STRONG CONVERGENCE OF KERNEL ESTIMATORS FOR PRODUCT DENSITIES OF ABSOLUTELY REGULAR POINT PROCESSES

LOTHAR HEINRICH\textsuperscript{a} and ECKHARD LIEBSCHER\textsuperscript{b}

\textsuperscript{a}Freiberg University of Mining and Technology, Institute of Stochastics, Bernhard-von-Cotta-Str. 2, D-09596 Freiberg, Germany; \textsuperscript{b}Technical University of Ilmenau, Institute of Mathematics, P.O. Box 327, D-98684 Ilmenau, Germany

(Received 13 December 1996; In final form 21 January 1997)

In this paper we prove almost sure convergence of kernel-type estimators of second-order product densities for stationary absolutely regular (\(\beta\)-mixing) point processes in \(\mathbb{R}^d\). This type of mixing condition can be verified for various classes of point processes under mild additional assumptions. We also obtain rates of convergence which mainly depend on the decay of the mixing coefficient and the choice of the bandwidth. In case of motion-invariant processes the behaviour of kernel estimators of the pair correlation function is considered separately. The results are applied to kernel-type renewal density estimators.

Keywords: Second-order point process; product density; pair correlation function; kernel-type estimator; absolute regularity; large deviation inequalities; estimation variance; rates of convergence; uniform strong convergence

AMS 1991 subject classifications: Primary: 60G55; Secondary: 62G20

1. INTRODUCTION

Point processes are suitable models for the description of randomly or irregularly distributed point-like objects. Their statistical analysis is mostly based on a single observation of a point pattern in a (sufficiently large, convex, compact) sampling window and consists in estimating the first and second moment measures. The crucial point
of this second-order analysis consists in assuming strict (or at least second-order) stationarity of the underlying point process. The translation invariance of the distribution (or at least of the first and second moment measure) reduces our statistical problem to the estimation of the intensity $\lambda$ (= mean number of points per unit area) and the reduced second moment measure $\alpha^{(2)}_{\text{red}}$ or its Lebesgue density function (called product density $\varrho^{(2)}$). The statistical analysis simplifies still further when isotropy is added to stationarity. In this case the measure $\alpha^{(2)}_{\text{red}}$ is completely determined by the reduced second moment function $K(r) = \alpha^{(2)}_{\text{red}}(b(o,r))/\lambda$, where $b(y,r)$ denotes the closed ball with radius $r \geq 0$ centred at $y \in \mathbb{R}^d$, and $\varrho^{(2)}(x)$ (which only depends on the Euclidean norm $||x||$) is replaced by the pair correlation function $g(r) = \varrho^{(2)}(x)/\lambda$ for $x \in \mathbb{R}^d$ with $||x|| = r$, satisfying $K(r) = d\omega_d \int_0^r g(s) s^{d-1}ds$, where $\omega_d = \nu_d(b(o,1))$.

For the sake of definiteness we recall some facts from point process theory; for further details and more about statistics of point processes the reader is referred to the monographs by Daley and Vere-Jones [3], Ripley [22], Karr [17] and Stoyan, Kendall and Mecke [24].

Let $\mathcal{B}^d (\mathcal{B}_d^d)$ be the $\sigma$-field (ring) of (bounded) Borel sets of $\mathbb{R}^d$ and $\nu_d$ denotes the $d$-dimensional Lebesgue measure. A point process $\Psi = \sum_{i \geq 1} \delta_{X_i}$ is defined to be a random locally finite counting measure on $[\mathbb{R}^d, \mathcal{B}^d]$ over certain hypothetical probability space $[\Omega, \mathcal{F}, \mathbb{P}]$. Throughout in this paper we assume that $\Psi$ is simple (i.e. $\mathbb{P}(\Psi(\{x\}) = 1$ for all $x \in \mathbb{R}^d = 1$) and strictly stationary (i.e. $\Psi(B + x)$ and $\Psi(B)$ have the same distribution for any $x \in \mathbb{R}^d$ and $B \in \mathcal{B}_d^d$). This means that the random closed (discrete) point set $s(\Psi) = \{X_i, i \geq 1\}$ determines $\Psi$ completely. For any $E \in \mathcal{B}^d$, let $\mathcal{F}_\Psi(E)$ denote the smallest sub-$\sigma$-field of $\mathcal{F}$ that contains all events of the form $\{\Psi(E \cap B) = k\}$ for $B \in \mathcal{B}_d^d$ and $k = 1, 2, \ldots$

We are now in a position to formulate a mixing condition $\Psi = \sum_{i \geq 1} \delta_{X_i}$ which is based on the following absolute regularity (or $\beta$-mixing) coefficient between any two sub-$\sigma$-fields $\mathcal{X}, \mathcal{Y} \subseteq \mathcal{F}$, see Volkonskii and Rozanov [27]:

$$
\beta(\mathcal{X}, \mathcal{Y}) = V(P_{\mathcal{X} \cap \mathcal{Y}} - P_{\mathcal{X}} \times P_{\mathcal{Y}})
= \frac{1}{2} \sup \sum_i \sum_j |P(A_i \cap B_j) - P(A_i)P(B_j)|.
$$
Here \( \mathcal{V}(\cdot) \) designates the total variation norm and the supremum is taken over all pairs of finite partitions \( \{A_1, \ldots, A_J\} \) and \( \{B_1, \ldots, B_J\} \) of \( \Omega \) such that \( A_i \in \mathcal{F}, i = 1 \ldots I \) and \( B_j \in \mathcal{F}, j = 1 \ldots J \). Define \( E_s := [-s, s]^d \) and \( E^c_i := \mathbb{R}^d \setminus E_i \).

**Condition \( \beta \)** Assume that the point process \( \Psi = \sum_{i \geq 1} \delta_{X_i} \) admits the estimate

\[
\beta(\overline{\delta}_\Psi(E_s), \overline{\delta}_\Psi(E^c_{s+t})) \leq \left( \frac{t \wedge s}{t} \right)^{d-1} \beta_\Psi(t) \text{ for all } s, t \geq 1,
\]

where \( \beta_\Psi(\cdot) \) is a non-increasing function — the \( \beta \)-mixing rate of \( \Psi \) — satisfying \( \lim_{t \to \infty} \beta_\Psi(t) = 0 \). \( \Psi \) is called \( m \)-dependent, if \( \beta_\Psi(t_0) = 0 \) for some \( t_0 \in [0, \infty) \).

Condition \( \beta \) can be verified for quite a few classes of point processes under mild additional assumptions. For example, let \( \Psi(\cdot) := \sum_{i \geq 1} (\Psi^{(i)}(\cdot) - X_i) \) be a homogeneous cluster process generated by a stationary process \( \Phi = \sum_{i \geq 1} \delta_{X_i} \) of cluster centres with intensity \( \lambda \) and a sequence \( \Psi^{(i)}_c, i \geq 1 \), of i.i.d. copies of a typical cluster \( \Psi_c \) having a \( \mathbb{P} \)-a.s. finite cluster radius

\[
R_c = \sup \{ \|x\| : \Psi_c(\{x\}) > 0 \}.
\]

In Heinrich [10] (see also Heinrich and Molchanov [11] for a generalization) the following estimate is derived:

\[
\beta(\overline{\delta}_\Psi(E_s), \overline{\delta}_\Psi(E^c_{s+t})) \leq \beta(\overline{\delta}_\Phi(E^c_{s+1/4}), \overline{\delta}_\Phi(E^c_{s+3t/4}))
\]

\[
+ \lambda d 2^{d+1} \left( \frac{4s}{t} + 1 \right)^{d-1} + \left( \frac{12s}{t} + 1 \right)^{d-1} \mathbb{E} R^d_c 1(R_c \geq t/4).
\]

In the special case of a **Poisson cluster process**, that is, \( \Phi \) is a stationary Poisson process with intensity \( \lambda \), we have

\[
\beta_\Psi(t) \leq c_1 \mathbb{E} R^d_c 1(R_c \geq t/4) \text{ with } c_1 = \lambda d 2^{d+1} (5^{d-1} + 13^{d-1}).
\]
Obviously, if in addition $P(R_c \leq t_0) = 1$, the Poisson cluster process is $m$-dependent. Similar estimates of the $\beta$-mixing rate are known for some classes of

(i) dependently thinned (Poisson) point processes (including hard- and soft-core processes as defined by B. Matérn and by D. & H. Stoyan, respectively, see [25]),

(ii) point processes associated with germ-grain models (e.g. positive and negative tangent points), see Heinrich and Molchanov [11],

(iii) point processes generated by a (Poisson-) Voronoi tessellation of $\mathbb{R}^d$ (e.g. vertices, midpoint of edges), see Heinrich [10].

(iv) Gibbsian point processes under Dobrushin’s uniqueness conditions, see e.g. Jensen [12] and Heinrich [9].

Condition $\beta$ with at least polynomial $\beta$-mixing rate is needed for proving asymptotic normality of empirical characteristics in point process statistics provided the sampling window expands unboundedly. Of course, Condition $\beta$ implies the ergodicity of the point process $\Psi$ which in turn is sufficient to prove strong consistency of the empirical (reduced) moment measures, see Karr [17]. By contrast, ergodicity does not suffice to obtain (neither local nor uniform) consistency of kernel estimators of the corresponding product densities.

At the end of the first section we recall the definitions of factorial moment measures and the product densities of any order. Provided that $E\psi^k(B) < \infty$ for all $B \in \mathcal{B}_0$, the $k$th-order factorial moment measure $\alpha^{(k)}$ is defined by

$$\alpha^{(k)}(B_1 \times \ldots \times B_k) = E \sum_{x_1, \ldots, x_k \in s(\Psi)} 1_{B_1}(x_1) \ldots 1_{B_k}(x_k)$$

for $B_1, \ldots, B_k \in \mathcal{B}_0$.

