Submitted to the Annals of Statistics

LONG–RANGE DEPENDENT TIME SERIES SPECIFICATION

By Jiti Gao and Qiying Wang

The University of Western Australia and The University of Sydney

Model specification of short–range dependent stationary time series has become a very active research field in both econometrics and statistics since about two decades ago. In the meantime, estimation of long–range dependent stationary time series models has also been quite active. To the best of our knowledge, however, model specification of stationary time series with long–range dependence (LRD) has not been discussed in the literature. This is probably due to unavailability of certain central limit theorems for weighted quadratic forms of stationary time series with LRD. In this paper we try to tackle such difficult issues by establishing a nonparametric model specification test for parametric time series with LRD. In order to establish asymptotic distributions of the proposed test statistic, we develop new central limit theorems for certain weighted quadratic forms of stationary time series with LRD. In order to implement the proposed test in practice, we develop a computer–intensive parametric bootstrap simulation procedure for finding simulated critical values. As a result, our finite–sample studies show that both the proposed theory and the simulation procedure work well and that the proposed test has little size distortion and reasonable power.

1. Introduction. Various specification test statistics based on nonparametric and semiparametric techniques have been proposed and studied extensively for both independent and short–range dependent cases during the last two decades. Thus, there is a very long list of research papers in both the econometrics and statistics literature. Most results can be found from recent survey papers and manuscripts, such as Tjøstheim (1994), Fan and Gijbels (1996), Härdle, Lütkepohl and Chen (1997), Hart (1997), Härdle, Liang and Gao (2000), Cai and Hong (2003), Fan and Yao (2003), Fan (2005), Gao (2006) and the references therein.

Recent studies show that some data sets may display LRD (see Beran 1994; Robinson 1994a; Baillie and King 1996, Anh and Heyde 1999; Robinson 2003; Gao 2004; and others). In addition, existing studies (Hidalgo 1997; Robinson 1997; Csörgó and Mielniczuk 1999; Mielniczuk and Wu 2004; and

AMS 2000 subject classifications: Primary 62E20; secondary 60F05.

Keywords and phrases: Central limit theorem, Gaussian process, linear process, long-range dependence, parametric time series regression, specification testing.

others) also discuss nonparametric regression analysis of data with LRD. Recently, some applied researches (Anh, *et al* 1999; Mikosch and Starica 2004; Gao and Hawthorne 2006) show that some real data in environment and finance with both LRD and nonlinearities may be modelled by

$$Y_t = m\left(\frac{t}{n}\right) + e_t, \tag{1.1}$$

where $m(\cdot)$ is an unknown and probably nonlinear trend function of time t, and $\{e_t\}$ is a sequence of time series errors with possible LRD. The key findings of such existing studies suggest that in order to avoid misrepresenting the mean function or the conditional mean function of a long–range dependent data, we should let the data 'speak' for themselves in terms of specifying the true form of the mean function or the conditional mean function. This is particularly important for data with LRD, because unnecessary nonlinearity or complexity in mean functions may cause erroneous LRD.

In order to address such issues, we propose to model data with possible LRD, nonlinearity and nonstationary using a general nonparametric trend model. The main objective of this paper is thus to specify the trend by constructing a nonparametric kernel–based test. Consider a nonlinear time series model of the form

$$Y_t = m(X_t) + e_t, \ t = 1, 2, \cdots, n, \tag{1.2}$$

where n is the number of observations, $\{X_t\}$ is either a sequence of fixed designs of $X_t = \frac{t}{n}$ or independent and identically distributed (i.i.d.) random designs, $m(\cdot)$ is an unknown function, and $\{e_t\}$ is a long–range dependent linear process with $E[e_t] = 0$ and $0 < E[e_t^2] = \sigma^2 < \infty$. $\{X_s\}$ and $\{e_t\}$ are assumed to be independent for all $s, t \geq 1$ when $\{X_t\}$ is a sequence of i.i.d. random designs. To achieve our purpose, we will develop a kernel-based test in this paper for the hypotheses:

$$H_0: \ m(x) = m_{\theta_0}(x) \quad \text{versus} \quad H_1: \ m(x) = m_{\theta_1}(x) + c_n \Delta(x) \tag{1.3}$$

for all $x \in R = (-\infty, \infty)$, where θ_0 and θ_1 are vectors of unknown parameters, $m_\theta(x)$ is a known parametric function of x indexed by a vector of unknown parameters, θ, $\Delta(x)$ is a known smooth function, and $\{c_n\}$ is a sequence of real numbers tending to zero when $n \to \infty$. Note that, under H_0, model (1.2) becomes a parametric model of the form

$$Y_t = m_{\theta_0}(X_t) + e_t, \tag{1.4}$$

which covers many important cases. For example, model (1.4) becomes a simple linear model with LRD as in (1.1) of Robinson and Hidalgo (1997) when $m_{\theta_0}(X_t) = \alpha_0 + \beta_0 X_t$. For a given set of long–range dependent data, the acceptance of H_0 suggested by a test statistic may indicate that the mean function of the LRD data should be specified parametrically. In the case of the Nile river data as analyzed in Anh, *et al* (1999), one will consider using a linear mean function of the form $m\left(\frac{t}{n}\right)\alpha_0 + \beta_0 \cdot \frac{t}{n}$ if a suitable test suggests the acceptance of $H_0 : \ m(x) = \alpha_0 + \beta_0\ x$. Similarly, if a proper test suggests accepting a second–order polynomial function of the form $m\left(\frac{t}{n}\right)\alpha_0 + \beta_0\left(\frac{t}{n}\right) + \gamma_0\left(\frac{t}{n}\right)^2$ as the true trend of a financial data set $\{Y_t\}$, we will need only to difference $\{Y_t\}$ twice to generate a stationary set of the data.

This paper is organized as follows. The proposed test for the hypothesis (1.3) will be presented in Section 2.1. To investigate the proposed specification test, the limit theorems for the leading term $M_n(h)$ of our test statistics are investigated in Section 2.2, where

$$M_n(h) = \sum_{s=1}^{n} \sum_{t=1,\neq s}^{n} e_s\ a_n(X_s, X_t)\ e_t, \tag{1.5}$$

with $a_n(X_s, X_t) = K\left(\frac{X_s - X_t}{h}\right)$, in which $K(\cdot)$ is a probability kernel function and h is a bandwidth parameter. We mention that for the case where $\{X_t\}$ is a sequence of either fixed or random designs but $\{e_t\}$ is a sequence of long–range dependent errors, the problem of establishing limiting distributions for $M_n(h)$ is quite difficult. Because there is an involvement of $h \to 0$ into the inside of $K(\cdot)$, existing central limit theorems for U–statistics of long–range dependent processes (see a latest result by Hsing and Wu 2004) are not applicable. The limit theorems in Section 2.2 therefore are interesting and useful in themselves. In Section 3, we discuss some important extensions and applications of the theory established in Section 2. A parametric bootstrap simulation procedure as well as some resulting properties are established in Section 4. Section 4 also provides an example to demonstrate how to implement the proposed test and the bootstrap simulation procedure in practice. Section 5 concludes the paper with some remarks on extensions. Mathematical details are relegated to Appendices A and B. Throughout the paper, we use the symbol $a_n \sim b_n$ for $\lim_{n\to\infty} a_n/b_n = 1$.

2. Asymptotic theory. The test statistic for the hypothesis (1.3) is proposed in Section 2.1. In order to investigate the proposed test statistic, some new limiting distributions of weighted quadratic forms of dependent processes with LRD are developed in Section 2.2. Their proofs, along with other proofs, are relegated to Appendix A.

2.1 Model specification test. Let K be a one-dimensional bounded probability density function and h be a smoothing bandwidth. When $\{Y_t\}$ is a sequence of long–range dependent time series, the conventional kernel estimator of $m(\cdot)$ is defined by

$$\widehat{m}(x) = \frac{\frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{x-X_t}{h}\right) Y_t}{\widehat{f}(x)}, \tag{2.1}$$

where $\widehat{f}(x) = \frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{x-X_t}{h}\right)$ is the density estimate of the marginal density function, $f(x)$, of $\{X_t\}$ when $\{X_t\}$ is a sequence of i.i.d. random variables. When $X_t = \frac{t}{n}$, $\widehat{f}(x) = \frac{1}{nh}\sum_{t=1}^{n} K\left(\frac{nx-t}{nh}\right)$ is a sequence of functions of x. Various asymptotic properties for $\widehat{m}(x)$ have been studied in the literature. See Robinson 1997, Anh, *et al* (1999) and others, for example. One might expect that the non-parametric test statistic for the hypothesis (1.3) is related to $\widehat{m}(x)$. However, as demonstrated in the model specification literature for both the independent and short–range dependent time series cases (Zheng 1996; Li and Wang 1998; Li 1999; Fan, Zhang and Zhang 2001; Fan and Linton 2003; Gao 2006), there are certain classes of nonparametric test statistics that have little involvement of nonparametric estimation. One of the advantages is that both large and finite–sample properties of such test statistics are much less sensitive to the use of individual nonparametric estimation as well as the resulting estimation biases. In order to test the hypothesis (1.3), we therefore propose a kernel–based test statistic of the form

$$\widehat{M}_n(h) = \sum_{t=1}^{n} \sum_{s=1,\neq t}^{n} \widehat{e}_s\ a_n(X_s, X_t)\ \widehat{e}_t \tag{2.2}$$

for the case where $\{X_t\}$ is a sequence of i.i.d. random variables, where $a_n(X_s, X_t) = K\left(\frac{X_s-X_t}{h}\right)$ and $\widehat{e}_t = Y_t - m_{\widetilde{\theta}}(X_t)$, in which $\widetilde{\theta}$ is a consistent estimator of θ_0 under H_0. For the case of fixed-design mean with LRD errors, we also suggest the same form of (2.2) with $X_t = \frac{t}{n}$. As pointed out earlier, the choice of (2.2), instead of those related to $\widehat{m}(x)$, is mainly because our experience shows that such a form does not involve biases caused by nonparametric estimation and thus works well both theoretically and practically.

We need the following assumptions on the error process $\{e_t\}$, the kernel function $K(\cdot)$ and the regression function $m_\theta(x)$ for the main results of this paper.

Assumption 2.1. (i) $\{e_t\}$ is a sequence of linear processes defined by $e_t = \sum_{j=-\infty}^{\infty} \psi_j \eta_{t-j}$, where the innovations η_j are i.i.d. random variables with $E[\eta_1] = 0$, $E[\eta_1^2] = 1$ and $E[\eta_1^4] < \infty$, and the co–variance $\gamma(k) = E[e_t e_{k+t}] = \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+k}$ satisfies that $\gamma(0) = \sum_{j=-\infty}^{\infty} \psi_j^2 < \infty$ and $\gamma(k) \sim \eta\ |k|^{-\alpha}$, as $k \to \infty$, where $\lambda = (\alpha, \eta)$ with $0 < \alpha < 1$ and $0 < \eta < \infty$ is a vector of unknown parameters. (ii) In addition, we have $\psi_j \geq 0$ and $E[\eta_1^6] < \infty$.

Assumption 2.2. Let $\lambda = (\alpha, \eta)$ be defined as in Assumption 2.1. There exists some $\widetilde{\lambda} = (\widetilde{\alpha}, \widetilde{\eta})$ such that $||\widetilde{\lambda} - \lambda|| = O_P(w_n^{-1})$, where $\{w_n\}$ is a sequence of positive numbers satisfying $\lim_{n\to\infty} w_n / \log n = \infty$.

Assumption 2.3. (i) $K(\cdot)$ is a bounded and symmetric probability kernel function over the real line R. (ii) There exists some $0 < \beta < \alpha - \frac{1}{2}$ for $\frac{1}{2} < \alpha < 1$ such that $K(x) = O\left(\frac{1}{1+|x|^{1-\beta}}\right)$.

Assumption 2.4. When $\{X_t\}$ is a sequence of i.i.d. random designs, $\{X_t\}$ and $\{e_s\}$ are independent for all $s \geq 1$ and $t \geq 1$, and the density function $f(x)$ of X_t is bounded and uniformly continuous.

Assumption 2.5 (Random Design). (i) Under the null hypothesis H_0, $||\widetilde{\theta} - \theta_0|| = o_P(n^{-\alpha/2})$, where $||\cdot||$ denotes the Euclidean norm. (ii) $0 < E\left[\left|\left|\frac{\partial m_\theta(X_1)}{\partial \theta}|_{\theta=\theta_0}\right|\right|^2\right] < \infty$ and there exists a $\epsilon_0 > 0$ such that $\frac{\partial m_\theta(x)}{\partial \theta}$ is continuous in both $x \in R$ and $\theta \in \Theta_0$, where $\Theta_0 = \{\theta : ||\theta - \theta_0|| \leq \epsilon_0\}$.

Assumption 2.6 (Fixed Design). (i) Under the null hypothesis H_0, $||\widetilde{\theta} - \theta_0|| = O_P(n^{-\alpha/2})$. (ii) There exists a positive constant C_{θ_0} depending on θ_0 only such that, for all $0 \leq x \leq 1$,

$$|m_\theta(x) - m_{\theta_0}(x)| \leq C_{\theta_0}\ ||\theta - \theta_0||.$$

Assumption 2.1 (i) is quite general. For instance, the case where $\{e_t\}$ is a sequence of Gaussian errors is included. Assumption 2.1 (ii) is required to establish Theorems 2.2 and 2.4 below. The positivity of ψ_j in Assumption 2.1(ii) may be replaced by less restrictive those like that ψ_j are eventually positive. Also notice that it is possible to involve a slow–varying function into the form of $\gamma(k)$. Since this is not very essential for both our theory and practice, we use the current Assumption 2.1 throughout this paper. As for the alternative conditions on the $\{\epsilon_t\}$ in this kind of study, we refer to Cheng and Robinson (1994), Robinson (1997) and Robinson and Hidalgo (1997).

Assumption 2.2 may be justified for certain w_n like $w_n = n^{2/5}/\log n$. See the discussion for the construction of $\widehat{\lambda}$ in Section 4.

Assumption 2.3(i) is a standard condition on the kernel function. To establish Theorems 2.2 and 2.4 below, we also need Assumptions 2.3(ii). It imposes some restrictions on α and $K(\cdot)$, but these restrictions are easily verifiable. For instance, Assumption 2.3(ii) is satisfied when $K(\cdot)$ is either the standard normal density function or belongs to a class of probability kernel functions with compact support. Note also that, under Assumption 2.3, $A_\alpha < \infty$ where

$$A_\alpha = \int_0^\infty \int_0^\infty \int_0^\infty x^{-\alpha} y^{-\alpha} \big[K(z)K(x+y-z) + K(z-x)K(z-y)\big]\, dxdydz \tag{2.3}$$

as shown in Lemma A.6.

Assumption 2.4 is a standard condition in this kind of problem.

Assumptions 2.5 and 2.6 require certain smoothness of $m_\theta(x)$ with respect to θ to ensure a certain rate of convergence. Assumptions 2.5(i) and 2.6(i) may be viewed as a counterpart of the conventional $\sqrt{n}$–rate of convergence in the short–range dependent case. These conditions are satisfied automatically when $\widetilde{\theta}$ achieves the conventional $\sqrt{n}$–rate of convergence as discussed in the literature. See for example, Beran and Ghosh (1998). It follows easily from the proofs of Theorems 2.1 in Appendix A that $||\widetilde{\theta} - \theta_0|| = o_P(n^{-\alpha/2})$ may be replaced by $||\widetilde{\theta} - \theta_0|| = O_P(n^{-\alpha/2})$ in Theorem 2.1(i). To avoid introducing repetitious arguments, we use current Assumption 2.5(i) for the statement of Theorem 2.1. Assumption 2.5(ii) basically imposes certain moment conditions on both the form $m_{\theta_0}(\cdot)$ and the design $\{X_t\}$. Assumption 2.6 is used only for the fixed–design case in Theorem 2.2.

We now state the main results of this paper. Theorem 2.1 considers the case where $\{X_t\}$ is a sequence of i.i.d. random designs. The case where $X_t = \frac{t}{n}$ is discussed in Theorem 2.2. All notation are defined as before. The proofs of these results are given in Appendix A.

THEOREM 2.1. *Suppose that Assumptions 2.1(i), 2.3(i) and 2.4–2.5 hold.*

(i) If $\lim_{n\to\infty} n^{2(1-\alpha)} h = 0$ *and* $\lim_{n\to\infty} nh = \infty$, *then under* H_0,

$$\widehat{L}_{1n}(h) \equiv \frac{\widehat{M}_n(h)}{\widehat{\sigma}_{1n}(h)} \to_D N(0,1) \quad \text{as } n \to \infty, \tag{2.4}$$

where $\widehat{\sigma}_{1n}^2(h) = 2n^2 h \int K^2(x)dx \left(\frac{1}{n}\sum_{t=1}^n \widehat{f}(X_t)\right)\left(\frac{1}{n}\sum_{t=1}^n \widehat{e}_t^2\right)^2$.

