ELSEVIER Stochastic Processes and their Applications 63 (1996) 153–174

stochastic processes and their applications

On central and non-central limit theorems in density estimation for sequences of long-range dependence

Hwai-Chung Ho

Institute of Statistical Science, Academia Sinica, Taipei, Taiwan 115, ROC

Received February 1995; revised October 1995

Abstract

This paper studies the asymptotic properties of the kernel probability density estimate of stationary sequences which are observed through some non-linear instantaneous filter applied to long-range dependent Gaussian sequences. It is shown that the limiting distribution of the kernel estimator can be, in quite contrast to the case of short-range dependence, Gaussian or non-Gaussian depending on the choice of the bandwidth sequences. In particular, if the bandwidth $h(N)$ for sample of size N is selected to converge to zero fast enough, the usual $\sqrt{Nh(N)}$ rate asymptotic normality still holds.

Keywords: Long-range dependence; Central limit theorem; Non-central limit theorem; Kernel density estimator; Instantaneous filter

AMS classification: Primary 62G07, 62G20; secondary 60G10, 60G18

1. Introduction

For years, the issue of smoothed nonparametric probability density estimation has been discussed quite extensively (see, e.g., Silverman, 1986 and reference therein). Most of the discussion has been formulated under the settings where the data are collected from iid sequences or, more generally, stationary sequences of short-range dependence such as ARMA models, Markov processes, and stationary sequences satisfying certain mixing conditions (see, e.g., Robinson, 1983; Hart, 1984; Roussas, 1969; Rosenblatt, 1970; Chanda, 1983; Castellana and Leadbetter, 1986 and for a review, see Györf et al., 1989 and Rosenblatt, 1991). Considerable evidence has indicated, however, that correlations of many empirical time series are seen to decay at rates much slower than that of short-range dependence (for a review, see Beran, 1992; Robinson, 1990). Stochastic processes showing this type of dependence feature constitute the basic model of this paper. Let $\{X_n\}$ be a stationary Gaussian sequence with zero mean, unit variance and covariance function

$$r(n) = EX_0X_n = |n|^{-\alpha}L(n),\ 0 < \alpha < 1, \tag{1.1}$$

where $L(x)$ is a slowly varying function. Such a sequence is said to exhibit long-range dependence in the sense that the sum of covariances $r(n)$'s diverge to positive infinity.

0304-4149/96/$15.00
PII S0304-4149(96)00072-5

Suppose that observations $\{Y_n\}$ are made by an instantaneous filter $G(x)$ applied to $\{X_n\}$, i.e.,

$$Y_n = G(X_n). \tag{1.2}$$

Let $f(y)$ be the common density of $\{Y_n\}$, which is to be estimated. In this article, we investigate the asymptotic properties of the kernel estimator of $f(y)$,

$$\hat{f}_N(y) = \frac{1}{Nh(N)} \sum_{n=1}^{N} K\left(\frac{y - G(X_n)}{h(N)}\right),$$

where $\{h(N)\}$ is the bandwidth sequence of positive numbers converging to zero, and the kernel function $K(x)$ satisfies $K(x) \geqslant 0$ and $\int K(x)\mathrm{d}x = 1$.

As a result of the persistent dependence displayed in (1.1), it is possible that the Nth partial sums of $\{Y_n\}$ will, after being normalized by a factor greater than root N rate, converge in distribution to a random variable which may not be Gaussian (Rosenblatt, 1961; Dobrushin and Major, 1979; Taqqu, 1979). This non-Gaussian domain of attraction phenomenon or non-central limit theorem is naturally expected to take place as one examines the limiting distributions for the kernel estimator $\hat{f}_N(y)$. Theorems 1–3 in Section 2 confirm this conjecture by showing that under long-range dependence both Gaussian and non-Gaussian limits are possible for the centered and normalized kernel estimate $Z(B_N, h(N), y) \equiv B_N(\hat{f}_N(y) - E\hat{f}_N(y))$. The theorems also contain some interesting properties which contrast noticeably to the short-range-dependent cases. First, the central limit theorem for $Z(B_N, h(N), y)$ may hold with various choices for the norming factor B_N; B_N can be the usual $\sqrt{Nh(N)}$ (Theorems 1 and 3) or some positive regularly varying function of N not depending on $h(N)$ (Theorem 2). Moreover, the limiting laws may shift from Gaussian to non-Gaussian even as the values of y differ. A more striking result given by Theorem 3 says that even when $Z(B_N, h(N), y)$ converges to a non-Gaussian limit with some norming factor other than $\sqrt{Nh(N)}$, one can still choose a new set of "narrower" bandwidth sequences $\{h'(N)\}$ satisfying

$$\lim_{N\to\infty} h'(N) \sum_{n=1}^{N} |r(n)| = 0 \tag{1.3}$$

to assure that the limit for $Z(\sqrt{Nh'(N)}, h'(N), y)$ as $N \to \infty$ is Gaussian. Notice that as $h'(N) \to 0$, condition (1.3) will be satisfied by short-range dependent sequences with absolutely summable covariances.

Hart and Hall (1990) studied another class of long-range dependent sequences that are modeled as infinite moving averages: $Y_n = \mu + \sum_j a_j \varepsilon_{n-j}$, in which $\sum |a_j| = \infty, \sum a_j^2 < \infty$, and $\{\varepsilon_j\}$ is iid and square-integrable. They computed the mean integrated square error of $\hat{f}_N$ and discovered an interesting "ceiling rate" phenomenon which is in spirit very close to part (A) of Proposition 2 below.

Under the same model as specified by (1.1) and (1.2), Rosenblatt (1991) focuses in his Lecture 9 on the case where the transformation $G(x)$ is continuously differentiable and monotone, and derives that the random vector $(Z(N^{\alpha/2}, h(N), y_i), i = 1, 2, \ldots, m)$ converges in distribution to a Gaussian vector with covariance matrix $[a(y_i)a(y_j)], i, j = 1, \ldots, m$, for some function $a(y)$, provided that $h(N)N^{1-\alpha} \to \infty$ as $N \to \infty$. Assuming

$G(x) = x$, Robinson (1987) gave results of marginal limit distribution as well as MSE under long range dependence. With more general $G(x)$, including two-dimensional non-instantaneous filters, Cheng and Robinson (1991) extend Robinson's (1987) results and obtain a broader class of limiting distributions. Studies mentioned above did not address, however, the question as to what if values of the abscissa variable y of $f(y)$ are in $\{G(x) \mid G'(x) = 0,\ x \text{ in } \Re\}$. The present paper fills this gap by assuming that the set $\{x \text{ in } \Re \mid G'(x) = 0\}$ is finite. Our main results are summarized in Theorems 1–3, and stated in Section 2. An example to express these theorems is also given in Section 2. Proofs are all given in Section 3.

2. Main theorems

We start with a list of conditions which we throughout assume most frequently.

(C1) The probability density function of any finite-dimensional joint distribution of $\{Y_n, n \in Z\}$ is continuously differentiable. Moreover, for each fixed $\boldsymbol{y} = (y_1, \ldots, y_m) \in \Re^m$, the joint density $f_{i_1,\ldots,i_m}(\boldsymbol{y})$ is uniformly bounded over all integral m-tuples $(i_1, \ldots, i_m)$ with distinct coordinates.

(C2) The instantaneous filter $G(x)$ is continuously differentiable, and the set $\{x \mid G'(x) = 0\}$ is finite.

(C3) The nonnegative kernel function $K(x)$ is bounded, has compact support and satisfies $\int K(x)\,\mathrm{d}x = 1$.

(C4) The positive bandwidth sequence $\{h(N)\}$ satisfies $h(N) \to 0$ and $Nh(N) \to \infty$ as $N \to \infty$.

(C5) The covariance function $r(n)$ satisfies $r(n) \to 0$ as $n \to \infty$.

Let X be an N(0,1) random variable. Set

$$\tilde{K}\left(\frac{y - G(X)}{h(N)}\right) = K\left(\frac{y - G(X)}{h(N)}\right) - EK\left(\frac{y - G(X)}{h(N)}\right)$$

and expand $\tilde{K}((y - G(X))/h(N))$ in terms of the Hermite polynomials $H_j(X)$'s as

$$\tilde{K}\left(\frac{y - G(X)}{h(N)}\right) = \sum_{j=1}^{\infty} \frac{a_{jh(N)}(y)}{j!} H_j(X). \tag{2.1}$$

The Hermite polynomials are defined by

$$H_j(x) = (-1)^j e^{x^2/2} \frac{\mathrm{d}^j}{\mathrm{d}x^j} e^{-x^2/2}.$$

The first few are $H_0(x) = 1$, $H_1(x) = x$ and $H_2(x) = x^2 - 1$. The coefficient $a_{jh(N)}(y)$'s in (2.1) are given by

$$a_{jh(N)}(y) = \int \tilde{K}\left(\frac{y - G(x)}{h(N)}\right) H_j(x)\phi(x)\,\mathrm{d}x, \tag{2.2}$$

where ϕ is the standard normal density. Write

$$Z_{Nh(N)}(y) \equiv \hat{f}_N(y) - E\hat{f}_N(y)$$

$$= \frac{1}{Nh(N)} \sum_{n=1}^{N} \tilde{K}\left(\frac{y - G(X_n)}{h(N)}\right)$$
$$= \sum_{j=1}^{\infty} \frac{a_{jh(N)}(y)}{j!h(N)} \left(\frac{1}{N}\sum_{n=1}^{N} H_j(X_n)\right). \tag{2.3}$$

We shall see later that the first index j_0 in (2.1) with $a_{j_0h(N)}(y) \neq 0$ plays a key role in determining the norming factor for $Z_{Nh(N)}(y)$. This property is slightly different from a direct analogy to what is noted by Dobrushin and Major (1979) and Taqqu (1979) as they deal with the partial sums of $G(X_n)$'s, because, under current situations, there is the bandwidth sequence $\{h(N)\}$ involved. Prior to giving our main theorems we need

Proposition 1. *Under* (C1)–(C3), *the limit*

$$\lim_{N\to\infty} \frac{a_{jh(N)}(y)}{h(N)} \equiv g_j(y) \tag{2.4}$$

exists for each j and y, and the function $g_j(y)$ satisfied, for fixed y

$$g_j(y) = 0 \quad \forall j \geqslant 1 \Leftrightarrow f(y) = 0. \tag{2.5}$$

Eq. (2.5) validates the following definition.