Here the sum $\sum$ is taken over all $k$-tuples of pairwise distinct points $x_1, \ldots, x_k \in s(\Psi)$. The stationarity assumption yields $\alpha^{(1)}(.) = \lambda \nu_d(.)$ and admits the desintegration of $\alpha^{(k)}$ for $k \geq 2$ w.r.t. $\nu_d$,

$$\alpha^{(k)}(B_1 \times \ldots \times B_k) = \lambda \int_{B_k} \alpha^{(k)}_{\text{red}} ((B_1 - x) \times \ldots \times (B_{k-1} - x)) \nu_d(dx),$$
where $\alpha^{(k)}_{\text{red}}$ is called the reduced $k$th-order factorial moment measure on $[\mathbb{R} d^{k-1}], \mathcal{B} d^{k-1}]$. If the measure $\alpha^{(k)}_{\text{red}}$ is absolutely continuous w.r.t. $\nu_{d^{k-1}}$, then the corresponding Radon-Nikodym density $\varrho^{(k)}(\cdot)$ on $[\mathbb{R} d^{k-1}]$ is called the (reduced) $k$th-order product density of $\Psi$. Note that for a ($\beta^{-}$) mixing point process the generalized Blackwell theorem holds which implies $\varrho^{(2)}(x) \to \lambda$ as $\|x\| \to \infty$, see [3], p. 488.

In analogy to Parzen-Rosenblatt probability density estimators in classical i.i.d. statistics, see e.g. [23], kernel product density estimators were proposed for “increasing domain statistics” of point processes by D. R. Brillinger [1] (for $d = 1$) and K. Krickeberg [18] for any $d \geq 1$. Under the basic assumption that the stationary point process is Brillinger-mixing, see [17], [7], E. Jolivet [13] derived asymptotically exact bounds of the bias of a kernel estimator for $\varrho^{(k)}$ and of all moments of the absolute difference between $\varrho^{(k)}$ and its kernel estimator. Edge-corrected kernel estimators of $\varrho^{(2)}(x)$ and $g(r)$ on a fixed sampling domain were considered by Fiksel [4], see [26] for a recent discussion. Asymptotic normality of kernel product density estimators could be proved for Brillinger-mixing point processes, see [14], and for Poisson cluster processes under minimal moment assumptions, see [6], [7]. The main goal of the present paper is to assess the quality of the kernel estimators for product densities by proving rates of strong (P-a.s.) consistency under mild moment conditions and Condition $\beta$. Rates of strong (uniform) convergence have been already obtained in kernel regression estimation, see Collomb and Härdle [2], and in kernel density estimation in the case of $\alpha - \text{ resp. } \beta$-mixing observations, see Liebscher [16] resp. Yu [29].

The paper is organized as follows: Section 2 contains the main asymptotic results on kernel estimators for second-order product densities of stationary absolutely regular point processes. This concerns firstly the asymptotic behaviour of mean and variance of the estimator under certain smoothness conditions and a minimal $\beta$-mixing rate and, secondly, the speed of pointwise as well as of uniform strong convergence of under various rates of decay of $\beta_q(t)$ as $t \to \infty$. Corresponding results for two types of kernel estimators for the pair correlation functions of isotropic point processes are given in Section 3. At the end of Section 3, the obtained general results are applied to obtain the rate of convergence of a renewal density kernel estimator. Section 4 deals with large deviation inequalities of the Nagaev-Fuk
type for absolutely regular random fields. Section 5 resp. 6 contains
the proofs of the results stated in section 2 resp. 3. Throughout, let
c, c_1, c_2 \ldots denote general constants numbered according their
appearance in the text.

2. RESULTS

Throughout in this paper, let \( \Psi = \sum_{i \geq 1} \delta_{x_i} \) be a stationary (second-
order) point process on \( \mathbb{R}^d \) possessing a second-order product density \( g^{(2)} \). In the following we consider
a kernel-type estimator \( \hat{g}^{(2)}_n \) for \( \lambda g^{(2)} \)
which is defined on the cubic sampling window \( W_n = [0, n]^d \) by

\[
\hat{g}^{(2)}_n(x) = \frac{1}{\nu_d(W_n)b_n^d} \sum_{x_1, x_2 \in s(\Psi)} 1_{W_n}(x_1) k\left( \frac{x_2 - x_1 - x}{b_n} \right), \quad x \in \mathbb{R}^d,
\]

where \( k: \mathbb{R}^d \rightarrow \mathbb{R}^1 \) denotes the kernel function and the sequence of
bandwidths \( b_n > 0, n \in \mathbb{N} \), is chosen such that

\[
\lim_{n \to \infty} b_n = 0 \quad \text{and} \quad \lim_{n \to \infty} n b_n = \infty.
\]

Next we list a set of conditions which are needed to obtain (rates of)
local and uniform strong convergence of the estimator (2.1) to the true
function \( \lambda g^{(2)}(x) \) under the basic Condition \( \beta \).

Condition \( \mathcal{K}(d, p) \) Suppose that, for some \( 0 < \beta < \infty \),

\[
k(x) = 0 \quad \text{for} \quad \|x\| > R, \quad \sup_{\|x\| \leq R} |k(x)| < \infty, \quad \int_{b(0,R)} k(x) \, dx = 1.
\]

and, for \( i_1, \ldots, i_l \in \{1, \ldots, d\}, l = 1, \ldots, p - 1 \) (provided that \( p \geq 2 \)),

\[
\int_{b(0,R)} x_{i_1}, \ldots, x_{i_l} k(x_1, \ldots, x_d) \, d(x_1, \ldots, x_d) = 0.
\]
Examples The following two kernel functions satisfy Condition $\mathcal{K}(d,2)$, see e.g. [20], [23]:

\[
k(x) = \frac{d(d + 1)\Gamma(d/2)}{2\pi d/2 R^d} \left(1 - \frac{\|x\|}{R}\right) I_{b(o,R)}(x) : \text{"conus" kernel},
\]

\[
k(x) = \frac{d(d + 2)\Gamma(d/2)}{4\pi d/2 R^d} \left(1 - \frac{\|x\|^2}{R^2}\right) I_{b(o,R)}(x) : \text{Epanechnikov kernel}.
\]

Condition $\mathcal{M}(\gamma)$ Suppose that $\mathbb{E} \psi^\gamma([0,1]^d) < \infty$ for some $\gamma \geq 4$ and the product densities $\varrho^{(2)}$, $\varrho^{(3)}$, and $\varrho^{(4)}$ exist.

Condition $\mathcal{P}(x)$ The product density $\varrho^{(2)}$ is continuous at $x \in \mathbb{R}^d$ and, for some $\varepsilon > 0$,

\[
\sup_{u,v \in b(x,\varepsilon) \cup b(-x,\varepsilon)} \varrho^{(3)}(u,v) < \infty \quad \text{and} \quad \sup_{u,v \in b(x,\varepsilon), w \in \mathbb{R}^d} \varrho^{(4)}(u,w,v+w) < \infty.
\]

(2.5)

Condition $\mathcal{P}(K)$ The product density $\varrho^{(2)}$ is continuous on a compact set $K \subset \mathbb{R}^d$ and, for some $\varepsilon > 0$ with $K_\varepsilon = \cup_{x, \|x\| \leq \varepsilon}(K+x)$,

\[
\sup_{u,v \in K_\varepsilon \cup (-K_\varepsilon)} \varrho^{(3)}(u,v) < \infty \quad \text{and} \quad \sup_{u,v \in K_\varepsilon, w \in \mathbb{R}^d} \varrho^{(4)}(u,w,v+w) < \infty.
\]

(2.6)

Proposition 2.1. Under Condition $\mathcal{K}(d,1)$ we have

\[
\lim_{n \to \infty} \mathbb{E} \hat{\varrho}_n^{(2)}(x) = \lambda \varrho^{(2)}(x)
\]

at any continuity point $x$ of $\varrho^{(2)}$. If Condition $\mathcal{K}(d,p)$ is satisfied and $\varrho^{(2)}$ has bounded partial derivatives of order $p$ in $b(x, \varepsilon)$ (for some $\varepsilon > 0$), then

\[
\mathbb{E} \hat{\varrho}_n^{(2)}(x) = \lambda \varrho^{(2)}(x) + O(b_n^p) \quad \text{as} \quad n \to \infty.
\]

(2.7)

Theorem 2.1. Let the Conditions $\beta$, $\mathcal{K}(d,1)$, $\mathcal{P}(x)$ and Condition $\mathcal{M}(4+\delta)$ for some $\delta > 0$ be satisfied such that $\lim_{n \to \infty} n^{2d(4+\delta)/\delta} \beta_n(n) = 0$. Then, for $x \neq o$,

\[
\nabla^2 \hat{\varrho}_n^{(2)}(x) = (n b_n)^{-d} \left(\lambda \varrho^{(2)}(x) \int_{\mathbb{R}^d} k^2(y)dy + o(1)\right) \quad \text{as} \quad n \to \infty.
\]

(2.8)
Relation (2.8) remains valid for $x = 0$ with $k(y)^2(k(y) + k(-y))$ instead of the integrand $k^2(y)$. In case of $\beta(0) = 0$ for some $t_0 \geq 0$, (2.8) holds already for $\delta = 0$.

**Theorem 2.2.** Suppose that the Conditions $\beta$, $M(2\gamma)$, $K(d, p)$ and $P(x)$ are satisfied for some $\gamma > 2$ and some $p \in \mathbb{N}$. Further, assume that $g^{(2)}$ has bounded partial derivatives of order $p$ in $b(x, \varepsilon)$ for some $\varepsilon > 0$.

