3.4. *Suppose that $c_n \to \infty$ and $n^{-1+\epsilon_0} c_n \to 0$ for some $\epsilon_0 > 0$. Then, for any $\beta > 0$, we have*

$$I_n := \sup_{t\in R} \sup_{s:|s-t|\le \epsilon_n} \Big|\frac{c_n}{n}\sum_{k=1}^n f_{t,s}(x_{k,n})\Big| = O_{a.s.}(\log^{-\beta} n) \tag{3.31}$$

where $\epsilon_n \le c_n n^{-\alpha}$ for some $\alpha > 0$.

Proof. It suffices to prove:

$$I_{1n} := \sup_{|t|\le c_n n^2} \sup_{s:|s-t|\le \epsilon_n} \Big|\frac{c_n}{n}\sum_{k=1}^n f_{t,s}(x_{k,n})\Big| = O_{a.s.}(\log^{-\beta} n). \tag{3.32}$$

Indeed it is readily seen from (3.32) that

$$\begin{aligned} I_n &\le I_{1n} + \frac{c_n}{n} \sup_{|t|\ge c_n n^2} \sup_{|s|\ge c_n n^2/2} \sum_{k=1}^n |f_{t,s}(x_{k,n})| I(|x_{k,n}| \le n^2/2) \\ &\quad + \frac{C\, c_n}{n} \sum_{k=1}^n I(|x_{k,n}| \ge n^2/2) \\ &\le O_{a.s.}(\log^{-\beta} n) + 2c_n \sup_{|t|>c_n n^2/4} |g(t)| + O_{a.s}(c_n/n) \\ &= O_{a.s.}(\log^{-\beta} n), \end{aligned}$$

where we have used the following fact: $\sup_{|t|>c_n n^2/4} |g(t)| \le (c_n n^2)^{-\rho} \le C/n$ due to Assumption 2.1 and $\rho \ge 1$, and

$$\begin{aligned} &P\Big(\sum_{k=1}^n I(|x_{k,n}| > n^2/2) > C,\ i.o.\Big) \\ &\le C \lim_{r\to\infty} \sum_{n=r}^\infty n^{-4} \sum_{k=1}^n E|x_{k,n}|^2 \\ &\le C \lim_{r\to\infty} \sum_{n=r}^\infty n^{-4+1} E|\epsilon_0|^2 \le C \lim_{r\to\infty} \sum_{n=r}^\infty n^{-3} = 0 \end{aligned}$$

which implies $\sum_{k=1}^n I(|x_{k,n}| > n^2/2) = O(1), a.s.$

To prove (3.32), we first introduce the following blocking scheme. Let $\eta_n = (n/c_n) \log^{-(\beta+1)} n$, $b_n = [n^{1-\nu}]$, for some $0 < \nu < \min\{\epsilon_0, \nu_0\}$,

$$\nu_0 = \begin{cases} \big(\frac{1}{2d}\big)\alpha, & if \quad 0 < d \le 2/3, \\ \big(\frac{1-d}{d^2}\big)\alpha, & if \quad 2/3 < d < 1, \end{cases} \tag{3.33}$$

and let T_n be the largest integer s such that $sb_n \le n$. Also let $-c_n n^2 = t_1 < ... < t_{q_{n1}} = c_n n^2$ and $-\epsilon_n = s_1 < ... < s_{q_{n2}} = \epsilon_n$, with $t_i - t_{i-1} \sim n^{-7}$ and $s_i - s_{i-1} \sim c_n n^{-10}$.

It is readily seen that

$$n/b_n \sim n^{\nu}, \quad T_n b_n \le n, \quad n - T_n b_n \le b_n, \quad q_{n1}, q_{n2} \le n^{10} \tag{3.34}$$

due to $c_n \to \infty$. Under these notation, to prove (3.32), by the local Lipschitz continuity of g, it suffices to prove that

$$\begin{aligned} & \max_{1\le i\le q_{n1}} \max_{1\le j\le q_{n2}} \Big| \sum_{k=1}^{n} f_{t_i, t_i+s_j}(x_{k,n}) \Big| \\ \le \; & \max_{1\le i\le q_{n1}} \max_{1\le j\le q_{n2}} \Big\{ \Big| \sum_{w=2}^{T_n-1} \Delta_{nw}(t_i, s_j) \Big| + \Delta_n(t_i, s_j) \Big\} + O_{a.s.}[(n/c_n)^{1/2}], \end{aligned} \tag{3.35}$$

where, for $w = 1, ..., T_n$,

$$\begin{aligned} \Delta_{nw}(t, s) &= \sum_{k=wb_n+1}^{(w+1)b_n} f_{t,t+s}(x_{k,n}), \\ \Delta_n(t, s) &\le \Big(\sum_{k=1}^{2b_n} + \sum_{k=T_n b_n}^{n} \Big) |f_{t,t+s}(x_{k,n})|. \end{aligned}$$

Recall $\eta_n = (n/c_n) \log^{-(\beta+1)} n$. Using Lemma 3.3 and (3.34), it is readily seen that

$$\begin{aligned} \max_{1\le i\le q_{n1}} \max_{1\le j\le q_{n2}} \Delta_n(t_i, s_j) &\le C\big[(b_n + |n - T_n b_n|)/c_n\big] \log n \\ &\le C\,(n/c_n)\, n^{-\nu} \le C\,\eta_n \log n, \quad a.s. \end{aligned}$$

This, together with (3.35), implies that (3.31) will follow if we prove

$$\max_{1\le i\le q_{n1}} \max_{1\le j\le q_{n2}} \Big(\Big| \sum_{\substack{w=2 \\ w\in even}}^{T_n} \Delta_{nw}(t_i, s_j) \Big| + \Big| \sum_{\substack{w=2 \\ w\in odd}}^{T_n} \Delta_{nw}(t_i, s_j) \Big| \Big) = O_{a.s.}(\eta_n \log n). \tag{3.36}$$

We only prove (3.36) for $w \in even$. The other is similar and hence the details are omitted. To this end, let $\mathcal{F}^*_{n,v} = \mathcal{F}_{n,(2v+1)b_n}$, $v \ge 0$, and $M_1 > 0$ is chosen later,

$$\begin{aligned} \Delta'_{nw}(t, s) &= \Delta_{n,2w}(t, s) I(|\Delta_{n,2w}(t, s)| \le M_1\, \eta_n), \\ \Delta^*_{nw}(t, s) &= \Delta'_{n,w}(t, s) - \mathbb{E}\big(\Delta'_{n,w}(t, s) \mid \mathcal{F}^*_{n,w-1}\big). \end{aligned}$$

Under these notation, to prove (3.36) for $w \in even$, it suffices to show

$$\lambda_{1n} := \max_{1\le i\le q_{n1}} \max_{1\le j\le q_{n2}} |\sum_{w=1}^{T_n/2} \Delta^*_{nw}(t_i, s_j)| = O_{a.s.}(\eta_n \log n), \tag{3.37}$$

$$\lambda_{2n} := \max_{1\le i\le q_{n1}} \max_{1\le j\le q_{n2}} |\sum_{w=1}^{T_n/2} \mathbb{E}\big(\Delta_{n,2w}(t_i, s_j) \mid \mathcal{F}^*_{n,w-1}\big)| = O_{a.s.}(\eta_n \log n), \tag{3.38}$$

$$\begin{aligned}\lambda_{3n} &:= \max_{1\le i\le q_{n1}} \max_{1\le j\le q_{n2}} |\sum_{w=1}^{T_n/2} \Big(\Delta_{n,2w}(t_i, s_j) I(|\Delta_{n,2w}(t_i, s_j)| > M_1\, \eta_n) \\ &\qquad + \mathbb{E}\Big[\Delta_{n,2w}(t_i, s_j) I(|\Delta_{n,2w}(t_i, s_j)| > M_1\, \eta_n) \mid \mathcal{F}^*_{n,w-1}\Big]\Big) \\ &= O_{a.s.}(\eta_n \log n).\end{aligned} \tag{3.39}$$