Definition. Fix $G(\cdot)$ and $K(\cdot)$. For each y with $f(y)>0$, define $k_{K,G,y} \equiv \min\{j \geqslant 1 | a_{jh(N)}(y)$ is nonzero infinitely often as $N \to \infty$ for some sequence $\{h(N)\}$ converging to zero$\}$. We shall call $k_{K,G,y}$ the Hermite rank of $(K(\cdot), G(\cdot), y)$, and occasionally abbreviate it by k if no confusion will be created.

Example 1. Suppose $G(x) = x$ and $K(x)$ is symmetric, bounded, and has compact support. It is easy to see that the Hermite rank $k_{K,G,y}$ for each y is

$$k_{K,G,y} = \begin{cases} 2 & \text{if } y = 0. \\ 1 & \text{if } y \neq 0. \end{cases}$$

Accordingly, the functions defined by (2.4) is $g_j(y) = -H_j(y)\phi(y)$, $j \geqslant 1$.

Remark 1. Let k_0 be the Hermite rank of $(K(\cdot), G(\cdot), y_0)$. Clearly, (2.5) gives $k_0 < \infty$, and by the definition of the Hermite rank k_0, there exists an $\varepsilon > 0$ such that $\forall j$, $1 \leqslant j \leqslant k_0 - 1$, $a_{jh(N)}(y_0) = 0$ if $|h(N)| < \varepsilon$. The expansion for $Z_{Nh(N)}(y_0)$ in (2.3) should be modified as

$$Z_{Nh(N)}(y_0) = \sum_{j=k_0}^{\infty} \frac{a_{jh(N)}(y_0)}{j!h(N)} \left(\frac{1}{N}\sum_{n=1}^{N} H_j(X_n)\right), \quad |h(N)| < \varepsilon. \tag{2.6}$$

Remark 2. Let $F(x)$ be the distribution function of Y_n, and let $F_N(x)$ be the empirical distribution function of the observations $Y_1, \ldots, Y_N$, i.e.,

$$F_N(x) = \frac{1}{N} \sum_{n=1}^{N} I_{\{G(X_n) \leqslant x\}}.$$

Expand $I_{\{G(X_n)\leqslant x\}} - F(x)$ as

$$I_{\{G(X_n)\leqslant x\}} - F(x) = \sum_{j=m}^{\infty} \frac{A_j(x)}{j!} H_j(X_n),$$

where $m = \inf\{j \geqslant 1 | A_j(x)$ is nonzero for at least one $x\}$ [see Dehling and Taqqu, 1989]. Then, formally, integration by parts gives

$$\begin{aligned} Z_{Nh(N)}(y) &= \frac{1}{h(N)} \int K\left(\frac{y-x}{h(N)}\right) \mathrm{d}\,[F_N(x) - F(x)] \\ &= \sum_{j=m}^{\infty} \left(\frac{\sum_{n=1}^{N} H_j(X_n)}{N}\right) \frac{1}{j!h(N)} \int A_j(x)\,\mathrm{d}K\left(\frac{y-x}{h(N)}\right). \end{aligned}$$

The integral $\int A_j(x)\,\mathrm{d}K((y-x)/h(N))$ is exactly the function we have denoted by $a_{jh(N)}(y)$ (cf. (2.6)). By arguments used in Example 3 of Dehling and Taqqu (1989), we see that for any $i \geqslant 1$ there are $G(\cdot)$'s such that the Hermite rank $k_{K,G,y}$ is greater than i.

As the first step toward finding the norming factor for $Z_{Nh(N)}(y)$, we compute in the following proposition the variance of $Z_{Nh(N)}(y)$ under various circumstances. Set

$$R(j,m) = \sum_{n=1}^{m} r(n)^j \text{ and } |R|(j,m) = \sum_{n=1}^{m} |r(n)|^j.$$

Proposition 2. *Assume* (C1)–(C5). *Given y with its Hermite rank k, then we have the following*:

(A) *As $N \to \infty$*

$$E(Z_{Nh(N)}(y))^2 = \mathrm{O}(N^{-1}|R|(k,N))\left(\frac{g_k^2(y)}{k!} + B_N + \mathrm{o}(1)\right) + (Nh(N))^{-1}\left(f(y)\int K^2(u)\,\mathrm{d}u + \mathrm{o}(1)\right), \tag{2.7}$$

where $g_k(y)$ is as defined in (2.4), *and the term B_N is bounded by one and is* $\mathrm{o}(1)$ *if* $\lim_{N\to\infty} |R|(k,N) = \infty$.

(B) *If the sequence $\{h(N)\}$ satisfies*

$$\lim_{N\to\infty} h(N)|R|(k,N) = 0, \tag{2.8}$$

then

$$\lim_{N\to\infty} E\left(\sqrt{Nh(N)}Z_{Nh(N)}(y)\right)^2 = f(y)\int K^2(u)\,\mathrm{d}u \tag{2.9}$$

(*note that under* $h(N)=\mathrm{o}(1)$, (2.8) *follows automatically if* $\lim_{N\to\infty} |R|(k,N) < \infty$).

(C) *Suppose $|R|(k,N) = L_1(N)$ is slowly varying and diverges to $+\infty$. Assume*

$$\lim_{N\to\infty} L_1^{-1}(N)R(k,N)$$

exists. Then, with $h(N)$ *satisfying* $\lim_{N\to\infty} h(N)L_1(N) = \infty$,

$$\lim_{N\to\infty} E\left(\sqrt{N}L_1^{-1/2}(N)Z_{Nh(N)}(y)\right)^2 = \frac{cg_k^2(y)}{k!}, \tag{2.10}$$

where

$$c = \lim_{N\to\infty} (NL_1(N))^{-1} \sum_{m,n=1}^{N} r^k(m-n). \tag{2.11}$$

(D) *Assume* (1.1). *If* $k\alpha < 1$ *and* $\lim_{N\to\infty} h(N)N^{1-k\alpha}L^k(N) = \infty$, *then*

$$\lim_{N\to\infty} E(N^{k\alpha/2}L^{-k/2}(N)Z_{Nh(N)}(y))^2 = \frac{2g_k^2(y)}{k!(1-k\alpha)(2-k\alpha)}. \tag{2.12}$$

Remark 3. From the right-hand side of (2.7), we see that there are two rates competing with each other, $N^{-1}|R|(k,N)$ and $(Nh(N))^{-1}$. The central idea upon which the following three theorems are based is to carefully control the bandwidth sequences $\{h(N)\}$ to have one rate dominate the other. More specifically, (2.8) is equivalent to $N^{-1}|R|(k,N) \prec (Nh(N))^{-1}$, and condition $\lim_{N\to\infty} h(N)|R|(k,N) = \infty$ in (C) and (D) of Proposition 2 holds if and only if $(Nh(N))^{-1} \prec N^{-1}|R|(k,N)$ ($A \prec B$ means $A/B = o(1)$). The former case leads to Theorem 1 and 3, and the later case is exploited in Theorem 2. Also interesting is the possibility that the two rates may be equally competitive, in other words, the ratio $|R|(k,N)/h(N)$ is bounded away from zero and infinity. The asymptotics under this circumstance will be discussed in a subsequent paper.

We now present below our main theorems concerning limiting distributions. Each of the theorems is, as pointed out in Remark 3, in relation to certain part of Proposition 2. In fact, the norming factors employed in Theorems 1 and 3, Case (a) of Theorem 2, and Case (b) of Theorem 2 are derived, respectively, in (B), (C) and (D) of Proposition 2. Throughout the following three theorems. we let y_i's, $1 \leqslant i \leqslant m$, be distinct real numbers such that $f(y_i) > 0$, and let $k = \min\{k_1, \ldots, k_m\}$, where k_i denotes the Hermite rank of $(K(\cdot), G(\cdot), y_i)$.

Theorem 1. *Assume conditions* (C1)–(C5) *are satisfied. If* (2.8) *holds, then as* $N \to \infty$

$$\sqrt{Nh(N)}(Z_{Nh(N)}(y_1), \ldots, Z_{Nh(N)}(y_m)) \xrightarrow{\text{d}} N\left(0, \int K^2(u)\text{d}u \ \text{diag}\{f(y_1), \ldots, f(y_m)\}\right). \tag{2.13}$$

Theorem 2. *Assume* (C1)–(C5) *hold, and the bandwidth sequence* $\{h(N)\}$ *satisfies*

$$\lim_{N\to\infty} h(N)|R|(k,N) = \infty.$$

Suppose either one of the cases below holds.

Case (a): *Let* $|R|(k,N) = L_1(N)$ *be the same as in* (C) *of Proposition* 2. *Assume*

$$\lim_{N\to\infty} L_1^{-1}(N)R(k,N)$$

exists. Choose $A_N = \sqrt{N}L_1^{-1/2}(N)$ *and let the random variable Z be* $\mathrm{N}(0,\sigma_k^2)$ *with*

$$\sigma_k^2 = k! \lim_{N\to\infty} (NL_1(N))^{-1} \sum_{m,n=1}^{N} r^k(m-n).$$

Case (b): *Suppose* (1.1) *holds and* $k\alpha < 1$. *Choose* $A_N = N^{k\alpha/2}L^{-k/2}(N)$ *and let the random variable Z be represented through a k-fold multiple Wiener–Itô integral*

$$Z = \left[2\Gamma(\alpha)\cos\left(\frac{\alpha\pi}{2}\right)\right]^{-k/2} \int \frac{e^{i(x_1+\cdots+x_k)}-1}{i(x_1+\cdots+x_k)} |x_1\cdots x_k|^{(\alpha-1)/2}\mathrm{d}W(x_1)\cdots\mathrm{d}W(x_k). \tag{2.14}$$

$W(x)$ *is the complex Gaussian white noise* (cf. Major, 1981). *Then*

$$A_N\left(Z_{Nh(N)}(y_1),\ldots,Z_{Nh(N)}(y_m)\right) \xrightarrow{\mathrm{d}} \left(\frac{g_k(y_1)}{k!},\ldots,\frac{g_k(y_m)}{k!}\right)Z. \tag{2.15}$$

(Note that $g_k(y_i) = 0$ if $k_{K,G,y_i} > k$.)