Case (i): If $\beta(t) = O(t^{-q})$ as $t \to \infty$ for some $q > 2d\gamma/(\gamma - 2)$, then, as $n \to \infty$,

$$
\hat{g}^{(2)}_n(x) - \lambda g^{(2)}(x) = O(\ln^{1/2} n \ (n b_n)^{-d/2} + \ln n \ n^{-\theta} b_n^{-d} + b_n^p) \ P-a.s.
$$

for any $\kappa_1 > (\gamma d + q)/(\gamma d + q)$, where $\theta = (q(\gamma d - d - 1) - \gamma d)/\gamma(d + q)$ (> 0).

Case (ii): If $\beta(t) = O(\exp\{-at^b\})$ as $t \to \infty$ for some $a, b > 0$, then, as $n \to \infty$,

$$
\hat{g}^{(2)}_n(x) - \lambda g^{(2)}(x) = O(\ln^{1/2} n \ (n b_n)^{-d/2} + \ln n \ n^{-(d\gamma - d - 1)/\gamma} b_n^{-d} + b_n^p) \ P-a.s.
$$

for any $\kappa_2 > (d\gamma - d + b)/b\gamma$.

Case (iii): If $\beta(0) = 0$ for some $t_0 \geq 0$, then, as $n \to \infty$,

$$
\hat{g}^{(2)}_n(x) - \lambda g^{(2)}(x) = O(\ln^{1/2} n \ (n b_n)^{-d/2} + \ln n \ n^{-(d\gamma - d - 1)/\gamma} b_n^{-d} + b_n^p) \ P-a.s.
$$

for any $\kappa_3 > 1/\gamma$. The latter relation remains valid for $\gamma = 2$.

By equating $b_n^p$ with each of the other two terms contributing to the rate of convergence we are able to determine the optimal bandwidth (up to a multiplicative constant) that leads to the best possible rates in the Cases (i) – (iii) considered in Theorem 2.2.
Corollary 2.1. Let the assumptions of Theorem 2.2 be satisfied. Then, as $n \to \infty$,
\[
\hat{\varrho}_n^{(2)}(x) - \lambda \varrho^{(2)}(x) = O(b_n^p) \quad \mathbb{P}\text{-a.s.},
\]
where the optimal bandwidths minimizing this rate of convergence are as follows:

\[
b_n = \begin{cases} 
\max\{\ln^{\kappa_1} n n^{-\theta}, (\ln n n^{-d})^{\frac{1}{d-\theta}}\} & \text{in Case (i)} \\
\max\{\ln^{\kappa_2} n n^{-\frac{d-1}{2\gamma_d}} (\ln n n^{-d})^{\frac{1}{d-\theta}}\} & \text{in Case (ii)} \\
\max\{\ln^{\kappa_3} n n^{-\frac{d-1}{2\gamma_d}} (\ln n n^{-d})^{\frac{1}{d-\theta}}\} & \text{in Case (iii)}.
\end{cases}
\]

The next theorem provides the corresponding rates of convergence of the uniform error
\[
\Delta_n(K) = \sup_{x \in K} |\hat{\varrho}_n^{(2)}(x) - \varrho^{(2)}(x)|.
\]

For this purpose a further smoothness condition on the kernel function is needed.

Condition $\mathcal{L}$  The kernel function $k$ is Lipschitz continuous, i.e. there is constant $L \geq 0$ such that
\[
|k(u) - k(v)| \leq L\|u - v\| \quad \text{for } u, v \in \mathbb{R}^d.
\]

Theorem 2.3. Suppose that the Conditions $\beta$, $\mathcal{M}(2\gamma)$, $\mathcal{K}(d, p)$, $\mathcal{L}$, and $\mathcal{P}(K)$ are satisfied for some $\gamma > 2$ and some $p \geq 1$. Further, assume that $\varrho^{(2)}$ has bounded partial derivatives of order $p$ on $K_\varepsilon$ for some $\varepsilon > 0$.

Case (i): If $\beta_{\varrho}(t) = O(t^{-q})$ as $t \to \infty$ for some $q > 2d\gamma/(\gamma - 2)$, then, as $n \to \infty$,
\[
\Delta_n(K) = O(\ln^{1/2} n (n b_n)^{-d/2} + \ln^{\kappa_1} n n^{-\hat{\theta}/(d+\gamma)} b_n^{-(d(1+\gamma))/(d+\gamma)} + b_n^p) \quad \mathbb{P}\text{-a.s.}
\]
for any $\kappa_1 > (q\gamma + d + q)/(d + \gamma)(d + q)$, where $\hat{\theta} = (q(\gamma d - d - 1) - \gamma d)/(d + q)(> 0)$.

Case (ii)  If $\beta_{\varrho}(t) = O(\exp\{-at\})$ as $t \to \infty$ for some $a, b > 0$, then, as $n \to \infty$,
\[
\Delta_n(K) = O(\ln^{1/2} n (n b_n)^{-d/2} + \ln^{\kappa_2} n n^{-(d\gamma - d - b)/(d+\gamma)} b_n^{-(d(1+\gamma))/(d+\gamma)} + b_n^p) \quad \mathbb{P}\text{-a.s.}
\]
for any $\kappa_2 > (d\gamma - d + b)/b(d + \gamma)$. 
Case (iii) If $\beta_0(t_0) = 0$ for some $t_0 \geq 0$, then, as $n \to \infty$,

$$\Delta_n(K) = O\left(\frac{\ln^{1/2}n (n b_n)^{-d/2}}{\ln^{1/2}n (n b_n)^{-d/2}} + \ln^{1/2}n n^{-(d/(d-1)+1/(d+\gamma))} b_n^{-d(1+\gamma)/(d+\gamma)} + b_n^p\right) \ P\text{-a.s.}$$

for any $\tilde{n}_3 > 1/(d + \gamma)$. The latter relation is also valid for $\gamma = 2$.

In quite analogy to the above Corollary 2.1 we get the following

**Corollary 2.2.** Let the assumptions of Theorem 2.3 be satisfied. Then, as $n \to \infty$,

$$\sup_{x \in \mathbb{R}} |\hat{\varphi}_n^{(2)}(x) - \lambda \varphi^{(2)}(x)| = O(b_n^p) \ P\text{-a.s.},$$

where the optimal bandwidths minimizing this uniform rate of convergence are as follows:

$$b_n = \begin{cases} 
\max\{\ln^{(d+\gamma)}n n^{-\hat{b}}, (\ln n n^{-d})^{\frac{1}{d+\gamma}}\} & \text{in Case (i)} \\
\max\{\ln^{(2d+\gamma)}n n^{-(d-1)(\gamma-1)} (\ln n n^{-d})^{\frac{1}{d+\gamma}}\} & \text{in Case (ii)} \\
\max\{\ln^{(d+\gamma)}n n^{-(d-1)(\gamma-1)} (\ln n n^{-d})^{\frac{1}{d+\gamma}}\} & \text{in Case (iii)}.
\end{cases}$$

**Remark 1** For $d=1$ the assumption "$\varphi^{(2)}$ has bounded partial derivative of order $p$" in Prop. 2.1, Theorem 2.2 and 2.3 can be replaced by "$\varphi^{(2)}$ has a Lipschitz continuous $(p-1)$st-order derivative".

**Remark 2** The above corollaries reveal that in Theorem 2.2 resp. Theorem 2.3 in both Cases (ii) and (iii) the optimal bandwidth is given by $b_n = (\ln n n^{-d})^{\frac{1}{2d+\gamma}}$ leading to

$$\hat{\varphi}_n^{(2)}(x) - \lambda \varphi^{(2)}(x) = O((\ln n n^{-d})^{\frac{1}{2d+\gamma}}) \ \text{resp.} \ \Delta_n(K) = O((\ln n n^{-d})^{\frac{1}{2d+\gamma}})$$

provided that $\gamma > 2 + \frac{2 + d}{d+1}$ resp. $\gamma > d + 2 + \frac{2d+1}{d}$ and $\gamma$ is large enough.

### 3. Kernel Estimators for the Pair Correlation Function and the Renewal Density

First in this section we relate the results of the previous section to motion-invariant (i.e. stationary and isotropic) point processes...
observed in $W_n$. As mentioned in Sect. 1 in this case statistical second-order analysis is based on the estimation of the pair correlation function $g(r), r \geq 0$. Two appropriate (univariate) kernel estimators $\hat{g}_n(r)$ and $\tilde{g}_n(r)$ for $\lambda^2 g(r)$ are the following:

$$\hat{g}_n(r) = \frac{1}{d \omega_d \nu_d(W_n) b_n r^{d-1}} \sum_{x_1, x_2 \in \mathcal{S}(\Psi)} \mathbf{1}_{W_n}(x_1) k\left(\frac{\|x_2 - x_1\| - r}{b_n}\right) \text{for } r > 0$$

(3.1)

and

$$\tilde{g}_n(r) = \frac{1}{d \omega_d \nu_d(W_n) b_n} \sum_{x_1, x_2 \in \mathcal{S}(\Psi)} \mathbf{1}_{W_n}(x_1) k\left(\frac{\|x_2 - x_1\| - r}{b_n}\right) \text{for } r \geq 0,$$

(3.2)

where $k: \mathbb{R}^d \mapsto \mathbb{R}^d$ is an univariate kernel function satisfying $K(1, 1)$ and $b_n$ obeys (2.2). In [4] and [26] the univariate Epanechnikov kernel (compare the examples in Section 2 for $d = 1$) is recommended.