First notice that, for any $2wb_n < k \le (2w+1)b_n$ and $|t-s| \le c_n n^{-\alpha}$,

$$|\mathbb{E}\big(\Delta_{n,2w}(t,s) \mid \mathcal{F}^*_{n,w-1}\big)| \le C\,|t-s| b_n c_n^{-2} (n/b_n)^{2d} \le C\, b_n c_n^{-1} n^{2d\nu-\alpha} \tag{3.40}$$

due to Lemma 3.1. It follows from (3.33) and (3.40) that

$$\begin{aligned}\lambda_{2n} &\le \sum_{w=1}^{T_n/2} \max_{1\le i\le q_{n1}} \max_{1\le j\le q_{n2}} |\mathbb{E}\big(\Delta_{n,2w}(t_i, s_j) \mid \mathcal{F}^*_{n,w-1}\big)| \\ &\le C\,(n/c_n)\, n^{2d\nu-\alpha} = O_{a.s.}(\eta_n \log n),\end{aligned} \tag{3.41}$$

which yields (3.38).

We next prove (3.39). Using Lemma 3.2 with $t_1 = 0, t_2 = 2sb_n + 1$ and $t_3 = (2s+1)b_n$, for any integer $m \ge 1$, we haev

$$\begin{aligned}\sup_{t,s} \mathbb{E}|\Delta_{n,2w}(t,s)|^m &\le H_0^m (m+1)!\,(n/c_n) \left\{1 + \left[(n/c_n)(n/b_n)^{d-1}\right]^{m-1}\right\} \\ &\le 2H_0^m (m+1)! (n/c_n)^m (n/b_n)^{(d-1)(m-1)},\end{aligned}$$

because $c_n/n^{1-\nu(1-d)} \le c_n/n^{1-\epsilon_0} \to 0$. By virtue of this fact, it follows that

$$\begin{aligned}E\lambda_{3n} &\le 2\sum_{i=1}^{q_{n1}} \sum_{j=1}^{q_{n2}} \sum_{w=1}^{T_n/2} \mathbb{E}|\Delta_{n,2w}(t_i, s_j)| I(|\Delta_{n,2w}(t_i, s_j)| > M_1\, \eta_n) \\ &\le q_{n1} q_{n2} T_n\, H_0^m (m+1)! (n/c_n) \Big[\frac{(n/c_n)(n/b_n)^{d-1}}{M_1\, \eta_n}\Big]^{m-1} \\ &\le C\, n^{22} (H_0/M_1)^m\,(m+1)! \log^{-(m-1)} n,\end{aligned}$$

due to (3.34) and $\nu > 0$. Now, by taking $m = \log n$ and letting $M_1 \ge 25H_0$, it follows from the Stirling approximation of $(m+1)!$ that for any $\epsilon > 0$,

$$\begin{aligned}
P[\lambda_{3n} > \epsilon, i.o.] &\le \lim_{s\to\infty} \sum_{n=s}^{\infty} \epsilon^{-1} E\lambda_{3n} \\
&\le C \lim_{s\to\infty} \sum_{n=s}^{\infty} \epsilon^{-1} n^{22} \log^5 n \exp\{-(M_1/H_0)\, \log n\} \\
&\le C \lim_{s\to\infty} \sum_{n=s}^{\infty} \epsilon^{-1} n^{-3} \log^5 n \to 0, \qquad (3.42)
\end{aligned}$$

which implies that $\lambda_{3n} = o_{a.s.}(1)$, and hence (3.39) follows.

We finally consider (3.37). First note that, by Lemma 3.1, for any $|t-s| \le c_n n^{-\alpha}$,

$$\begin{aligned}
&E[\Delta_{nw}^{*2}(t,s)|\mathcal{F}_{n,w-1}^{*}] \le 2E[\Delta_{n,2w}^{2}(t,s)|\mathcal{F}_{n,(2w-1)b_n}] \\
&\le \sum_{k=2wb_n+1}^{(2w+1)b_n} E\Big(f_{s,t}^2(x_{k,n}) \,|\, \mathcal{F}_{n,(2w-1)b_n}\Big) \\
&\quad + 2 \sum_{2wb_n+1\le k<j\le(2w+1)b_n} \Big|E\Big(f_{s,t}(x_{k,n}) f_{s,t}(x_{j,n}) \Big| \mathcal{F}_{n,(2w-1)b_n}\Big)\Big| \\
&\le C\,(n/c_n)(n/b_n)^{d-1} + 2 \sum_{2wb_n+1\le k<j\le(2w+1)b_n} E\Big(|f_{s,t}(x_{k,n})|\,|I_{k,j}| \,\Big|\, \mathcal{F}_{n,(2w-1)b_n}\Big) \\
&\le C\,(n/c_n)(n/b_n)^{d-1} + C\, n^{2d}\, c_n^{-2}\, b_n^{-d} \sum_{2wb_n+1\le k<j\le(2w+1)b_n} (j-k)^{-d} \min\{n^{d-\alpha}(j-k)^{-d}, 1\} \\
&\le C\,(n/c_n)(n/b_n)^{d-1} + C\, n^{2d}\, c_n^{-2}\, b_n^{1-d} \sum_{k=1}^{b_n} k^{-d} \min\{n^{-\alpha}(n/k)^{d}, 1\} \\
&\le C\,(n/c_n)(n/b_n)^{d-1}\big[1 + n^{1-\eta_0}/c_n\big], \qquad (3.43)
\end{aligned}$$

where $I_{k,j} = E\big[f_{t,s}(x_{j,n})|\mathcal{F}_{n,k}\big]$ and

$$\eta_0 = \begin{cases} \alpha + \nu(1-2d), & if \quad 0 < d < 1/2, \\ \alpha/4, & if \quad d = 1/2, \\ \big(\frac{1-d}{d}\big)\alpha, & if \quad 1/2 < d < 1. \end{cases} \qquad (3.44)$$

and we have used the fact: for $0 < d < 1$, letting $\zeta = \alpha/d$,

$$\sum_{k=1}^{b_n} k^{-d} \min\{n^{-\alpha}(n/k)^{d}, 1\} \;\le\; \sum_{k=1}^{n^{1-\zeta}} k^{-d} + n^{d-\alpha} \sum_{k=n^{1-\zeta}+1}^{b_n} k^{-2d} \le C\, n^{1-d-\eta_0}.$$