Remark 4. In Case (b) of Theorem 2, instead of a weaker condition

$$|R|(k,N) = N^{\lambda}L_2(N), \quad 0 < \lambda < 1, \tag{2.16}$$

for some slowing varying function $L_2(\cdot)$, we assume (1.1). This is because (2.16) does not guarantee the existence of the limiting random variable Z (Dobrushin and Major, 1979, Remark 4.2). The limiting random variable Z as shown in (2.14) is non-Gaussian for $k\geqslant 2$ (Major, 1981, p. 68).

Theorem 3. *Assume conditions* (C1)–(C5) *hold. Suppose* $|R|(1,N) = N^{\beta}L_3(N)$ *is regularly varying with exponent* β, $0\leqslant\beta\leqslant 1$. *When* $\beta = 1$ *we further assume* $L_3(N)\to 0$ *as* $N\to\infty$. *If* $h(N)$ *satisfies* $h(N)|R|(1,N) = \mathrm{o}(1)$, *then* (2.13) *still holds.*

Remark 5. In Theorem 3, the condition on $|R|(1,N)$ is given mainly to assure that (C4) and $h(N)|R|(1,N) = 0$ can simultaneously hold for some $\{h(N)\}$. Also note that, when $\beta = 1$, (C5) implies $L_3(N) = \mathrm{o}(1)$. Theorem 3 has the advantage that one can assure the asymptotic normality (2.13) without needing to know the Hermite rank. This property is more useful, especially under non-parametric settings, when the instantaneous filter $G(x)$ is unknown and the calculation of the Hermite rank cannot be executed. Even when the decaying rate α, under (1.1), is not given, (2.13) remains to hold, if the bandwidth sequence $h(N)$ is chosen properly so that the conditions (C4) and $h(N)|R|(1,N) = \mathrm{o}(1)$ are both satisfied, e.g., $h(N) = L_4(N)/N$. The slowing varying function $L_4(N)$ tends to ∞ as $N\to\infty$.

Example 2. Under (1.1), let $G(x)$ and $K(x)$ be the same as given in Example 1. Recall that the Hermite rank $k_{K,G,y} = 1$ if $y\neq 0$, and 2 if $y = 0$. If the bandwidth sequence $\{h(N)\}$ satisfies $\lim_{n\to\infty} h(N)N^{1-\alpha}L(N) = \infty$, then by Case (b) of Theorem 2, $N^{\alpha/2}L^{-1/2}(N)Z_{Nh(N)}(y)$ tends to $(-H_1(y)\phi(y))Z$ with $y\neq 0$, where Z is Gaussian as shown in (2.14) with $k = 1$. Suppose $0 < \alpha < \frac{1}{2}$ and $\lim_{N\to\infty} h(N)N^{1-2\alpha}L(N)$

$=\infty$. By Case (b) of Theorem 2 again, $N^{\alpha}L^{-1}(N)Z_{Nh(N)}(0)$ converges in distribution to $(2\sqrt{2\pi})^{-1}Z'$. Z' is as specified in (2.14) with $k = 2$ and is thus non-Gaussian (see Remark 3). We have just demonstrated that the limiting distribution of properly normalized $Z_{Nh(N)}(y)$ may, under long-range dependence, change dramatically as y moves from 0 to any non-zero real number. When $\alpha = \frac{1}{2}, |R|(2,N) = L_5(N)$ is slowly varying, and by Case (a) of Theorem 2, $\sqrt{N}L_5^{-1/2}(N)Z_{Nh(N)}(0)$ is asymptotically $(2!\sqrt{2\pi})^{-1}\mathrm{N}(0,\sigma^2)$ with

$$\sigma^2 = 2! \lim_{N\to\infty} (NL_5(N))^{-1} \sum_{m,n=1}^{N} r^2(m-n),$$

provided

$$\lim_{N\to\infty} h(N)L_5(N) = \infty.$$

In the case of $1/2 < \alpha < 1$, Theorem 1 gives under (C4) that $\sqrt{Nh(N)}Z_{Nh(N)}(0)$ converges in distribution to $\mathrm{N}(0,(\sqrt{2\pi})^{-1}\int K^2(x)\,\mathrm{d}x)$. Without considering the Hermite rank $k_{K,G,y}$, we still have from Theorem 3 that the limit of $\sqrt{Nh(N)}$ $(Z_{Nh(N)}(y_1),\dots,Z_{Nh(N)}(y_m))$ for distinct y_i's is $N(0,\int K^2(x)\,\mathrm{d}x\ \mathrm{diag}\{f(y_1),\dots,f(y_m)\})$, if

$$\lim_{N\to\infty} h(N)N^{1-\alpha}L(N) = 0.$$

Remark 6. We may in above theorems consider the kernel functions which take negative values. A kernel function is said of high order $m\geqslant 2$, if

$$\int u^j K(u)\,\mathrm{d}u = 0, 1\leqslant j\leqslant m-1, \text{ and } \int u^m K(u)\,\mathrm{d}u \neq 0.$$

Suppose $f(y)\in C^m$, and the kernel $K(\cdot)$ is of high order m, then

$$E\hat{f}_N(y) - f(y) \sim \frac{(-h(N))^m f^{(m)}(y)}{m!}\int u^m K(u)\,\mathrm{d}u.$$

Let the estimator $\hat{f}_N(y)$ be centered at the true density $f(y)$ instead of $E\hat{f}_N(y)$, and set $Z'_{Nh(N)}(y) = \hat{f}_N(y) - f(y)$. The same conclusions as in Theorem 1, Case (b) of Theorem 2, and Theorem 3 are still true for $A_N Z'_{Nh(N)}(y)$ (A_N the norming factor), because there allows the bandwidth to be selected to satisfy $h^m(N)A_N \to 0$ as $N\to\infty$.

3. Proofs

During the course of our proofs, we use C, for convenience, to denote generic positive constant whose value may differ from one place to another.

Proof of Proposition 1. If the set $E = \{x \mid G'(x) = 0\}$ is empty, the assertion holds trivially. Suppose $E = \{x_1,\dots,x_d\}, x_i < x_{i+1}$. Set $B_0 = (-\infty, x_1], B_1 = [x_1,x_2],\dots,$ and $B_d = [x_d,+\infty)$. Define

$$G_i^{-1}(y) = \begin{cases} x & \text{if } y\in G(B_i) \text{ and } G(x) = y, x\in B_i,\\ 0 & \text{if } y\notin G(B_i),\end{cases}$$

and $\mu(B) = \int_B \phi(x)\,\mathrm{d}x$, for all Borel $B \in \mathscr{B}(\Re)$. Clearly,

$$\int_A f(y)\,\mathrm{d}y = \sum_{i=0}^{d} \mu(G_i^{-1}(A)), \quad \forall A \in \mathscr{B}(\Re). \tag{3.1}$$

Let $I_{y,\Delta}$ denote the interval $(y, y+\Delta)$ if $\Delta > 0$, or $(y+\Delta, y)$ if $\Delta < 0$. Then, with $\lambda_i(y,\Delta) \equiv \mu(G_i^{-1}(I_{y,\Delta} \cap G(B_i)))$, the limit

$$f_i(y) \equiv \lim_{\Delta \to 0} |\Delta|^{-1} \lambda_i(y,\Delta)$$

exists, for otherwise (3.1) leads to the following contradiction:

$$\begin{aligned} f(y) &= \lim_{\Delta\to 0} |\Delta|^{-1} \int_{I_{y,\Delta}} f(u)\,\mathrm{d}u = \sum_{i=1}^{d} \liminf_{\Delta\to 0} |\Delta|^{-1} \lambda_i(y,\Delta) \\ &< \sum_{i=1}^{d} \limsup_{\Delta\to 0} |\Delta|^{-1} \lambda_i(y,\Delta) = \lim_{\Delta\to 0} |\Delta|^{-1} \int_{I_{y,\Delta}} f(u)\,\mathrm{d}u = f(y). \end{aligned}$$

Moreover, the functions $f_i(y), 0 \leqslant i \leqslant d$, defined above satisfy

$$f(y) = \sum_{i=0}^{d} f_i(y), \quad y \in \Re, \tag{3.2}$$

and

$$f_i(y) = \begin{cases} \phi(G_i^{-1}(y)) \left|[G_i^{-1}(y)]'\right|, & y \in \operatorname{Int} G(B_i), \\ 0 & y \notin G(B_i). \end{cases} \tag{3.3}$$

From (3.3) and (C2) we see that each $f_i(y)$ is continuous on $\Re - \partial G(B_i)$. We are going to show that $f_i(y)$ is continuous on $\Re$. Taking $\limsup_{y\to y_0}$ and $\liminf_{y\to y_0}$ on both sides of (3.2) proves the limits

$$\lim_{y\to y_0} f_i(y)$$

exist $\forall y_0 \in \Re$ and $0 \leqslant i \leqslant d$. In particular, by (3.3),

$$\lim_{y\to y_i} f_i(y) = 0, y_i \in \partial G(B_i), \quad 0 \leqslant i \leqslant d. \tag{3.4}$$

(set, if $A = [c,\infty)$ or $(-\infty, c], \partial A = \{c\}$). For any fixed y, define $J(y)$ to be the subset of $\{0, 1, \ldots, d\}$ such that $i \in J(y)$ if and only if $y \in \operatorname{Int} G(B_i)$. Use (3.2) to obtain

$$[f(y+\delta) - f(y)] - \sum_{i\in J(y)} [f_i(y+\delta) - f_i(y)] = \sum_{i \notin J(y)} [f_i(y+\delta) - f_i(y)]. \tag{3.5}$$

Let $\delta \to 0$ on both sides of (3.5). The left-hand side is clearly zero due to the continuity of $f(y)$ and $f_i(y)$ with $i \in J(y)$. For the right-hand side,

$$\lim_{\delta\to 0} \sum_{i\notin J(y)} f_i(y+\delta) = 0$$

by (3.3) and (3.4). Hence, $f_i(y) = 0 \ \forall i \notin J(y)$, and, in particular, $f_i(y) = 0$ if $y \in \partial G(B_i)$. This implies $f_i(y)$ is continuous $\forall y \in \Re$ and $0 \leqslant i \leqslant d$. Define

$$g_{ji}(y) = H_j(G_i^{-1}(y)) f_i(y), \quad 0 \leqslant i \leqslant d.$$

Since $f_i(y)$ is continuous and $f_i(y)=0\ \forall y\notin \text{Int } G(B_i), g_{ji}(y)$ is continuous on $\Re$ at least for $1\leqslant i\leqslant d-1$. When $i=0$ or d, $g_{ji}(y)$ is continuous on Int $G(B_i)$ by (3.3). We now show that $g_{ji}(\cdot)$ is also continuous on $\partial G(B_i, i=0$ or $d)$ and need only to focus on the case that $y=c$ for some finite number c is the asymptotic line of the graph of $G(x)$, i.e.,

$$\lim_{x\to x^\star} G(x)=c,(x^\star=+\infty \text{ or } -\infty).$$

For convenience, set $x^\star=+\infty$ (i.e. $i=d$, the case of $i=0$ is similar). First, $g_{jd}(c)=0$. Recall (3.3) and by change of variable,

$$\lim_{\substack{y\to c\\ y\in G(B_d)}} |g_{jd}(y)| = \lim_{x\to\infty}\left|\frac{H_j(x)[\phi(x)]'}{x}\right| = 0 = \lim_{\substack{y\to c\\ y\notin G(B_d)}} g_{jd}(y).$$