**Proposition 3.1.** Let $k$ be a symmetric kernel function satisfying $K(1, 1)$. If the one-sided limits $g(r - 0)$ and $g(r + 0)$ exist for some $r > 0$, then

$$\lim_{n \to \infty} E\hat{g}_n(r) = \lim_{n \to \infty} E\tilde{g}_n(r) = \frac{\lambda^2}{2}(g(r + 0) + g(r - 0)).$$

(3.3)

If Condition $K(1, p)$ is satisfied and $g$ has a Lipschitz-continuous $(p - 1)$st-order derivative $g^{(p-1)}$ on $[r_1 - \varepsilon, r_2 + \varepsilon]$ for $0 < r_1 < r_2 < \infty$ and some $\varepsilon > 0$, then

$$E\hat{g}_n(r) = \lambda^2 g(r) + (1 + r^{-(d-1)}) \left(\max_{0 \leq k \leq p-1} |g^{(k)}(r)| + L_{p-1}\right) O(b_n^p)$$

(3.4)

and

$$E\tilde{g}_n(r) = \lambda^2 g(r) + O(b_n^p) \text{ as } n \to \infty,$$

(3.5)

where $L_{p-1}$ denotes the Lipschitz constant of $g^{(p-1)}$. The rate of convergence $O(b_n^p)$ in (3.4) and (3.5) is uniform in $r \in [r_1, r_2]$. 
Next we formulate the counterparts of the Theorem 2.1 and 2.3 for the empirical pair correlation function \( \hat{g}_n(r) \) and \( \hat{g}_n^*(r) \) in a somewhat abridged form. The following two theorems provide the behaviour of the variance \( \Delta^2 \hat{g}_n(r) \) and give the speeds of convergence of the uniform error

\[
\delta_n(r_1, r_2) = \sup_{r_1 \leq r \leq r_2} |\hat{g}_n(r) - \lambda^2 g(r)|
\]

together with the optimal choice of \( b_n \) for the Cases (i) – (iii) considered in the Theorems 2.2 and 2.3. We only mention that the Theorem 3.1 and 3.2 remain in force verbatim when \( \hat{g}_n(r) \) is replaced by \( \hat{g}_n^*(r) \). Only the constants hidden behind the symbols “\( o \)” and “\( O \)” will change.

**Theorem 3.1.** Let the kernel function \( k: \mathbb{R}^1 \rightarrow \mathbb{R}^1 \) satisfy Condition \( K(1, 1) \). Further, let the Conditions \( \beta, \mathcal{P}(x) \) and \( \mathcal{M}(4 + \delta) \) for some \( \delta > 0 \) be satisfied such that \( \lim_{n \to \infty} n^{2(d+4+\delta)/\beta} \beta g(n) = 0 \). Then, for \( r = ||x|| > 0 \),

\[
\Delta^2 \hat{g}_n(r) = \frac{1}{n^d b_n} \left( \frac{\lambda^2 g(r)}{d\omega_d r^{d-1}} \int_{\mathbb{R}^1} k^2(y)dy + o(1) \right) \text{ as } n \to \infty.
\]

**Theorem 3.2.** Suppose that the Conditions \( \beta, \mathcal{M}(2\gamma), \mathcal{K}(1,p), \mathcal{L}, \mathcal{P}(K) \) are satisfied for some \( \gamma > 2, p \geq 1 \) and \( K = \{x \in \mathbb{R}^d : r_1 \leq ||x|| \leq r_2 \} \) with \( 0 < r_1 < r_2 < \infty \). In addition, assume that \( g \) has a Lipschitz-continuous \((p - 1)\)st-order derivative \( g^{(p-1)} \) on \([r_1 - \varepsilon, r_2 + \varepsilon]\) for some \( \varepsilon > 0 \). Then the best possible rate of convergence of \( \delta_n(r_1, r_2) \) is equal to \( O(b_n^p) \), where the corresponding optimal bandwidths \( b_n \) are as follows:

**Case (i)** If \( \beta g(t) = O(t^{-q}) \) as \( t \to \infty \) for some \( q > 2d\gamma/(\gamma - 2) \), then

\[
b_n = \max\{\ln n n^{-\theta}, (\ln n n^{-d})^{1/(q-1)}\} \quad \text{with} \quad \theta = \frac{q(\gamma d - d - 1) - \gamma d}{(1 + \gamma)(d + q)(p + 1)}
\]

for any \( \eta_1 > (\gamma d + q)/(1 + \gamma)(d + q)(p + 1) \).

**Case (ii)** If \( \beta g(t) = O(\exp\{-at^b\}) \) as \( t \to \infty \) for some \( a > 0, 0 < b \leq 2d \), then

\[
b_n = \max\{\ln n n^{-\theta}, (\ln n n^{-d})^{1/(q-1)}\}
\]

for any \( \eta_2 > (d\gamma - d + b)/b(1 + \gamma)(1 + p) \).
Case (iii) If $\beta(\tau_0) = 0$ for some $\tau_0 \geq 0$, then
\[
 b_n = \max\{\ln^{\eta_3} n n^{-\frac{d+1}{p+1}}, (\ln n n^{-d})^{\gamma+1}\}
\]
for any $\eta_3 > 1/(1 + \gamma)(1 + p)$. The latter relation is also valid for $\gamma = 2$.

In particular, we have
\[
\delta_n(r_1, r_2) = O((\ln n n^{-d})^{\frac{p}{p+1}}),
\]
in the Cases(ii) and (iii) whenever $\gamma > 3 + \frac{2}{d} + \frac{2}{p} + \frac{1}{dp}$. This convergence rate remains also valid in Case (i) if $q > d + 2 + (d + 1)/p$ and $\gamma$ is large enough.

Remark 3 In general, even if $d = 1$, the pair correlation function $g(r)$ has a discontinuity (of the first kind or a pole) at $r = 0$. For this reason the relations (3.4) and (3.5) of Prop. 3.1 and the Theorems 3.1–3.3 do not hold for $r_1 = 0$.

In the second part of this section we specify the foregoing results to stationary renewal processes. For this end, let $\xi_1, \xi_2, \xi_3, \ldots$ be independent non-negative random variables on $[0, \infty)$ with distribution functions $F(x) = P(\xi_i \leq x)$ for $i \geq 2$ and $P(\xi_1 \leq x) = \int_0^x (1 - F(y))dy/m$, where $m = \xi_2 = \int_0^\infty (1 - F(x))dx < \infty$. The partial sums $X_i = \xi_1 + \ldots + \xi_i, i \geq 1$, define a (stationary) simple point process $\Psi = \sum_{i \geq 1} \delta_{X_i}$ on $[0, \infty)$ - a so-called (stationary) renewal process associated with $F$, see [3] for details. It is well-known that the intensity of $\Psi$ equals $1/m$ and the Lebesgue density $h(x)$ of the measure $\alpha^{(2)}_{\text{red}}([0, \infty) \cap (\cdot))$ which coincides with the product density $g^{(2)}(x) = g(x)/m, x \geq 0$, is called (Palm) renewal density function and takes the form
\[
 h(x) = \sum_{k=1}^\infty f^{*k}(x) = f(x) + \int_0^x f(x-t)h(t)dt, \quad (3.6)
\]
where $f$ denotes the probability density of $F$ and $f^{*k}$ is the $k$-th convolution power of $f$.

Therefore, according to (3.1) and (3.2)
\[
 \hat{h}_n(x) = \frac{1}{2n b_n} \sum_{i,j \geq 1}^{\ast} 1_{[0,\infty]}(X_i)k \left( \frac{|X_i - X_j| - x}{b_n} \right)
\]
is an appropriate kernel estimator for $h(x)/m$. 


Theorem 3.3. Let the kernel function \( k: \mathbb{R}^1 \to \mathbb{R}^1 \) satisfy the Conditions \( L \) and \( K(1,p) \) for some \( p \geq 1 \), and let \( \int_0^\infty x^{s-1} (1 - F(x)) \, dx < \infty \) for some \( s > (5p + 2)/p \). Further, assume that the density of \( F \) has a Lipschitz-continuous \((p - 1)\)st-order derivative on \([0, a]\), \(0 < a < \infty\). Then, for \( b_n = (\ln n/n)^{(2p+1)} \) and \( 0 < r_1 < r_2 < a \),

\[
\sup_{r_1 \leq x \leq r_2} |\hat{h}_n(x) - h(x)/m| = O((\ln n/n)^{-p/(2p+1)}) \quad \text{P-a.s.} \quad (3.8)
\]

as \( n \to \infty \).

4. Tail Probabilities for Sums of Absolutely Regular Random Fields

Let \( \Psi \) be a stationary point process satisfying Condition \( \beta \). Define \( W_t := [0, t]^d \) for a real number \( t \geq 1 \). Further, let \( f_n: \mathbb{R}^d \to \mathbb{R}^1 \) be a sequence of measurable functions with uniformly bounded support, i.e., there exists a fixed number \( s > 0 \) such that \( f_n(y) = 0 \) for \( \|y\| > s \).

Define

\[
S_n^{(s)} = \sum_{x_1, x_2 \in s(\Psi)} 1_{W_t}(x_1) f_n(x_2 - x_1) \quad \text{for } n \in \mathbb{N}, s \geq 1.
\]

Obviously, the kernel estimator \( \hat{S}_n^{(2)}(x) \) coincides with \( S_n^{(n)} \) for \( f_n(y) = (nb_n)^{-d} k((y - x)/b_n) \) and \( s = \|x\| + R b_n \).

The following lemma provides an exponential inequality for the sum \( S_n^{(n)} \) similar to that obtained for sums of absolutely regular random sequences, see Yoshihara [28] and Yu [29].

Lemma 4.1. Let \( \Psi \) be a stationary point process satisfying Condition \( \beta \) \( \mathbb{E}(\Psi([0, 1]^d)^{2\gamma} < \infty \) for some \( \gamma \geq 2 \). Further, let \( m, n \), be positive integers such that \( 2m \leq n \) and put \( t = n/2m \). Then

\[
\mathbb{P}(|S_n^{(n)} - \mathbb{E} S_n^{(n)}| \geq \tau) \leq 2 \cdot \frac{1}{d+1} \exp \left\{ - \frac{c_1 \tau^2}{2^d D^2 S_n^{(t)}} \left( \frac{n}{t} \right)^{-d} \right\} 
+ c_2 2^{d+1} \tau^{-\gamma} \left( \frac{n}{t} \right)^{d} \mathbb{E} S_n^{(t)} - \mathbb{E} S_n^{(t)} \right\} 
+ \left( \frac{n}{t} \right)^{d} \beta_{\Psi}(t - 2s) \quad (4.1)
\]

for any \( \tau > 0 \), where \( c_1 = 2(\gamma + 2)^{-1} e^{-\gamma} \) and \( c_2 = 2(1 + 2/\gamma)^{\gamma} \).
For proving (4.1) we need an exponential inequality for sums of independent and not necessarily bounded random variables which goes back to Fuk and Nagaev [5], see also Petrov [21].