It follows from this estimate that

$$
\begin{aligned}
& \max_{1\le i\le q_{n1}} \max_{1\le j\le q_{n2}} \sum_{w=1}^{T_n/2} \mathbb{E}[\Delta_{nw}^{*2}(t_i, y_j) \mid \mathcal{F}_{n,w-1}^*] \\
\le\ & C\,(n/c_n)(n/b_n)^d \big[1 + n^{1-\eta_0}/c_n\big] \le \begin{cases} C(n/c_n)^2\, n^{d\nu-\eta_0}, & if \quad \eta_0 \le \epsilon_0, \\ C(n/c_n)\, n^{d\nu}, & if \quad \eta_0 > \epsilon_0 \end{cases} \\
\le\ & C\,\eta_n^2 \log n, \quad a.s.
\end{aligned}
$$

due to $d\nu - \eta_0 < 0$ by simple calculation and $(n/c_n)n^{-d\nu} < n^{1-\epsilon_0}/c_n \to 0$. This, together with the facts that $|\Delta_{nw}^*(t_i, y_j)| \le \eta_n$ and for each i, j, $\{\Delta_{nw}^*(t_i, s_j), \mathcal{F}_{n,w}^*\}$ forms a martingale difference, and the well-known martingale exponential inequality (see, e.g., de la Pena (1999)) implies that there exists a $M_0 \ge 22$ such that, as $n \to \infty$,

$$
\begin{aligned}
& P[\lambda_{1n} \ge M_0 \eta_n \log n, i.o.] \\
\le\ & P\Big[\lambda_{1n} \ge M_0\eta_n \log n, \ \max_{1\le i\le q_{n1}} \max_{1\le j\le q_{n2}} \sum_{w=1}^{T_n/2} \mathbb{E}[\Delta_{ns}^{*2}(t_i, y_j) \mid \mathcal{F}_{n,w-1}^*] \le C\,\eta_n^2 \log n, i.o.\Big] \\
\le\ & \lim_{s\to\infty} \sum_{n=s}^{\infty} P\Big[\lambda_{1n} \ge M_0\eta_n \log n, \ \max_{1\le i\le q_{n1}} \max_{1\le j\le q_{n2}} \sum_{w=1}^{T_n/2} \mathbb{E}[\Delta_{ns}^{*2}(t_i, y_j) \mid \mathcal{F}_{n,w-1}^*] \le C\,\eta_n^2 \log n\Big] \\
\le\ & \lim_{s\to\infty} \sum_{n=s}^{\infty} \sum_{i=1}^{q_{n1}} \sum_{j=1}^{q_{n2}} P\Big[\sum_{w=1}^{T_n/2} \Delta_{nw}^*(t_i, y_j) \ge M_0\eta_n \log n, \ \sum_{w=1}^{T_n/2} \mathbb{E}[\Delta_{nw}^{*2}(t_i, y_j) \mid \mathcal{F}_{n,w-1}^*] \le C\,\eta_n^2 \log n\Big] \\
\le\ & \lim_{s\to\infty} \sum_{n=s}^{\infty} q_{n1}q_{n2} \exp\Big\{-\frac{M_0^2 \log^2 n}{2C\log n + 2M_0 \log n}\Big\} \\
\le\ & \lim_{s\to\infty} \sum_{n=s}^{\infty} q_{n1}q_{n2} \exp\{-M_0 \log n\} = 0,
\end{aligned} \tag{3.45}
$$

where the last inequality follows from (3.34). This yields $\lambda_{1n} = O_{a.s.}\big(\eta_n \log n\big)$. Combining (3.41)-(3.71), we establish (3.36), and also completes the proof of Lemma 3.5. □

The following lemma is an extension of Theorem 2.1 in Wang and Chan (2014), where we use the notation $||x|| = \max_{1\le i\le d} |z_i|$ if $x = (z_1, ..., z_d)$.

Lemma 3.5. *Suppose that*

(a) $(e_k, z_k), k \ge 1$, *is a sequence of random vectors on* $R \times R^d$, $d \ge 1$. $\{e_t, \mathcal{F}_t\}_{t\ge1}$ *is a martingale difference, where* $\mathcal{F}_t = \sigma(z_1, ..., z_{t+1}, e_1, ..., e_t)$, *satisfying* $\sup_{t\ge1} E(|e_t|^{2p} \mid \mathcal{F}_{t-1}) < \infty, a.s.$, *for some* $p \ge 1$;

(b) $f_n(x,y), n \geq 1$, *is a sequence of real functions on* $R^d \times R^{d_1}$, *where* $d, d_1 \geq 1$, *satisfying* $\sup_{n,x,y} |f_n(x,y)| < \infty$ *and there exists an* $\alpha > 0$ *such that, whenever* $||a||$ *is sufficiently small,*

$$\sup_{n,x,y} |f_n(x, y+a) - f_n(x,y)| \leq C\, n^{\alpha}\, \|a\|; \tag{3.46}$$

(c) *there exist positive constant sequences* $\gamma_n \to \infty$ *and* $b_n = O(n^k)$ *for some* $k > 0$ *such that*

$$\sup_{\|y\| \leq b_n} \sum_{k=1}^{n} f_n^2(z_k, y) \quad = \quad O_P(\gamma_n). \tag{3.47}$$

Then, for any $n\, \gamma_n^{-p} \log^{p-1} n = O(1)$, *where* p *is given in (a), we have*

$$\sup_{\|y\| \leq b_n} \Big| \sum_{k=1}^{n} e_t\, f_n(z_k, y) \Big| \quad = \quad O_P\big[(\gamma_n \log n)^{1/2}\big]. \tag{3.48}$$

Proof. It is similar to Theorem 2.1 of Wang and Chan (2014), which is restated in Appendix for convenience of reading.

Lemma 3.6. *Under the conditions of Theorem 2.3, we have*

$$\Psi_n \quad := \quad \sup_{y \in R} \Big| \sum_{k=1}^{n} \frac{\sigma(x_k)}{\sigma(y)} K_j\big(\frac{x_k - y}{h}\big)\, u_k \Big| = O_P\big[(nh/d_n)^{1/2} \log^{1/2} n\big], \tag{3.49}$$

where $K_j(x) = x^j K(x)$. *If in addition* $E(u_t^2 \mid \mathcal{F}_{t-1}) \to 1, a.s.$ *and* $\sup_{t \geq 1} E(|u_t|^{4p} \mid \mathcal{F}_{t-1}) < \infty$, *where* $p \geq 1 + 1/\epsilon_0$ *for some* $\epsilon_0 > 0$, *result (3.49) still holds when* u_k *is replaced by* $u_k^2 - 1$.

Proof. Let $f_n(x,y) = \frac{\sigma(x)}{\sigma(y)} K_j\big(\frac{x-y}{h}\big)$. First note that, for some η_0 sufficiently small,

$$C_0 := \sup_{y, |x-y| \leq \eta_0} \frac{\sigma(x)}{\sigma(y)} < \infty,$$

due to (2.23). This, together with $K_j(z) = 0$ for $|z| \geq A$, implies that

$$|f_n(x,y)| \leq C_0\, \big|K_j\big(\frac{x-y}{h}\big)\big|, \tag{3.50}$$

uniformly for n, x and y, whenever h is sufficiently small. Similarly, by noting

$$K_j\big(\frac{x-y-a}{h}\big) - K_j\big(\frac{x-y}{h}\big) = 0$$

if $|x-y| \geq hA + |a|$, it follows from (2.23) and Assumption 2.7 that

$$
\begin{aligned}
|f_n(x, y+a) - f_n(x,y)| &\leq \frac{\sigma(x)}{\sigma(y+a)} \big| K_j\big(\frac{x-y-a}{h}\big) - K_j\big(\frac{x-y}{h}\big)\big| \\
&\qquad + |f_n(x,y)| \frac{|\sigma(y) - \sigma(y+a)|}{\sigma(y+a)} \\
&\leq C(h^{-1}+1)|a| \leq C\, n|a|, \qquad (3.51)
\end{aligned}
$$

uniformly for n, x and y, whenever $nh \to \infty$ and $|a|$ are sufficiently small.