(ii) If $\lim_{n\to\infty} h = 0$, $\lim_{n\to\infty} n^{2(1-\alpha)} h = \infty$ *and Assumption 2.2 hold, then under* H_0,

$$\widehat{L}_{2n}(h) \equiv \frac{\widehat{M}_n(h)}{\widehat{\sigma}_{2n}(h)} \to_D \chi^2(1) \quad \text{as } n \to \infty, \tag{2.5}$$

where $\chi^2(1)$ *is the chi–square distribution with degree one of freedom, and*

$$\widehat{\sigma}_{2n}(h)\frac{2\,n^{2-\widetilde{\alpha}}\,h\,\widetilde{\eta}}{(1-\widetilde{\alpha})(2-\widetilde{\alpha})}\,I_{(0<\widetilde{\alpha}<1)}\,\frac{1}{n}\sum_{t=1}^{n}\widehat{f}(X_t).$$

THEOREM 2.2. *Suppose that Assumptions 2.1–2.3 and 2.6 hold. Let* w_n *be given as in Assumption 2.2. If* $\lim_{n\to\infty} h = 0$ *hold,* $\lim_{n\to\infty} w_n\, h^{1/2}/\log n = \infty$ *and* $nh \to \infty$*, then under* H_0*,*

$$\begin{aligned}\widehat{L}_{3n}(h) \quad &\equiv \quad \frac{\sum_{t=1}^{n}\sum_{s=1,\neq t}^{n} b_n(s,t)\left(\widehat{e}_s\widehat{e}_t - \widehat{\gamma}(s-t)\right)}{\widehat{\sigma}_{3n}(h)} \\ &\to_D \quad N(0,1) \qquad \text{as } n\to\infty,\end{aligned} \tag{2.6}$$

where $\widehat{\sigma}_{3n}^2(h) = 8\,\widetilde{\eta}^2\, n\,(nh)^{3-2\widetilde{\alpha}}\, A^*_{\widetilde{\alpha}}$ *with*

$$\begin{aligned}A^*_{\widetilde{\alpha}} \quad = \quad &\int_{1/n}^{n}\int_{1/n}^{n}\int_{1/n}^{n} x^{-\widetilde{\alpha}} y^{-\widetilde{\alpha}}\big[K(z)K(x+y-z) \\ &\qquad + K(z-x)K(z-y)\big]\,dxdydz,\end{aligned}$$

$b_n(s,t) = K\left(\frac{s-t}{nh}\right)$ *and*

$$\widehat{\gamma}(k) \quad = \quad \begin{cases} \frac{1}{n}\sum_{i=1}^{n-|k|}\widehat{e}_i\widehat{e}_{i+|k|}, & \text{for } |k| \le (nh)^{1/3}, \\ \\ \widetilde{\eta}\,|k|^{-\widetilde{\alpha}}, & \text{for } (nh)^{1/3} < |k| \le n-1. \end{cases}$$

REMARK 2.1. (i) As expected, the limiting distributions of the test statistic $\widehat{M}_n(h)$ (under certain normalization) for the hypothesis (1.3) depend on the value of α and the choice of the bandwidth h. We require $1/2 < \alpha < 1$, together with $h^{-1} = o(n)$ and $h = o(n^{-2(1-\alpha)})$, in Theorem 2.1(i). By contrast, Theorem 2.1(ii) allows $0 < \alpha < 1$, but we have to restrict $h = o(1)$ and $h^{-1} = o(n^{2(1-\alpha)})$. These facts imply that, to make the limiting distribution of $\widehat{M}_n(h)$ (under certain normalization) normal, the conditions that $1/2 < \alpha < 1$ and $h = o(n^{-2(1-\alpha)})$ are essentially necessary for the case where $\{X_t\}$ is a sequence of i.i.d. random designs.

(ii) By tackling the proof of Theorem 2.1(i), routine modifications show that the conclusion (2.4) also holds under H_0 as $h \to 0$ and $nh \to \infty$ if we replace Assumption 2.1 (i) by a short-range dependent linear process (that is, $e_t = \sum_{k=-\infty}^{\infty}\psi_k\eta_{t-k}$ with $\sum_{k=-\infty}^{\infty}|\psi_k| < \infty$). The short-range dependent cases have been investigated in Li and Wang (1998), Li (1999) and Chapter 3 of Gao 2006. It is interesting to notice that the techniques used in this paper are different from those mentioned above.

(iii) By checking the proofs of Theorems 2.1 and 2.3, we may show that the conclusions of Theorems 2.1 and 2.3 still hold if we replace the $\{X_t\}$ satisfying Assumption 2.4 by a sequence of strictly stationary and α–mixing random variables (under certain conditions). In this case, we need to replace Lemma A.4 listed in the appendix by a more general one, such as Lemma B.1 of Gao and King (2005). Since there are no essential differences in the main steps of the proof but much more technicalities are involved, we focus on the current $\{X_t\}$ for reading convenience.

(iv) For the fixed-design case of $X_t = \frac{t}{n}$, we also require $1/2 < \alpha < 1$ for the asymptotical normality of $\widehat{M}_n(h)$ (under certain normalization) in Theorem 2.2. Furthermore, the range of the bandwidth h depends on the accuracy of $||\widetilde{\lambda} - \lambda|| = O_P(w_n^{-1})$. As it may be justified that $w_n = n^{2/5}/\log n$ (see Theorem 4.2), the optimal bandwidth $h \sim C\, n^{-1/(4+\alpha)}$ in theory is included in Theorem 2.2. We also mention that the $\widehat{\gamma}(k)$ defined in Theorem 2.2 provides a consistent estimate of $\gamma(k)$ for each fixed k, but it is not possible to replace $\widehat{\gamma}(k) = \tilde{\eta}|k|^{-\tilde{\alpha}}$ by $\widehat{\gamma}(k) = \frac{1}{n}\sum_{i=1}^{n-|k|} \widehat{e}_i \widehat{e}_{i+|k|}$ when $(nh)^{1/3} < k \leq n-1$ because (A.69) is not true in the latter case. The limiting distribution of $\widehat{M}_n(h)$ (under certain normalization) for $0 < \alpha \leq 1/2$ in the fixed-design case is an open problem.

(v) As suggested by a referee, it should be possible to establish some corresponding results of Theorem 2.1 for the case where both $\{X_i\}$ and $\{e_i\}$ exhibit LRD. In this case, existing studies (Hidalgo 1997; Csörgó and Mielniczuk 1999; Mielniczuk and Wu 2004) in nonparametric estimation have already shown that while similar techniques would be used to establish and prove such corresponding results, the corresponding conditions and proofs are more technical than those involved in the current paper. We therefore wish to leave such extensions for future research.

In order to motivate the necessity of establishing some limit theorems for general quadratic forms of dependent processes with LRD, we observe that

$$\begin{aligned} \widehat{M}_n(h) &= \sum_{t=1}^{n} \sum_{s=1,\neq t}^{n} \widehat{e}_s\, a_n(X_s, X_t)\, \widehat{e}_t \\ &= M_n(h) + 2R_{1n}(h) + R_{2n}(h), \end{aligned} \tag{2.7}$$

where $M_n(h) = \sum_{t=1}^n \sum_{s=1,\neq t}^n e_s\, a_n(X_s, X_t)\, e_t$,

$$R_{1n}(h) \;=\; \sum_{t=1}^{n}\sum_{s=1,\neq t}^{n} a_n(X_s, X_t)\, e_s \left(m(X_t) - m_{\widetilde{\theta}}(X_t)\right),$$

$$R_{2n}(h) \;=\; \sum_{t=1}^{n}\sum_{s=1,\neq t}^{n} a_n(X_s, X_t) \left(m(X_s) - m_{\widetilde{\theta}}(X_s)\right)\left(m(X_t) - m_{\widetilde{\theta}}(X_t)\right).$$

It will be shown in Appendix A below that $2R_2(h) + R_2(h) = o_P(\sigma_{in})$ for $i = 1$ and 2 under the corresponding conditions of Theorem 2.1. Thus, in order to prove Theorems 2.1, we need to establish limit theorems for $M_n(h)$, which is a weighted quadratic form of $\{e_t\}$. Similar arguments also work for Theorem 2.2. Since such limit theorems are interesting and useful in themselves, we formally establish them in Section 2.2 below.

2.2 Limit theorems for quadratic forms. As both the conditions and results for the case of random designs are different from those for the case of fixed designs, we establish the following two theorems separately. Their proofs are given in Appendix A.

THEOREM 2.3. *Suppose that Assumptions 2.1(i), 2.3(i) and 2.4 hold.*

(i) If $\lim_{n\to\infty} n^{2(1-\alpha)} h = 0$ *and* $\lim_{n\to\infty} nh = \infty$*, then*

$$\frac{\sum_{t=1}^n \sum_{s=1,\neq t}^n e_s\ a_n(X_s, X_t)\ e_t}{\sigma_{1n}(h)} \to_D N(0,1), \quad \text{as } n \to \infty, \tag{2.8}$$

where $\sigma_{1n}^2(h) = n^2 h A_{1\alpha}^2$*, in which* $A_{1\alpha}^2 = 2\,\gamma^2(0) \int K^2(x)dx \int f^2(y)dy$.

(ii) *If* $\lim_{n\to\infty} h = 0$ *and* $\lim_{n\to\infty} n^{2(1-\alpha)} h = \infty$*, then*

$$\frac{\sum_{t=1}^n \sum_{s=1,\neq t}^n e_s\ a_n(X_s, X_t)\ e_t}{\sigma_{2n}(h)} \to_D \chi^2(1), \quad \text{as } n \to \infty, \tag{2.9}$$

where $\sigma_{2n}(h) = n^{2-\alpha}\, h\, A_{2\alpha}$*, in which* $A_{2\alpha} = \frac{2\eta}{(1-\alpha)(2-\alpha)} \left(\int f^2(y)dy\right)$.

THEOREM 2.4. *Suppose that Assumptions 2.1 and 2.3 hold. If* $\lim_{n\to\infty} h = 0$ *and* $\lim_{n\to\infty} nh = \infty$*, then*

$$\frac{\sum_{t=1}^n \sum_{s=1,\neq t}^n b_n(s,t)\,(e_s e_t - \gamma(s-t))}{\sigma_{3n}(h)} \to_D N(0,1), \text{ as } n \to \infty, \tag{2.10}$$

where $\sigma_{3n}^2(h) = 8\eta^2\, n\, (nh)^{3-2\alpha} A_\alpha$*, in which* A_α *is defined as in (2.3).*

REMARK 2.2. (i) Theorem 2.3 extends the existing limit theorems for both i.i.d. and short–range dependent cases (see Hjellvik, Yao and Tjøstheim

1998, Li 1999, and others) to the situation where $\{e_t\}$ is a long–range dependent linear process. Unlike these existing results, our theorem shows that random functions, the limiting distribution of the random weighted quadratic forms can be either a standard normal distribution or a chi-square distribution.

(ii) The related results on quadratic forms of long–range dependent time series can be found in Fox and Taqqu (1987), Avram (1988), Giraitis and Surgailis (1990), Giraitis and Taqqu (1997), Ho and Hsing (1996, 1997, 2003), Hsing and Wu (2004), and others. Since the weighted coefficients in these existing results (see Hsing and Wu 2004, for example) are non–random and independent of n, they are not applicable for the establishment of Theorems 2.3 and 2.4. Both Theorems 2.3 and 2.4 therefore provide some sorts of extensions of various existing results.

3. Extensions and applications. This section show that the leading term of many existing kernel–based test statistics may be represented by a quadratic form similar to (1.5). Theorems 2.1-2.4 suggest the feasibility of the corresponding results based on the long-range dependent errors for these existing test statistics. In order to avoid some repetitious arguments, we only state some main steps. Further details will be omitted.

3.1. *Existing kernel–based tests for conditional mean.* A very simple idea for constructing a kernel test for H_0 is to compare the L_2-distance between a nonparametric kernel estimator of $m(\cdot)$ and a parametric counterpart. Let us denote the nonparametric estimator of $m(\cdot)$ by $\widehat{m}(\cdot)$ as in (2.1) and the parametric estimator of $m_{\theta_0}(\cdot)$ by $\widetilde{m}_{\widetilde{\theta}}(\cdot)$ given by

$$\widetilde{m}_{\widetilde{\theta}}(x) = \frac{\sum_{i=1}^n K_h(x - X_i) m_{\widetilde{\theta}}(X_i)}{\sum_{i=1}^n K_h(x - X_i)}, \tag{3.1}$$

where $\widetilde{\theta}$ is a consistent estimator of θ_0 as defined before and $K_h(\cdot) = \frac{1}{h} K\left(\frac{\cdot}{h}\right)$. Härdle and Mammen (1993) proposed a test statistic of the form

$$L_{n1}(h) \quad = \quad n\sqrt{h} \int \left\{ \widehat{m}(x) - \widetilde{m}_{\widetilde{\theta}}(x) \right\}^2 w(x) dx, \tag{3.2}$$

where $w(x)$ is a non-negative weight function. Recall the model (1.4). Under

H_0, it is readily seen that

$$\begin{aligned} L_{n1}(h) &= n\sqrt{h}\int \Big(\frac{\left[\sum_{i=1}^n K_h(x-X_i)\left(e_i+m_{\theta_0}(X_i)-m_{\widetilde{\theta}}(X_i)\right)\right]^2}{n^2\widehat{f}^2(x)}\Big)w(x)dx \\ &= n\sqrt{h}\sum_{i=1}^n\sum_{j=1}^n\left(\int \frac{K_h(x-X_i)K_h(x-X_j)}{n^2\widehat{f}^2(x)}w(x)dx\right)e_ie_j \\ &\quad + \text{remainder term,} \end{aligned} \tag{3.3}$$

in which the leading term is similar to (1.5). In a related work, Kreiss, Neuman and Yao (2002) proposed the following simplified version of $L_{n1}(h)$:

$$L_{n2}(h) = n\sqrt{h}\int\left(\sum_{i=1}^n K_h(x-X_i)\left[Y_i - m_{\widetilde{\theta}}(X_i)\right]\right)^2 w(x)dx. \tag{3.4}$$

As an alternative to $L_{n1}(h)$, Horowitz and Spokoiny (2001) established a test statistic based on a discretised version of the form

$$L_{n3}(h) = \sum_{i=1}^n\left(\widehat{m}(X_i) - \widetilde{m}_{\widetilde{\theta}}(X_i)\right)^2, \tag{3.5}$$

in which $\{X_i\}$ is only a sequence of fixed designs. In order to avoid introducing a secondary estimation procedure required for estimating $\sigma^2(\cdot)$ when the variance function is not constant, Chen, Härdle and Li (2003) construct a test statistic based on empirical likelihood ideas. As shown in their paper, the first–order approximation of their test is asymptotically equivalent to

$$L_{n4}(h) = nh\int\left[\widehat{m}(x) - \widetilde{m}_{\widetilde{\theta}}(x)\right]^2 w(x)dx. \tag{3.6}$$

It can be easily shown that all statistics $L_{nj}(h), j = 2, 3, 4$ have a similar decomposition like (3.3) in which the leading term is similar to (1.5).

3.2. *Testing conditional mean with conditional variance.* Since the main objective of this paper is to specify parametrically the form of $m(\cdot)$, we have assumed that the variance or conditional variance σ^2 is an unknown parameter. As can be seen from (1.5), we may easily replace e_t by $e_t = \sigma(X_t)\epsilon_t$ with $\{\epsilon_t\}$ being a sequence of long–range dependent linear process. In this case, the leading term of $\widehat{M}_n(h)$ in (2.2) becomes

$$L_{n5}(h) = \sum_{s=1}^n\sum_{t=1,\neq s}^n \epsilon_s\ \sigma(X_s)\ K\left(\frac{X_s-X_t}{h}\right)\ \sigma(X_t)\ \epsilon_t, \tag{3.7}$$

which is also a quadratic form of (X_t, ϵ_t).

3.3. *Testing conditional mean in additive form.* When $X_t = (X_{t1}, \cdots, X_{td})$ in (1.2) is a vector of d–dimensional designs, we may consider a hypothesis problem of the form

$$H_0' : \ m(x) = \sum_{i=1}^{d} m_{i\theta_0}(x_i) \ \text{ versus } \ H_1' : \ m(x) = \sum_{i=1}^{d} m_{i\theta_1}(x_i) + c_n \sum_{i=1}^{d} \Delta_i(x_i), \tag{3.8}$$

where each $m_{i\theta_0}(\cdot)$ is a known function indexed by θ_0 and $\Delta_i(\cdot)$ is also a known function over $I\!R$. Various additive models have been discussed in the literature (see Fan, Härdle and Mammen 1998; Sperlich, Tjøstheim and Yang 2002; Gao, Lu and Tjøstheim 2006 for example). The construction of $\widehat{M}_n(h)$ suggests a test statistic of the form:

$$L_{n6}(h) = \sum_{j=1}^{n} \sum_{i=1}^{n} \widetilde{Y}_i \ \prod_{k=1}^{d} K\left(\frac{X_{ik} - X_{jk}}{h}\right) \ \widetilde{Y}_j, \tag{3.9}$$

for the hypothesis (3.8), where $\widetilde{Y}_i = \left(Y_i - \sum_{k=1}^{d} m_{k\widetilde{\theta}_0}(X_{ik})\right)$. Clearly $L_{n6}(h)$ also has a leading term that is similar to (1.5).

As mentioned before, some corresponding results of Theorems 2.1–2.4 may be established accordingly for $L_{ni}(h)$, $1 \le i \le 6$. Instead of providing some repetitious arguments and statements about such corresponding results, in Section 4 below we propose using simulation and implementation procedures to ensure that the main theory and the proposed test statistic established in Section 2 are applicable in practice.

4. Simulation and implementation procedures. In this section, we are interested in the implementation of the proposed test statistics. To do so, we need to develop our simulation procedure for the choice of a simulated critical value and then propose an estimation procedure for the parameter α involved in the proposed test. Section 4.1 establishes a novel simulation procedure for implementing our test in practice. An estimation procedure for α is briefly mentioned in Section 4.2. Section 4.3 gives an example of implementation to check whether the theory works well in practice.