**Proposition 4.1.** Let $Y_1,\ldots,Y_n$ be independent random variables satisfying $\mathbb{E}|Y_i|^{\gamma} < \infty$, $i = 1,\ldots,n$, for some $\gamma \geq 2$. Then

$$
P\left(\left|\sum_{i=1}^{n}(Y_i - \mathbb{E}Y_i)\right| \geq \tau\right) \leq 2\exp\left\{-c_1\frac{\tau^2}{\sum_{i=1}^{n}D^2Y_i}\right\} + c_2\tau^{-\gamma}\sum_{i=1}^{n}\mathbb{E}|Y_i - \mathbb{E}Y_i|^\gamma$$

holds for all $\tau > 0$, where the constants $c_1, c_2$ as in Lemma 4.1.

Let $m$, $n$ and $t$ be chosen as in Lemma 4.1. We split $W_n$ into $2^dm^d$ cubes $W_t + 2z + y$, for $z \in I_m = \{0,1,\ldots,m-1\}^d$ and $y \in I_2 = \{0,1\}^d$.

Further, define the sets $E_{\Psi}^y(s) = \bigcup_{x:||x|| \leq s}(W_t + 2z + y + x)$ and the product measure $\tilde{P} = \times_{z \in I_m}P_{\Psi}(E_{\Psi}^y(s))$, where $P_{\Psi}$ stands for the restriction of $P$ to some $\sigma$-field $\mathcal{F}$

Put

$$S_n^{(y)} = \sum_{y \in I_2}Z_{n,y}, \quad Z_{n,y} = \sum_{z \in I_m}X_{n,z}^y, \quad X_{n,z}^y = \sum_{x_1,x_2 \in s(\Psi)}1_{W_t+2z+(x_1 + f_n(x_2-x_1)).}$$

**Proposition 4.2.** Assume that the stationary point process $\Psi$ satisfies Condition $\beta$. Then

$$\sup_{A \in \mathcal{B}, x \in I_m}P_{\Psi}(A) - P_{\Psi}(A) \leq (m^d - 1)\beta_{\Psi}(t - 2s) \quad (4.3)$$

for any $y \in I_2$ and $t \geq 6s$, $s \geq 1/4$.

**Proof of Proposition 4.2.** Let $E_{\Psi}^y(s) = \bigcup_{u \in I_m \setminus \{z\}}E_{\Psi}^u(s)$. Then, by the stationarity $\Psi$ and Condition $\beta$, we get

$$\sup_{A \in \mathcal{B}, x \in I_m}P_{\Psi}(A) - (P_{\Psi}(E_{\Psi}^y(s)) \times P_{\Psi}(E_{\Psi}(s)))(A)$$

$$\leq \beta(\tilde{\Psi}(E_{1/2-s}), \tilde{\Psi}(E_{3t/2-s})) \leq \beta_{\Psi}(t - 2s)$$
for \( t \geq 6s, s \geq 1/4 \) and all \( z \in I_m, y \in I_2 \), where \( \mathcal{W} = \mathcal{F}_Y(E^y_x(s)) \otimes \mathcal{F}_Y(E^z_x(s)) \). Using the properties of the total variation norm \( V \), a successive application of the latter inequality leads directly to (4.3).

**Proof of Lemma 4.1.** From Prop. 4.2 we obtain that

\[
P(\|Z^y_n - \mathbb{E}Z^y_n\| \geq \tau) \leq P\left( \sum_{z \in I_m} (\bar{X}^y_{n,z} - E\bar{X}^y_{n,z}) \geq \tau \right) + m^d \beta_\Psi(t - 2s), \quad y \in I_2,
\]

where \( \bar{X}^y_{n,z}, z \in I_m \) are i.i.d. random variables having the common distribution as \( S^{(t)}_n \) for any \( y \in I_2 \). Thus, we are in a position to apply Prop. 4.1 to each of the sums \( Z^y_n \) yielding

\[
P(\|Z^y_n - \mathbb{E}Z^y_n\| \geq \tau) \leq 2 \exp \left\{ -c_1 \frac{\tau^2}{m^d \mathbb{D}^2 S^{(t)}_n} \right\} + c_2 \tau^{-\gamma} m^d \mathbb{E}|S^{(t)}_n - \mathbb{E}S^{(t)}_n|^{\gamma} + m^d \beta_\Psi(t - 2s).
\]

Finally, remembering that \( (2m)^d = (n/t)^d \), the inequality

\[
P(\|S^{(n)}_n - \mathbb{E}S^{(n)}_n\| \geq \tau) \leq \sum_{y \in I_2} P(\|Z^y_n - \mathbb{E}Z^y_n\| \geq \tau/2^d)
\]

completes proof of Lemma 4.1. \( \square \)

5. PROOFS OF THE RESULTS OF SECTION 2

5.1. Proof of Proposition 2.1

By definition of the second factorial moment measure and the second product density \( \rho^{(2)} \) the expectation of the kernel estimator \( \hat{\rho}^{(2)}_n \) is seen to be equal to

\[
\mathbb{E}\hat{\rho}^{(2)}_n(x) = \frac{\lambda}{\nu_n(W_n) b_n^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} 1_{W_n}(x_1) k\left( \frac{x_2 - x_1 - x}{b_n} \right) \rho^{(2)}(x_2 - x_1) dx_2 dx_1
= \lambda \int_{\mathbb{R}^d} k(y) \rho^{(2)}(x + yb_n) dy
\]
From Condition $\mathcal{K}(d, 1)$ and Lebesgue's convergence theorem the first assertion of Prop. 2.1 follows immediately. By Taylor expansion of the integrand in $b(x, \epsilon)$ (with $p$th-order remainder term), see Nadaraya [20], Chapt. II, and making use of Condition $\mathcal{K}(d, p)$ and the differentiability properties of $\varphi^{(2)}$, we get the asserted rate of convergence in (2.7).

5.2. Proof of Theorem 2.1

For the sake of generality we determine the asymptotic behaviour of the variance of

$$
\vartheta_n^{(2)}(x) = \frac{1}{\nu_d(W_n) b_n^d} \sum_{x_1, x_2 \in s(\Psi)} 1_{W_n(x_1) k\left(\frac{x_2 - x_1 - x}{b_n}\right)}; \quad x \in \mathbb{R}^d,
$$

where $t = t_n$ is a non-decreasing sequence of real numbers with $1 \leq t_n \leq n$. In this proof we consider only large $n$ for which $Rb_n \leq 1$.

**Proposition 5.1.** Under the conditions of Theorem 2.1 with $\varphi^{(2)}(x) > 0$ we have

$$
D^2 \vartheta_n^{(2)}(x) = \left(\frac{1}{n}\right)^d D^2 \vartheta_n^{(2)}(x)(1 + o(1)) \text{ as } n \to \infty \quad (5.5)
$$

where the asymptotic behaviour of $D^2 \vartheta_n^{(2)}(x)$ is given in (2.8).

A straightforward, but lengthy calculation (see also Heinrich [7], Jolivet [13]) leads to

\[
\begin{align*}
D^2 \left( \sum_{x_1, x_2 \in s(\Psi)} g_n(x_1, x_2) \right) &= \int_{\mathbb{R}^{2d}} g_n(x_1, x_2) (g_n(x_1, x_2) + g_n(x_2, x_1)) \\
& \times \alpha^{(2)}(dx_1, dx_2) + \int_{\mathbb{R}^{2d}} g_n(x_1, x_2) g_n(x_1, x_3) \\
& + g_n(x_3, x_1) + g_n(x_2, x_3) + g_n(x_3, x_2)) \\
& \times \alpha^{(3)}(dx_1, dx_2, dx_3) + \int_{\mathbb{R}^{3d}} g_n(x_1, x_2) g_n(x_3, x_4) \\
& \times (\alpha^{(4)} - \alpha^{(2)} \times \alpha^{(2)})(dx_1, dx_2, dx_3, dx_4)
\end{align*}
\]
Taking into account the definition of $g^{(k)}$, $k = 2, 3, 4$, and inserting $g_n(x_1, x_2) = 1_{W_1}(x_1)k((x_2 - x_1 - x)/b_n)$ into the latter formula we find that

$$D^2 g_n^{(2)}(x) = \lambda(n b_n)^{-2d}(A_{n1} + \cdots + A_{n7})$$

with

$$A_{n1} = \int_{\mathbb{R}^d} 1_{W_1}(x_1)k^2 \left(\frac{x_2 - x}{b_n}\right) g^{(2)}(x_2)dx_1dx_2,$$

$$A_{n2} = \int_{\mathbb{R}^d} 1_{W_1}(x_1)1_{W_1}(x_2 + x_1)k \left(\frac{x_2 - x}{b_n}\right) k \left(\frac{-x_2 - x}{b_n}\right) g^{(2)}(x_2)dx_1dx_2,$$

$$A_{n3} = \int_{\mathbb{R}^d} 1_{W_1}(x_1)k \left(\frac{x_2 - x}{b_n}\right) k \left(\frac{x_1 - x}{b_n}\right) g^{(3)}(x_2, x_3)dx_1dx_2dx_3,$$

$$A_{n4} = \int_{\mathbb{R}^d} 1_{W_1}(x_1)1_{W_1}(x_3 + x_1)k \left(\frac{x_2 - x}{b_n}\right) k \left(\frac{-x_3 - x}{b_n}\right)$$

$$\times g^{(3)}(x_2, x_3)dx_1dx_2dx_3,$$

$$A_{n5} = \int_{\mathbb{R}^d} 1_{W_1}(x_1 + x_2)1_{W_1}(x_2)k \left(\frac{-x_1 - x}{b_n}\right) k \left(\frac{x_3 - x}{b_n}\right)$$

$$\times g^{(3)}(x_1, x_3)dx_1dx_2dx_3,$$

$$A_{n6} = \int_{\mathbb{R}^d} 1_{W_1}(x_1 + x_2)1_{W_1}(x_3)k \left(\frac{-x_1 - x}{b_n}\right) k \left(\frac{-x_3 - x}{b_n}\right)$$

$$\times g^{(3)}(x_1, x_3)dx_1dx_2dx_3,$$

$$A_{n7} = \lambda^{-1} \int_{\mathbb{R}^d} 1_{W_1}(x_1)1_{W_1}(x_3)k \left(\frac{x_2 - x_1 - x}{b_n}\right) k \left(\frac{x_4 - x_3 - x}{b_n}\right)$$

$$\times (\alpha^{(4)}(dx_1, dx_2, dx_3, dx_4) - \alpha^{(2)}(dx_1, dx_2)\alpha^{(2)}(dx_3, dx_4)).$$