By virtue of (3.50), $\sup_{n,x,y} |f_n(x,y)| < \infty$ and (3.47) holds with $\gamma_n = nh/d_n$ due to Corollary 2.2. Similarly, we have (3.46) with $\alpha = 1$ due to (3.51). Furthermore, by recalling $p > 1/\epsilon_0$, we have

$$
n\, \gamma_n^{-p} \log^{p-1} n \leq (n^{1-\epsilon_0} h/d_n)^{-p}\, n^{\epsilon_0 p - 1} \log^{p-1} n = o(1).
$$

Now it follows from Lemma 3.5 that, for any $\eta > 0$, there exists a $M_0 > 0$ such that

$$
\begin{aligned}
& P\big(\Psi_n \geq M_0 (nh/d_n)^{1/2} \log^{1/2} n\big) \\
\leq\; & P\big(\max_{1\leq k\leq n} |x_k| \geq n^4/2\big) + P\Big(\sup_{\|y\|\leq n^4} \Big| \sum_{k=1}^n e_t\, f_n(z_k, y)\Big| \geq M_0(nh/d_n)^{1/2} \log^{1/2} n\Big) \\
\leq\; & n^{-4} \sum_{k=1}^n E|x_k| + \eta \leq 2\eta,
\end{aligned}
$$

when n is sufficiently large, where we have used the facts that $E|x_k| \leq nE|\xi_1|$ and, whenever $\max_{1\leq k\leq n} |x_k| \leq n^4/2$,

$$
\sup_{\|y\|>n^4} \Big| \sum_{k=1}^n e_t\, f_n(z_k, y)\Big| = 0,
$$

due to $K_j(z) = 0$ for $|z| \geq A$.

This proves (3.49). Under the additional conditions on u_k, the proof of (3.49) is similar, and hence the details are omitted. □

3.2 Proof of Theorem 2.1. First consider $\tau = \int g(x)dx \neq 0$ and, without loss of generality, assume $\tau = 1$. Define $\bar{g}(x) = g(x) I\{|x| \leq n^{\zeta}/2\}$, where $0 < \zeta < 1 - \delta/\rho$ is small enough such that $n^{\zeta}/c_n \leq n^{-\delta}$, where ρ and δ are given in Assumption 2.1 and 2.2 respectively. Further let $\varepsilon = n^{-\alpha}$ with $0 < \alpha < \min\{\delta/2, \zeta(\rho-1)\}$, and for a fixed x_0, define a triangular function

$$
g_{x_0\varepsilon}(y) = \begin{cases} 0, & |y - x_0| > \varepsilon, \\ \frac{y - x_0 + \varepsilon}{\varepsilon^2}, & x_0 - \varepsilon \leq y \leq x_0, \\ \frac{x_0 + \varepsilon - y}{\varepsilon^2}, & x_0 \leq y \leq x_0 + \varepsilon. \end{cases}
$$

It suffices to show that

$$\Phi_{1n} := \sup_{x\in R}\Big|\frac{c_n}{n}\sum_{j=1}^n \big\{g(c_n(x_{j,n}+x)) - \bar g(c_n(x_{j,n}+x))\big\}\Big| = o_P(\log^{-\beta} n), \tag{3.52}$$

$$\begin{aligned}\Phi_{2n} &:= \sup_{x\in R}\Big|\frac{c_n}{n}\sum_{j=1}^n \bar g(c_n x_{j,n} - c_n x - n^{\zeta}/2) - \frac{1}{n}\sum_{j=1}^n g_{x\varepsilon}(x_{j,n})\Big| \\ &= o_P(\log^{-\beta} n),\end{aligned} \tag{3.53}$$

$$\Phi_{3n} := \sup_{x\in R}\Big|\frac{1}{n}\sum_{j=1}^n g_{x\varepsilon}(x_{j,n}) - L_{G_n}(1,x)\Big| = o_P(\log^{-\beta} n), \tag{3.54}$$

$$\Phi_{4n} := \sup_{x\in R}\Big|L_{G_n}(1,x) - L_{G_n}(1, x+n^{\zeta}/(2c_n))\Big| = o_P(\log^{-\beta} n). \tag{3.55}$$

Indeed it follows from (3.52)–(3.55) that

$$\begin{aligned}&\sup_{x\in R}\Big|\frac{c_n}{n}\sum_{j=1}^n g\big[c_n(x_{k,n}+x)\big] - L_{G_n}(1,-x)\Big| \\ =\ &\sup_{x\in R}\Big|\frac{c_n}{n}\sum_{j=1}^n g\big(c_n x_{k,n} - c_n x - n^{\zeta}/2\big) - L_{G_n}(1, x+n^{\zeta}/(2c_n))\Big| \\ \le\ &\Phi_{1n}+\Phi_{2n}+\Phi_{3n}+\Phi_{4n} = o_P(\log^{-\beta} n),\end{aligned}$$

which yields the required (2.14).

The proof of (3.52) is simple. It follows from $\sup_x |x|^{\rho}\,|g(x)| < \infty$ that

$$\Phi_{1n} \ \le\ c_n \sup_{|x|\ge n^{\zeta}/2} |g(x)| I\{|x| > n^{\zeta}/2\} \le C\, n^{-\zeta\,\rho} c_n = o(\log^{-\beta} n),$$

as $n^{\zeta}/c_n \le n^{-\delta}$ and $\rho > \delta/(1-\zeta)$.

Recall $\{G_n(t); 0\le t\le 1\} =_D \{G(t); 0\le t\le 1\}$ for all $n\ge 1$, by Assumption 2.2. For any $\epsilon>0$ and $\beta>0$, we have

$$\begin{aligned}&P(|\Phi_{4n}| \ge \epsilon\,\log^{-\beta} n) \\ =\ &P\big(\sup_{x\in R}\big|L_G(1,x) - L_G(1, x+n^{\zeta}/(2c_n))\big| \ge \epsilon\,\log^{-\beta} n\big) \\ \to\ &0, \quad \text{as } n\to\infty,\end{aligned}$$

due to (2.3) and $n^{\zeta}/c_n \le n^{-\delta}$. This yields (3.55).

Recalling the definition of $g_{x\varepsilon}(y)$ and $\int_{-\infty}^{\infty} g_{x\varepsilon}(y)dy = 1$, it follows again from (2.3) that

$$\begin{aligned}&\Big|\int_0^1 g_{x\varepsilon}(G(t))dt - L_G(1,x)\Big| \\ =\ &\Big|\int_{-\infty}^{\infty} g_{x\varepsilon}(y) L_G(1,y)dy - L_G(1,x)\Big|\end{aligned}$$

$$\leq \int_{-\infty}^{\infty} g_{x\varepsilon}(y)|L_G(1,y) - L_G(1,x)|dy = O_{a.s.}(\epsilon^\beta)$$

uniformly for all $x \in R$. This implies $\sup_x \big| \int_0^1 g_{x\varepsilon}(G_n(t))dt - L_{G_n}(1,x)\big| = O_P(\epsilon^\beta)$ by using the similar arguments as in the proof of (3.55). Hence it follows from (**??**)in Assumption 2.2 that

$$\begin{aligned}
&\Big|\frac{1}{n}\sum_{j=1}^n g_{x\varepsilon}(x_{j,n}) - L_{G_n}(1,x)\Big| \\
\leq\ & \Big|\int_0^1 g_{x\varepsilon}(x_{[nt],n})dt - \int_0^1 g_{x\varepsilon}(G_n(t))dt\Big| + 2/(\varepsilon n) + \Big|\int_0^1 g_{x\varepsilon}(G_n(t))dt - L_{G_n}(1,x)\Big| \\
=\ & O_P\big[\varepsilon^{-2}n^{-\delta} + 2/(\varepsilon n) + \varepsilon^\beta\big] \\
=\ & O_P\big[n^{2\alpha-\delta} + 2n^{\alpha-1} + n^{-\alpha\beta/2}\big] = o_P(\log^{-\beta} n)
\end{aligned}$$

uniformly for all $x \in R$, as $\alpha < \delta/2$, which implies (3.54).