Thus, $g_{jd}(y)$ is continuous on the whole $\Re$. Observe that

$$\begin{aligned}\frac{a_{jh(N)}(y)}{h(N)} &= \sum_{i=0}^{d}\int K(u)\,g_{ji}(y-h(N)u)\,\mathrm{d}u\\ &\to \sum_{i=0}^{d} g_{ji}(y) \quad \text{as } h(N)\to 0. \end{aligned} \tag{3.6}$$

The interchange of the integral and the limit is valid because $g_{ji}(y)$ is continuous and $K(u)$ has compact support. Set $g_j(y)=\sum_{i=0}^{d} g_{ji}(y)$ to conclude (2.4). To show (2.5), recall that

$$g_j(y)=\sum_{i=0}^{d} H_j(G_i^{-1}(y))f_i(y) \quad \text{and} \quad f(y)=\sum_{i=0}^{d} f_i(y).$$

The if part of (2.5) is obvious. Conversely, we construct a function $H(x)$ such that $H(x)>0,\ \forall x\in\Re$, and the Fourier–Hermite expansion

$$H(x)=\sum_{j=1}^{\infty} b_j H_j(x)$$

converges absolutely. Sum up $g_j(y)b_j$ over $j\geqslant 1$, $f(y)=0$ follows immediately. □

Proof of Proposition 2. (A) As $EH_i(X_m)H_j(X_n)=\delta(i-j)j!r^j(m-n)$ (cf. Taqqu, 1977, Lemma 3.2), we get from (2.3)

$$\begin{aligned} E\left(Z_{Nh(N)}(y)\right)^2 &= (Nh(N))^{-2}\sum_{j=k}^{\infty}\frac{a^2_{jh(N)}(y)}{j!}\left(\sum_{m,n=1}^{N} r^j(m-n)\right)\\ &= N^{-1}\left(\frac{a^2_{kh(N)}(y)}{k!\,h^2(N)}\right)\left(1+2N^{-1}\sum_{n=1}^{N-1}R(k,n)\right)\\ &\quad +(Nh(N))^{-1}\sum_{j=k+1}^{\infty}\frac{a^2_{jh(N)}(y)}{j!\,h(N)}+N^{-1}\sum_{j=k+1}^{\infty}\frac{a^2_{jh(N)}(y)}{j!\,h^2(N)}\left(2N^{-1}\sum_{n=1}^{N-1}R(j,n)\right)\\ &\equiv N^{-1}S_1+(Nh(N))^{-1}S_2+N^{-1}S_3. \end{aligned}$$

Note that

$$N^{-1}S_1 = E\left[\frac{a_{kh(N)}(y)}{k!h(N)}\left(\frac{1}{N}\sum_{n=1}^{N}H_k(X_n)\right)\right]^2$$

and

$$(Nh(N))^{-1}S_2 + N^{-1}S_3 = E\left[\sum_{j=k+1}^{\infty}\frac{a_{jh(N)}(y)}{j!h(N)}\left(\frac{1}{N}\sum_{n=1}^{N}H_j(X_n)\right)\right]^2.$$

As $h(N) \to 0$,

$$S_2 = \int K^2(x)f(y-h(N)x)\,\mathrm{d}x - \frac{a^2_{kh(N)}(y)}{k!h(N)} \longrightarrow f(y)\int K^2(x)\,\mathrm{d}x \tag{3.7}$$

by (2.1) and (2.4), and

$$S_1 = \mathrm{O}(|R|(k,N))\left(\frac{g_k^2(y)}{k!} + \mathrm{o}(1)\right) \tag{3.8}$$

by (2.4). For Hermite polynomials $H_j(x)$, the following upper bound (Lukacs, 1970, p.78)

$$|H_j(x)| \leqslant \mathrm{e}^{x^2/2}2^{j/2}\pi^{-1/2}\Gamma\left(\frac{j+1}{2}\right) \tag{3.9}$$

and the recursive relation (Major, 1981, p. 38)

$$xH_j(x) = H_{j+1}(x) + jH_{j-1}(x)$$

jointly imply

$$\sup_{x\in\Re}|H_j(x)[\phi(x)]'| \leqslant \pi^{-1}2^{j/2}j\Gamma\left(\frac{j}{2}\right). \tag{3.10}$$

We are now ready to provide an asymptotic bound for S_3. First, by relation (3.6),

$$|a_{jh(N)}(y)| \leqslant h(N)\sum_{i=0}^{d}\int K(u)\,|g_{ji}(y-h(N)u)|\,\mathrm{d}u. \tag{3.11}$$

For each i, $0 < i < d$, let $G_i^{-1}(\cdot)$ be as defined in the beginning of Proof of Proposition 1, then $\sup_{u\in\Re}\left|G_i^{-1}(u)\right| < \infty$, and by (3.9) and Stirling's formula

$$\begin{aligned}
&\frac{1}{\sqrt{j!}}\left|\int K(u)g_{ji}(y-h(N)u)\,\mathrm{d}u\right| \\
&\leqslant \frac{1}{\sqrt{j!}}\int K(u)\left|H_j(G_i^{-1}(y-h(N)u))f_i(y-h(N)u)\right|\,\mathrm{d}u \qquad (3.12)\\
&\leqslant C(j+1)^{-1/4}.
\end{aligned}$$

Suppose, when $i = 0$ or d, $\sup_{u\in\Re}|G_i^{-1}(u)| = \infty$. Set, for convenience, $i = d$. Then (3.10) and Stirling's formula assure

$$\begin{aligned} &\frac{1}{\sqrt{j!}}\left|\int K(u)g_{jd}(y-h(N)u)\,\mathrm{d}u\right| \\ &\leqslant \frac{1}{\sqrt{j!}}\int K(u)\frac{\left|H_j(G_d^{-1}(y-h(N)u))\left[\phi(G_d^{-1}(y-h(N)u))\right]'\right|}{\left|(G_d^{-1}(y-h(N)u))\right|}\,\mathrm{d}u \\ &\leqslant C(j+1)^{1/4}. \end{aligned} \tag{3.13}$$

Note that

$$N^{-1}\sum_{n=1}^{N-1} R(j,m) = \sum_{n=1}^{N-1}\left(1-\frac{n}{N}\right)r^j(n).$$

Applying (3.12) and (3.13) to relation (3.11), we obtain

$$\begin{aligned} |S_3| &\leqslant C\sum_{n=1}^{N}\sum_{j=k+1}^{\infty} j^{1/4}|r(n)|^j \\ &= C\sum_{n=1}^{N}|r(n)|^k\sum_{j=k+1}^{\infty} j^{1/4}|r(n)|^{j-k} \\ &= B_N\,\mathrm{O}(|R|(k,N))\ \ (\text{by (C5)}), \end{aligned} \tag{3.14}$$

where B_N is bounded by one and is o(1) if $\lim_{N\to\infty}|R|(k,N) = \infty$. Note that (C5) implies $\sup_{n\neq 0}|r(n)| < 1$. Relation (2.7) then follows by (3.7), (3.8) and (3.14).

(B) (2.8) implies $N^{-1}|R|(k,N) = \mathrm{o}(Nh(N))^{-1}$. This and (2.7) give (2.9).

(C) First, with $h(N)$ satisfied $\lim_{N\to\infty} h(N)|R|(k,N) = \infty$,

$$(Nh(N))^{-1} = \mathrm{o}\left(N^{-1}|R|(k,N)\right) \tag{3.15}$$

We then have (2.10) from

$$\begin{aligned} &\lim_{N\to\infty} E\left(\sqrt{N}|R|^{-1/2}(k,N)Z_{Nh(N)}(y)\right)^2 \\ &= \lim_{N\to\infty}\frac{a^2_{kh(N)}(y)}{k!h^2(N)}(N|R|(k,N))^{-1}\frac{1}{k!}E\left(\sum_{n=1}^{N}H_k(X_n)\right)^2 \ (\text{by (2.7) and (3.15)}), \\ &= \frac{cg_k^2(y)}{k!}\ (\text{by (2.4) and (2.11)}). \end{aligned} \tag{3.16}$$

(2.11) is justified by the assumption that the limit $\lim_{N\to\infty}|R|^{-1}(k,N)R(k,N)$ exists.