The asymptotic behaviour of the leading term $A_{n1}$ is obtained by applying the dominated convergence theorem:

$$A_{n1} = (t b_n)^d \int_{\mathbb{R}^d} k^2(y)g^{(2)}(x + y b_n)dy$$

$$= (t b_n)^d \left( g^{(2)}(x) \int_{\mathbb{R}^d} k^2(y)dy + o(1) \right) \quad (5.6)$$

as $n \to \infty$. 

Likewise, for $x = o$,

$$A_{n2} = b_n^d \int_{\mathbb{R}^d} \nu_d(W_t \cap (W_t + yb_n))k(y)k(-y)\varrho^{(2)}(yb_n)dy$$

$$= \left( t b_n \right)^d \left( \varrho^{(2)}(o) \int_{\mathbb{R}^d} k(y)k(-y)dy + o(1) \right)$$  \hspace{1cm} (5.7)

as $n \to \infty$.

For brevity, let $M = \sup_{y: ||y|| \leq R} |k(y)|$. In case of $x \neq o$ it is easily seen from (2.3) that

$$|A_{n2}| \leq t^d M^2 \int_{\mathbb{R}^d} 1_{b(x, Rb) \cap b(-x, Rb)}(y)\varrho^{(2)}(y)dy = 0 \text{ if } b_n < ||x||/R.$$  \hspace{1cm} (5.8)

Integrating over the indicator functions and substituting the other two variables in the integral expressions appearing in $A_{n3}, \ldots, A_{n6}$ we get the following estimate:

$$\sum_{k=3}^{6} |A_{nk}| \leq t^d b_n^d \int_{\mathbb{R}^d} |k(y_2)k(y_3)|\left( \varrho^{(3)}(y_2b_n + x, y_3b_n + x) + \varrho^{(3)}(y_2b_n + x, -y_3b_n - x) + \varrho^{(3)}(-y_2b_n - x, y_3b_n + x) + \varrho^{(3)}(-y_2b_n - x, -y_3b_n - x) \right)dy_2dy_3.$$  

Hence, from (2.3) and (2.5),

$$\sum_{k=3}^{6} |A_{nk}| \leq c_4(t b_n^2)^d.$$  \hspace{1cm} (5.9)

Now we treat the convariance term $A_{n7}$. For convenience, without losing generality, we may assume that $t = t_n \in \{1, \ldots, n\}$ and $t_n \uparrow \infty$ as $n \to \infty$. Put $E_z = [0, 1)^d + z$ and $F_z(x) = E_z \cup (E_z + b(x, Rb_n))$ for $z \in I_t = \{0, 1, \ldots, t - 1\}^d$ and choose an integer $a \geq 2(|x| + Rb_n + 1)$ (with $a \leq t$) implying $F_z(x) \cap F_z(x) = \emptyset$ if $|y - z| > a$, where $|x|$ denotes the maximum norm of $x \in \mathbb{R}^d$. We may decompose $A_{n7} = \lambda^{-1}(B_{n1} + B_{n2})$ with

$$B_{n1} = \lambda \sum_{z \in I_t} \sum_{y \in I_z \cap [-|z|]} \int_{\mathbb{R}^d} 1_{E_z}(x)1_{E_y}(x)k\left( \frac{x_2 - x_1 - x}{b_n} \right)k\left( \frac{x_4 - x_3 - x}{b_n} \right)$$

$$\times \left( \varrho^{(4)}(x_2 - x_1, x_3 - x_1, x_4 - x_1) - \lambda \varrho^{(2)}(x_2 - x_1) \varrho^{(2)}(x_4 - x_3) \right)$$

$$\times dx_1dx_2dx_3dx_4$$
and

$$B_{n2} = \sum_{z \in I} \sum_{y \in I_{|z - y|} > a} (E_{U_{n,z}} U_{n,y} - E_{U_{n,z}} E U_{n,y}),$$

where $U_{n,z} = \sum_{x_1, x_2 \in s(U)} 1_E(x_1) k^{(\frac{xy - x_2}{b_n})}$.

It is easy to see that

$$|B_n| \leq \lambda t^d ((2a + 1)^d \wedge t^d) b_n^{2d}$$

$$\times \sup_{y \in I_{|y - z|} \leq a} \int_{\mathbb{R}^d} \nu_d(E_x \cap (E_{y} + y_3)) |k(y_2)||k(y_4)|$$

$$\times g^{(4)}(y_2 b + x, y_3, y_3 + y_4 b_n + x)$$

$$- \lambda g^{(2)}(y_2 b_n + x) g^{(4)}(y_4 b_n + x)|dy_2 dy_3 dy_4.$$ 

Using (2.3) and (2.5) we can bound the integral on the rhs by some constant. Therefore,

$$B_n \leq c_4 (a \wedge t)^d t^d b_n^{2d}. \quad (5.10)$$

The choice of the sequence $a = a_n \uparrow \infty$ will be specified after estimating $B_{n2}$. For doing this we apply an estimate of the covariance of two random variables in terms of their absolute regularity coefficient, see Yoshihara [28] or Heinrich [10]. In view of Condition $M(4 + \delta)$ and Condition $\beta$ this estimate yields for $0 < \eta \leq \delta/2$ the following:

$$|E U_{n,x} U_{n,y} - E U_{n,z} E U_{n,y}| \leq 2(E U_{n,o}|^{2+\eta})^{2/(2+\eta)} \beta^{\eta/2}(\beta(F_x(x)), \beta(F_z(x)))$$

$$\leq 2M^2 (E U_{n,o} \beta(E_{o} + b(x, Rb_n)))^{2/(2+\eta)}$$

$$\times \beta^{\eta/2}(\beta(F_x(x)), \beta(F_z(x)))$$

if $|y - z| > 3(|x| + Rb_n + 1)$. Setting $\eta = \delta/2$ the latter shows that

$$|B_{n2}| \leq 2M^2 t^d (E U_{n,o} \beta(E_{o} + b(x, Rb_n)))^{4/(4+\delta)}$$

$$\times \sum_{k=a+1}^{\infty} ((2k + 1)^d - (2k - 1)^d) \beta^{\delta/2}(k - 2(|x| + Rb_n + 1))$$

$$\quad (5.11)$$

for $a \geq 3(|x| + Rb_n + 1)$. Since $\beta(n) = o(n^{-2d(4+\delta)/\delta})$ as $n \to \infty$, by standard arguments from analysis we find some null sequence $\varepsilon_n$ such
that
\[ b_n^{-d} \sum_{k = [\varepsilon_n/n]}^{\infty} k^{d-1} \beta^{\delta/(4+\delta)}(k) \to 0 \text{ as } n \to \infty. \tag{5.12} \]

Now, we put \( a = a_n = [\varepsilon_n/n] \). Thus, from (5.10), \( B_{n1} = O(\varepsilon_n^n b_n^d) \).

On the other hand, (5.11) and (5.12) imply that \( B_{n2} = o(t^d b_n^d) \) if \( t \geq a_n \) (and \( B_{n2} = 0 \) if \( t < a_n \)). Hence, \( A_n = o(t^d b_n^d) \) as \( n \to \infty \). Combining the latter with (5.6) – (5.9), we obtain in the particular case \( t = n \) the assertion of Theorem 2.1, which in turn, since \( q^{(2)}(x) > 0 \), shows that
\[ \lambda \sum_{k=1}^{T} A_{nk} = (ntb_n^2)^d \hat{\varrho}_n^{(2)}(x)(1 + o(1)) \text{ as } n \to \infty. \]

Provided that \( \Psi \) is \( m \)-dependent the term \( B_{n2} \) disappears so that the existence of \( \mathbb{E}[\Psi^d(E_0)] \) is sufficient. This completes the proof of Proposition 5.1 and Theorem 2.1.

**Lemma 5.1.** Let \( \Psi \) be a stationary point process satisfying \( \mathbb{E}[\Psi([0,1]^d)]^{2\gamma} < \infty \) for some \( \gamma \geq 1 \). Then, for \( 1 \leq t \leq n \) and \( x \in \mathbb{R}^d \),
\[ \mathbb{E}[|\hat{\varrho}_n^{(2)}(x) - \hat{\varrho}_n^{(2)}(x)|^\gamma] \leq c_5 \left( \frac{t}{n b_n} \right)^{\gamma d} \tag{5.13} \]
with some constant \( c_5 \) not depending on \( t \), \( n \) or \( x \).