4.1 *Simulation scheme and asymptotic properties.* In the construction of simulation, the covariance structure $\gamma_\lambda(k) = E[e_t e_{t+k}]$ needs to be replaced by an estimated version $\gamma_{\widetilde{\lambda}}(k)$ with $\widetilde{\lambda} = (\widetilde{\alpha}, \widetilde{\eta})$ being a pair of consistent estimators of the pair $\lambda = (\alpha, \eta)$. We assume the existence of $\widetilde{\lambda}$ at the moment. Its construction will be briefly mentioned at the end of this section.

SIMULATION PROCEDURE 4.1: Let $T_n(h)$ be either $\widehat{L}_{1n}(h)$, $\widehat{L}_{2n}(h)$ or $\widehat{L}_{3n}(h)$ as defined in (2.4), (2.5) or (2.6). Let l_r $(0 < r < 1)$ be the $1-r$ quantile of the exact finite–sample distribution of $T_n(h)$. Because l_r may not be evaluated in practice, we suggest an approximate r–level critical value l_r^* to replace it by using the following bootstrap procedure:

- Generate $Y_i^* = m_{\widetilde{\theta}}(X_i) + e_i^*$ for $1 \le i \le n$, where the original sample $\mathcal{X}_n = (X_1, \cdots, X_n)$ acts in the resampling as a fixed design even when the X_i are random, $\{e_i^*\}$ is a sequence of stationary LRD Gaussian errors drawn from a stationary LRD Gaussian process with the covariance structure being given by $\gamma_{\widetilde{\lambda}}(k) \sim \widetilde{\eta}\, |k|^{-\widetilde{\alpha}}$.
- Use the data set $\{(X_i, Y_i^*) : 1 \le i \le n\}$ to estimate $\widehat{\theta}$ by $\widetilde{\theta}^*$ and to compute $\widehat{T}_n^*(h)$, where $\widehat{T}_n^*(h)$ is the corresponding version of $T_n(h)$ under H_0 with $\{(X_i, Y_i) : 1 \le i \le n\}$ and $(\theta_0, \widehat{\theta})$ being replaced by $\{(X_i, Y_i^*) : 1 \le i \le n\}$ and $(\widehat{\theta}, \widetilde{\theta}^*)$.
- Repeat the above step M times and produce M versions of $\widehat{T}_n^*(h)$ denoted by $\widehat{T}_{n,m}^*(h)$ for $m = 1, 2, \ldots, M$. Use the M values of $\widehat{T}_{n,m}^*(h)$ to construct their empirical distribution function. The bootstrap distribution of $\widehat{T}_n^*(h)$ given $\mathcal{W}_n = (X_1, \cdots, X_n; Y_1, \cdots, Y_n)$ is defined by $P^*\left(\widehat{T}_n^*(h) \le x\right) = P\left(\widehat{T}_n^*(h) \le x | \mathcal{W}_n\right)$. Then let l_r^* $(0 < r < 1)$ satisfy $P^*\left(\widehat{T}_n^*(h) \ge l_r^*\right) = r$ and estimate l_r by l_r^*.

REMARK 4.1. In the simulation procedure, we generate a sequence of bootstrap resamples $\{e_i^*\}$ from a stationary Gaussian process with LRD even though $\{e_i\}$ may not be Gaussian. As discussed in Li and Wang (1998), Bühlmann (2002), Franke, Kreiss and Mammen (2002), Horowitz (2003) and others, we may also use a wild bootstrap to generate a sequence of resamples for $\{e_i^*\}$. Since the proposed simulation procedure works well in Section 5 below, we avoid further discussions in such issues.

To investigate asymptotic properties of l_r^* and $\widehat{T}_n^*(h)$, we need the following assumption.

ASSUMPTION 4.1. (i) Let H_0 be false. Assume that either Assumption 2.5 or Assumption 2.6 holds with θ_0 replaced by θ_1. (ii) $\lim_{n\to\infty} n^{\alpha} h^{\alpha - \frac{1}{2}}\, c_n^2 = \infty$ for $\frac{1}{2} < \alpha < 1$ and $0 < \int_0^1 \Delta^2(x)dx < \infty$, where c_n and $\Delta(x)$ are as defined in (1.3).

ASSUMPTION 4.2. Let H_0 be true. Assume that either Assumption 2.5 or Assumption 2.6 holds in probability with respect to the joint distribution of $\mathcal{W}_n$ when θ_0 is replaced by $\widetilde{\theta}^*$.

Assumption 4.1 requires some conditions under the alternative to ensure that $T_n(h)$ has some power. Assumption 4.2 is the bootstrap version of either Assumption 2.5 or 2.6. We now have the following theorem. Its proof is similar to Theorems 2.1 and 2.2 and will be outlined in Appendix B.

THEOREM 4.1. (i) *If, in addition to the conditions of either Theorem 2.1 or Theorem 2.2, Assumption 4.2 holds, then under* H_0,

$$\sup_{x \in R} \left| P^*(\widehat{T}_n^*(h) \le x) - P(T_n(h) \le x) \right| = o_P(1) \tag{4.1}$$

and

$$\lim_{n \to \infty} P\left(T_n(h) > l_r^*\right) = r. \tag{4.2}$$

(ii) *If, in addition to the conditions of either Theorem 2.1 or Theorem 2.2, Assumptions 4.1–4.2 hold, then under* H_1

$$\lim_{n \to \infty} P\left(T_n(h) > l_r^*\right) = 1. \tag{4.3}$$

Theorem 4.1 shows that the bootstrap approximation works well asymptotically. For the independent errors case, Li and Wang (1998) established a result similar to (4.1). To the best of our knowledge, Theorem 4.1 is new in this kind of long–range dependent time series specification.

Note that l_r^* is a function of h. A natural problem raised in simulation is the choice of a suitable bandwidth h. To solve this problem, define the size and power functions of $T_n(h)$ as

$$\gamma_n(h) = P\left(T_n(h) > l_r^* | H_0 \text{ true}\right) \text{and } \beta_n(h) = P\left(T_n(h) > l_r^* | H_0 \text{ false}\right). \tag{4.4}$$

Clearly, a reasonable selection procedure for a suitable bandwidth is such that the size function $\gamma_n(h)$ is controlled by a significance level, but the power function $\beta_n(h)$ is maximized over such bandwidths that make $\gamma_n(h)$ is controllable. This suggests using an optimal bandwidth of the form

$$\widehat{h}_{\text{test}} = \arg \max_{h \in \mathcal{H}_n} \beta_n(h) \quad \text{with} \quad \mathcal{H}_n = \left\{h : \gamma_n(h) \le r\right\}. \tag{4.5}$$

Theoretically, we have not been able to study $\widehat{h}_{\text{test}}$ asymptotically. In Example 4.1 below, we instead combine the proposed Simulation Procedure 4.1 and the Implementation 4.1 below to numerically approximate $\widehat{h}_{\text{test}}$.

IMPLEMENTATION PROCEDURE 4.1: Use $\widehat{l}_r^* = l_r^*(\widehat{h}_{\text{test}})$ as the simulated critical value to compute the sizes and power values of $T_n(\hat{h}_{\text{test}})$.

4.2. *LRD parameter estimation.* As briefly mentioned at the beginning of Section 4, we need to estimate $\lambda = (\alpha, \eta)$ when λ is unknown. We now give an outline of our estimation procedure. Let $u_i = Y_i - m_{\widetilde{\theta}}(X_i)$, $\omega_j = \frac{2\pi j}{n}$ and $I_u(\omega_j) = \frac{1}{2\pi n}\left|\sum_{s=1}^n u_s e^{is\omega_j}\right|^2$ for $1 \le j \le N$, where N is the number of frequencies and is chosen as the largest integer part of $C\ n^{\frac{4}{5}}$ with some positive $C > 0$. In practice, a data–driven choice of N may be used as proposed in Robinson (1994b). Introduce an objective function of the form

$$\Gamma_u(\lambda) = \frac{1}{N}\sum_{j=1}^N \left(\log\left(\phi(\omega_j;\lambda) + \frac{I_u(\omega_j)}{\phi(\omega_j;\lambda)}\right)\right), \tag{4.6}$$

where $\phi(\omega;\lambda)$ is the spectral density function of $\{u_i\}$ satisfying as $\omega \downarrow 0$,

$$\phi(\omega;\lambda) \sim \frac{\eta}{2\Gamma(\alpha)\sin\left(\frac{1-\alpha}{2}\pi\right)}\frac{1}{\omega^{1-\alpha}}. \tag{4.7}$$

We then estimate λ by $\widetilde{\lambda} = \arg\min_\lambda \Gamma_u(\lambda)$. Define

$$\Sigma = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\frac{\partial\log(\phi(\omega;\lambda)}{\partial\omega}\right)\left(\frac{\partial\log(\phi(\omega;\lambda)}{\partial\omega}\right)^\tau d\omega. \tag{4.8}$$

The following Theorem 4.2 establishes an asymptotic consistency result for $\tilde{\lambda}$. Its proof will be briefly mentioned in Appendix B.

THEOREM 4.2. *Assume that the conditions of either Theorem 2.1 or Theorem 2.2 except Assumption 2.2 hold. If Σ^{-1} exists, then*

$$||\widetilde{\lambda} - \lambda|| = o_P\left(\log n/n^{2/5}\right). \tag{4.9}$$

Theorem 4.2 shows that Assumption 2.2 may be justified for $w_n = n^{2/5}/\log n$.

4.3. *An example of implementation.* In this section, we implement the proposed simulation procedure to show how to assess the finite–sample properties of the proposed test $T_n(h)$ through using a simulated example.

EXAMPLE 4.1. Consider a linear model of the form

$$Y_i = \alpha_0 + \beta_0 X_i + e_i,\ t = 1, \cdots, n, \tag{4.10}$$

where (α_0, β_0) is a pair of unknown parameters to be estimated, $\{X_i\}$ is a sequence of i.i.d. random designs drawn from uniform U$[0,1]$, and $\{e_i\}$ is a sequence of dependent errors given by $e_i = \sum_{j=-\infty}^{\infty} b_j\ u_{i-j}$, in which $\{u_k : k = 0, \pm1, \cdots\}$ is a sequence of independent observations drawn from $N(0,1)$, and $b_s = c(\alpha)\ |s|^{-\frac{1+\alpha}{2}}$ for $\frac{1}{2} < \alpha < 1$, in which $c(\alpha) > 0$ is chosen such that $E[e_i^2] = 1$.

The parameters α_0 and β_0 are estimated by the ordinary least squares estimators $\widetilde{\alpha}_0$ and $\widetilde{\beta}_0$. The parameter α is estimated using (4.7). Throughout this section, we use the standard Normal kernel function $K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

Since the conditions of Theorem 2.1(i) are satisfied, we can now implement $\widehat{L}_{1n}(h)$ in this example. In order to compute the sizes and power values of $\widehat{L}_{1n}(h)$, we generate $\{Y_i\}$ from

$$H_0: \ Y_i = \alpha_0 + \beta_0 X_i + e_i \quad \text{or} \quad H_1: \ Y_i = \alpha_0 + \beta_0 X_i + \gamma_0 X_i^2 + e_i, \tag{4.11}$$

where the parameters (α_0, β_0) are estimated by $(\widetilde{\alpha}_0, \widetilde{\beta}_0)$ under H_0, and the parameters $(\alpha_0, \beta_0, \gamma_0)$ is estimated by the ordinary least squares estimators $(\widetilde{\alpha}_0, \widetilde{\beta}_0, \widetilde{\gamma}_0)$ under H_1. When we generate $\{Y_i\}$, the initial values are $\alpha_0 = \beta_0 \equiv 1$ and $\gamma_0 \neq 0$ is to be chosen for computing power values in various cases.

We first apply both the Simulation Procedure 4.1 and the Implementation Procedure 4.1 to find the optimal bandwidth $\widehat{h}_{\text{test}}$. To assess the variability of both the size and power with respect to various bandwidth values, we then consider a set of bandwidth values of the form: $h_i = \frac{i}{2}\, \widehat{h}_{\text{test}}$ for $1 \leq i \leq 5$.

In order to examine the finite sample properties of the maximized version of the form:

$$L_{\max} = \max_{h = h_i : 1 \leq i \leq 5} \widehat{L}_{1n}(h), \tag{4.12}$$

we also produce the simulated critical value for each case. To simplify the notation, we introduce

$$L_{\text{test}} = \widehat{L}_{1n}(h_2). \tag{4.13}$$

In the implementation of the Simulation Procedure 4.1, we consider cases where the number of replications of was $M = 1000$, each with $B = 250$ number of bootstrapping resamples, and the simulations were done for data sets of sizes $n = 250, 500$ and 750. The corresponding simulated critical values, sizes and power values for $L_{\max}$ and L_{test} are given in Tables 4.1 and 4.2 below.

Table 4.1. Simulated size values at the 5% level

Observation	$L_{\max}$		L_{test}	
n	size	critical value	size	critical value
250	0.099	2.087	0.040	1.787
500	0.095	2.091	0.057	1.591
750	0.084	1.928	0.048	1.673

Table 4.2. Simulated power values at the 5% level

	$n = 250$		$n = 500$		$n = 750$	
γ_0	$L_{\max}$	L_{test}	$L_{\max}$	L_{test}	$L_{\max}$	L_{test}
0.10	0.132	0.132	0.122	0.123	0.237	0.236
0.20	0.271	0.274	0.538	0.540	0.741	0.743
0.30	0.574	0.576	0.829	0.832	0.884	0.885
0.40	0.814	0.815	0.897	0.901	0.987	0.990

Table 4.1 shows that the sizes of L_{test} tend to converge to 5% when n increases from 250 to 750. As expected, the power values of the maximized version $L_{\max}$ are always larger than the corresponding sizes of L_{test}. With respect to power values, our finite sample results in Table 4.2 show that the sizes of L_{test} in almost all cases are greater than those of $L_{\max}$. This shows that the selection criterion proposed in (4.4) works well numerically.

5. Conclusion. We have proposed a new nonparametric test for the parametric specification of the mean function of long–range dependent time series. Asymptotic distributions of the proposed test for both the fixed and random design cases have been established. In addition, we have also proposed both the Simulation Procedure 4.1 and the Implementation Procedure 4.1 to implement the proposed test in practice. The finite–sample results show that the proposed test as well as the two procedures are practically applicable and implementable. Further topics including how to represent $\widehat{h}_{\text{test}}$ explicitly are left for future research.

6. Acknowledgments. The authors would like to thank the Co–Editor, the Associate Editor and the referees for their constructive and very helpful comments and suggestions on the original submission of this paper.

Earlier versions of this paper were presented at The University of Zheijiang in China in July 2005, the International Conference on Financial Engineering and Risk Management in Shanghai in July 2005, and the 49th Australian Mathematical Society Meeting on Financial Mathematics in Perth in September 2005. The authors would like to thank the seminar and conference participants, in particular, Vo Anh, Maxwell King, Nicolai Leonenko, Zhengyan Lin, Eckhard Platen and Qiman Shao for some constructive suggestions and comments.

Thanks also go to Isabel Casas and Gowry Sriananthakumar for their computing assistance and the Australian Research Council Discovery Grants Program for its financial support.

Appendix A. This appendix provides technical details for the asymptotic theory in Section 2. Appendix A.1 establishes several necessary preliminary lemmas. Appendix A.2 gives the proofs of Theorem 2.3 and 2.4. Theorems 2.1 and 2.2 are proved in Appendix A.3. Throughout the section, we denote constants by $C, C_1, ...$, which may have different values at each appearance.

A.1. *Technical lemmas.*

LEMMA A.1. *Let $\{e_t\}$ be a linear process defined by $e_t = \sum_{j=-\infty}^{\infty} \psi_j \eta_{t-j}$, where η_j are i.i.d. random variables with $E[\eta_1] = 0$, $E[\eta_1^2] = 1$ and $E[\eta_1^4] < \infty$, and $\gamma(0) < \infty$ where $\gamma(k) = E[e_t e_{k+t}] = \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+k}$. Then, for all j, k, s and t,*

$$\begin{aligned} E[e_j\, e_k\, e_s\, e_t] &= (E\eta_1^4 - 3) \sum_{m=-\infty}^{\infty} \psi_{j-m}\psi_{k-m}\psi_{s-m}\psi_{t-m} \\ &\quad + \gamma(j-k)\gamma(s-t) + \gamma(j-s)\gamma(k-t) + \gamma(j-t)\gamma(k-s). \end{aligned} \tag{A.1}$$

In particular, we have that $E\left[e_1^4\right] < \infty$,

$$\left|E\left[e_j^2 e_k^2\right] - \gamma^2(0)\right| \leq C\,\gamma^2(j-k), \tag{A.2}$$

$$\left|E\left[e_j^2\, e_k\, e_s\right]\right| \leq C\left[\gamma(j-k) + \gamma(j-s) + \gamma(k-s)\right], \tag{A.3}$$

for all $j \neq k \neq s$, and if in addition $\psi_k \geq 0$, then for all j, k, s and t,

$$\begin{aligned} &\left|E\left[\left(e_j e_{j+s} - \gamma(s)\right)\left(e_k e_{k+t} - \gamma(t)\right)\right]\right| \\ &\quad \leq C\,\gamma(j-k)\gamma(j-k+s-t) + \gamma(j-k+s)\gamma(j-k-t). \end{aligned} \tag{A.4}$$

PROOF. We only prove (A.1). By (A.1), others are simple. For all j, k, s and t, we may write

$$e_j\, e_k\, e_s\, e_t \sum_{m,n,g,l=-\infty}^{\infty} \psi_{j-m}\psi_{k-n}\psi_{s-g}\psi_{t-l}\, \eta_m\, \eta_n\, \eta_g\, \eta_l.$$

Recall that η_k are i.i.d. random variables with $E[\eta_1] = 0$, $E[\eta_1^2] = 1$ and $E[\eta_1^4] < \infty$.