(D) For $k\alpha < 1$, we can apply Karamata's theorem (Feller, 1971, p. 281) to (1.1) and get as $N \to \infty$

$$R(k,N) = |R|(k,N) \sim \frac{N^{1-k\alpha}L^k(N)}{1-k\alpha} \tag{3.17}$$

$$\sum_{m,n=1}^{N} r^k(m-n) \sim \frac{2N^{2-k\alpha}L^k(N)}{(1-k\alpha)(2-k\alpha)}. \tag{3.18}$$

$C_N \sim C'_N$ means that $C_N/C'_N \to 1$ as $N \to \infty$. We also have the same asymptotic relation as (3.15) from (3.17) and the condition $\lim_{N\to\infty} h(N)N^{1-k\alpha}L^k(N) = \infty$. (2.12) then follows by

$$\lim_{N\to\infty} E(N^{k\alpha/2}L^{-k/2}(N)Z_{Nh(N)}(y))^2$$
$$= \lim_{N\to\infty}\left[\frac{a^2_{kh(N)}(y)}{k!h^2(N)}\right]\left[N^{k\alpha}L^{-k}(N)E\left(\frac{1}{k!}\sum_{n=1}^{N} H_k(X_n)\right)^2\right] \text{(by (2.7) and (3.15))},$$
$$= \frac{2g_k^2(y)}{k!(1-k\alpha)(2-k\alpha)} \text{(by (2.4) and (3.18)).} \tag{3.19}$$

Proof of Theorems 1 and 3. Our proof is based upon the method of moments. Consider a linear combination

$$\sum_{i=1}^{m} b_i Z_{Nh(N)}(y_i)$$

of $Z_{Nh(N)}(y_i)$'s. If, for any positive integer p,

$$E\left[\sqrt{Nh(N)}\left(\sum_{i=1}^{m} b_i Z_{Nh(N)}(y_i)\right)\right]^p$$
$$\overset{N\to\infty}{\longrightarrow} E\left[N\left(0, \sum_{i=1}^{m} b_i^2 f(y_i)\right)\int K^2(u)\,\mathrm{d}u\right]^p, \tag{3.20}$$

then by the Cramér–Wold device, (2.13) follows. To show (3.20), we begin with the following notations. Fix positive integer p and define

$$Q(p) = \left\{ \boldsymbol{t} = (t_1,\ldots,t_j), 1 \leqslant j \leqslant p \,\middle|\, t_i \in Z^+, \sum_{i=1}^{j} t_i = p, t_i \leqslant t_{i+1} \right\}$$

and

$$Q^\star(p) = \{\boldsymbol{t} = (t_1,\ldots,t_j) \in Q(p), 1 \leqslant j \leqslant p \,|\, t_i \geqslant 2, 1 \leqslant i \leqslant j\}.$$

For each given partition $\boldsymbol{t} = (t_1,\ldots,t_s) \in Q(p)$, set $|\boldsymbol{t}| \equiv s$ and, for fixed m and N,

$$N(s) \equiv \{(n_1,\ldots,n_s) \,|\, 1 \leqslant n_i \leqslant N \text{ and } n_i\text{'s}, 1 \leqslant i \leqslant s, \text{ are all distinct, i.e., } n_i \neq n_j \text{ if } i \neq j\},$$
$$N_T(s) \equiv \{(n_1,\ldots,n_s) \in N(s) \,|\, |n_i - n_j| > T \;\forall i \neq j, 1 \leqslant i, j \leqslant s\}$$

and

$$m(\boldsymbol{t}) \equiv \{\boldsymbol{l} = (l_{1,1},\ldots,l_{1,t_1}, l_{2,1},\ldots,l_{2,t_2},\ldots,l_{s,1},\ldots,l_{s,t_s}) \,|\, 1 \leqslant l_{i,j} \leqslant m \text{ for each pair } (i,j) \text{ with } 1 \leqslant i \leqslant s \text{ and } 1 \leqslant j \leqslant t_i\}.$$

For every element $\boldsymbol{l} \in m(\boldsymbol{t})$, the (i,j)th component $l_{i,j}$ of $\boldsymbol{l}$ is denoted by $\boldsymbol{l}(i,j)$. Use the notations given above to simplify

$$
\begin{aligned}
&E\left[\sqrt{Nh(N)}\sum_{i=1}^{m} b_i Z_{Nh(N)}(y)\right]^p \\
&= (Nh(N))^{-p/2} E\left[\sum_{n=1}^{N}\sum_{i=1}^{m} b_i \tilde{K}\left(\frac{y_i - G(X_n)}{h(N)}\right)\right]^p \\
&= (Nh(N))^{-p/2} \sum_{\boldsymbol{t}\in Q(p)} c(\boldsymbol{t}) \sum_{\boldsymbol{l}\in m(\boldsymbol{t})} \sum_{\boldsymbol{n}=(n_1,\dots,n_{|\boldsymbol{t}|})\in N(|\boldsymbol{t}|)} \\
&\quad \times E\prod_{i=1}^{|\boldsymbol{t}|}\prod_{j=1}^{t_i} b_{\boldsymbol{l}(i,j)} \tilde{K}\left(\frac{y_{\boldsymbol{l}(i,j)} - G(X_{n_i})}{h(N)}\right) \\
&= (Nh(N))^{-p/2}\left(\sum_{\substack{\boldsymbol{t}\in Q(p)\\ |\boldsymbol{t}|<p/2}} + \sum_{\substack{\boldsymbol{t}\in Q^\star(p)\\ |\boldsymbol{t}|=p/2}} + \sum_{\substack{\boldsymbol{t}\in Q(p)-Q^\star(p)\\ |\boldsymbol{t}|\geq p/2}}\right) \\
&\equiv (Nh(N))^{-p/2}\left(\textstyle\sum_1 + \sum_2 + \sum_3\right).
\end{aligned}
$$

$c(\boldsymbol{t})$ are constants associated with the partition $\boldsymbol{t}$ of p. Note that the sum $\sum_2$ above is null for odd p, and covers only one term with $\boldsymbol{t} = (2,\dots,2)$ ($p/2$ components) if p is even. It is easy to check that the general formula for coefficient $c(\boldsymbol{t})$ with $\boldsymbol{t} = (t_1,\dots,t_s) \in Q(p)$ is

$$c(\boldsymbol{t}) = (s!t_1!\cdots t_s!)^{-1}p!.$$

In particular, when p is even and the partition $\boldsymbol{t}$ is $\boldsymbol{t} = (2,\dots,2)$ ($p/2$ terms)

$$c(\boldsymbol{t}) = \frac{p!}{2^{p/2}(\frac{p}{2})!}. \tag{3.21}$$

We shall conclude (3.20) by showing as $N \to \infty$

$$
(Nh(N))^{-p/2}\left(\textstyle\sum_1 + \sum_2\right) \to
\begin{cases}
\frac{p!}{2^{\frac{p}{2}}(p/2)!}\left[\int K^2(u)\,\mathrm{d}u\left(\sum_{i=1}^{m} b_i^2 f(y_i)\right)\right]^{p/2} & \text{if } p \text{ is even,}\\
0 & \text{if } p \text{ is odd,}
\end{cases}
\tag{3.22}
$$

and

$$(Nh(N))^{-p/2}\textstyle\sum_3 \longrightarrow 0. \tag{3.23}$$

First of all, we show that

$$\lim_{N\to\infty} N^{-s} \sum_{(n_1,\dots,n_s)\in N(s)} f_{n_1,\dots,n_s}(y_1,\dots,y_s) = \prod_{t=1}^{s} f(y_t). \tag{3.24}$$

Set $I_t(\triangle) = (y_t - \triangle, y_t + \triangle), \triangle > 0$. Fix a sufficiently large T so that for each $(n_1,\dots,n_s) \in N_T(s)$, $|EX_{n_i}X_{n_j}| \leqslant \sup_{n\geqslant T}|r(n)| \equiv \theta_T < 1/(s-1)$, $i \neq j$. Then by

Lemma 3.3 of Taqqu (1977)

$$
\begin{aligned}
&P(Y_{n_t} \in I_t(\triangle), 1 \leqslant t \leqslant s)/(2\triangle)^s \\
&= \prod_{t=1}^{s} \left[P(Y_{n_t} \in I_t(\triangle))/(2\triangle)\right] \\
&\quad + \sum_{q=1}^{\infty} \sum_{\substack{k_1+\cdots+k_s=2q \\ 0\leqslant k_1,\ldots,k_s\leqslant q}} EH_{k_1}(X_{n_1})\cdots H_{k_s}(X_{n_s}) \\
&\quad \times \prod_{t=1}^{s} (2\triangle)^{-1} \int_{I_t(\triangle)} \frac{H_{k_t}(x_t)}{k_t!} \phi(x_t)\,\mathrm{d}x_t \\
&\equiv R_1(\triangle, y_t, 1\leqslant t\leqslant s) + R_2(n_t, y_t, 1\leqslant t\leqslant s).
\end{aligned}
$$

By mean value theorem and continuity of $f'(\cdot)$ (assumed in (C1))

$$R_1(\triangle, y_t, 1\leqslant t\leqslant s) \to \prod_{t=1}^{s} f(y_t) \tag{3.25}$$

as $\triangle \to 0$. Applying similar argument used in (3.12) and (3.13) to the integral in $R_2(n_t, y_t, 1\leqslant t\leqslant s)$, we obtain

$$(2\triangle)^{-1} \int_{I_t(\triangle)} \left| \frac{H_{k_t}(x_t)}{\sqrt{k_t!}} \right| \phi(x_t)\,\mathrm{d}x_t = \mathrm{O}(k_t^{1/4}). \tag{3.26}$$

For each $(n_1, \ldots, n_s) \in N_T(s)$, we also have

$$|EH_{k_1}(X_{n_1})\cdots H_{k_s}(X_{n_s})| \leqslant \prod_{t=1}^{s} \left([\theta_T(s-1)]^{k_t} k_t!\right)^{1/2} \tag{3.27}$$

(see Taqqu, 1977, the second display in p. 214). Then (3.26) and (3.27) jointly imply

$$
\begin{aligned}
|R_2(n_t, y_t, 1\leqslant t\leqslant s)| &\leqslant C \left[\sum_{j=1}^{\infty} (\theta_T(s-1))^{j/2} \right]^s \\
&= C \left[\frac{(\theta_T(s-1))^{1/2}}{1-(\theta_T(s-1))^{1/2}} \right]^s,
\end{aligned} \tag{3.28}
$$

where the constant C is independent of T and any particular s-tuple $(n_1, \ldots, n_s) \in N_T(s)$. Note that $\theta_T \to 0$ as $T \to \infty$ (by (C5)). Immediately from (3.25) and (3.28)

$$\lim_{T\to\infty} \sup_{(n_1,\ldots,n_s)\in N_T(s)} \left| f_{n_1,\ldots,n_s}(y_1, \ldots, y_s) - \prod_{t=1}^{s} f(y_t) \right| = 0. \tag{3.29}$$

Write

$$
\begin{aligned}
&\sum_{(n_1,\ldots,n_s)\in N(s)} \left| f_{n_1,\ldots,n_s}(y_1, \ldots, y_s) - \prod_{t=1}^{s} f(y_t) \right| \\
&= \sum_{(n_1,\ldots,n_s)\in N_T(s)} + \sum_{\substack{(n_1,\ldots,n_s)\in N(s) \\ |n_i-n_j|\leqslant T,\ \text{for some } (i,j),\ i\neq j}} \\
&\equiv \textstyle\sum_{1,T}^{\star} + \sum_{2,T}^{\star}.
\end{aligned}
$$