**Proof of Lemma 5.1.** Using the inequality
\[ |X_1 + \cdots + X_m|^\gamma \leq m^{\gamma-1}(|X_1|^\gamma + \cdots + |X_m|^\gamma), \quad \gamma \geq 1, \quad m \geq 1, \]
and the stationarity of \( \Psi \) we find that
\[ \mathbb{E}\left[|\hat{\varrho}_n^{(2)}(x) - \hat{\varrho}_n^{(2)}(x)|^\gamma\right] \leq (2 t^d)^\gamma \mathbb{E}|\hat{\varrho}_n^{(2)}(x)|^\gamma. \]

Further, with \( M = \sup_{x:|x| \leq R}|k(x)| \),
\[ \mathbb{E}|\hat{\varrho}_n^{(2)}(x)|^\gamma \leq \left( \frac{M}{n b_n^d} \right)^\gamma \mathbb{E}\left( \sum_{x_1 \in s(\Psi) \cap [0,1]^d} (\Psi - \delta_{x_1})(b(x + x_1, Rb_n)) \right)^\gamma \]
\[ \leq \left( \frac{M}{n b_n^d} \right)^\gamma \mathbb{E}(\Psi([0,1]^d + b(o, Rb_n)))^{2\gamma}. \]
The previous lemma combined with Lemma 4.1 and Theorem 2.1 yields

**Proposition 5.2.** Let the Conditions $\beta, K (d, l), P(x),$ and $M(2\gamma)$ for some $\gamma > 2$ be satisfied. Then

\[
P(\{\tilde{\varrho}_n^{(2)}(x) - E\tilde{\varrho}_n^{(2)}(x) \geq \tau\})
\leq 2^{d+1}\exp\{-c_6\tau^2n^d b_n^d\} + c_7\tau^{-\gamma}n^{-\gamma d + d} t^{-d+\gamma d} b_n^{-\gamma d}
+ n^d t^{-d} \beta\psi(t - 2(\|x\| + R b_n))
\]

for any $\tau > 0$ and sufficiently large $t \leq n$. Moreover, if $\beta\psi(t_0) = 0$ for some $t_0 > 0$, then

\[
P(\{\tilde{\varrho}_n^{(2)}(x) - E\tilde{\varrho}_n^{(2)}(x) \geq \tau\})
\leq 2^{d+1}\exp\{-c_6\tau^2n^d b_n^d\} + c_8\tau^{-\gamma}n^{-\gamma d + d} b_n^{-\gamma d}
\]

for $\gamma \geq 2$.

**5.3. Proof of Theorem 2.2**

According to the definition of $\mathcal{P}$-a.s. convergence (or by employing the Borel-Cantelli lemma) the relation

\[
\tilde{\varrho}_n^{(2)}(x) - E\tilde{\varrho}_n^{(2)}(x) = O(a_n) \quad \text{P-a.s. as } n \to \infty \tag{5.14}
\]

holds for a positive (null) sequence $a_n$, whenever

\[
\sum_{n \geq 1} P(\{\tilde{\varrho}_n^{(2)}(x) - E\tilde{\varrho}_n^{(2)}(x) \geq \tau a_n\}) < \infty
\]

for some fixed $\tau > 0$.

From Proposition 5.1 we have

\[
P(\{\tilde{\varrho}_n^{(2)}(x) - E\tilde{\varrho}_n^{(2)}(x) \geq \tau a_n\})
\leq 2^{d+1}\exp\{-c_6\tau^2a_n^2(n b_n)^d\} + c_7\tau^{-\gamma}a_n^{-\gamma d + d} t^{-d+\gamma d} b_n^{-\gamma d}
+ n^d t^{-d} \beta\psi(t - 2(\|x\| + R b_n))
\]
Now, in accordance with the rate of decay of $\beta_\psi(t)$ we choose $t = t_n$ for the Cases (i), (ii) and (iii) in the following way:

$$t_n = \begin{cases} 
(n^{d+1} \ln n (\ln \ln n)^2)^{1/(d+q)} & \text{Case (i)} \\
((d+2) \ln n/a)^{1/b} & \text{Case (ii)} \\
t_0 + 2(||x|| + Rb_n) & \text{Case (iii)}. 
\end{cases}$$  \hspace{1cm} (5.15)

This choice guarantees that $\sum_{n \geq 1} n^d t_n^{-d} \beta_\psi(t_n - 2(||x|| + Rb_n)) < \infty$. It remains to find the sequence $a_n$ such that, for some $\tau > 0$,

$$\sum_{n \geq 1} \exp\{-\tau^2 a_n^2 (n b_n)^d\} < \infty \quad \text{and} \quad \sum_{n \geq 1} a_n^{-\gamma} n^{-\gamma d + d} t_n^{-d+\gamma d} b_n^{-\gamma d} < \infty.$$

It is easily checked that appropriate sequences are

$$a_n = \begin{cases} 
\ln^{1/2} n (n b_n)^{-d/2} + \ln^\kappa n n^{-d+(d+1)/\gamma} t_n^{-d+\gamma d} b_n^{-d} & \text{for } \kappa > 1/\gamma: \text{Cases (i) and (ii)} \\
\ln^{1/2} n (n b_n)^{-d/2} + \ln^\kappa n n^{-d+(d+1)/\gamma} b_n^{-d} & \text{for } \kappa > 1/\gamma: \text{Cases (iii)}. 
\end{cases}$$

By (2.7),

$$\hat{g}_n^{(2)}(x) - \lambda g^{(2)}(x) = O(a_n + b_n^p) \ \ P\text{-a.s. as } n \to \infty.$$

Finally, after inserting the sequences $t_n$ in the above formulae we obtain the desired asymptotic relations for each of the Cases (i), (ii) and (iii). This completes the proof of Theorem 2.2. \hfill \Box

5.4. Proof of Theorem 2.3

In analogy to (4.2) define

$$\hat{g}_n^{(2)}(x) = \sum_{j \in I_2} \sum_{z \in I_m} X_{ny}(x),$$

$$X_{ny}(x) = \frac{1}{(n b_n)^d} \sum_{x_1, x_2 \in s(\Psi)} 1_{W_{i;2;3}}(x_1) k\left(\frac{x_2 - x_1 - x}{b_n}\right).$$
The compact set \( K \) can be covered by \((d\)-dimensional\) cubes \( C_1, \ldots, C_N \) having edges of length \( h \) and midpoints \( u_1, \ldots, u_N \) such that \( N \leq c_8 h^{-d} \). Put \( C_i = C_i + b(o, Rb_n) \) and

\[
Y_{nz}^x = \frac{1}{(n b_n)^d} \sum_{x_1, x_2 \in \Phi} \mathbf{1}_{W_{z,2z-1}((x_1))} \mathbf{1}_{C_i(x_2 - x_1)}.
\]

To ease the notation we do not indicate the dependence on \( i \in \{1, \ldots, N\} \) on the lhs. From Condition \( \mathcal{L} \), we deduce that

\[
\sup_{u,v \in C_i} |X_{nz}^x(u) - X_{nz}^x(v)| \leq L \sup_{u,v \in C_i} \left\| \frac{u - v}{b_n} \right\| Y_{nz}^x \leq \frac{L \sqrt{d} h}{b_n} Y_{nz}^{\hat{x}}. \tag{5.16}
\]

By the definition of the reduced second-order factorial moment measure and the continuity of \( \phi^{(2)} \) on \( K \), it follows that

\[
E Y_{nz}^x \leq \frac{\lambda}{n^d b_n^d} \nu_d(W) \alpha_{red}^{(2)}(\tilde{C}_i) \leq c_9 \left( \frac{t h + 2Rb_n}{n b_n} \right)^d. \tag{5.17}
\]

and, in analogy to the proof of Lemma 5.1 using \( E \Psi([0, 1]^d)^{2\gamma} < \infty \),

\[
E(Y_{nz}^x)^\gamma \leq \left( \frac{t}{n b_n} \right)^{d\gamma} E \left( \sum_{x_1 \in \Phi \cap [0,1]^d} (\Psi - \delta_{x_1})(\tilde{C}_i + x_1) \right)^\gamma \leq c_{10} \left( \frac{t}{n b_n} \right)^{d\gamma}. \tag{5.18}
\]

Without loss of generality we may assume that \( a_n \to 0 \) as \( n \to \infty \). Put \( h = \varepsilon \cdot b_n \cdot \tau a_n \), where \( \varepsilon > 0 \) is chosen such that \( 2^d L \sqrt{d} c_9 \varepsilon (\varepsilon \tau a_n + 2R)^d \leq 1/6 \) which, by (5.16), implies that

\[
\sum_{z \in I_m} E \sup_{u,v \in C_i} |X_{nz}^x(u) - X_{nz}^x(v)| \leq L \sqrt{d} \frac{h}{b_n} \sum_{z \in I_m} E Y_{nz}^x \leq \frac{\tau a_n}{6 \cdot 2^d}. \tag{5.19}
\]
Applying Prop. 4.2 in a similar way as in the proof of (4.4) we obtain the following estimates:

\[
P\left( \sup_{x \in K} |\hat{\theta}^{(2)}_n(x) - \hat{\theta}^{(2)}_n(x) - E \hat{\theta}^{(2)}_n(x)| \geq \tau a_n \right) \\
\leq \sum_{y \in I_2} \sum_{i=1}^N \left( \sup_{x \in C_i} \left| \sum_{z \in I_m} \left( \bar{X}^{(1)}_{n_2}(x) - E \bar{X}^{(1)}_{n_2}(x) \right) \right| \geq \frac{\tau a_n}{2^d} \right) \\
+ \left( \frac{n}{t} \right)^d \sup_{x \in K} \beta \phi (t - 2(\|x\| + Rb_n)) \\
\leq \sum_{y \in I_2} \sum_{i=1}^N \left( \sum_{z \in I_m} \left( \bar{X}^{(1)}_{n_2}(x) - E \bar{X}^{(1)}_{n_2}(x) \right) \right) \geq \frac{\tau a_n}{3 \cdot 2^d} \\
+ \sum_{y \in I_2} \sum_{i=1}^N \left( \sup_{x \in C_i} \left| \sum_{z \in I_m} \left( \bar{X}^{(1)}_{n_2}(x) - \bar{X}^{(1)}_{n_2}(u_i) \right) \right| \geq \frac{\tau a_n}{3 \cdot 2^d} \right) \\
+ \left( \frac{n}{t} \right)^d \sup_{x \in K} \beta \phi (t - 2(\|x\| + Rb_n))
\]