We finally prove (3.53). Let $\bar{g}_{x\varepsilon n}(z)$ be the step function which takes the value $g_{x\varepsilon}(x + kn^\zeta/c_n)$ for $z \in [x + kn^\zeta/c_n, x + (k+1)n^\zeta/c_n), k \in \mathbb{Z}$. It suffices to show that, uniformly for all $x \in R$, (letting $\bar{g}_j(y) = \bar{g}(c_n x_{j,n} - y - n^\zeta/2)$),

$$\Delta_{1n}(x) := \Big|\frac{1}{n}\sum_{j=1}^n g_{x\varepsilon}(x_{j,n}) - \frac{1}{n}\sum_{j=1}^n \bar{g}_{x\varepsilon n}(x_{j,n}) \int_{-\infty}^{\infty} \bar{g}_j(y)dy\Big| = o_P(\log^{-\beta} n) \tag{3.56}$$

$$\Delta_{2n}(x) := \Big|\frac{1}{n}\sum_{j=1}^n \bar{g}_{x\varepsilon n}(x_{j,n}) \int_{-\infty}^{\infty} \bar{g}_j(y)dy - \int_{-\infty}^{\infty} \frac{1}{n}\sum_{j=1}^n g_{x\varepsilon}(y/c_n)\bar{g}_j(y)dy\Big|$$
$$= o_P(\log^{-\beta} n), \tag{3.57}$$

$$\Delta_{3n}(x) := \Big|\int_{-\infty}^{\infty} \frac{1}{n}\sum_{j=1}^n g_{x\varepsilon}(y/c_n)\bar{g}_j(y)dy - \frac{c_n}{n}\sum_{j=1}^n \bar{g}(c_n x_{j,n} - c_n x - n^\zeta/2)\Big|$$
$$= o_P(\log^{-\beta} n), \tag{3.58}$$

(3.56) first. Note that $|g_{x\varepsilon}(y) - g_{x\varepsilon}(z)| \leq \varepsilon^{-2}|y-z|$, and

$$\begin{aligned}
|\bar{g}_{x\varepsilon\, n}(y) - g_{x\varepsilon}(z)| &\leq |\bar{g}_{x\varepsilon\, n}(y) - g_{x\varepsilon}(y)| + |g_{x\varepsilon}(y) - g_{x\varepsilon}(z)| \\
&\leq C\varepsilon^{-2}(n^\zeta/c_n + |y-z|).
\end{aligned} \tag{3.59}$$

It follows that, uniformly for all $j = 1, ..., n$ and $x \in R$,

$$\begin{aligned}
&\Big|g_{x\varepsilon}(x_{j,n}) - \bar{g}_{x\varepsilon n}(x_{j,n}) \int_{-\infty}^{\infty} \bar{g}_j(y)dy\Big| \\
&\leq \Big|g_{x\varepsilon}(x_{j,n}) - \bar{g}_{x\varepsilon n}(x_{j,n})\Big| + |\bar{g}_{x\varepsilon n}(x_{j,n})|\Big|1 - \int_{-\infty}^{\infty} \bar{g}_j(y)dy\Big| \\
&\leq C\varepsilon^{-2}n^\zeta/c_n + C_1 n^{-\zeta(\rho-1)} = o_P(\log^{-\beta} n).
\end{aligned}$$

where we have used the fact that (recalling $\int g(y)dy = 1$),

$$\Big|1 - \int_{-\infty}^{\infty} \bar{g}_j(y)dy\Big| \leq \Big| \int_{-\infty}^{\infty} g(y)I\{|y| > n^{\zeta}/2\}dy\Big| \leq C\ n^{-\zeta(\rho-1)}$$

due to $\sup_y |y|^{\rho}|g(y)| < \infty$ and $\rho > 1$.

(3.57) next. By (3.59) and the definition of $\bar{g}_j(y)$, we have

$$\begin{aligned}
&\int_{-\infty}^{\infty} |\bar{g}_{x\varepsilon n}(x_{j,n})\bar{g}_j(y) - g_{x\varepsilon}(y/c_n)\bar{g}_j(y)|dy \\
&\leq \Big(\int_{-\infty}^{\infty} g(y)dy\Big)\Big(\sup_y \big|\bar{g}_{x\varepsilon n}(x_{j,n}) - g_{x\varepsilon}(y/c_n)\big|\, I\{|c_n x_{j,n} - y - n^{\zeta}/2| \leq n^{\zeta}/2\}\Big) \\
&\leq C \sup_y \Big[\varepsilon^{-2}(n^{\zeta}/c_n + |x_{j,n} - y/c_n|)\, I\big\{\big|x_{j,n} - y/c_n - n^{\zeta}/(2c_n)\big| \leq n^{\zeta}/(2c_n)\big\}\Big] \\
&\leq C\varepsilon^{-2}(n^{\zeta}/c_n) = o_P(\log^{-\beta} n).
\end{aligned}$$

uniformly for all $j = 1, ..., n$ and $x \in R$.

Finally for (3.58). Using Lemma 3.4, we have

$$\begin{aligned}
\Delta_{3n} &= \Big| \int_{-\infty}^{\infty} \frac{1}{n}\sum_{j=1}^{n} g_{x\varepsilon}(y/c_n)\bar{g}(c_n x_{j,n} - y - n^{\zeta}/2)dy \\
&\qquad\qquad - \frac{c_n}{n}\sum_{j=1}^{n} \bar{g}(c_n x_{j,n} - c_n x - n^{\zeta}/2)\Big| \\
&\leq \sup_{|y - c_n x| \leq c_n\varepsilon} \Big|\frac{c_n}{n}\sum_{j=1}^{n}\big\{\bar{g}(c_n x_{j,n} - y - n^{\zeta}/2) - \bar{g}(c_n x_{j,n} - c_n x - n^{\zeta}/2)\big\}\Big| \\
&\qquad\qquad \times \Big(\frac{1}{c_n}\int_{-\infty}^{\infty} g_{x\varepsilon}(y/c_n)dy\Big) \\
&= o_P(\log^{-\beta} n). \qquad (3.60)
\end{aligned}$$

uniformly in $x \in R$. The proof of (2.6) for $\tau \neq 0$ is now complete.