It is clear that $|N_T(s)| = O(N^s)$ and $|N(s) - N_T(s)| = O(N^{s-1})$. Hence, (3.29) implies that for any $\varepsilon > 0$ we can find a large T such that $\lim_{N\to\infty} N^{-s}\sum_{1,T}^{\star} < \varepsilon$ and, using the uniform boundedness of $f_{n_1,\ldots,n_s}(\cdot)$ assumed in (C1), $\lim_{N\to\infty} N^{-s}\sum_{2,T}^{\star} = 0$. (3.24) is then evident. Fix $\boldsymbol{t} \in Q(p), \boldsymbol{l} \in m(\boldsymbol{t})$, and $\boldsymbol{n} = (n_1, \ldots, n_{|\boldsymbol{t}|}) \in N(|\boldsymbol{t}|)$. With i fixed, let $\boldsymbol{l}(i) \in \{\boldsymbol{l}(i,j), 1 \leqslant j \leqslant t_i\}$ be the positive integer such that

$$y_{\boldsymbol{l}(i)} = \min_{1\leqslant j\leqslant t_i} \{y_{\boldsymbol{l}(i,j)}\},$$

and set $\Delta_{\boldsymbol{l}(i,j)} = y_{\boldsymbol{l}(i,j)} - y_{\boldsymbol{l}(i)}, 1 \leqslant j \leqslant t_i$. Applying mean value theorem to $f_{n_1,\ldots,n_{|\boldsymbol{t}|}}$ gives

$$\begin{aligned}
&E(\boldsymbol{t},\boldsymbol{l},\boldsymbol{n}) \\
&\equiv E\prod_{i=1}^{|\boldsymbol{t}|}\prod_{j=1}^{t_i} b_{\boldsymbol{l}(i,j)}\tilde{K}\left(\frac{y_{\boldsymbol{l}(i,j)} - G(X_{n_i})}{h(N)}\right) \\
&= (h(N))^{|\boldsymbol{t}|}\int\left\{\prod_{i=1}^{|\boldsymbol{t}|}\prod_{j=1}^{t_i} b_{\boldsymbol{l}(i,j)}\tilde{K}\left(\frac{\Delta_{\boldsymbol{l}(i,j)}}{h(N)} + u_i\right)\right\} \\
&\quad\times f_{n_1,\ldots,n_{|\boldsymbol{t}|}}\left(y_{\boldsymbol{l}(1)} - h(N)u_1, \ldots, y_{\boldsymbol{l}(|\boldsymbol{t}|)} - h(N)u_{|\boldsymbol{t}|}\right) \mathrm{d}u_1\cdots \mathrm{d}u_{|\boldsymbol{t}|} \qquad (3.30)\\
&= \begin{cases} (h(N))^{|\boldsymbol{t}|} f_{n_1,\ldots,n_{|\boldsymbol{t}|}}(y_{\boldsymbol{l}(1)},\ldots,y_{\boldsymbol{l}(|\boldsymbol{t}|)}) & \\ \quad\times\left(\prod_{i=1}^{|\boldsymbol{t}|}\int[b_{\boldsymbol{l}(i)}\tilde{K}(u_i)]^{t_i}\mathrm{d}u_i\right)(1+O(h(N))) & \text{if all } \Delta_{\boldsymbol{l}(i,j)} = 0, \\ o(h^{|\boldsymbol{t}|}(N)) & \text{otherwise.}\end{cases}
\end{aligned}$$

Note that as $h(N) \to 0$,

$$\int \tilde{K}^t(u)\,\mathrm{d}u \to \int K^t(u)\,\mathrm{d}u. \tag{3.31}$$

When $\boldsymbol{t} \in Q(p)$ satisfies $|\boldsymbol{t}| < p/2$, by (3.24) and $Nh(N) \to \infty$,

$$\begin{aligned}(Nh(N))^{-p/2}\textstyle\sum_1 &= (Nh(N))^{-(p/2-|\boldsymbol{t}|)}\left[(Nh(N))^{-|\boldsymbol{t}|}\sum_{\boldsymbol{n}\in N(|\boldsymbol{t}|)} E(\boldsymbol{t},\boldsymbol{l},\boldsymbol{n})\right] \\ &\to 0 \text{ as } N \to \infty. \end{aligned}\tag{3.32}$$

Suppose p is even and $\boldsymbol{t} = (2,\ldots,2)$ ($p/2$ terms), then, by noting (3.21),

$$\begin{aligned}
&(Nh(N))^{-p/2}\textstyle\sum_2 \\
&= \frac{p!}{2^{p/2}(\frac{p}{2})!}\left\{\sum_{\substack{\boldsymbol{l}\in m(\boldsymbol{t})\\ \Delta_{\boldsymbol{l}(i,j)}=0}}\left(\prod_{i=1}^{p/2}\int\left[b_{\boldsymbol{l}(i)}\tilde{K}(u_i)\right]^2\mathrm{d}u_i\right)\right. \\
&\quad\left.\times\left(N^{-p/2}\sum_{\boldsymbol{n}\in N(|\boldsymbol{t}|)} f_{n_1,\ldots,n_{p/2}}(y_{\boldsymbol{l}(1)},\ldots,y_{\boldsymbol{l}(p/2)})\right)\right\} + o(1) \qquad \text{(by (3.30))} \\
&\overset{N\to\infty}{\longrightarrow} \frac{p!}{2^{p/2}(\frac{p}{2})!}\left(\int K^2(u)\,\mathrm{d}u\right)^{p/2}
\end{aligned}$$

$$\times\left(\sum_{\substack{l\in m(t)\\ l(i,1)=l(i,2)}}\prod_{i=1}^{p/2} b_{l(i)}^2 f(y_{l(i)})\right) \text{ (by (3.24) and (3.31))}$$

$$=\frac{p!}{2^{p/2}(\frac{p}{2})!}\left[\int K^2(u)\mathrm{d}u\left(\sum_{i=1}^{m} b_i^2 f(y_i)\right)\right]^{p/2}.$$

This and (3.32) gives (3.22). It remains to verify (3.23) which requires a delicate analysis of the growth of $\sum_3$. The main task to achieve this is to compute $E(\boldsymbol{t},\boldsymbol{l},\boldsymbol{n})$ (defined in (3.30)), $\boldsymbol{n}=(\boldsymbol{t})$, with respect to the joint p.d.f. $\phi_{\boldsymbol{t}}$ of $(X_{n_1},\dots,X_{n_{|t|}})$. The first step is to split $\phi_{\boldsymbol{t}}$ as the product of the joint p.d.f. of $(X_{n_{i_1}},\dots,X_{n_{i_j}})$ and the joint conditional p.d.f. of $(X_{n_{i_1^\star}},\dots,X_{n_{i^\star_{|t|-j}}})$ given $(X_{n_{i_1}},\dots,X_{n_{i_j}})$, then carry out the integration. To this end, we have to introduce another set of notations. Fix $p\geqslant 2$ and $\boldsymbol{t}\in Q(p)-Q^\star(p)$ with $|\boldsymbol{t}|\geqslant p/2$. Set

$$S(|\boldsymbol{t}|)=\{1,2,\dots,|\boldsymbol{t}|\},\quad \mathscr{F}(|\boldsymbol{t}|)=\text{ family of all the subsets of } S(|\boldsymbol{t}|).$$

For each $A=\{i_1,\dots,i_{|A|}\}\in\mathscr{F}(|\boldsymbol{t}|)$, its complement is denoted by $A^{\mathrm{c}}=S(|\boldsymbol{t}|)-A=\{i_1^\star,\dots,i^\star_{|A^{\mathrm{c}}|}\}$. Fix $T>0$ and $A\in\mathscr{F}(|\boldsymbol{t}|)$, and define

$$N_{T,A}(|\boldsymbol{t}|)=\{\boldsymbol{n}=(n_1,\dots,n_{|\boldsymbol{t}|})\in N(|\boldsymbol{t}|)\,|\,\forall i,j\in A, i\neq j, |n_i-n_j|>T;$$
$$\text{for each } i\in A^{\mathrm{c}}\ \exists j\in S(|\boldsymbol{t}|)\ni |n_i-n_j|\leqslant T\},$$

with the convention

$$N_{T,\emptyset}(|\boldsymbol{t}|)=\{\boldsymbol{n}=(n_1,\dots,n_{|\boldsymbol{t}|})\in N(|\boldsymbol{t}|)\,|\,|n_i-n_j|\leqslant T\ \forall i\leqslant j, 1\leqslant i,j\leqslant|\boldsymbol{t}|\}$$

and $N_{T,|A|}(|\boldsymbol{t}|)=\emptyset$ if $|A|=1$. Recall that

$$N_T(|A|)=\{\boldsymbol{n}=(n_1,\dots,n_{|A|})\in N(|A|)\,|\,|n_i-n_j|>T\ \forall i\neq j, 1\leqslant i,j\leqslant |A|\},$$

and, for any given $\boldsymbol{v}\in N_T(|A|)$, define

$$N_{T,A}(|\boldsymbol{t}|)/\boldsymbol{v}=\{\boldsymbol{n}=(n_1,\dots,n_{|\boldsymbol{t}|})\in N_{T,A}(|\boldsymbol{t}|)\,|(n_{i_1},\dots,n_{i_{|A|}})=\boldsymbol{v}\}.$$

Clearly,

$$N(|\boldsymbol{t}|)\subset\bigcup_{A\in\mathscr{F}(|\boldsymbol{t}|)}N_{T,A}(|\boldsymbol{t}|), \tag{3.33}$$

$$N_{T,A}(|\boldsymbol{t}|)=\sum_{\boldsymbol{v}\in N_T(|A|)}\left(N_{T,A}(|\boldsymbol{t}|)/\boldsymbol{v}\right),\quad |A|\geqslant 2. \tag{3.34}$$

The inclusive relation (3.33) is due to that some members of $N(|\boldsymbol{t}|)$ may belong to $N_{T,A}(|\boldsymbol{t}|)$ for several different A's. The multiplicity of the overcounts for all $\boldsymbol{n}\in N(|\boldsymbol{t}|)$ have a uniform upper bound which does not depend on N. For example, $|\boldsymbol{t}|=5$, $T=2$, and $\boldsymbol{n}=(6,3,2,4,7)$. Then $\boldsymbol{n}\in N_{2,A}(5)$ for $A=\{1,2\},\{4,5\},\{1,3\},\{2,5\}$ or$\{3,5\}$. Define

$$J(\boldsymbol{t},\boldsymbol{l},A,T)\equiv\sum_{\boldsymbol{n}\in N_{T,A}(|\boldsymbol{t}|)}E(\boldsymbol{t},\boldsymbol{l},\boldsymbol{n}).$$