Routine calculations imply that

$$\begin{aligned}
E[e_j\, e_k\, e_s\, e_t] &= E\eta_1^4 \sum_{m=-\infty}^{\infty} \psi_{j-m}\psi_{k-m}\psi_{s-m}\psi_{t-m} \\
&\quad + \sum_{\substack{m,g=-\infty\\ m\neq g}}^{\infty} \psi_{j-m}\psi_{k-m}\psi_{s-g}\psi_{t-g} + \sum_{\substack{m,g=-\infty\\ m\neq g}}^{\infty} \psi_{j-m}\psi_{k-g}\psi_{s-m}\psi_{t-g} \\
&\quad + \sum_{\substack{m,g=-\infty\\ m\neq g}}^{\infty} \psi_{j-m}\psi_{k-g}\psi_{s-g}\psi_{t-m} \\
&= (E\eta_1^4 - 3) \sum_{m=-\infty}^{\infty} \psi_{j-m}\psi_{k-m}\psi_{s-m}\psi_{t-m} \\
&\quad +\gamma(j-k)\gamma(s-t) + \gamma(j-s)\gamma(k-t) + \gamma(j-t)\gamma(k-s). \qquad \text{(A.5)}
\end{aligned}$$

This yields (A.1) and the proof of Lemma A.1 is complete.

LEMMA A.2. *Let $1/2 < \alpha < 1$ and $0 < \beta < \alpha - 1/2$. Then for all $k \geq 3$ and as $n \to \infty$,*

$$\begin{aligned}
I_n &= \frac{1}{n^{k/2}} \int_1^n \int_1^n \cdots \int_1^n |x_1 - x_2|^{-\alpha}|x_2 - x_3|^{\beta-1} \cdots \\
&\qquad |x_{2k-1} - x_{2k}|^{-\alpha}|x_{2k} - x_1|^{\beta-1}\, dx_1 dx_2 \cdots dx_{2k} \;\to\; 0. \qquad \text{(A.6)}
\end{aligned}$$

PROOF. Let $\alpha(x)$ and $\beta(x)$ be integrable real symmetric functions on $[-\pi, \pi]$ having the Fourier series:

$$\alpha(x) \sim \alpha_0 + \sum_{k=1}^{\infty} k^{-\alpha}\cos(kx) \quad \text{and} \quad \beta(x) \;\sim\; \beta_0 + \sum_{k=1}^{\infty} k^{\beta-1}\cos(kx).$$

where $\alpha_0 = \frac{1}{\pi}\int_0^\pi \alpha(x)dx$ and $\beta_0 = \frac{1}{\pi}\int_0^\pi \beta(x)dx$. Let R_n be a matrix with entries $(R_n)_{j,j} = \alpha_0$ and $(R_n)_{j,k} = |j-k|^{-\alpha}$ for $j \neq k$, and A_n be a matrix with entries $(A_n)_{j,j} = \beta_0$ and $(A_n)_{j,k} = |j-k|^{\beta-1}$ for $j \neq k$. Let $\mathrm{Tr}(M)$ denote the trace of matrix M. Recall $1/2 < \alpha < 1$ and $0 < \beta < \alpha - 1/2$. It is readily seen that, for each $\delta > 0$, as $x \to 0$,

$$\alpha(x) = O(|x|^{\alpha-1-\delta}) \quad \text{and} \quad \beta(x) = O(|x|^{-\beta-\delta}).$$

Now, by noting that $\alpha(x)$ and $\beta(x)$ have the Fourier coefficients:

$$r(0) = \alpha_0,\; r(k) = |k|^{-\alpha},\; |k| \geq 1, \quad \text{and} \quad a(0) = \beta_0,\; a(k) = |k|^{\beta-1},\; |k| \geq 1,$$

respectively, it follows easily from Theorem 1 of Fox and Taqqu (1987) (The regularity condition in Theorem 1 of Fox and Taqqu 1987 is not necessary. This has been mentioned in Giraitis and Surgailis 1990) that for all $k \geq 3$,

$$\begin{aligned}
I_n &\sim \frac{1}{n^{k/2}} \sum_{j_1, j_2, \ldots, j_{2k}=1}^{n} r(j_1 - j_2)\, a(j_2 - j_3) \cdots r(j_{2k-1} - j_{2k})\, a(j_{2k} - j_1) \\
&= \frac{1}{n^{k/2}} \mathrm{Tr}(R_n A_n)^k \;\leq\; C\, \max\{n^{-1/2}, o\big(n^{(\beta-\alpha+1/2)k+\epsilon}\big)\} \qquad \text{(A.7)}
\end{aligned}$$

for every arbitrarily small $\epsilon > 0$. This implies that $I_n \to 0$ as $0 < \beta < \alpha - 1/2$, and thus completes the proof of Lemma A.2.

In Lemma A.3 below, let $\{X_i, i \geq 1\}$ be a sequence of i.i.d. random variables with density function $f(x)$ and set $g_n(X_i, X_j) = K\left(\frac{X_i - X_j}{h}\right) - E\left[K\left(\frac{X_i - X_j}{h}\right)\right]$,

$$g_{1n}(X_i) = E\big[g_n(X_i, X_j)|X_i\big], \quad g_{2n}(X_i, X_j) = g_n(X_i, X_j) - g_{1n}(X_i) - g_{1n}(X_j). \tag{A.8}$$

LEMMA A.3. *Let $K(x)$ satisfy the Assumption 2.3(i). If $f(x)$ is a bounded and uniformly continuous function on $I\!R$, then*

$$E\left[K\big[(X_1 - X_2)/h\big]\right] \sim c_1\, h, \tag{A.9}$$

$$E\left[g_{2n}^2(X_1, X_2)\right] \sim E\left[K^2\big[(X_1 - X_2)/h\big]\right] \sim c_2\, h, \tag{A.10}$$

$$E\left[g_{2n}^4(X_1, X_2)\right] \sim E\left[K^4\big[(X_1 - X_2)/h\big]\right] \sim c_4\, h, \tag{A.11}$$

where $c_j = \int_{-\infty}^{\infty} K^j(s)ds \int_{-\infty}^{\infty} f^2(y)dy$ for $j = 1, 2, 4$. Furthermore,

$$E\left[g_{1n}^2(X_1)\right] \sim d_1\, h^2, \tag{A.12}$$

$$E\left[g_{2n}(X_1, X_3)\, g_{2n}(X_1, X_4)\, g_{2n}(X_2, X_3)\, g_{2n}(X_2, X_4)\right] \sim d_2\, h^3, \tag{A.13}$$

where

$$d_1 = \int_{-\infty}^{\infty} \left[f(x) - \int_{-\infty}^{\infty} f^2(y)dy\right]^2 f(x)dx, \tag{A.14}$$

$$d_2 = \iiint K(s)K(t)K(x+s)K(x+t)dsdtdx \int_{-\infty}^{\infty} f^4(y)dy. \tag{A.15}$$

PROOF. We only prove (A.13), as the others are similar. Write

$$\eta(s,t) = E\left[K\left(\frac{X_1 - X_2}{h}\right)\right] + g_{1n}(s) + g_{1n}(t).$$

Recall that $g_{1n}(s) = E\left[g_n(X_1, X_2)|X_1 = s\right]$ and $f(x)$ is a bounded and uniformly continuous density function. It is readily seen that, for all $(s,t) \in I\!R^2$,

$$\begin{aligned} |\eta(s,t)| &\leq \iint K\big[(x-y)/h\big] f(x) f(y) dxdy \\ &\quad + \int_{-\infty}^{\infty} \Big\{K\big[(s-y)/h\big] + K\big[(t-y)/h\big]\Big\} f(y)dy \\ &\leq 3\, h \sup_x f(x). \end{aligned} \tag{A.16}$$

This implies that, uniformly for (x, y) on $I\!R^2$,

$$
\begin{aligned}
f(x,y) &\equiv E\left[g_{2n}(X_1,x)\, g_{2n}(X_1,y)\right] \\
&= \int_{-\infty}^{\infty} \Big[K\big(\frac{s-x}{h}\big) - \eta(s,x)\Big]\Big[K\big(\frac{s-y}{h}\big) - \eta(s,y)\Big] f(s) ds \\
&= \int_{-\infty}^{\infty} K\big(\frac{s-x}{h}\big)\, K\big(\frac{s-y}{h}\big)\, f(s)\, ds + R_{1n} \\
&= h \int_{-\infty}^{\infty} K(s)\, K\big(\frac{x-y}{h} + s\big)\, f(x+sh)\, ds + R_{1n}, \qquad \text{(A.17)}
\end{aligned}
$$

where

$$
\begin{aligned}
|R_{1n}| &\leq \int_{-\infty}^{\infty} |\eta(s,x)| K\big(\frac{s-y}{h}\big)\, f(s)\, ds + \int_{-\infty}^{\infty} |\eta(s,y)| K\big(\frac{s-x}{h}\big)\, f(s)\, ds \\
&\quad + \int_{-\infty}^{\infty} |\eta(s,x)|\, |\eta(s,y)|\, f(s)\, ds \leq 15\, h^2 \sup_x f^2(x). \qquad \text{(A.18)}
\end{aligned}
$$

It is now readily seen that

$$
\begin{aligned}
&E\left[g_{2n}(X_1,X_3)\, g_{2n}(X_1,X_4)\, g_{2n}(X_2,X_3)\, g_{2n}(X_2,X_4)\right] = Ef^2(X_1,X_2) \\
&= h^2 \iint K(s)K(t)\, E\Big[K\big(\frac{X_1-X_2}{h} + s\big)\, K\big(\frac{X_1-X_2}{h} + t\big) \\
&\qquad \times f(X_1+sh) f(X_1+th)\Big] ds dt \;+\; R_{2n} \\
&\sim h^3 \iiint K(s)K(t)K(x+s)K(x+t) ds dt dx \int_{-\infty}^{\infty} f^4(y) dy, \qquad \text{(A.19)}
\end{aligned}
$$

where we have used the facts: under the conditions on $f(x)$ and $K(x)$,

$$
\begin{aligned}
&E\Big[K\big(\frac{X_1-X_2}{h} + s\big)\, K\big(\frac{X_1-X_2}{h} + t\big) f(X_1+sh) f(X_1+th)\Big] \\
&= \iint K\big(\frac{x-y}{h} + s\big)\, K\big(\frac{x-y}{h} + t\big) f(x+sh) f(x+th) f(x) f(y) dx dy \\
&\sim h \int_{-\infty}^{\infty} K(x+s) K(x+t) dx \int_{-\infty}^{\infty} f^4(y) dy, \qquad \text{(A.20)}
\end{aligned}
$$

and

$$
\begin{aligned}
|R_{2n}| &\leq 30h^3 \int_{-\infty}^{\infty} K(s)\, E\left[K\big(\frac{X_1-X_2}{h} + s\big)\, f(X_1+sh)\right]\, ds \\
&\quad + E\left[R_{1n}^2\right] \;=\; O(h^4).
\end{aligned}
$$

This proves (A.13), and hence completes the proof of Lemma A.3.

Our next lemma establishes a Berry–Esseen–type bound for random weighted U–statistics. This lemma is interesting and useful in itself.

LEMMA A.4. *Let $\{\epsilon_k, k \ge 1\}$ be a sequence of i.i.d. random variables. Let $\{a_{nij}\}$ be a sequence of constants with $a_{nij} = a_{nji}$ for all $n \ge 1$. Let $\{\varphi_n(x,y)\}$ be a sequence of symmetric Borel-measurable functions such that for all $n \ge 1$,*

$$E\left[\varphi_n^2(\epsilon_1,\epsilon_2)\right] > 0, \quad E\left[\varphi_n(\epsilon_1,\epsilon_2) \mid \epsilon_1\right] = 0. \tag{A.21}$$

Then there exists an absolute constant $A > 0$ such that

$$\sup_x \left|P\left(B_n^{-1}S_n \le x\right) - \Phi(x)\right| \le \; A\, B_n^{-4/5}\left(A_{1n}E\varphi_n^4(\epsilon_1,\epsilon_2) + A_{2n}\,\mathcal{L}_n\right)^{1/5},$$

where $S_n = \sum_{1\le i<j\le n} a_{nij}\,\varphi_n(\epsilon_i,\epsilon_j)$, $B_n^2 = \sum_{1\le i<j\le n} a_{nij}^2 E\varphi_n^2(\epsilon_1,\epsilon_2)$,

$$A_{1n} = \sum_{i=2}^{n}\Big(\sum_{j=1}^{i-1} a_{nij}^2\Big)^2, \quad A_{2n} = \sum_{i=2}^{n-1}\sum_{j=i+1}^{n}\Big(\sum_{k=1}^{i-1} a_{nik}\,a_{njk}\Big)^2, \tag{A.22}$$

$$\mathcal{L}_n = E\big[\varphi_n(\epsilon_1,\epsilon_3)\,\varphi_n(\epsilon_1,\epsilon_4)\,\varphi_n(\epsilon_2,\epsilon_3)\,\varphi_n(\epsilon_2,\epsilon_4)\big]. \tag{A.23}$$

PROOF. In the proof of Lemma A.4, we omit the subscripts n in a_{nij} and φ_n for convenience. Set, for $i = 2, 3, ..., n$,

$$Z_i = \sum_{k=1}^{i-1} a_{ik}\varphi(\epsilon_i,\epsilon_k), \quad \mathcal{F}_i = \sigma(\epsilon_1, ..., \epsilon_i)$$

It is readily seen that $S_n = \sum_{i=2}^n Z_i$ and $E(Z_i \mid \mathcal{F}_{i-1}) = 0, i = 2, 3, ..., n$, by (A.21). This implies that $\{S_j, \mathcal{F}_j, 2 \le j \le n\}$ forms a martingale sequence. Hence it follows from Theorem 3.9 with $\delta = 1$ in Hall and Heyde (1980) that there exists an absolute constant $A > 0$ such that

$$\sup_x \left|P\left(B_n^{-1}S_n \le x\right) - \Phi(x)\right| \le AB_n^{-4/5}\,M_n^{1/5}, \tag{A.24}$$

where $U_n^2 = \sum_{i=2}^n Z_i^2$ and $M_n = \sum_{i=2}^n E[Z_i^4] + E(U_n^2 - B_n^2)^2$.

Next we will show that

$$M_n \le 10\,A_{1n}\,E\varphi_n^4(\epsilon_1,\epsilon_2) + 4\,A_{2n}\,\mathcal{L}_n, \tag{A.25}$$

and then (A.22) follows immediately. In fact, by noting $B_n^2 = E[U_n^2]$,

$$M_n = \sum_{i=2}^{n} E[Z_i^4] + EU_n^4 - B_n^4 = 2\sum_{i=2}^{n} E[Z_i^4] + 2\sum_{2\le i<j\le n} E\left[Z_i^2 Z_j^2\right] - B_n^4. \tag{A.26}$$

Since the moments $E\big[\varphi(\epsilon_i,\epsilon_k)\varphi(\epsilon_i,\epsilon_{k_1})\varphi(\epsilon_j,\epsilon_l)\varphi(\epsilon_i,\epsilon_{l_1})\big]$, in which k, k_1, l, l_1 appear only once (not equal to i and j), always equal zero in view of (A.21), we obtain readily that, for all $i < j$,

$$
\begin{aligned}
E\left[Z_i^2 Z_j^2\right] &= \sum_{k,k_1=1}^{i-1} \sum_{l,l_1=1}^{j-1} a_{ik} a_{ik_1} a_{jl} a_{jl_1} E\big[\varphi(\epsilon_i,\epsilon_k)\varphi(\epsilon_i,\epsilon_{k_1})\varphi(\epsilon_j,\epsilon_l)\varphi(\epsilon_j,\epsilon_{l_1})\big] \\
&= \sum_{k=1}^{i-1}\sum_{l=1}^{j-1} a_{ik}^2 a_{jl}^2 E\big[\varphi^2(\epsilon_i,\epsilon_k)\,\varphi^2(\epsilon_j,\epsilon_l)\big] \\
&\quad + 2\sum_{\substack{k,l=1\\ k\neq l}}^{i-1} a_{ik}a_{il}a_{jk}a_{jl}\, E\big[\varphi(\epsilon_i,\epsilon_k)\,\varphi(\epsilon_i,\epsilon_l)\,\varphi(\epsilon_j,\epsilon_k)\,\varphi(\epsilon_j,\epsilon_l)\big] \\
&\quad + 2\sum_{k=1}^{i-1} a_{ik}^2 a_{ji} a_{jk}\, E\big[\varphi^2(\epsilon_i,\epsilon_k)\,\varphi(\epsilon_j,\epsilon_i)\,\varphi(\epsilon_j,\epsilon_k)\big] \\
&= R_{1ij} + R_{2ij} + R_{3ij}, \qquad \text{(A.27)}
\end{aligned}
$$

where

$$
\begin{aligned}
R_{1ij} &= \sum_{k=1}^{i-1}\sum_{\substack{l=1\\ l\neq k}}^{j-1} a_{ik}^2 a_{jl}^2 \left(E\left[K^2(\epsilon_1,\epsilon_2)\right]\right)^2 \\
R_{2ij} &= 2\sum_{k,l=1}^{i-1} a_{ik}a_{il}a_{jk}a_{jl}\, E\big[\varphi(\epsilon_1,\epsilon_3)\varphi(\epsilon_1,\epsilon_4)\varphi(\epsilon_2,\epsilon_3)\varphi(\epsilon_2,\epsilon_4)\big] \\
|R_{3ij}| &\le \sum_{k=1}^{i-1} a_{ik}^2 \big(3a_{jk}^2 + 2|a_{ji}a_{jk}|\big) E\left[\varphi^4(\epsilon_1,\epsilon_2)\right] \\
&\le 4\sum_{k=1}^{i-1} a_{ik}^2\big(a_{jk}^2 + a_{ji}^2\big) E\left[\varphi^4(\epsilon_1,\epsilon_2)\right].
\end{aligned}
$$

Similarly, for all $2 \le i \le n$,

$$
\begin{aligned}
E[Z_i^4] &= \sum_{j=1}^{i-1} a_{ij}^4 E[\varphi^4(\epsilon_1,\epsilon_2)] + \sum_{1\le j\neq k\le i-1} a_{ij}^2 a_{ik}^2 E\big[\varphi^2(\epsilon_1,\epsilon_2)\varphi^2(\epsilon_1,\epsilon_3)\big] \\
&\le \Big(\sum_{j=1}^{i-1} a_{ij}^2\Big)^2 E\left[\varphi^4(\epsilon_1,\epsilon_2)\right]. \qquad \text{(A.28)}
\end{aligned}
$$

By virtue of (A.27) and (A.28), it is readily seen that $\sum_{i=2}^n E\left[Z_i^4\right] \le A_{1n} E\left[\varphi^4(\epsilon_1,\epsilon_2)\right]$ and

$$
2\sum_{2\le i<j\le n} E\left[Z_i^2 Z_j^2\right] \le B_n^4 + 4A_{2n}\mathcal{L}_n + 8\,A_{1n}\,E\left[\varphi^4(\epsilon_1,\epsilon_2)\right].
$$

Taking these estimates back into (A.26), we obtain the inequality (A.25). The proof of Lemma A.4 is now complete.