The proof of (2.6) for $\tau = 0$ is similar except more simpler. Indeed, under the notation above, we have

$$\sup_{x\in R}\Big|\frac{c_n}{n}\sum_{j=1}^{n} g\big[c_n(x_{k,n} + x)\big]\Big| \ \leq\ \Phi_{1n} + \Delta_{2n} + \Delta_{3n} + \widetilde{\Delta}_{1n}, \qquad (3.61)$$

where

$$\widetilde{\Delta}_{1n} = \sup_{x\in R}\Big|\frac{1}{n}\sum_{j=1}^{n}\bar{g}_{x\varepsilon n}(x_{j,n})\int_{-\infty}^{\infty}\bar{g}_j(y)dy\Big|.$$

Recalling $|\bar{g}_{x\varepsilon n}(x)| \leq \epsilon^{-1} = n^{\alpha}$, $\alpha < \zeta(\rho - 1)$ and

$$|\int_{-\infty}^{\infty}\bar{g}_j(y)dy| \ \leq\ |\int_{-\infty}^{\infty} g(y)I\{|y| > n^{\zeta}/2\}dy\Big| \leq C\ n^{-\zeta(\rho-1)}$$

due to $\int g(x)dx = 0$, $\sup_y |y|^\rho |g(y)| < \infty$ and $\rho > 1$, it is readily seen that

$$\widetilde{\Delta}_{1n} \le C\, n^{\alpha-\zeta(\rho-1)} = O(\log^{-\beta} n),$$

for any $\beta > 0$. Taking this estimates, (3.52), (3.57) and (3.58) into (3.61), we obtain the claim required. This completes the proof of Theorem 2.1. □

3.3. Proof of Theorem 2.3. Let $V_n(x) = \sum_{i=1}^n w_i(x)$. It is readily seen that

$$\widehat{m}_n(x) - m(x) \;=\; \Gamma_{1n}(x) + \Gamma_{2n}(x), \tag{3.62}$$

where $\Gamma_{1n}(x) = V_n^{-1}(x) \sum_{i=1}^n w_i(x)\sigma(x_i)u_i$ and

$$\Gamma_{2n}(x) \;=\; V_n^{-1}(x) \sum_{i=1}^n w_i(x)\big[m(x_i) - m(x)\big].$$

Recall $V_{n,j}(x) = \sum_{i=0}^n K_j\big(\frac{x_i - x}{h}\big)$ with $K_j(x) = x^j K(x)$, $j = 0, 1, 2$. It follows from (2.15) in Corollary 2.2 with $g(x) = K_1(x)$ that

$$\sup_{x\in R} |V_{n,1}(x)| = \sup_{x\in R} \Big|\sum_{i=1}^n K_1\Big(\frac{x_i - x}{h}\Big)\Big| = O_P\big[(nh/d_n)\, \log^{-\beta} n\big], \tag{3.63}$$

for any $\beta > 0$. Similarly, by Corollary 2.2 with $g(x) = K_j(x)$, we get

$$\sup_{x\in R} |V_{n,j}(x)| \;=\; O_P(nh/d_n), \quad \big\{\inf_{|x|\le b_n} |V_{n,j}(x)|\big\}^{-1} = O_P\big[d_n/(nh)\big],$$

for $j = 0$ and 2. It follows from these facts that

$$\begin{aligned}
\Big\{\inf_{|x|\le b_n} |V_n(x)|/V_{n,2}(x)\Big\}^{-1} &\le \Big\{\inf_{|x|\le b_n} \Big|V_{n,0}(x) - \frac{V_{n,1}^2(x)}{V_{n,2}(x)}\Big|\Big\}^{-1} \\
&= \Big\{\inf_{|x|\le b_n} |V_{n,0}(x)| - o_P(nh/d_n)\Big\}^{-1} \\
&= O_P\big[d_n/(nh)\big]
\end{aligned}$$

and $\sup_{|x|\le b_n} |V_{n,1}(x)|/V_{n,2}(x) = o_P(\log^{-\beta} n)$ for any $\beta > 0$. Now it follows from Lemma 3.6 that

$$\begin{aligned}
\sup_{|x|\le b_n} \frac{|\Gamma_{1n}(x)|}{\sigma(x)} &\le \Big\{\inf_{|x|\le b_n} \frac{|V_n(x)|}{V_{n,2}(x)}\Big\}^{-1} \Big\{\sup_{|x|\le b_n} \Big|\sum_{i=1}^n K\Big(\frac{x_i - x}{h}\Big)\frac{\sigma(x_i)}{\sigma(x)} u_i\Big| \\
&\qquad + \sup_{|x|\le b_n} \frac{|V_{n,1}(x)|}{V_{n,2}(x)} \Big|\sum_{i=1}^n K_1\Big(\frac{x_i - x}{h}\Big)\frac{\sigma(x_i)}{\sigma(x)} u_i\Big|\Big\} \\
&= O_P\Big[\Big(\frac{d_n}{nh}\Big)^{1/2} \log^{1/2} n\Big].
\end{aligned} \tag{3.64}$$

To consider $\Gamma_{2n}(x)$, note that $\sum_{i=1}^n w_i(x)(x_i - x) = 0$ and by Assumption 2.8,

$$\begin{aligned}|m(y) - m(x) - m'(x)(y-x)| &= \Big| \int_x^y m'(s) - m'(x) ds \Big| \\ &\le m_0(x) \int_x^y |s-x|^\tau ds = m_0(x)|y-x|^{\tau+1}.\end{aligned}$$

It follows from these facts and Assumption 2.7 that

$$\begin{aligned}
\sup_{|x|\le b_n} \frac{|\Gamma_{2n}(x)|}{\sigma(x)} &= \sup_{|x|\le b_n} \frac{|\sum_{i=1}^n w_i(x)\big[m(x_i) - m(x) - m'(x)(x_i - x)|\big]}{\sigma(x) V_n(x)} \\
&\le \sup_{|x|\le b_n} \frac{|m_0(x)|}{2\sigma(x)} \frac{\sum_{i=1}^n |w_i(x)||x_i - x|^{\tau+1}}{V_n(x)} \\
&\le C\,\delta_n \sup_{|x|\le b_n} \Big| \frac{\sum_{i=1}^n |x_i - x|^{\tau+1} K[(x_i - x)/h]}{V_{n,2}^{-1}(x) V_n(x)} + \\
&\qquad \frac{[\sum_{i=1}^n |x_i - x|^{\tau+2} K[(x_i - x)/h]]}{V_{n,2}^{-1}(x) V_n(x)} \Big\{ \frac{|V_{n,1}(x)|}{V_{n,2}(x)} \Big\} \Big| \\
&\le \frac{C d_n\, h^{\tau+1} \delta_n}{nh} \sup_{|x|\le b_n} \sum_{i=1}^n K\Big(\frac{x_i - x}{h}\Big) \\
&\le C\, h^{\tau+1} \delta_n. \qquad (3.65)
\end{aligned}$$

Taking (3.64) and (3.65) into (3.62), we prove (2.24). □

3.4. Proof of Theorem 2.4. First note that, due to Assumption 2.9 and $K(s) = 0$ if $|s| \ge A$,

$$\frac{|\sigma^i(x_k) - \sigma^i(x)|}{\sigma^i(x)} K\big[(x_k - x)/h\big] \;\le\; C\, h\, K\big[(x_k - x)/h\big],$$

for $i = 1, 2$, all $x \in R$ and h sufficiently small. Similarly, whence $b_n \ge Ah$, we have

$$\begin{aligned}&|\widehat{m}_n(x_k) - m(x_k)|^i\, K\big[(x_k - x)/h\big]/\sigma^i(x) \\ &\qquad \le \frac{\sigma^i(x_k)}{\sigma^i(x)} K\big[(x_k - x)/h\big] \sup_{|x|\le 2b_n} \big\{ |\widehat{m}_n(x) - m(x)|^i / \sigma^i(x) \big\},\end{aligned}$$

for $i = 1, 2$, $|x| \le b_n$ and h sufficiently small. By virtue of these estimates, simple calculations show that (recalling $V_{n,0}(x) = \sum_{k=1}^n K\big[(x_k - x)/h\big]$)