It is clear from (3.33) that we need only to show

$$(Nh(N))^{-p/2} J(\boldsymbol{t},\boldsymbol{l},A,T) \to 0 \quad \text{as } N \to \infty,$$

to have (3.23). As mentioned before, the set A for $J(\boldsymbol{t},\boldsymbol{l},A,T)$ can only be empty or $|A| \geqslant 2$. Recall that $p \geqslant 2$, $|\boldsymbol{t}| \geqslant p/2$ and $\boldsymbol{t} \in Q(p) - Q^\star(p)$. When $A = \emptyset$, from expression (3.30)

$$(Nh(N))^{-p/2} |J(\boldsymbol{t},\boldsymbol{l},A,T)| = \mathrm{O}\left(N^{1-p/2}(h(N))^{|\boldsymbol{t}|-p/2}\right) \to 0 \text{ as } N \to \infty,$$

because $\boldsymbol{t} \in Q(p) - Q^\star(p)$ assures $|\boldsymbol{t}| = 2$ when $p = 2$. Hence, it suffices to concentrate on those $J(\boldsymbol{t},\boldsymbol{l},A,T)$'s with $|A| \geqslant 2$. Fix $A = \{i_1,\ldots,i_{|A|}\}$ with $|A| \geqslant 2$ and $A^c = \{i_1^\star,\ldots,i_{|A^c|}^\star\}$, and for each $\boldsymbol{a} = (n_{i_1},\ldots,n_{i_{|A|}}) \in N_T(|A|)$ and $\boldsymbol{n}' = (n_1,\ldots,n_{|\boldsymbol{t}|}) \in N_{T,A}(|\boldsymbol{t}|)/\boldsymbol{a}$, we adopt the following abbreviation: $\phi_{\boldsymbol{a}}(\cdot)$, the joint p.d.f. of $(X_{n_{i_1}},\ldots,X_{n_{i_{|A|}}})$; and $\phi_{\boldsymbol{n}',\boldsymbol{a}}(\cdot)$, the joint conditional p.d.f. of $(X_{n_{i_1^\star}},\ldots,X_{n_{i_{|A^c|}^\star}})$ given $(X_{n_{i_1}},\ldots,X_{n_{i_{|A|}}})$. Use (3.34) to see

$$\begin{aligned} &J(\boldsymbol{t},\boldsymbol{l},A,T)\\ &= \sum_{\boldsymbol{a}\in N_T(|\boldsymbol{t}|)} \int \left\{ \prod_{s=1}^{|A|}\prod_{j=1}^{t_{i_s}} b_{\boldsymbol{l}(i_s,j)} \tilde{K}\left(\frac{y_{l(i_s,j)} - G(u_s)}{h(N)}\right)\right\} \phi_{\boldsymbol{a}}(u_1,\ldots,u_{|A|})\mathrm{d}u_1\cdots\mathrm{d}u_{|A|}\\ &\times \int \sum_{\boldsymbol{n}'\in N_{T,A}(|\boldsymbol{t}|)/\boldsymbol{a}} \left\{ \prod_{s=1}^{|A^c|}\prod_{j=1}^{t_{i_s^\star}} b_{l(i_s^\star,j)} \tilde{K}\left(\frac{y_{l(i_s^\star,j)} - G(v_s)}{h(N)}\right)\right\}\\ &\times \phi_{\boldsymbol{n}',\boldsymbol{a}}\left(v_1,\ldots,v_{|A^c|}\,\middle|\, u_1,\ldots,u_{|A|}\right)\mathrm{d}v_1\cdots\mathrm{d}v_{|A^c|} \end{aligned}$$

For sufficiently large T such that $|r(n)| < (1/(p-1))\wedge(e(|A|-1))^{-1}$, $\forall n \geqslant T$, Lemma 3.3 and 3.4 in Taqqu's (1977) paper ensure

$$\begin{aligned} J(\boldsymbol{t},\boldsymbol{l},A,T) &= \sum_{q=1}^{\infty} \sum_{\substack{k_1+\cdots+k_s=2q\\ 0\leqslant k_1,\ldots,k_s\leqslant q}} \sum_{\substack{a=(n_{i_1},\ldots,n_{i_{|A|}})\\ \in N_T(|A|)}} EH_{k_1}\left(X_{n_{i_1}}\right)\cdots H_{k_{|A|}}\left(X_{n_{i_{|A|}}}\right)\\ &\times \left\{ \prod_{s=1}^{|A|} \int \frac{1}{k_s!}\left[\prod_{j=1}^{t_{i_s}} b_{l(i_s,j)}\tilde{K}\left(\frac{y_{l(i_s,j)} - G(u_s)}{h(N)}\right)\right] H_{k_s}(u_s)\phi(u_s)\,\mathrm{d}u_s\right\}\\ &\times \int \sum_{\boldsymbol{n}'\in N_{T,A}(|\boldsymbol{t}|)/\boldsymbol{a}} \left\{ \prod_{s=1}^{|A^c|}\prod_{j=1}^{t_{i_s^\star}} b_{l(i_s^\star,j)} \tilde{K}\left(\frac{y_{l(i_s^\star,j)} - G(v_s)}{h(N)}\right)\right\}\\ &\times \phi_{\boldsymbol{n}',\boldsymbol{a}}\left(v_1,\ldots,v_{|A^c|}\mid u_1,\ldots,u_{|A|}\right)\mathrm{d}v_1\cdots\mathrm{d}v_{|A^c|}. \end{aligned} \tag{3.35}$$

Fix s and k_s, $1 \leqslant s \leqslant |A|$, the first integral on the right-hand side of (3.35) is bounded by

$$\begin{aligned} &\int \prod_{j=1}^{t_{i_s}} \left| b_{l(i_s,j)}\tilde{K}\left(\frac{y_{l(i_s,j)} - G(u_s)}{h(N)}\right)\right| |H_{k_s}(u_s)\phi(u_s)|\,\mathrm{d}u_s\\ &\leqslant Ch(N)\sqrt{k_s!}(k_s+1)^{1/4}, \end{aligned} \tag{3.36}$$

which follows by arguments similar to that used to derive (3.12) and (3.13). (The product $\prod_{j=1}^{t_{i_s}} b_{l(i_s,j)}\tilde{K}(\cdot)$, instead of a single term $K(\cdot)$, inside the integral in (3.36) does not make much difference.) For fixed $\boldsymbol{a} \in N_T(|A|)$ and $N_{T,A}(|\boldsymbol{t}|)/\boldsymbol{a}$, we now try to bound $|N_{T,A}(|\boldsymbol{t}|)/\boldsymbol{a}|$. Given any decomposition of A^c

$$A^c = B_0 + \sum_{i=1}^{I} B_i$$

with $|B_i| = 0$ or $\geqslant 2$,$\forall 1 \leqslant i \leqslant I$ (set $I \equiv 0$ if $B_0 = A^c$), we define $\mathscr{D}(B_0, B_1, \ldots, B_I) \subset N_{T,A}(|\boldsymbol{t}|)/\boldsymbol{a}$ as the collection of all the $\boldsymbol{n} = (n_1, \ldots, n_{|\boldsymbol{t}|}) \in N_{T,A}(|\boldsymbol{t}|)/\boldsymbol{a}$ such that

(i) For each $b \in B_0, \exists j \in A \ni |n_b - n_j| \leqslant T$, and $|n_b - n_a| > T \;\forall a \in A^c - \{b\}$.

(ii) For each $b \in B_i, 1 \leqslant i \leqslant I, \exists b' \in B_i \ni b' \neq b$ and $|n_b - n_{b'}| \leqslant T$.

(iii) For any b and b' such that $b \in B_i$ and $b' \in B_j$ with $i \neq j, 1 \leqslant i, j, \leqslant I$, then $|n_b - n_{b'}| > T$.

Clearly,

$$|\mathscr{D}(B_0, B_1, \ldots, B_I)| \leqslant CN^I,$$

and for each $\boldsymbol{n} \in N_{T,A}(|\boldsymbol{t}|)/\boldsymbol{a}$ there exists a decomposition $(B'_0, B'_1, \ldots, B'_{I'})$ of A^c such that $\boldsymbol{n} \in \mathscr{D}(B'_0, B'_1, \ldots, B'_{I'})$. Therefore,

$$|N_{T,A}(|\boldsymbol{t}|)/\boldsymbol{a}| \leqslant CN^{[|A^c|/2]}, \tag{3.37}$$

since $I \leqslant [|A^c|/2]$ and the total number of decompositions of A^c depends only on $|A^c|$ and is thus bounded for fixed p. The last constant C can be made to be independent of $\boldsymbol{a}$ and N. For $\boldsymbol{t} = (t_1, \ldots, t_{|\boldsymbol{t}|}) \in Q(p) - Q^\star(p)$ with $|\boldsymbol{t}| \geqslant p/2$, and $A = \{i_1, \ldots, i_{|A|}\}$, put

$$\begin{aligned}\tau &\equiv \text{number of 1's in } \{t_{i_1}, \ldots, t_{i_{|A|}}\} \subset \{t_1, \ldots, t_{|\boldsymbol{t}|}\} \\ &= \text{the least number of functions in } \left\{ \prod_{j=1}^{t_{i_s}} \tilde{K}\left(\frac{y_{l(i_s,j)} - G(u_s)}{h(N)}\right), 1 \leqslant s \leqslant |A| \right\} \\ &\quad \text{whose Hermite rank is no less than } k.\end{aligned}$$