LEMMA A.5. *Let $\{\epsilon_k, k \geq 1\}$ be a sequence of i.i.d. random variables with $E\epsilon_1 = 0$ and $E\epsilon_1^6 < \infty$. Let $\{a_{nij}\}$ be a sequence of real numbers with $a_{nij} = a_{nji}$ and $||A||^2 \equiv \sum_{i,j=-\infty}^{\infty} a_{nij}^2 < \infty$ for all $n \geq 1$. If there exists an absolute constant $b_1^2 > 0$ such that $1 - \frac{V^2}{||A||^2} \geq b_1^2$ with $V^2 = \sum_{i=-\infty}^{\infty} a_{nii}^2$, then*

$$\sup_x \Big| P\big(S_n/B_n \leq x\big) - \Phi(x) \Big| \quad \leq \quad C \, \frac{\big\{\mathrm{Tr}(A^4)\big\}^{1/4}}{||A||}, \tag{A.29}$$

where $S_n = \sum_{i,j=-\infty}^{\infty} a_{nij}\, (\epsilon_i\, \epsilon_j - E[\epsilon_i \epsilon_j])$, $B_n^2 = 2(||A||^2 - V^2)\mu_2^2 + V^2(\mu_4 - \mu_2^2)$ with $\mu_j = E|\epsilon_1|^j$, and A is the infinite matrix with a_{nij} as its (i,j)th element.

The proof of Lemma A.5 follows immediately from Theorem 1.1 of Götze and Tikhomirov (2002), together with (2.4) and Remark 1.8 in the same paper. We omit the details.

LEMMA A.6. *Let $K(x)$ be a non-negative symmetric integrable function satisfying $K(x) = O\big[(1 + |x|^{1-\beta})^{-1}\big]$, where $0 < \beta \leq \alpha - 1/2$ and $1/2 < \alpha < 1$. Then,*

$$\Delta_0 \quad \equiv \quad \int_0^\infty x^{-\alpha} K(x) dx < \infty, \tag{A.30}$$

$$A_\alpha \quad \equiv \quad \int_0^\infty\!\!\int_0^\infty\!\!\int_0^\infty x^{-\alpha}\, y^{-\alpha} \big[I_1(x,y,w) + I_2(x,y,w)\big]\, dxdydw < \infty, \tag{A.31}$$

and as $h \to 0$,

$$\begin{aligned} \Delta_1 \quad \equiv \quad & \int_0^{1/h}\!\!\int_0^{1/h}\!\!\int_x^{1/h} x^{-\alpha}\, y^{-\alpha} \, \max\{w, y\} \big[I_1(x,y,w) \\ & \qquad\qquad\qquad + I_2(x,y,w)\big]\, dxdydw \quad = \quad o(1/h), \end{aligned} \tag{A.32}$$

where $I_1(x,y,w) = K(w)K(x+y-w)$ and $I_2(x,y,w) = K(w-x)K(w-y)$.

PROOF. The proof of (A.30) is easy. We now prove (A.31) and (A.32). Note that $\beta - 2\alpha < -\alpha - 1/2 < -1$, and for any $u \in I\!\!R$,

$$\int_0^\infty K(w)K(w+u)\, dw \leq C/(1 + |u|^{1-\beta}). \tag{A.33}$$

It is readily seen that

$$\begin{aligned} A_\alpha \quad & \leq \quad \int_0^\infty\!\!\int_0^\infty\!\!\int_0^\infty x^{-\alpha}\, y^{-\alpha}\, K(w)\big[K(w+x-y) + K(w-x-y)\big]\, dxdydw \\ & \leq \quad C + C_1 \int_1^\infty\!\!\int_1^\infty x^{-\alpha}\, y^{-\alpha} \frac{dxdy}{1 + |x-y|^{1-\beta}} \\ & \leq \quad C + C_2 \int_1^\infty x^{\beta - 2\alpha}\, dx < \infty, \end{aligned} \tag{A.34}$$

which implies (A.31). Similarly, it follows from (A.33) and $1+\beta-2\alpha<0$ that

$$\begin{aligned}
\Delta_1 &\le C+\int_1^{1/h}\int_1^{1/h}\int_1^{1/h} x^{-\alpha}\,y^{-\alpha}\,(w+y)\,K(w) \\
&\qquad\qquad \times [K(w+x-y)+K(w-x-y)]\,dxdydw \\
&\le C+C_1\,(1/h)^{\beta}\int_1^{1/h}\int_1^{1/h} x^{-\alpha}\,y^{-\alpha}dxdy \\
&\qquad\qquad + C_2\int_1^{1/h}\int_1^{1/h} x^{-\alpha}\,y^{1-\alpha}\frac{dxdy}{1+|x-y|^{1-\beta}} \\
&\le C+C_3\,(1/h)^{2-2\alpha+\beta} = o(1/h), \qquad \text{(A.35)}
\end{aligned}$$

which yields (A.32). This also completes the proof of Lemma A.6.

A.2. *Proofs of Theorems 2.3 and 2.4.*

Proof of Theorem 2.3. We may write

$$\sum_{1\le i\ne j\le n} e_i\,e_j\left[K\left(\frac{(X_i-X_j)}{h}\right)\right]=\widetilde{Q}_{2n}^{(1)}+\widetilde{Q}_{2n}^{(2)}+\widetilde{Q}_{2n}^{(3)},$$

where

$$\widetilde{Q}_{2n}^{(1)} = \sum_{1\le i\ne j\le n} e_i\,e_j\,E\left[K\left(\frac{(X_i-X_j)}{h}\right)\right], \qquad \text{(A.36)}$$

$$\widetilde{Q}_{2n}^{(2)} = \sum_{1\le i\ne j\le n} e_i\,e_j\,[g_{1n}(X_i)+g_{1n}(X_j)], \qquad \text{(A.37)}$$

$$\widetilde{Q}_{2n}^{(3)} = \sum_{1\le i\ne j\le n} e_i\,e_j\,g_{2n}(X_i,X_j), \qquad \text{(A.38)}$$

where $g_{1n}(X_i)$ and $g_{2n}(X_i,X_j)$ are defined as in (A.8). Theorem 2.3 now follows easily if we prove: whenever $h\to 0$,

$$\left(A_{2\alpha}n^{2-\alpha}\,h\right)^{-1}\widetilde{Q}_{2n}^{(1)} \to_D \chi^2(1), \qquad \text{(A.39)}$$

$$\widetilde{Q}_{2n}^{(2)} = o_P\left(\max\left\{n^{2-\alpha}\,h, n\sqrt{h}\right\}\right), \qquad \text{(A.40)}$$

and if in addition $nh\to\infty$, then

$$\left(A_{1\alpha}\,n\,\sqrt{h}\right)^{-1}\widetilde{Q}_{2n}^{(3)} \to_D N(0,1). \qquad \text{(A.41)}$$

Actually, if $h\to 0$ and $\sqrt{h}\,n^{1-\alpha}\to\infty$, then $\widetilde{Q}_{2n}^{(2)}+\widetilde{Q}_{2n}^{(3)}=o_P\left(n^{2-\alpha}h\right)$ by virtue of (A.40) and (A.41). This, together with (A.39), yields Theorem 2.3(ii). Similarly, if $\sqrt{h}\,n^{1-\alpha}\to 0$ and $nh\to\infty$, then $\widetilde{Q}_{2n}^{(1)}+\widetilde{Q}_{2n}^{(2)}=o_P\left(n\sqrt{h}\right)$ by virtue of (A.39) and (A.40). This, together with (A.41), yields Theorem 2.3(i).

We now prove (A.39)-(A.41). (A.39) first. By (A.9),

$$\left(A_{2\alpha} n^{2-\alpha}\, h\right)^{-1} \widetilde{Q}_{2n}^{(1)} = (1+o_P(1))\Big[\Big(\frac{1}{d_n}\sum_{j=1}^n e_j\Big)^2 - \frac{1}{d_n^2}\sum_{j=1}^n e_j^2\Big],$$

where $d_n^2 = \frac{2\eta}{(1-\alpha)(2-\alpha)} n^{2-\alpha}$. It is readily seen that $\frac{1}{d_n^2}\sum_{j=1}^n e_j^2 \to 0$ a.s. by the stationary ergodic theorem. This, together with (A.42) and the continuous mapping theorem, yields that (A.39) will follow if we prove

$$\frac{1}{d_n}\sum_{j=1}^n e_j \to_D N(0,1). \tag{A.42}$$

In fact, by letting $\nu_{jn} = \sum_{t=1}^n \psi_{t-j}$ and recalling $\gamma(k) = \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+k} \sim \eta |k|^{-\alpha}$, simple calculations show that, $\sum_{j=1}^n e_j = \sum_{j=-\infty}^{\infty} \nu_{jn}\eta_j$,

$$\sum_{j=-\infty}^{\infty} \nu_{jn}^2 = \sum_{i,j=1}^n \gamma(i-j) = n\gamma(0) + 2\sum_{k=1}^{n-1}(n-k)\gamma(k) \sim d_n^2, \tag{A.43}$$

and $\max_j |\nu_{jn}| \le \sqrt{n}\,\gamma^{1/2}(0) = o(d_n)$. Equation (A.42) now follows from Lemma 1 of Robinson (1997). This proves (A.39).

Second, we prove (A.40). In fact, by (A.12), independence of e_i and X_i and (A.3),

$$\begin{aligned} E\left(\widetilde{Q}_{2n}^{(2)}\right)^2 &= E\Big(\sum_{i=1}^n g_{1n}(X_i)\, e_i \sum_{1\le j\ne i\le n} e_j\Big)^2 \\ &= [d_1\, h^2 + o(h^2)]\left(\sum_{1\le i\ne j\le n} E\left[e_i^2\, e_j^2\right] + \sum_{1\le i\ne k\ne j\le n} E\left[e_i^2 e_j e_k\right]\right) \\ &\le C\, h^2 \left[n^2\, E e_1^4 + n\sum_{i,j=1}^n \gamma(i-j)\right] \le C\, h^2\left(n^2 + n^{3-\alpha}\right). \end{aligned}$$

Thus, equation (A.40) follows immediately from the Markov's inequality.

Finally, we prove (A.41). Write

$$B_n^2 = \sum_{1\le i<j\le n} e_i^2\, e_j^2\, E\left[g_{2n}^2(X_i,X_j)\right] = \frac{1}{2}\, E\left[g_{2n}^2(X_1,X_2)\right]\cdot\left[\left(\sum_{i=1}^n e_i^2\right)^2 - \sum_{i=1}^n e_i^4\right].$$

By (A.10) and the stationary ergodic theorem which yields that $\frac{1}{n}\sum_{i=1}^n e_i^2 \to E[e_1^2] = \gamma(0)$ a.s. and $\frac{1}{n}\sum_{i=1}^n e_i^4 \to E[e_1^4] < \infty$ a.s., it is readily seen that as $n\to\infty$

$$4\, A_{1\alpha}^{-2}\, n^{-2}\, h^{-1}\, B_n^2 \to 1, \quad a.s., \tag{A.44}$$

where $A_{1\alpha}$ is defined as in Theorem 2.1(i). So, to prove (A.41), it suffices to show that

$$\widetilde{Q}_{2n}^{(3)}/(2B_n) \to_D N(0,1). \tag{A.45}$$

Lemma A.4 will be used to establish (A.45). In fact, by noting that

$$\widetilde{Q}_{2n}^{(3)} = 2 \sum_{1\le i<j\le n} e_i\, e_j\, g_{2n}(X_i, X_j)$$

and $E\big[g_{2n}(X_i, X_j)|e_i\big] = 0$ for all $i \neq j$, it follows from the independence of e_i and X_i, Lemma A.4, (A.11) and (A.13) that

(A.46)
$$\sup_x \big|P\big(\widetilde{Q}_{2n}^{(3)}/2B_n \le x|e_1, ..., e_n\big) - \Phi(x)\big| \le\; A\, B_n^{-4/5}\big(c_4\, h\, A_{1n} + d_2\, h^3\, A_{2n}\big)^{1/5},$$

where A is an absolute constant, c_4 and d_2 are defined as in (A.11) and (A.13), and

$$\begin{aligned} A_{1n} &= \sum_{i=2}^{n}\left(\sum_{j=1}^{i-1}(e_ie_j)^2\right)^2 \le \sum_{i=2}^{n} e_i^4\left(\sum_{j=1}^{n} e_j^2\right)^2, \\ A_{2n} &= \sum_{i=2}^{n-1}\sum_{j=i+1}^{n}\left(\sum_{k=1}^{i-1} e_i\, e_j\, e_k^2\right)^2 \le \left(\sum_{1\le i<j\le n} e_i^2\, e_j^2\right)^2. \end{aligned}$$

By the stationary ergodic theorem again, for n large enough,

$$\frac{1}{n^3}A_{1n} \le 2E[e_1^4]\cdot(E[e_1^2])^2 \quad a.s. \quad \text{and} \quad \frac{1}{n^4}A_{2n} \le 2\left(E[e_1^2]\right)^4 \quad a.s.$$

This, together with (A.44) and (A.46), implies that, for n large enough,

$$\sup_x \left|P\left(\widetilde{Q}_{2n}^{(3)}/(2B_n) \le x|e_1, ..., e_n\right) - \Phi(x)\right| \le C\left(\frac{1}{nh} + h\right)^{1/5}, \quad a.s.$$

Now, if $h \to 0$ and $nh \to \infty$, then

$$\begin{aligned} &\lim_{n\to\infty}\sup_x \big|P\big(\widetilde{Q}_{2n}^{(3)}/(2B_n) \le x\big) - \Phi(x)\big| \\ &\qquad \le\; E\left[\lim_{n\to\infty}\sup_x\left|P\left(\widetilde{Q}_{2n}^{(3)}/(2B_n) \le x \mid e_1, ..., e_n\right) - \Phi(x)\right|\right] = 0. \end{aligned}$$

This proves (A.45), and hence also completes the proof of Theorem 2.3.

Proof of Theorem 2.4. Let

$$\widetilde{Q}_{1n} = K(0)\textstyle\sum_{i=1}^n e_i^2 \quad \text{and} \quad Q_{1n} = \sum_{i=1}^n\sum_{j=1, j\ne i}^n e_i\, e_j\, b_n(i,j).$$

We have that

$$\begin{aligned} \widetilde{Q}_n &\equiv Q_{1n} - EQ_{1n} + \widetilde{Q}_{1n} - E\widetilde{Q}_{1n} = \sum_{i,j=1}^{n}\big(e_i\, e_j - \gamma(i-j)\big)\, b_n(i,j) \\ &= \sum_{k,l=-\infty}^{\infty} a_{nkl}\left(\eta_k\eta_l - E\left[\eta_k\eta_l\right]\right), \qquad \text{(A.47)} \end{aligned}$$

where $a_{nkl} = \sum_{i,j=1}^{n} \psi_{i-k}\psi_{j-l}\, b_n(i,j)$ and we have used the facts that $E[\eta_k \eta_l] = 0$ for $k \neq l$, $E\eta_k^2 = 1$ and

$$\sum_{k=-\infty}^{\infty} a_{nkk} = \sum_{i,j=1}^{n} K\left(\frac{i-j}{nh}\right) \sum_{k=-\infty}^{\infty} \psi_{i-k}\psi_{j-k} = \sum_{i,j=1}^{n} \gamma(i-j)\; b_n(i,j).$$

By Lemma A.5, in order to prove Theorem 2.4, it suffices to show as $n \to \infty$

$$\frac{(\widetilde{Q}_{1n} - E\widetilde{Q}_{1n})}{\tau_n} \quad \to_P \quad 0, \tag{A.48}$$

$$2\,||A||^2 \;\equiv\; 2 \sum_{k,l=-\infty}^{\infty} a_{nkl}^2 \quad \sim \quad A_0^2\, \tau_n^2, \tag{A.49}$$

$$V^2 = \sum_{k=-\infty}^{\infty} a_{nkk}^2 \quad = \quad o(\tau_n^2), \tag{A.50}$$

$$\mathrm{Tr}(A^4) \quad = \quad o(\tau_n^4), \tag{A.51}$$

where $\tau_n = n^{2-\alpha}\, h^{3/2-\alpha}$ and $A_0^2 = 8\eta^2 A_\alpha$ with A_α being defined in (2.3). Indeed, by virtue of (A.49)–(A.51), it follows from Lemma A.5 that

$$(\widetilde{Q}_n - E[\widetilde{Q}_n])/\tau_n \to_D A_0\, N(0,1).$$

This, together with (A.48), yields Theorem 2.4.