$$
\begin{aligned}
\frac{|\widehat{\sigma}^2(x) - \sigma^2(x)|}{\sigma^2(x)} &\le V_{n,0}^{-1}(x)\Big|\sum_{k=1}^n K\big(\frac{x_k - x}{h}\big)(u_k^2 - 1)\Big| \\
&\quad + V_{n,0}^{-1}(x)\sum_{k=1}^n K\big(\frac{x_k - x}{h}\big)\Big\{\frac{|\widehat{m}_n(x_k) - m(x_k)|^2}{\sigma^2(x)} \\
&\qquad + \frac{|\sigma^2(x_k) - \sigma^2(x)|}{\sigma^2(x)}u_k^2 + 2\frac{|\widehat{m}_n(x_k) - m(x_k)|}{\sigma(x)}\frac{\sigma(x_k)}{\sigma(x)}|u_k|\Big\} \\
&\le V_{n,0}^{-1}(x)\Big|\sum_{k=1}^n K\big(\frac{x_k - x}{h}\big)(u_k^2 - 1)\Big| \\
&\quad + V_{n,0}^{-1}(x)\Delta_n\sum_{k=1}^n \big[1 + \frac{\sigma^2(x_k)}{\sigma^2(x)}\big](1 + u_k^2)\big]K\big(\frac{x_k - x}{h}\big),
\end{aligned}
$$

for $|x| \le b_n$ and h sufficiently small, where

$$
\Delta_n = h + \sup_{|x|\le 2b_n}\frac{|\widehat{m}_n(x) - m(x)|}{\sigma(x)} + \sup_{|x|\le 2b_n}\frac{|\widehat{m}_n(x) - m(x)|^2}{\sigma^2(x)}.
$$

Consequently, by using Lemmas 3.6, we get

$$
\sup_{|x|\le b_n}\frac{|\widehat{\sigma}^2(x) - \sigma^2(x)|}{\sigma^2(x)} = O_P\Big\{h + \big(nh/d_n\big)^{-1/2}\log^{1/2} n + h^{1+\tau}\,\delta_n\Big\},
$$

which yields (2.25).

It follows from this estimate that

$$
\begin{aligned}
\sup_{|x|\le b_n}\left|\frac{\sigma(x)}{\widehat{\sigma}(x)} - 1\right| &\le \sup_{|x|\le b_n}\frac{\sigma(x)}{\widehat{\sigma}(x)}\,\frac{|\widehat{\sigma}^2(x) - \sigma^2(x)|}{\sigma^2(x)} \\
&= o_P(1)\sup_{|x|\le b_n}\frac{\sigma(x)}{\widehat{\sigma}(x)} = o_P(1).
\end{aligned}
$$

This, together with (2.24), implies (2.26), and also completes the proof of Theorem 2.4. □

REFERENCES

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

Bandi, F. (2004). On Persistence and nonparametric estimation (with an application to stock return predictability). *Unpublished manuscript.*

Berkes, I., and Horváth, L., 2006. Convergence of integral functionals of stochastic processes. Econometric Theory, 22(2), 304–322.

Bingham, N. H. (1989). *Regular Variation.* Cambridge University Press.

Borodin, A. N. and Ibragimov, I. A. (1995). *Limit theorems for functionals of random walks.* Proceedings of the Steklov Institute of Mathematics, American Mathematical Society

Cai, Z., Li, Q. and Park, J. Y. (2009). Functional-coefficient models for nonstationary time series data. *Journal of Econometrics*, **148**, 101–113.

Chan, N. and Wang, Q. (2014). Uniform convergence for non-parametric estimators with non-stationary data. *Econometric Theory*, *online.*

Chen, J., Gao, J. and Li, D. (2010). Semiparametric Regression Estimation in Null Recurrent Nonlinear Time Series. *Unpublished manuscript.*

Choi, I. and Saikkonen, P. (2004). Testing linearity in cointegrating smooth transition regressions. *The Econometrics Journal,* **7**, 341–365.

Choi, I. and Saikkonen, P. (2009). Tests for nonlinear cointegration. *Econometric Theory*, **26**, 682–709.

Csörgö, M. and Révész, P. (1981). *Strong approximations in probability and statistics.* Probability and Mathematical Statistics. Academic Press, Inc., New York-London

Duffy, J. (2014 a). A Uniform law for the convergence to local time, *unpublished manuscript,* Yale University.

Duffy, J. (2014 b). Uniform in bandwidth converbgence rates, on a maximal domain, in structure nonparametric cointegrating regression, *unpublished manuscript,* Yale University.

Gao, J., Maxwell, K., Lu, Z., Tjøstheim, D. (2009a). Nonparametric specification testing for nonlinear time series with nonstationarity. *Econometric Theory*, **25**, 1869–1892.

Gao, J., Maxwell, K., Lu, Z., Tjøstheim, D. (2009b). Specification testing in nonlinear and nonstationary time series autoregression. *The Annals of Statistics*, **37**, 3893–3928.

Gao, J., Li, D. and Tjøstheim, D. (2011). Uniform consistency for nonparametric estimates in null recurrent time series. *Working paper series No. 0085,* , The University of Adelaide, School of Economics.

Geman, D. and Horowitz, J. (2004). Occupation densities. *Annals of Probability*, **8**, 1–67.

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

Hansen, B. E. (2008). Uniform convergence rates for kernel estimation with dependent data. *Econometric Theory* **24**, 726–748.

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.

Kasparis, I. and Phillips, P. C. B. (2012). Dynamic misspecification in nonparametric cointegrating regression. *Journal of Econometrics*, **168(2)**, 270–284.

Karlsen, H. A., Myklebust, T. and Tjøstheim, D. (2007). Nonparametric estimation in a nonlinear cointegration model. *The Annals of Statistics*, **35**, 252–299.

Linton, O. and Wang, Q. (2013). Non-parametric transformation regression with non-stationary data, *Econometric Theory,* Accepted.

Marmer, V. (2008). Nonlinearity, nonstationarity, and spurious forecasts. *Journal of Econometrics*, **142**, 1–27.

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

Park, J. Y. and Phillips P. C. B. (2001). Nonlinear regressions with integrated time series. *Econometrica*, **69**, 117-161.

Phillips, P. C. B. and Park J. Y. (1998). Nonstationary Density Estimation and Kernel Autoregression. *Cowles Foundation discuss paper No. 1181.*

de la Pena, V. H. (1999). A General Class of Exponential Inequalities for Martingales and Ratios. *Annals of Probability*, **27**, 537–564.

Revuz, D. and Yor, M. (1994) *Continuous Martingales and Brownian Motion.* Fundamental Principles of Mathematical Sciences 293. Springer-Verlag

Wang, Q. (2014). Martingale limit theorems revisited and non-linear cointegrating regression. *Econometric Theory*, online.

Wang, Q. and Chan, N. (2014). Uniform convergence rates for a class of martingales with application in non-linear co-integrating regression. *Bernoulli,* **1**, 207230.

Wang, Q., Lin, Y. X., Gulati, C. M., (2003). Strong approximation for long memory processes with applications. *Journal of Theoretical Probability*, **16**, 377–389.

Wang, Q. and Phillips, P. C. B., (2009a). Asymptotic theory for local time density estimation and nonparametric cointegrating regression. *Econometric Theory*, **25**, 710-738.

Wang, Q. and Phillips, P. C. B., (2009b). Structural nonparametric cointegrating regression. *Econometrica*, **77**, 1901-1948.

Wang, Q. and Phillips P. C. B. (2011). Asymptotic theory for zero energy functionals with nonparametric regression applications. *Econometric Theory*, **27**, 235-259.