Since $f(y_i) > 0, 1 \leqslant i \leqslant m$, all the inverse images of $G^{-1}(y_i), 1 \leqslant i \leqslant m$, are interior points of the support of $f(\cdot)$ and thus finite real numbers. Therefore, by (C2) and (C3), $\int K((y - G(x))/h(N))\mathrm{d}x = \mathrm{O}(h(N))$. This together with $EK((y - G(x))/h(N)) = \mathrm{O}(h(N))$ and (3.35)–(3.37) imply

$$\begin{aligned}&|J(\boldsymbol{t}, \boldsymbol{l}, A, T)| \\ &\leqslant Ch^{|\boldsymbol{t}|}(N)N^{[|A^c|/2]} \sum_{q=k\tau-[k\tau/2]}^{\infty} \sum_{\substack{k_1+\cdots+k_{|A|}=2q \\ 0 \leqslant k_1, \ldots, k_{|A|} \leqslant q}} \frac{[(k_1+1)\cdots(k_{|A|}+1)]^{1/4}}{\sqrt{k_1!}\cdots\sqrt{k_{|A|}!}} \\ &\times \left(\sum_{\boldsymbol{a} \in N_T(|A|)} \left| EH_{k_1}\left(X_{n_{i_1}}\right) \cdots H_{k_{|A|}}\left(X_{n_{i_{|A|}}}\right) \right| \right),\end{aligned} \tag{3.38}$$

where the last sum is bounded by (Taqqu, 1977, Lemma 4.5 and its proof)

$$\sum_{\boldsymbol{a}\in N_T(|A|)} \left|EH_{k_1}\left(X_{n_{i_1}}\right)\cdots H_{k_{|A|}}\left(X_{n_{i_{|A|}}}\right)\right|$$
$$\leqslant C\varepsilon^{\frac{(k_1+\cdots+k_{|A|})}{2}} k_1!\cdots k_{|A|}!\left|Z(k_1,\ldots,k_{|A|})\right| N^{|A|-\tau/2}(|R|(k,N))^{\tau/2}. \tag{3.39}$$

The positive number ε is such that $|r(T)|\leqslant \varepsilon < (1/(p-1))\wedge(e(|A|-1))^{-1}$, and the quantity $Z(k_1,\ldots,k_{|A|})$ satisfies (Taqqu, 1977, Corollary 4.2)

$$EH_{k_1}\left(X_{n_{i_1}}\right)\cdots H_{k_{|A|}}\left(X_{n_{i_{|A|}}}\right) = k_1!\cdots k_{|A|}!Z(k_1,\ldots,k_{|A|})$$

with X standard Gaussian random variable. Combining

$$\left|Z(k_1,\ldots,k_{|A|})\right| \leqslant \prod_{s=1}^{|A|}\frac{(|A|-1)^{k_s/2}}{\sqrt{k_s!}}$$

(Taqqu, 1977, Lemma 3.1) with (3.38) and (3.39), we get by noting $(k_i+1)^{1/2k_i}\leqslant e$

$$|J(\boldsymbol{t},\boldsymbol{l},A,T)| \leqslant C\,(h(N))^{|\boldsymbol{t}|}\,N^{|A|-\tau/2+[\frac{|A^c|}{2}]}|R|^{\tau/2}(k,N)\left(\sum_{j=0}^{\infty}(e(|A|-1)\varepsilon)^{j/2}\right)^{|A|},$$

which gives

$$(Nh(N))^{-p/2}\,|J(\boldsymbol{t},\boldsymbol{l},A,T)|$$
$$\leqslant C(h(N))^{|A^c|/2}(h(N)|R|(k,N))^{\tau/2}(Nh(N))^{|\boldsymbol{t}|-(p/2)-(|A^c|+\tau)/2} \tag{3.40}$$

as $|A|$ is replaced by $|\boldsymbol{t}|-|A^c|$. Denote by τ' the number of 1's in the coordinates of $\boldsymbol{t}=(t_1,\ldots,t_{|\boldsymbol{t}|})$. Observe that

$$|A^c|+\tau\geqslant\tau',$$
$$p=\sum_{i=1}^{|\boldsymbol{t}|}t_i\geqslant 1\cdot\tau'+2(|\boldsymbol{t}|-\tau')=2|\boldsymbol{t}|-\tau',$$

and hence

$$\frac{|A^c|+\tau}{2}\geqslant|\boldsymbol{t}|-p/2.$$

In turn, this implies by (3.40) and the assumption $Nh(N)\to\infty$,

$$(Nh(N))^{-p/2}\,|J(\boldsymbol{t},\boldsymbol{l},A,T)| \leqslant C(h(N))^{|A^c|/2}\,(h(N)|R|(k,N))^{\tau/2}$$
$$\leqslant C(h(N))^{|A^c|/2}\,(h(N)|R|(1,N))^{\tau/2}. \tag{3.41}$$

Viewing (3.41), we then have, by (2.8) or condition $h(N)|R|(1,N)=\mathrm{o}(1)$,

$$(Nh(N))^{-p/2}\,|J(\boldsymbol{t},\boldsymbol{l},A,T)|\to 0 \text{ as } N\to\infty,$$

$\forall\, A\in\mathscr{F}(|\boldsymbol{t}|)$ with $|A|\geqslant 2$, since $|A^c|+\tau\geqslant 1$ always holds if $\boldsymbol{t}\in Q(p)-Q^{\star}(p)$ with $\boldsymbol{t}\geqslant p/2$. The proof is completed. □

Proof of Theorem 2. Let us consider the finite linear combination $\sum_{i=1}^{m} b_i Z_{Nh(N)}(y_i)$ and write

$$\sum_{i=1}^{m} b_i Z_{Nh(N)}(y_i) = \left(\sum_{i=1}^{m} \frac{b_i a_{kh(N)}(y_i)}{k!h(N)}\right)\left(\frac{1}{N}\sum_{n=1}^{N} H_k(X_n)\right) + \sum_{j=k+1}^{\infty}\left(\sum_{i=1}^{m} \frac{b_i a_{jh(N)}(y_i)}{j!h(N)}\right)\left(\frac{1}{N}\sum_{n=1}^{N} H_j(X_n)\right) \equiv V_N + V_N'.$$

We, for convenience, denote by A_N the norming factor which is understood to be of the form as specified in Cases (a) and (b), i.e.,

$$A_N = \begin{cases} \sqrt{N}|R|^{-1/2}(k,N) & \text{for Case (a)}, \\ N^{k\alpha/2}L^{-k/2}(N) & \text{for Case (b)}. \end{cases}$$

Under Case (a), we have, by (2.4) and Theorem 1' of Breuer and Major (1983),

$$A_N V_N \xrightarrow{\mathrm{d}} \left(\sum_{i=1}^{m} \frac{b_i g_k(y_i)}{k!}\right) \mathrm{N}(0,\sigma^2)$$

with

$$\sigma^2 = k! \lim_{N\to\infty} (N|R|(k,N))^{-1} \sum_{i,j=1}^{N} r^k(i-j).$$

In Case (b), (2.4) and Theorem 1 of Dobrushin and Major (1979) imply

$$A_N V_N \xrightarrow{\mathrm{d}} \left(\sum_{i=1}^{m} \frac{b_i g_k(y_i)}{k!}\right) Z$$

with Z as specified in (2.16). Furthermore, (3.16) and (3.19) individually assure

$$\lim_{N\to\infty} E(A_N V_N')^2 = 0.$$

This shows (2.15) after the Cramér–Wold device is employed.

Acknowledgements

The author is very grateful to the referee for careful reading the manuscript and many valuable suggestions which substantially improve the quality of the paper.

References

J. Beran, Statistical methods for data with long-range dependence, Statist. Sci. 7 (1992) 404–420.

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

P. Breuer and P. Major, Central limit theorems for nonlinear functionals of Gaussian fields, J. Multivariate Anal. 13 (1983) 144–155.

J.V. Castellana and M.R. Leadbetter, On smoothed probability density estimation for stationary processes, Stochastic Processes Appl. 21 (1966) 179–193.

K.C. Chanda, Density estimation for linear process, Ann. Inst. Statist. Math. 35 (1983) 439–446.

B. Cheng and P.M. Robinson, Density estimation in strongly dependent non-linear time series, Statist. Sinica 1 (1991) 335–359.

H. Dehling and M.S. Taqqu, The empirical process of some long-range dependent sequences with an application to U-statistics, Ann. Statist. 17 (1989) 1767–1783.

R.L. Dobrushin and P. Major, Non-central limit theorems for nonlinear functionals of Gaussian fields, Z. Wahrsch. Verw. Gebiete 50 (1979) 27–52.

W. Feller, An Introduction to Probability Theory and its Application, Vol. 2 (Wiley, New York, 2nd ed., 1971).

L. Györfi, W. Härdle, P. Sarda and P. Vieu, Nonparametric Curve Estimation from Time Series, Lecture Notes in Statistics, Vol. 60 (Springer, New York, Berlin, 1989).

J.D. Hart, Efficiency of a kernel estimator under autoregressive dependence model, J. Amer. Statist. Assoc. 79 (1984) 110–117.

J.D. Hart and P. Hall, Convergence rates in density estimation for data from infinite-order moving average processes, Probab. Theory Related Fields 87 (1990) 253–274.

E. Lukacs, Characteristic Functions, (Griffin, London, 2nd ed., 1970).

P. Major, Multiple Wiener–Itô Integrals, Lecture Notes in Math, Vol. 849 (Springer, New York, 1981).

P.M. Robinson, Nonparametric estimators for time series, J. Time Ser. Anal. 4 (1983) 185–207.

P.M. Robinson, Time series with strong dependence, in: Adv. Econometrics, 6th World Congress (Cambridge Univ. Press, London, 1990).

P.M. Robinson, Nonparametric function estimates for long-memory time series, in: W. Barnett, J. Powell and G. Tauchen, eds., Nonparametric and Semiparametric Methods in Econometrics and Statistics (Cambridge Univ. Press, Cambridge, 1991) pp. 437–457.

M. Rosenblatt, Independence and dependence, Proc. 4th Berkeley Symp. on Math. Statis. Probab. (Univ. California Press, Berkeley, CA, 1961) pp. 431–443.

M. Rosenblatt, Density estimates and Markov sequences, in: M. Puri, ed., Non-parametric Techniques in Statistical Inference (Cambridge Univ. Press, London, 1970) pp. 199–210.

M. Rosenblatt, Stochastic Curve Estimation (Institute of Mathematical Statististics, Harvard, CA, 1991).

G.G. Roussas, Nonparametric estimation in Markov process, Ann. Inst. Statist. Math. 21 (1969) 73–87.

B.W. Silverman, Density Estimation for Statistics and Data Analysis (Chapman and Hall, London, 1986).

M.S. Taqqu, Law of the iterated logarithm for sums of non-linear functions of Gaussian variables that exhibit a long range dependence, Z. Wahrsch. Verw. Gebiete 40 (1977) 203–238.

M.S. Taqqu, Convergence of integrated processes of arbitrary Hermite rank, Z. Wahrsch. Verw. Gebiete 50 (1979) 53–83.