In the following, we give the proofs of (A.48)-(A.51). (A.48) first. Recall $1/2 < \alpha < 1$ and $\gamma(k) \sim \eta |k|^{-\alpha}$. By virtue of (A.2), it is readily seen that

$$\begin{aligned} E\left[\left(\widetilde{Q}_{1n} - E\widetilde{Q}_{1n}\right)^2\right] &= K^2(0) E\Big[\sum_{k=1}^{n} (e_k^2 - Ee_k^2)\Big]^2 \\ &\leq C \sum_{j,k=1}^{n} \gamma^2(j-k) \leq C \sum_{k=1}^{n} (n-k)\gamma^2(k) \leq C\, n. \end{aligned}$$

This, together with the Markov's inequality and $nh \to \infty$, yields (A.48).

Secondly, we prove (A.49). We have

$$\begin{aligned} ||A||^2 &= \sum_{k,l=-\infty}^{\infty} \left(\sum_{i,j=1}^{n} \psi_{i-k}\psi_{j-l}\; K\left(\frac{i-j}{nh}\right) \right)^2 \\ &= \sum_{i,j,s,t=1}^{n} \sum_{k,l=-\infty}^{\infty} \psi_{i-k}\psi_{j-l}\psi_{s-k}\psi_{t-l}\, K\left(\frac{i-j}{nh}\right)\; K\left(\frac{s-t}{nh}\right) \\ &= \sum_{i,j,s,t=1}^{n} K\left(\frac{i-j}{nh}\right)\; K\left(\frac{s-t}{nh}\right)\; \gamma(i-s)\,\gamma(t-j). \end{aligned}$$

Write $f_n(x,y;\,z,w) = K\left(\frac{x-z}{nh}\right)\,K\left(\frac{y-w}{nh}\right) + K\left(\frac{y-z}{nh}\right)\,K\left(\frac{x-w}{nh}\right)$. Clearly, $f_n(\cdot,\cdot;\cdot,\cdot)$ has the following symmetry in its indexes:

$$f_n(x,y;\,z,w) = f_n(y,x;\,z,w) = f_n(x,y;\,w,z) = f_n(y,x;\,w,z).$$

Also $f_n(x,y;\,z,w) = f_n(z,w;\,x,y)$. By noting that for any function $g(x,y)$ and symmetric function $b(x)$,

$$\sum_{i,j=1}^{n} b(i-j)g(i,j) = b(0)\sum_{i=1}^{n} g(i,i) + \sum_{i=1}^{n-1}\sum_{j=1}^{n-i} b(i)\left[g(j,j+i) + g(j+i,j)\right],$$

some algebras show that (noting $\gamma(k)=\gamma(-k)$)

$$\begin{aligned}
2\,||A||^2 &= \sum_{i_1,i_2,j_1,j_2=1}^{n} \gamma(i_1-i_2)\,\gamma(j_1-j_2)\,f_n(i_1,i_2;j_i,j_2) \\
&= \gamma^2(0)\sum_{i,j=1}^{n} f_n(i,i;\,j,j) + 4\,\gamma(0)\sum_{i=1}^{n}\sum_{j=1}^{n-1}\sum_{k=1}^{n-j}\gamma(i)\,f_n(i,i;\,k,j+k) \\
&+ 4\sum_{i_1=1}^{n-1}\sum_{i_2=1}^{n-i_1}\sum_{j_1=1}^{n-1}\sum_{j_2=1}^{n-j_1}\gamma(i_1)\,\gamma(j_1)\,f_n(i_2,i_1+i_2;\,j_2,j_1+j_2) \\
&\equiv \Delta_{1n} + 4\,\Delta_{2n} + 4\,\Delta_{3n}.
\end{aligned} \tag{A.52}$$

Recalling that $K(x)$ is a probability density function and $\gamma(x)\sim\eta\ x^{-\alpha}$ for $0<\alpha<1$ and $x>0$, we have that

$$\Delta_{3n} \sim \int_1^{n-1}\!\!\int_1^{n-1}\!\!\int_1^{n-x}\!\!\int_1^{n-y} \gamma(x)\,\gamma(y)\,f_n(z,x+z;\,w,y+w)\,dxdydzdw. \tag{A.53}$$

Write $g_n(x,y) = \int_1^{n-x}\int_1^{n-y} f_n(z,x+z;\,w,y+w)\,dzdw$. By the Fubini's theorem,

$$\begin{aligned}
g_n(x,y) &= \int_1^{n-x}\!\!\int_{1-z}^{n-y-z} f_n(z,x+z;\,w+z,y+w+z)\,dwdz \\
&= \int_{-(n-x-1)}^{n-y-1}\!\!\int_{\max\{1,1-w\}}^{\min\{n-x,n-y-w\}} f_n(0,x;\,w,y+w)\,dzdw \\
&= \int_0^{n-y-1}\big(n-1-\max\{x,y+w\}\big)\,f_n(0,x;\,w,y+w)\,dw \\
&\qquad + \int_0^{n-x-1}\big(n-1-\max\{x+w,y\}\big)\,f_n(0,x;\,-w,y-w)\,dw \\
&= \int_y^{n-1}\big(n-1-\max\{x,w\}\big)\,f_n(0,x;\,w-y,w)\,dw \\
&\qquad + \int_x^{n-1}\big(n-1-\max\{w,y\}\big)\,f_n(0,x;\,x-w,x+y-w)\,dw.
\end{aligned}$$

Taking this into (A.53), simple calculations show that if $h \to 0$ and $nh \to \infty$, then

$$\begin{aligned}
\Delta_{3n} &\sim \int_1^{n-1}\int_1^{n-1} \gamma(x)\,\gamma(y)\,g_n(x,y)\,dxdy \\
&\sim 2\int_1^{n-1}\int_1^{n-1}\int_x^{n-1} \gamma(x)\,\gamma(y)\big(n-1-\max\{w,y\}\big) \\
&\qquad \left[K\left(\frac{w}{nh}\right)K\left(\frac{x+y-w}{nh}\right) + K\left(\frac{w-x}{nh}\right)K\left(\frac{w-y}{nh}\right)\right] dxdydw \\
&\sim 2\,n\,(nh)^{3-2\alpha}\,\eta^2 \int_0^{1/h}\int_0^{1/h}\int_x^{1/h} x^{-\alpha}\,y^{-\alpha}\left(1-h\max\{w,y\}\right) \\
&\qquad \times\big[K(w)K(x+y-w) + K(w-x)K(w-y)\big]\,dxdydw \\
&\sim 2\,n\,(nh)^{3-2\alpha}\,\eta^2 \int_0^{\infty}\int_0^{\infty}\int_x^{\infty} x^{-\alpha}\,y^{-\alpha} \\
&\qquad \times\big[K(w)K(x+y-w) + K(w-x)K(w-y)\big]\,dxdydw \\
&= 2\,n\,(nh)^{3-2\alpha}\,\eta^2\,A_\alpha = n\,(nh)^{3-2\alpha}\,A_0^2/4, \qquad \text{(A.54)}
\end{aligned}$$

where we have used the facts that $K(x)$ is symmetric, $A_\alpha^2 < \infty$ and (A.32).

By a similar argument, if $h \to 0$ and $nh \to \infty$, then

$$\Delta_{1n} + 4\,\Delta_{2n} = O(n^{3-\alpha}h^2) = o(\Delta_{3n}). \qquad \text{(A.55)}$$

By virtue of (A.52), (A.54) and (A.55), we obtain the proof of (A.49).

Thirdly, we prove (A.50). Let

$$h(i,j,s,t) = \sum_{k=-\infty}^{\infty} \psi_{i-k}\psi_{j-k}\psi_{s-k}\psi_{t-k} = \sum_{k=-\infty}^{\infty} \psi_k\psi_{j-i+k}\psi_{s-i+k}\psi_{t-i+k}.$$

By $\psi_j \geq 0$ and $K(x) \geq 0$, it is readily seen that, for any $j \geq 0$, s and t,

$$\sum_{i=1}^{n} h(i,j+i,s,t) \leq \sum_{k=-\infty}^{\infty} \psi_k\psi_{j+k} \sum_{i=1}^{n} \psi_{s-i+k}\psi_{t-i+k} \leq \gamma(j)\,\gamma(t-s).$$

Therefore, as in (A.52)–(A.54), it follows from (A.30) that

$$
\begin{aligned}
V^2 &= \sum_{k=-\infty}^{\infty} \left(\sum_{i,j=1}^{n} \psi_{i-k}\psi_{j-k} K\left(\frac{i-j}{nh}\right)\right)^2 \\
&= \sum_{i,j,s,t=1}^{n} K\left(\frac{i-j}{nh}\right) K\left(\frac{s-t}{nh}\right) h(i,j,s,t) \\
&\le K(0) \sum_{i,s,t=1}^{n} K\left(\frac{s-t}{nh}\right) h(i,i,s,t) \\
&\qquad + 2\sum_{j=1}^{n} K\left(\frac{j}{nh}\right) \sum_{i,s,t=1}^{n} K\left(\frac{s-t}{nh}\right) h(i,j+i,s,t) \\
&\le \left[K(0)\gamma(0) + 2\sum_{j=1}^{n} K\left(\frac{j}{nh}\right) \gamma(j)\right] \sum_{s,t=1}^{n} K\left(\frac{s-t}{nh}\right) \gamma(s-t) \\
&\le C\, n \left(K(0)\gamma(0) + 2\int_1^n x^{-\alpha} K\left(\frac{x}{nh}\right) dx \right) \int_1^n x^{-\alpha} K\left(\frac{x}{nh}\right) dx \\
&\le C\, n^{3-2\alpha}\, h^{2-2\alpha} = o(\tau_n^2) \qquad \text{(A.56)}
\end{aligned}
$$

since $nh \to \infty$. This proves (A.50).

Finally, we prove (A.51). Tedious but simple calculations show that

$$
\begin{aligned}
\mathrm{Tr}(A^4) &= \sum_{i,j,l,m=-\infty}^{\infty} a_{nij}\, a_{njl}\, a_{nlm}\, a_{nmi} \\
&= \sum_{i,j,l,m=-\infty}^{\infty} \sum_{j_1,j_2,\cdots,j_7,j_8=1}^{n} \psi_{j_1-i}\psi_{j_2-j}\psi_{j_3-j}\psi_{j_4-l}\psi_{j_5-l}\psi_{j_6-m}\psi_{j_7-m}\psi_{j_8-i} \\
&\qquad \times K\left(\frac{j_1-j_2}{nh}\right) K\left(\frac{j_3-j_4}{nh}\right) K\left(\frac{j_5-j_6}{nh}\right) K\left(\frac{j_7-j_8}{nh}\right) \\
&= \sum_{j_1,j_2,\cdots,j_7,j_8=1}^{n} K\left(\frac{j_1-j_2}{nh}\right)\gamma(j_2-j_3)\cdots K\left(\frac{j_7-j_8}{nh}\right)\gamma(j_8-j_1).
\end{aligned}
$$

Recall that $K(x) = O[(1+|x|^{1-\beta})^{-1}]$. Similarly to the proof of (A.49), it follows from Lemma A.2 that,

$$\begin{aligned}\mathrm{Tr}(A^4) &\sim \int_1^n\int_1^n\cdots\int_1^n K\left(\frac{x_1-x_2}{nh}\right)\gamma(x_2-x_3)\cdots K\left(\frac{x_7-x_8}{nh}\right)\gamma(x_8-x_1)\\ &\qquad dx_1dx_2\cdots dx_7dx_8\\ &\sim \eta^4\ (nh)^{4(2-\alpha)}\int_0^{1/h}\int_0^{1/h}\cdots\int_0^{1/h}K(x_1-x_2)|x_2-x_3|^{-\alpha}\\ &\qquad K\big(x_7-x_8\big)\,|x_8-x_1|^{-\alpha}dx_1dx_2\cdots dx_7dx_8\\ &= O(1)\,(nh)^{4(2-\alpha)}\int_0^{1/h}\int_0^{1/h}\cdots\int_0^{1/h}|x_1-x_2|^{\beta-1}\,|x_2-x_3|^{-\alpha}\cdots\\ &\qquad |x_7-x_8|^{\beta-1}\,|x_8-x_1|^{-\alpha}dx_1dx_2\cdots dx_7dx_8\\ &= o(1)\,(nh)^{4(2-\alpha)}\,(1/h)^2\ =\ o(1)\,\tau_n^4.\end{aligned}$$

This yields (A.51) and thus completes the proof of Theorem 2.4.

A.3. *Proofs of Theorems 2.1 and 2.2.*

PROOF OF THEOREM 2.1. Note that

$$\left|n^{(\alpha-\hat\alpha)}-1\right|\le|\alpha-\hat\alpha|\,\log n\,\exp\{|\alpha-\hat\alpha|\,\log n\}. \tag{A.57}$$

It follows easily from (2.7), Theorems 2.3 and Assumption 2.2 that Theorems 2.1 will follow if we prove

$$\frac{1}{n}\sum_{t=1}^n\hat f(X_t) = \frac{1}{n^2h}\sum_{i,j=1}^n K\Big(\frac{X_i-X_j}{h}\Big)\to_P\int_{-\infty}^{\infty}f^2(x)dx, \tag{A.58}$$

$$\frac{1}{n}\sum_{t=1}^n\hat e_t^2 = \frac{1}{n}\sum_{t=1}^n(\epsilon_t+m_{\theta_0}(X_t)-m_{\tilde\theta}(X_t))^2\to_P\gamma(0), \tag{A.59}$$

and under the corresponding conditions of Theorem 2.1,

$$2R_{1n}(h)+R_{2n}(h) = o_P(\sigma_{in}(h)),\quad i=1,\ 2. \tag{A.60}$$

Recall that $K\big(\frac{X_1-X_2}{h}\big)\sim h\int_{-\infty}^{\infty}f^2(x)dx$ by (A.9). The proof of (A.58) follows from a standard method and hence the details are omitted. By the stationary ergodic theorem, $\frac{1}{n}\sum_{t=1}^n e_t^2\to\gamma(0)$, a.s. This implies that (A.59) will follow if we have

$$\frac{1}{n}\sum_{t=1}^n\Big[2\epsilon_t(m_{\theta_0}(X_t)-m_{\tilde\theta}(X_t))+(m_{\theta_0}(X_t)-m_{\tilde\theta}(X_t))^2\Big]\to_P 0. \tag{A.61}$$

Since the proof of (A.61) is similar to (A.60), we only prove (A.60) in the following.