Wang, Q. and Phillips, P. C. B., (2012). A specification test for nonlinear nonstationary models. *The Annals of Statistics*, **40**, 727-758.

Wang, Q. and Wang, R. (2013). Non-parametric cointegrating regression with NNH errors. *Econometric Theory*, **29**, 1–27.

Appendix

Proof of Lemma 3.5. We split the set $A_n = \{y : \|y\| \le b_n\}$ into m_n balls of the form

$$A_{nj} \;=\; \{y : \|y - y_j\| \le 1/m_n'\}$$

where $m_n' = [n^{1+\alpha}/(\gamma_n \log n)^{1/2}]$, $m_n = (b_n m_n')^d$ and y_j are chosen so that $A_n \subset \bigcup A_{nj}$. It follows that

$$\begin{aligned}
& \sup_{\|y\|\le b_n} \Big| \sum_{t=1}^n e_t\, f_n(z_t, y) \Big| \\
\le\; & \max_{0\le j\le m_n} \sup_{y\in A_{nj}} \sum_{t=1}^n |e_t| \, \big| f_n(z_t, y) - f_n(z_t, y_j) \big| \\
& \qquad + \max_{0\le j\le m_n} \Big| \sum_{t=1}^n e_t\, f_n(z_t, y_j) \Big| \\
:=\; & \lambda_{1n} + \lambda_{2n}. \qquad (3.66)
\end{aligned}$$

Recalling (3.46) and $\frac{1}{n}\sum_{k=1}^n |e_k| = O_P(1)$ due to $\sup_{t\ge1} E(|e_t||\mathcal{F}_{t-1}) < \infty$, it is readily seen that

$$\begin{aligned}
\lambda_{1n} \;&\le\; \sum_{t=1}^n |e_t| \max_{0\le j\le m_n} \sup_{y\in A_{nj}} \big| f_n(z_t, y) - f_n(z_t, y_j) \big| \\
&\le\; C\,(n^\alpha m_n')^{-1} \sum_{t=1}^n |e_t| \\
&\le\; C\,(\gamma_n \log n)^{1/2}\, \frac{1}{n} \sum_{t=1}^n |e_t| = O_P[(\gamma_n \log n)^{1/2}]. \qquad (3.67)
\end{aligned}$$

In order to investigate λ_{2n}, write $e_t' = e_t I[|e_t| \le (\gamma_n/\log n)^{1/2}]$ and $e_t^* = e_t' - E(e_t' \mid \mathcal{F}_{t-1})$. Recalling $E(e_t \mid \mathcal{F}_{t-1}) = 0$ and $\sup_{n,x,y} |f_n(x,y)| < \infty$, we have

$$\begin{aligned}
\lambda_{2n} \;&\le\; \max_{0\le j\le m_n} \Big| \sum_{t=1}^n e_t^*\, f_n(z_t, y_j) \Big| \\
&\qquad + \max_{0\le j\le m_n} \Big| \sum_{t=1}^n \big[|e_t - e_t'| + E(|e_t - e_t'| \mid \mathcal{F}_{t-1}) \big] f_n(z_t, y_j) \Big| \\
&\le\; \max_{0\le j\le m_n} \Big| \sum_{t=1}^n e_t^*\, f_n(z_t, y_j) \Big| + C \sum_{t=1}^n \big[|e_t - e_t'| + E(|e_t - e_t'| \mid \mathcal{F}_{t-1}) \big] \\
&:=\; \lambda_{3n} + \lambda_{4n}. \qquad (3.68)
\end{aligned}$$

Routine calculations show that, under $\sup_{t\ge 1} E(|e_t|^{2p} \mid \mathcal{F}_{t-1}) < \infty$ and $n\,\gamma_n^{-p}\,\log^{p-1} n = O(1)$,

$$
\begin{aligned}
\lambda_{4n} &\le \sum_{t=1}^{n} \Big[|e_t| I\{|e_t| > (\gamma_n/\log n)^{1/2}\} + E\big(|e_t| I\{|e_t| > (\gamma_n/\log n)^{1/2}\}|\mathcal{F}_{t-1}\big)\Big] \\
&\le C\,\big(\frac{\gamma_n}{\log n}\big)^{(1-2p)/2} \sum_{t=1}^{n} \big[|e_t|^{2p} + E(|e_t|^{2p} \mid \mathcal{F}_{t-1})\big] \\
&\le C\,(\gamma_n \log n)^{1/2}\,\frac{1}{n} \sum_{t=1}^{n} \big[|e_t|^{2p} + E(|e_t|^{2p} \mid \mathcal{F}_{t-1})\big] \\
&= O_P\big[(\gamma_n \log n)^{1/2}\big], \qquad (3.69)
\end{aligned}
$$

Next consider λ_{3n}. As $E[(e_t^*)^2 \mid \mathcal{F}_{t-1}] \le 2(E[|e_t|^{2p} \mid \mathcal{F}_{t-1}])^{1/p}, a.s.$, Conditions (a) and (c) imply that

$$
\max_{0\le j\le m_n} \sum_{t=1}^{n} f_n^2(z_t, y_j) E[(e_t^*)^2 \mid \mathcal{F}_{t-1}] = O_P(\gamma_n). \qquad (3.70)
$$

Hence, for any $\eta > 0$, there exists a $M_0 > 0$ such that

$$
P\big(\max_{0\le j\le m_n} \sum_{t=1}^{n} \sigma_{tj}^2 \ge M_0\gamma_n\big) \le \eta.
$$

where $\sigma_{tj}^2 = f_n^2(z_t, y_j)E[(e_t^*)^2 \mid \mathcal{F}_{t-1}]$, whenever n is sufficiently large. This, together with $|e_t^*| \le 2(\gamma_n/\log n)^{1/2}$ and the well-known martingale exponential inequality (see, e.g., de la Pana (1999)), implies that, for any $\eta > 0$, there exists a $M_0 \ge 6d(k+2+\alpha)$ (k is as in condition c and α is given in (3.46)) such that, whenever n is sufficiently large,

$$
\begin{aligned}
& P[\lambda_{3n} \ge M_0(\gamma_n\,\log n)^{1/2}] \\
\le\ & P\Big[\lambda_{3n} \ge M_0(\gamma_n\,\log n)^{1/2},\ \ \max_{0\le j\le m_n} \sum_{t=1}^{n} \sigma_{tj}^2 \le M_0\gamma_n\Big] + \eta \\
\le\ & \sum_{j=0}^{m_n} P\Big[\sum_{t=1}^{n} e_t^* f_n(z_k, y_j) \ge M_0(\gamma_n\,\log n)^{1/2},\ \ \sum_{t=1}^{n} \sigma_{tj}^2 \le M_0\,\gamma_n\Big] + \eta \\
\le\ & m_n\,\exp\big\{-\frac{M_0^2\,\gamma_n \log n}{6\,M_0\gamma_n}\big\} + \eta \le m_n n^{-M_0/6} + \eta \le 2\eta, \qquad (3.71)
\end{aligned}
$$

where we have used the following fact:

$$
m_n \ \le\ C[n^{k+1+\alpha}/(\gamma_n\,\log n)^{1/2}]^d \le C_1\, n^{(k+1+\alpha)d},
$$

as $\gamma_n \to \infty$. This yields $\lambda_{3n} = O_P\big[(\gamma_n \log n)^{1/2}\big]$. Combining (3.66)-(3.71), we establish (3.48), and also complete the proof of Lemma 3.5. □