For $\forall\epsilon>0$, write $\Omega_n=\{\widetilde\theta:||\widetilde\theta-\theta_0||\le\epsilon\,n^{-\alpha/2}\}$. Also, let

$$\begin{aligned}J_1(s,t) &= K\left(\frac{X_s-X_t}{h}\right)\Big(m_{\theta_0}(X_s)-m_{\widetilde\theta}(X_s)\Big),\\ J_2(s,t) &= K\left(\frac{X_s-X_t}{h}\right)\Big(m_{\theta_0}(X_s)-m_{\widetilde\theta}(X_s)\Big)\Big(m_{\theta_0}(X_t)-m_{\widetilde\theta}(X_t)\Big).\end{aligned}$$

Recall K is bounded and f is bounded and uniformly continuous. It follows easily from Assumption 2.4 that, under H_0, for all $s \neq s_1 \neq t$ and for n large enough such that $\Omega_n \subseteq \Theta_0$,

$$\begin{aligned} E\left[J_1^2(s,t)I(\widetilde{\theta} \in \Omega_n)\right] &\leq C\,\epsilon^2\, n^{-\alpha} E\left[K^2\Big(\frac{X_1 - X_2}{h}\Big)\Big|\Big|\frac{\partial m_\theta(X_1)}{\partial\theta}|_{\theta=\theta_0}\Big|\Big|^2\right] \\ &\leq C_1\,\epsilon^2\, n^{-\alpha} h E\left[\Big|\Big|\frac{\partial m_\theta(X_1)}{\partial\theta}|_{\theta=\theta_0}\Big|\Big|^2\right] \\ &\leq C_2\,\epsilon^2\, n^{-\alpha} h, \end{aligned}$$

$$\begin{aligned} E\left[|J_1(s,t)J_1(s_1,t)|I(\widetilde{\theta} \in \Omega_n)\right] &\leq C\,\epsilon^2\, n^{-\alpha} E\left[K\Big(\frac{X_1 - X_2}{h}\Big)K\Big(\frac{X_1 - X_3}{h}\Big)\right] \\ &\quad \times \left[\Big|\Big|\frac{\partial m_\theta(X_2)}{\partial\theta}|_{\theta=\theta_0}\Big|\Big|\Big|\Big|\frac{\partial m_\theta(X_3)}{\partial\theta}|_{\theta=\theta_0}\Big|\Big|\right] \\ &\leq C_1\,\epsilon^2\, n^{-\alpha} h^2 E\left[\Big|\Big|\frac{\partial m_\theta(X_1)}{\partial\theta}|_{\theta=\theta_0}\Big|\Big|^2\right] \\ &\leq C_2\,\epsilon^2\, n^{-\alpha}\, h^2. \end{aligned}$$

These facts imply that for any $1 \leq t \leq n$,

$$E\Big[\sum_{s=1,\neq t}^{n} J_1(s,t)I(\widetilde{\theta} \in \Omega_n)\Big]^2 \leq C\,\epsilon^2\,(n^{1-\alpha}h + n^{2-\alpha}h^2) \leq 2C\,\epsilon^2\, n^{2-\alpha}h^2,$$

since $nh \to \infty$, and hence by the independence between e_t and X_s,

$$\begin{aligned} E\left[R_{1n}^2(h)I(\widetilde{\theta} \in \Omega_n)\right] &= E\Big[\sum_{t=1}^{n} e_t \sum_{s=1,\neq t}^{n} J_1(s,t)I(\widetilde{\theta} \in \Omega_n)\Big]^2 \\ &\leq 2C\,\epsilon^2\, n^{2-\alpha}h^2 \sum_{t_1,t_2=1}^{n} E\left[e_{t_1}e_{t_2}\right] \leq C_1\,\epsilon^2\, n^{4-2\alpha}h^2, \end{aligned} \tag{A.62}$$

where we have used the fact [see (A.43)] that

$$\sum_{t_1,t_2=1}^{n} E\left[e_{t_1}e_{t_2}\right] = E\left(\sum_{t=1}^{n} e_t\right)^2 \leq Cn^{2-\alpha}.$$

In view of (A.62), Assumption 2.5(i) and the Markov's inequality, we obtain that, for $\forall\epsilon > 0$ and n sufficient large, (i) if $nh \to \infty$ and $n^{2(1-\alpha)}h \to 0$, then

$$\begin{aligned} P\left(|R_{1n}(h)| \geq \epsilon^{1/2}\sigma_{1n}\right) &\leq P\left(||\widetilde{\theta} - \theta_0|| > \epsilon\, n^{-\alpha/2}\right) \\ &\qquad + C\epsilon^{-1}(n^2h)^{-1}E\left[R_{1n}^2(h)I(\widetilde{\theta} \in \Omega_n)\right] \\ &\leq P\left(||\widetilde{\theta} - \theta_0|| > \epsilon\, n^{-\alpha/2}\right) + Cn^{2(1-\alpha)}h\epsilon \leq C_1\,\epsilon. \end{aligned} \tag{A.63}$$

(ii) If $h \to 0$ and $n^{2(1-\alpha)}h \to \infty$, then

$$\begin{aligned} P\left(|R_{1n}(h)| \geq \epsilon^{1/2}\sigma_{2n}\right) &\leq P\left(||\widetilde{\theta}-\theta_0|| > \epsilon\, n^{-\alpha/2}\right) \\ &\quad + C\epsilon^{-1}(n^{4-2\alpha}h^2)^{-1}E\left[R_{1n}^2(h)I(\widetilde{\theta}\in\Omega_n)\right] \\ &\leq P\left(||\widetilde{\theta}-\theta_0|| > \epsilon\, n^{-\alpha/2}\right) + C\,\epsilon \leq C_1\,\epsilon. \end{aligned} \tag{A.64}$$

(A.63) and (A.64) yield that $R_{1n}(h) = o(\sigma_{jn})$, $1 \leq j \leq 2$, under the corresponding conditions of Theorems 2.1. Similarly, by noting

$$\begin{aligned} E\left[|J_2(1,2)|I(\widetilde{\theta}\in\Omega_n)\right] &\leq C\,\epsilon\, n^{-\alpha} E\Big[K\Big(\frac{X_1-X_2}{h}\Big)\Big|\Big|\frac{\partial m_\theta(X_1)}{\partial\theta}|_{\theta=\theta_0}\Big|\Big| \\ &\qquad\qquad \times\Big|\Big|\frac{\partial m_\theta(X_2)}{\partial\theta}|_{\theta=\theta_0}\Big|\Big|\Big] \\ &\leq C_1\,\epsilon\, n^{-\alpha}h\ E\left[\Big|\Big|\frac{\partial m_\theta(X_1)}{\partial\theta}|_{\theta=\theta_0}\Big|\Big|^2\right] \leq C_2\,\epsilon\, n^{-\alpha}\,h, \end{aligned}$$

we obtain that for $1 \leq j \leq 2$ and n sufficiently large,

$$\begin{aligned} P(|R_{2n}(h)| \geq \epsilon^{1/2}\sigma_{jn}) &\leq P\left(||\widetilde{\theta}-\theta_0|| > \epsilon\, n^{-\alpha/2}\right) \\ &\quad + C\epsilon^{1/2}(\sigma_{jn})^{-1}n^2\, E\left[|J_2(1,2)|I(\widetilde{\theta}\in\Omega_n)\right] \leq C\,\epsilon^{1/2}, \end{aligned} \tag{A.65}$$

which implies that $R_{2n}(h) = o(\sigma_{jn})$ hold for $1 \leq j \leq 2$. We now prove (A.60) and hence also complete the proof of Theorem 2.1.

Proof of Theorem 2.2. As in (2.7), under H_0, we may write

$$\begin{aligned} &\sum_{t=1}^{n}\sum_{s=1,\neq t}^{n} b_n(s,t)\left(\widehat{e}_s\widehat{e}_t - \widehat{\gamma}(s-t)\right) \\ &\quad = M_n^*(h) + 2R_{1n}^*(h) + R_{2n}^*(h) + R_{3n}(h), \end{aligned} \tag{A.66}$$

where $M_n^*(h) = \sum_{t=1}^{n}\sum_{s=1,\neq t}^{n} b_n(s,t)\left[e_s\, e_t - \gamma(s-t)\right]$,

$$\begin{aligned} R_{1n}^*(h) &= \sum_{t=1}^{n}\sum_{s=1,\neq t}^{n} b_n(s,t)\, e_s\left(m_{\theta_0}(\frac{t}{n}) - m_{\widetilde{\theta}}(\frac{t}{n})\right), \\ R_{2n}^*(h) &= \sum_{t=1}^{n}\sum_{s=1,\neq t}^{n} b_n(s,t)\left(m_{\theta_0}(\frac{s}{n}) - m_{\widetilde{\theta}}(\frac{s}{n})\right)\left(m_{\theta_0}(\frac{t}{n}) - m_{\widetilde{\theta}}(\frac{t}{n})\right), \end{aligned}$$

and by the symmetries of $K(x)$, $\gamma(k)$ and $\hat{\gamma}(k)$,

$$\begin{aligned} R_{3n}(h) &= \sum_{t=1}^{n}\sum_{s=1,\neq t}^{n} b_n(s,t)\left[\widehat{\gamma}(s-t) - \gamma(s-t)\right] \\ &= 2\sum_{s=1}^{n-1}(n-s)K\big(\frac{s}{nh}\big)\left[\widehat{\gamma}(s) - \gamma(s)\right]. \end{aligned} \tag{A.67}$$

It follows easily from (A.66) and Theorem 2.4 that Theorem 2.2 will follow if we prove

$$2R^*_{1n}(h) + R^*_{2n}(h) = o_P\left[\sigma_{3n}(h)\right], \tag{A.68}$$
$$R_{3n}(h) = o_P\left[\sigma_{3n}(h)\right], \tag{A.69}$$
$$\frac{\widehat{\sigma}_{3n}(h)}{\sigma_{3n}(h)} \to_P 1. \tag{A.70}$$

We first prove (A.68). For $\forall\epsilon > 0$ and n sufficiently large, write $\Omega_{1n} = \{\widetilde{\theta} : ||\widetilde{\theta} - \theta_0|| \le C_0\, n^{-\alpha/2}\}$, where C_0 is chosen such that $P\left(||\widetilde{\theta} - \theta_0|| > C_0\, n^{-\alpha/2}\right) < \epsilon$. Since $||\widetilde{\theta} - \theta_0|| = O_P(n^{-\alpha/2})$, C_0 exists, under H_0. Also, write

$$J_3(s,t) = K\left(\frac{s-t}{nh}\right)\left(m_{\theta_0}(\frac{s}{n}) - m_{\widetilde{\theta}}(\frac{s}{n})\right). \tag{A.71}$$

It follows easily from Assumption 2.6 that, under H_0,

$$\begin{aligned} E\Big(\sum_{s=1,\neq t}^{n} J_3(s,t) I(\widetilde{\theta} \in \Omega_{1n})\Big)^2 &\le C_{\theta_0}^2\, E\left[||\widetilde{\theta} - \theta_0||^2 I(\widetilde{\theta} \in \Omega_{1n})\right]\Big[\sum_{s=1,\neq t}^{n} K(\frac{s-t}{nh})\Big]^2 \\ &\le C n^{2-\alpha} h^2 \end{aligned} \tag{A.72}$$

for all $1 \le t \le n$. As in the proof of (A.72),

$$\begin{aligned} E\left[R^{*2}_{1n}(h) I(\widetilde{\theta} \in \Omega_{1n})\right] &= E\left[\sum_{t=1}^{n} e_t \sum_{s=1,\neq t}^{n} J_3(s,t) I(\widetilde{\theta} \in \Omega_{1n})\right]^2 \\ &\le C\, n^{2-\alpha} h^2 \sum_{t_1,t_2=1}^{n} E\left[e_{t_1} e_{t_2}\right] \le C_1\, n^{4-2\alpha} h^2 = o(\sigma_{3n}(h)^2), \end{aligned} \tag{A.73}$$

since $h^2 = o(h^{3-2\alpha})$ for $1/2 < \alpha < 1$. Therefore, for $\forall\epsilon > 0$ and n sufficiently large,

$$\begin{aligned} P\left(|R^*_{1n}(h)| \ge \epsilon\sigma_{3n}(h)\right) &\le P\left(||\widetilde{\theta} - \theta_0|| > C_0\, n^{-\alpha/2}\right) \\ &\quad + \epsilon^{-2}\sigma_{3n}(h)^{-2} E\left[R^{*2}_{1n}(h) I(\widetilde{\theta} \in \Omega_{1n})\right] \le C\,\epsilon. \end{aligned} \tag{A.74}$$

Similarly, we have

$$\begin{aligned} P\left(|R^*_{2n}(h)| \ge \epsilon\sigma_{3n}(h)\right) &\le P\left(||\widetilde{\theta} - \theta_0|| > C_0\, n^{-\alpha/2}\right) \\ &\quad + \epsilon^{-1}\,\sigma_{3n}(h)^{-1}\, E\left[|R^*_{2n}(h)| I(\widetilde{\theta} \in \Omega_{1n})\right] \\ &\le \epsilon + C\, n^{-2}\, h^{\alpha-3/2} \sum_{t=1}^{n} \sum_{s=1,\neq t}^{n} K\left(\frac{s-t}{nh}\right) \\ &\le \epsilon + C\, h^{\alpha-1/2} \le 2\epsilon. \end{aligned}$$

This, together with (A.74), implies that $2R_{1n}^*(h)+R_{2n}^*(h)=o_P\big[\sigma_{3n}(h)\big]$, hence we complete the proof of (A.68).

We next prove (A.69). Write $\pi_n=(nh)^{1/3}$. Recalling (A.67), it suffices to show that

$$R_{3n}^{(1)}(h) \;:=\; \sum_{s=1}^{\pi_n}(n-s)K\big(\frac{s}{nh}\big)\left[\widehat{\gamma}(s)-\gamma(s)\right]=o_P\big[\sigma_{3n}(h)\big], \tag{A.75}$$

$$R_{3n}^{(2)}(h) \;:=\; \sum_{s=\pi_n+1}^{n-1}(n-s)K\big(\frac{s}{nh}\big)\left[\widehat{\gamma}(s)-\gamma(s)\right]=o_P\big[\sigma_{3n}(h)\big]. \tag{A.76}$$

To prove (A.75), by recalling that $\widehat{\gamma}(k)=\frac{1}{n}\sum_{i=1}^{n-|k|}\widehat{e}_i\widehat{e}_{i+|k|}$ for $|k|\le\pi_n$, we may write

$$R_{3n}^{(1)}(h) \;=\; R_{3n}^{(1)*}(h)+R_{3n}^{(2)*}(h)+R_{3n}^{(3)*}(h), \tag{A.77}$$

where

$$\begin{aligned}
R_{3n}^{(1)*}(h) &= \frac{2}{n}\sum_{s=1}^{\pi_n}(n-s)K\big(\frac{s}{nh}\big)\sum_{i=1}^{n-s}(e_ie_{i+s}-Ee_ie_{i+s}),\\
R_{3n}^{(2)*}(h) &= \frac{2}{n}\sum_{s=1}^{\pi_n}(n-s)K\big(\frac{s}{nh}\big)\sum_{i=1}^{n-s}\Big[e_i\Big(m_{\theta_0}(\frac{i+s}{n})-m_{\widetilde{\theta}}(\frac{i+s}{n})\Big)\\
&\qquad\qquad + e_{i+s}\Big(m_{\theta_0}(\frac{i}{n})-m_{\widetilde{\theta}}(\frac{i}{n})\Big)\Big],\\
R_{3n}^{(3)*}(h) &= \frac{2}{n}\sum_{s=1}^{\pi_n}(n-s)K\big(\frac{s}{nh}\big)\sum_{i=1}^{n-s}\Big(m_{\theta_0}(\frac{i}{n})-m_{\widetilde{\theta}}(\frac{i}{n})\Big)\\
&\qquad\qquad \times\Big(m_{\theta_0}(\frac{i+s}{n})-m_{\widetilde{\theta}}(\frac{i+s}{n})\Big).
\end{aligned}$$

Using the same arguments as in the proof of (A.68), we have $R_{3n}^{(2)*}(h)+R_{3n}^{(3)*}(h)=o_P\big[\sigma_{3n}(h)\big]$. As for $R_{3n}^{(1)*}(h)$, it follows easily from (A.4) and $\pi_n=(nh)^{1/3}$ that

$$\begin{aligned}
E\big[R_{3n}^{(1)*}(h)\big]^2 &\le 4\sum_{s,t=1}^{\pi_n}K\big(\frac{s}{nh}\big)K\big(\frac{t}{nh}\big)\sum_{i,j=1}^{n}\big|E(e_ie_{i+s}-Ee_ie_{i+s})(e_je_{j+t}-Ee_je_{j+t})\big|\\
&\le C\,n\sum_{s,t=1}^{\pi_n}K\big(\frac{s}{nh}\big)K\big(\frac{t}{nh}\big)\sum_{i=1}^{n}\left[\gamma(i)\gamma(i+s-t)+\gamma(i+s)\gamma(i-t)\right]\\
&\le C\,n\sum_{s,t=1}^{\pi_n}K\big(\frac{s}{nh}\big)K\big(\frac{t}{nh}\big)\\
&\le C\,n\,\pi_n^2 \;=\; o\big[\sigma_{3n}^2(h)\big],
\end{aligned}$$

where we have used the result that, whenever $1/2<\alpha<1$, $\gamma(0)<\infty$ and $\gamma(k)\sim\eta k^{-\alpha}$, $\sum_{i=1}^{n}\gamma(i+x)\gamma(i-y)\le C<\infty$, for all $x,y\ge 0$. These facts, together with (A.77), yield (A.75).

In order to prove (A.76), without loss of generality, we assume that $\gamma(k) = \eta\,|k|^{-\alpha}$ for all $\pi_n \le |k| \le n-1$. Otherwise the proof follows from some routine modifications. We first notice that, whenever $\tilde\lambda \in \Lambda_n := \{\tilde\lambda : ||\tilde\lambda - \lambda|| \le C_0\, w_n^{-1}\}$, where C_0 is chosen such that $P\big(||\tilde\lambda - \lambda|| > C_0\, w_n^{-1}\big) < \epsilon$,

$$\max_{1\le k\le n} \Big|\frac{\hat\gamma(k)}{\gamma(k)} - 1\Big| \le \big|\frac{\widetilde\eta}{\eta} - 1\big| + \big|\frac{\widetilde\eta}{\eta}\big| \max_{1\le k\le n} \Big||k|^{\alpha-\widetilde\alpha} - 1\Big| \le C\, w_n^{-1}\, \log n. \tag{A.78}$$

Therefore it follows from $w_n\, h^{1/2}/\log n \to \infty$ that, for $\forall\epsilon > 0$ and n sufficiently large ,

$$\begin{aligned} P\big[R_{3n}^{(2)}(h) \ge \epsilon\sigma_{3n}(h)\big] \ &\le\ P(R_{3n}(h) \ge \epsilon\sigma_{3n}(h), \tilde\lambda \in \Lambda_n) + P(||\tilde\lambda - \lambda|| > C_0\, w_n^{-1}) \\ &\le\ 2n(\epsilon\sigma_{3n}(h))^{-1} \sum_{s=1}^{n} K(\frac{s}{nh})\gamma(s) E \max_{1\le k\le n} \Big|\frac{\hat\gamma(k)}{\gamma(k)} - 1\Big| I_{(\tilde\lambda\in\Lambda_n)} + \epsilon \\ &\le\ C\, w_n^{-1}\, h^{-1/2}\, \log n + \epsilon\ \le C\,\epsilon, \end{aligned}$$

which yields (A.76). The proof of (A.69) is now complete.

We finally prove (A.70). Recall (A.57). It suffices to show that $A^*_{\underset{\sim}{\alpha}} - A_\alpha = o_P(1)$. In fact, by recalling $A_\alpha < \infty$ and noting for $\tilde\lambda \in \Lambda_n$,

$$\sup_{1/n\le x,y\le n} |(xy)^{\alpha-\hat\alpha} - 1| \le |e^{2|\alpha-\hat\alpha|\,\log n} - 1| + |e^{-2|\alpha-\hat\alpha|\,\log n} - 1| \le C\, w_n^{-1}\, \log n,$$

it is readily seen that, for $\forall\epsilon > 0$, when $\tilde\lambda \in \Lambda_n$ and n sufficiently large,

$$\begin{aligned} |A^*_{\underset{\sim}{\alpha}} - A_\alpha| \ &\le\ \ \epsilon + \int_{1/n}^{n}\int_{1/n}^{n}\int_{1/n}^{n} \big|(xy)^{-\widetilde\alpha} - (xy)^{-\alpha}\big|\big[K(z)K(x+y-z) \\ &\qquad\qquad + K(z-x)K(z-y)\big]dxdydz \\ &\le\ \ \epsilon + C \sup_{1/n\le x,y\le n} |(xy)^{\alpha-\hat\alpha} - 1| \le 2\epsilon. \end{aligned}$$

This implies that, for any $\epsilon > 0$ and n sufficiently large,

$$\begin{aligned} P(|A^*_{\underset{\sim}{\alpha}} - A_\alpha| \ge \epsilon^{1/2}) \ &\le\ \ P(|A^*_{\underset{\sim}{\alpha}} - A_\alpha| \ge \epsilon^{1/2}, \tilde\lambda \in \Lambda_n) + P(||\tilde\lambda - \lambda|| > C_0\, n^{-1/2}) \\ &\le\ \ \epsilon^{-1/2} E\big[|A^*_{\underset{\sim}{\alpha}} - A_\alpha| I_{(\tilde\lambda\in\Lambda_n)}\big] + \epsilon\ \le\ \epsilon + 2\epsilon^{1/2}, \end{aligned}$$

which yields $A^*_{\underset{\sim}{\alpha}} - A_\alpha = o_P(1)$. This proves (A.70) and hence completes the proof of Theorem 2.2.

Appendix B. This appendix provides technical details for the asymptotic theory in Section 4. Appendix B.1 proves Theorem 4.1. The proof of Theorem 4.2 is given in B.2. Since the proofs are similar for both the random and fixed designs, we provide only an outline of the proof in the fixed design situation. In this case, $T_n(h) = \widehat L_{3n}(h)$.

B.1. *Proof of Theorem 4.1.* We first prove (4.1). In view of Theorem 2.2, it suffices to show that

$$\sup_{x\in R}\left|P^*(\widehat{T}_n^*(h)\le x)-\Phi(x)\right| = o_P(1). \tag{B.1}$$

As in (A.66), we may rewrite $\widehat{T}_n^*(h)$ as

$$\widehat{T}_n^*(h) \;=\; \frac{1}{\widehat{\sigma}_{3n}(h)}\left[M_n^{**}(h)+2R_{1n}^{**}(h)+R_{2n}^{**}(h)\right], \tag{B.2}$$

where $M_n^{**}(h)=\sum_{t=1}^n\sum_{s=1,\neq t}^n b_n(s,t)\left[e_s^*\,e_t^*-\widehat{\gamma}(s-t)\right]$,

$$R_{1n}^*(h) \;=\; \sum_{t=1}^n\sum_{s=1,\neq t}^n b_n(s,t)\,e_s^*\left(m_{\widetilde{\theta}^*}(\frac{t}{n})-m_{\widetilde{\theta}}(\frac{t}{n})\right),$$

$$R_{2n}^*(h) \;=\; \sum_{t=1}^n\sum_{s=1,\neq t}^n b_n(s,t)\,\left(m_{\widetilde{\theta}^*}(\frac{s}{n})-m_{\widetilde{\theta}}(\frac{s}{n})\right)\left(m_{\widetilde{\theta}^*}(\frac{t}{n})-m_{\widetilde{\theta}}(\frac{t}{n})\right),$$

Since $\{e_i^*\}$ is drawn from a stationary Gaussian process with covariance structure $\gamma_{\widetilde{\lambda}}(k)\sim\widetilde{\eta}\,|k|^{-\widetilde{\alpha}}$, similarly to the proof of Theorem 2.4, we have

$$\sup_{x\in R}\left|P^*\Big(\frac{M_n^{**}(h)}{\sigma_{3n}(h)}\le x\Big)-\Phi(x)\right| = o_P(1). \tag{B.3}$$

On the other hand, similarly to the proof of (A.68), for any $\epsilon>0$,

$$P^*\Big(|2R_{1n}^{**}(h)+R_{2n}^{**}(h)|\ge\epsilon\sigma_{3n}(h)\Big) = o_P(1). \tag{B.4}$$

The facts (B.2)-(B.4), together with (A.70), yield (B.1).

We next prove (4.2). In view of Theorem 2.2, it suffices to show that

$$l_r^*-l_r = o_P(1). \tag{B.5}$$

In fact, by recalling the definitions of l_r^* and l_r, it is readily seen from (B.1) and Theorem 2.2 that $\Phi(l_r^*)-\Phi(l_r)=o_P(1)$, which implies (B.5) since $\Phi(x)$ is a bounded continuous function.

We finally prove (4.3). In view of (B.5), it suffices to show that under H_1,

$$P(\widehat{L}_{3n}(h)\ge l_r) = 1, \tag{B.6}$$

with l_r satisfying $\Phi(l_r)=1-r+o(1)$ with $0<r<1$. In order to prove (B.6), as in (A.66), under H_1, we may rewrite $\widehat{L}_{3n}(h)$ as

$$\widehat{L}_{3n}(h) \;=\; \frac{1}{\widehat{\sigma}_{3n}(h)}\left[S_n(h)+2Q_{1n}(h)+Q_{2n}(h)+R_{3n}(h)\right], \tag{B.7}$$

where

$$
\begin{aligned}
S_n(h) &= \sum_{t=1}^{n} \sum_{s=1,\neq t}^{n} b_n(s,t) \left[\zeta_s \, \zeta_t - \gamma(s-t)\right] \\
Q_{1n}(h) &= c_n \sum_{t=1}^{n} \sum_{s=1,\neq t}^{n} b_n(s,t) \, \zeta_s \, \Delta(\frac{t}{n}), \\
Q_{2n}(h) &= c_n^2 \sum_{t=1}^{n} \sum_{s=1,\neq t}^{n} b_n(s,t) \, \Delta(\frac{s}{n}) \, \Delta(\frac{t}{n}),
\end{aligned}
$$

in which $\zeta_s = \epsilon_t + \left[m_{\theta_1}(\frac{t}{n}) - m_{\widetilde{\theta}}(\frac{t}{n})\right]$, and $R_{3n}(h)$ is defined as in (A.66). Simple calculations show that

$$
\begin{aligned}
Q_{2n}(h) &\sim c_n^2 \int_1^n \int_1^n K(\frac{x-y}{nh}) \Delta(\frac{x}{n}) \Delta(\frac{y}{n}) dx dy \\
&\sim C\,[1+o(1)] c_n^2 \,(nh)^2 = C \, d_n \, \sigma_{3n}(h),
\end{aligned}
\tag{B.8}
$$

where $d_n = c_n^2 n^\alpha \, h^{\alpha - 1/2} \to \infty$. By the similar arguments as in the proof of Theorem 2.2, it follows easily that under Assumption 4.1,

$$
S_n(h)/\sigma_{3n}(h) \to_D N(0,1), \tag{B.9}
$$

and $Q_{1n}(h) = O_P[c_n n^{-\alpha/2} (nh)^2] = o_P[Q_{2n}(h)]$. These facts, together with (A.69) and (A.70), yield (B.6) since l_r is finite for $0 < r < 1$. The proof of Theorem 4.1 is now complete.

B.2. *Proof of Theorem 4.2*: The first step is to show that as $N \to \infty$

$$
\sqrt{N} \, \left(\widetilde{\lambda} - \lambda\right) \to_D N\left(0, \Sigma^{-1}\right), \tag{B.10}
$$

where Σ is a positive definite covariance matrix as in (4.8). The proof of (B.10) is standard in this kind of problem (see Theorem 2 of Robinson 1995 or Theorem 2.1(ii) of Gao 2004). It follows from (B.10) and $N \sim C n^{4/5}$ that

$$
\frac{n^{2/5}}{\log(n)} \, \left(\widetilde{\lambda} - \lambda\right) \sim \frac{C}{\log n} \, \sqrt{N} \, \left(\widetilde{\lambda} - \lambda\right) \to_P 0, \tag{B.11}
$$

which implies (4.9). This also completes an outline of the proof.

REFERENCES

Anh, V., Heyde, C., (ed.) 1999. Special Issue on Long-Range Dependence. *Journal of Statistical Planning & Inference* **80** 1.

Anh, V., Wolff, R., Gao, J., Tieng, Q., 1999. Local kernel regression with long–range dependent errors. *Australian & New Zealand Journal of Statistics* **41** 463–480.

Avram, F., 1988. On bilinear forms in Gaussian random variables and Toeplitz matrices.

Probability Theory and Related Fields **79** 37–45.
Baillie, R., King, M. L., (ed.) 1996. Special Issue of Journal of Econometrics. *Annals of Econometrics* **73** 1.
Beran, J., 1996. *Statistics for Long–Memory Processes*. Chapman & Hall, London.
Beran, J., Ghosh, S., 1998. Root–n–consistent estimation in partial linear models with long–memory errors. *Scandinavian Journal of Statistics* **25** 345–357.
Bühlmann, P., 2002. Bootstraps for time series. *Statistical Science* **17** 52–72.
Cai, Z., Hong, Y., 2003. Nonparametric methods in continuous-time finance: A selective review, In *Recent Advances and Trends in Nonparametric Statistics* (M.G. Akritas and D.M. Politis, eds.), 283–302.
Csörgö, S., Mielniczuk, J., 1999. Random-design regression under long–range dependence. *Bernoulli* **5** 209–224.
Chen, S. X., Härdle, W., Li, M., 2003. An empirical likelihood Goodness–of–fit test for time series. *Journal of the Royal Statistical Society Series B* **65** 663–678.
Fan, J., 2005. A selective overview of nonparametric methods in financial econometrics. *Statistical Science* **20**, 317-357.
Fan, J., Gijbels, I., 1996. *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London.
Fan, J., Härdle, W., Mammen, E., 1998. Direct estimation of low dimensional components in additive models. *Annals of Statistics* **26** 943–971.
Fan, J., Yao, Q., 2003. *Nonlinear Time Series*: nonparametric and parametric methods. Springer, New York.
Fan, J., Zhang, C. M., Zhang, J., 2001. Generalized likelihood ratio statistics and Wilks phenomenon. *Annals of Statistics* **29** 153–193.
Fan, Y., Linton, O., 2003. Some higher–theory for a consistent nonparametric model specification test. *Journal of Statistical Planning and Inference* **109** 125–154.
Fox, R., Taqqu, M. S., 1985. Noncentral limit theorems for quadratic forms in random variables having long–range dependence. *Annals of Probability* **13** 428–446.
Fox, R., Taqqu, M. S., 1987. Central limit theorems for quadratic forms in random variables having long–range dependence. *Probability Theory and Related Fields* **74** 213–240.
Franke, J., Kreiss, J. P., Mammen, E., 2002. Bootstrap of kernel smoothing in nonlinear time series. *Bernoulli* **8** 1–38.
Gao, J., 2004. Modelling long–range dependent Gaussian processes with application in continuous–time financial models. *Journal of Applied Probability* **41** 467–482.
Gao, J., 2006. *Nonlinear Time Series*: semiparametric methods. Chapman & Hall/CRC.
Gao, J., Anh, V., 1999. Semiparametric regression under long–range dependent errors. *Journal of Statistical Planning & Inference* **80** 37–57.
Gao, J., Hawthorne, K., 2006. Semiparametric estimation and testing of temperature series. *Econometrics Journal* **9** 332–355.
Gao, J., King, M. L., 2005. Model specification testing in nonparametric and semiparametric regression. Working paper available at www.maths.uwa.edu.au/˜jiti/jems.pdf.
Gao, J., Lu, Z., Tjøstheim, D., 2006. Estimation in semiparametric spatial regression. *Annals of Statistics* **34** 1395–1435.
Giraitis, L., Surgailis, D., 1990. A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate. *Probability Theory and Related Fields* **86** 87–104.
Giraitis, L., Taqqu, M. S., 1997. Limit theorems for bivariate Appell polynomials. I. Central limit theorems. *Probability Theory and Related Fields* **107** 359–381.
Götze, F., Tikhomirov, A., 2002. Asymptotic Distribution of Quadratic forms and applications. *Journal of Theoretical Probability* **15** 423-438.

Hall, P., Heyde, C., 1980. *Martingale Limit Theory and Its Applications*. Academic Press, New York.
Härdle, W., Liang, H., Gao, J., 2000. *Partially Linear Models*. Springer Series: Contributions to Statistics. Physica–Verlag, New York.
Härdle, W., Lütkepohl, H., Chen, R., 1997. A review of nonparametric time series analysis. *International Statistical Review* **65** 49–72.
Härdle, W., Mammen, E., 1993. Comparing nonparametric versus parametric regression fits. *Annals of Statistics* **21** 1926–1947.
Hart, J., 1997. *Nonparametric Smoothing and Lack-of-Fit Tests*. Springer, New York.
Hidalgo, F. J., 1997. Nonparametric estimation with strongly dependent multivariate time series. *Journal of Time Series Analysis* **8** 95–122.
Hjellvik, V., Yao, Q., Tjøstheim, D., 1998. Linearity testing using local polynomial approximation. *Journal of Statistical Planning and Inference* **68** 295–321.
Ho, H. C., Hising, T., 1996. On the asymptotic expansion of the empirical processes of long–memory moving averages. *Annals of Statistics* **24** 992–1024.
Ho, H. C., Hising, T., 1997. Limit theorems for functional of moving averages. *Annals of Probability* **25** 1636–1669.
Ho, H. C., Hising, T., 2003. A decomposition for generalized U-statistics of long-memory linear processes. *Theory and Applications of Long–Range Dependence* 143–155, Birkhuser Boston, Boston, MA.
Horowitz, J., 2003. Bootstrap methods for Markov processes. *Econometrica* **71** 1049–1082.
Horowitz, J., Spokoiny, V., 2001. An adaptive rate–optimal test of a parametric mean–regression model against a nonparametric alternative. *Econometrica* **69** 599–631.
Hsing, T., Wu, W. B., 2004. On weighted U–statistics for stationary processes. *Annals of Probability* **32** 1600–1631.
Kreiss, J. P., Neumann, M. H., Yao, Q., 2002. Bootstrap tests for simple structures in nonparametric time series regression. Private communication.
Li, Q., 1999. Consistent model specification tests for time series econometric models. *Journal of Econometrics* **92** 101–147.
Li, Q., Wang, S., 1998. A simple consistent bootstrap tests for a parametric regression functional form. *Journal of Econometrics* **87** 145–165.
Mielniczuk, J., Wu, W. B., 2004. On random–design model with dependent errors. *Statistica Sinica* **14** 1105–1126.
Mikosch, T., Starica, C., 2004. Nonstationarities in financial time series, the long–range dependence and the IGARCH effects. *Review of Economics and Statistics* **86** 378–390.
Robinson, P. M., 1994a. Time series with strong dependence. *In Advances In Econometrics*. Six World Congress (C. A. Sims, ed.) **1** 47–96. Cambridge University Press.
Robinson, P. M., 1994b. Rates of convergence and optimal spectral bandwidth for long range dependence. *Probability Theory and Related Fields* **99** 443–473.
Robinson, P. M., 1995. Gaussian semiparametric estimation of long–range dependence. *Annals of Statistics* **23** 1630–1661.
Robinson, P. M., 1997. Large–sample inference for nonparametric regression with dependent errors. *Annals of Statistics* **25** 2054–2083.
Robinson, P. M., (ed.) 2003. Time Series with Long Memory. *Advanced Texts in Econometrics*. Oxford University Press, Oxford.
Robinson, P. M., Hidalgo, F. J., 1997. Time series regression with long–range dependence. *Annals of Statistics* **25** 77–104.
Sperlich, S., Tjøstheim, D., Yang, L., 2002. Nonparametric estimation and testing of interaction in additive models. *Econometric Theory* **18** 197–251.
Taqqu, M. S., 1975. Weak convergence to fractional Brownian motion and to the Rosen-

blatt process. *Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete* **31** 287–302.

Tjøstheim, D., 1994. Nonlinear time series: a selective review. *Scandinavian Journal of Statistics* **21** 97–130.

Zheng, J. X., 1996. A consistent test of functional form via nonparametric estimation techniques. *Journal of Econometrics* **75** 263–289.

Jiti Gao
School of Mathematics and Statistics
The University of Western Australia
Crawley WA 6009, Australia
e-mail: jiti.gao@uwa.edu.au

Qiying Wang
School of Mathematics and Statistics
The University of Sydney
NSW 2006, Australia
e-mail: qiying@maths.usyd.edu.au