for any \( \tau > 0 \), where the random variables \( \bar{X}^{(1)}_{n_2}(x), z \in I_m, y \in I_2 \), are independent and \( \bar{X}^{(1)}_{n_2}(x) \) has the same distribution as \( X^{(1)}_{n_2}(x) \) for all \( z \in I_m, y \in I_2 \) and any \( x \in K \). The first sum on the rhs can be estimated by applying Prop. 4.1, Lemma 5.1 and Theorem 2.1 (as in the proof of Prop. 5.1) so that its upper bound is as follows:

\[
N 2^{d+1} \exp \left\{ -c_{11} \tau^2 a_n^2 (n b_n)^d \right\} + c_{12} N \tau^{-\gamma} a_n^{-\gamma} \left( \frac{t}{n} \right)^{d(\gamma-1)} b_n^{-\gamma d}. (5.20)
\]

By using (5.17), (5.19) and Prop. 4.1 we may write that

\[
P\left( \sup_{x \in C_i} \left| \sum_{z \in I_m} \left( \bar{X}^{(1)}_{n_2}(x) - \bar{X}^{(1)}_{n_2}(u_i) \right) \right| \geq \frac{\tau a_n}{3 \cdot 2^d} \right) \\
\leq P\left( \sum_{y \in I_2} \sum_{i=1}^N \bar{Y}^{(1)}_{n_2} \geq \frac{\tau a_n}{3 \cdot 2^d} \right)
\]
\[
\begin{align*}
&\leq P \left( \left| \sum_{z \in I_m} (\bar{Y}_{n_z}^y - \mathbb{E} \bar{Y}_{n_z}^y) \right| \geq \frac{1}{6L \sqrt{d \cdot 2^d \varepsilon}} \right) \\
&\leq 2 \exp \left\{ - \frac{c_{13}}{\varepsilon^2} \sum_{z \in I_m} \mathbb{E}^2 Y_{n_z}^y \right\} + c_{14} \varepsilon^\gamma \sum_{z \in I_m} \mathbb{E} (Y_{n_z}^y)^\gamma
\end{align*}
\]
for \( i = 1, \ldots, N \) and \( y \in I_2 \), where the random variables \( \bar{Y}_{n_z}^y, z \in I_m \), \( y \in I_2 \), are independent having the same distribution as \( \bar{Y}_{n_z}^y, z \in I_m \), \( y \in I_2 \). Applying Prop. 5.1. to \( Y_{n_z}^y \) with the kernel function \( k(x) = c_{15} \mathbb{E}_{Y/n_0}(x) \) gives \( \mathbb{E}^2 Y_{n_z}^y \leq c_{16} (t/n_0 b_0)^d \) which together with (5.18) leads to the following upper bound of second sum:

\[
N 2^{d+1} \exp \left\{ - c_{17} \varepsilon^{-2} (n b_n)^d \right\} + c_{18} \varepsilon^\gamma \left( \frac{t}{n b_n} \right)^\gamma \left( \frac{n}{t} \right)^d.
\]

Clearly, this expression is asymptotically smaller than (5.20). Therefore, since \( N \leq c_8 (\varepsilon \tau a_n b_n)^{-d} \), the convergence of the series

\[
\sum_{n=1}^\infty P \left( \sup_{x \in K} \left| \hat{\varphi}^{(2)}_n(x) - \mathbb{E} \hat{\varphi}^{(2)}_n(x) \right| \geq \tau a_n \right)
\]

for a (sufficiently large) \( \tau > 0 \) is guaranteed whenever \( a_n^2 n^{d} b_n^d \geq \ln n \) and

\[
a_n^{-d(\gamma-1)} \left( \frac{t_n}{n} \right)^{d(\gamma-1)} b_n^{-d(\gamma+1)} \leq n^{-\kappa} \ln n \text{ for some } \kappa > 1.
\]

These inequalities hold for

\[
a_n = \ln^{1/2} n (n b_n)^{-d/2} + \ln^{\kappa} n t_n^{d(\gamma-1)/(\gamma+1)} n^{-d(\gamma-1)/(\gamma+1)} b_n^{-d(\gamma+1)/(\gamma+1)}
\]

for some \( \kappa > 1/(\gamma + d) \), where the sequence \( t_n = t \) is given by (5.15). Finally, together with a uniform variant of (2.7) we obtain the rates of convergence of \( \Delta_n(K) \) as stated in the Cases (i), (ii) and (iii) of Theorem 2.3.
6. PROOFS OF THE RESULTS OF SECTION 3

6.1. Proof of Proposition 3.1

We argue in a similar way as above to prove Prop. 2.1. Since $\Psi$ is additionally isotropic we have

$$E\tilde{g}_n(r) = \lambda^2 \int_{-r/b_n}^{\infty} k(x)g(r + xb_n)dx$$

(6.1)

$$E\tilde{g}_n(r) = \lambda^2 \int_{-r/b_n}^{\infty} k(x)g(r + xb_n)\left(1 + \frac{xb_n}{r}\right)^{d-1}dx$$

(6.2)

As in proving Lemma 1 in Liebscher [15] we split the integral $\int_{-r/b_n}^{\infty} = \int_{-r/b_n}^{0} + \int_{0}^{\infty}$. Then, letting $n \to \infty$ yields

$$\lim_{n \to \infty} E\tilde{g}_n(r) = \lim_{n \to \infty} E\tilde{g}_n(r)$$

$$= \lambda^2 \left(\int_{-\infty}^{0} k(x) dx \ g(r - 0) + \int_{0}^{\infty} k(x) dx \ g(r + 0)\right)$$

whence, since $k(x) = k(-x)$, it follows (3.3).

Inserting the Taylor expansion

$$g(r + \vartheta xb_n) = \sum_{l=0}^{p-2} \frac{(xb_n)^l}{l!} g^{(l)}(r) + \frac{(xb_n)^{p-1}}{(p-1)!} g^{(p-1)}(r + \vartheta xb_n), \quad 0 < \vartheta < 1$$

in (6.1) and remembering that

$$\int_{-r/b_n}^{R} x^l k(x)dx = 0$$

and

$$\left| \int_{-r/b_n}^{R} x^{p-1} g^{(p-1)}(r + \vartheta xb_n)k(x)dx \right| \leq b_n L_{p-1} \int_{-R}^{R} \left|x\right|^p \left|k(x)\right|dx$$

for $r \geq Rb_n$, we obtain (3.5). Likewise, additionally expanding $(r + xb_n)^{d-1}$, one can prove (3.4). \qed
6.2. Proof of Theorem 3.2

We follow the line of the proof of Theorem 2.3 in Section 5. Therefore, it suffice to concentrate on estimates which are different from those in 5.4. We split up $\hat{g}_n(r)$ as follows:

$$\hat{g}_n(r) = \sum_{y \in I_y} \sum_{z \in I_z} \hat{X}^y_{nz}(r),$$

$$\hat{X}^y_{nz}(r) = \sum_{x_1, x_2 \in s(\Psi)} \frac{1_{W_{l_1, l_2}}(x_1)}{d\omega d r^{d-1} n^d b_n} k\left(\frac{\|x_2 - x_1\| - r}{b_n}\right).$$

Clearly, $[r_1, r_2]$ is covered by the intervals $C_i = [s_i - 1, s_i]$, where $s_i = r_1 + i(r_2 - r_1)/N, i = 1, \ldots, N$ with some $N$ which will be specified later. By Condition $L$,

$$\sup_{s_i - 1 \leq u, v \leq s_i} |\hat{X}^y_{nz}(u) - \hat{X}^y_{nz}(v)| \leq \frac{c_19}{Nb_n} \hat{Y}^y_{nz},$$

where

$$\hat{Y}^y_{nz} = \frac{1}{d\omega d r^{d-1} n^d b_n} \sum_{x_1, x_2 \in s(\Psi)} 1_{W_{l_1, l_2}}(x_1) 1_{[s_i - 1 - Rb_n, s_i + Rb_n]}(\|x_2 - x_1\|).$$

Quite similar as in 5.4 one can verify the estimates

$$\mathbb{E} \hat{Y}^y_{nz} \leq c_{20} \left(\frac{t}{n}\right)^d (1 + (Nb_n)^{-1}), \quad \mathbb{D}^2 \hat{Y}^y_{nz} \leq \frac{c_{22}}{n^d b_n} \left(\frac{t}{n}\right)^d (1 + (Nb_n)^{-1})$$

and $\mathbb{E}(\hat{Y}^y_{nz})^\gamma \leq c_{23} b_n^\gamma (t/n)^{d\gamma}$ for $\gamma > 2$.

Moreover, for large enough $\tau > 0$ and any positive null sequence $a_n$, we obtain that

$$\mathbb{P} \left( \sup_{r_1 \leq r \leq r_2} |\hat{g}_n(r) - \mathbb{E}\hat{g}_n(r)| \geq \tau a_n \right) \leq 2^{d+1} N \exp \left\{-c_{24} \tau^{-2} a_n^2 n^d b_n \right\}$$

$$\times \exp \left\{-c_{26} \varepsilon^{-2} n^d b_n \right\} + c_{27} N \left(\frac{\varepsilon}{b_n}\right)^\gamma \left(\frac{t}{n}\right)^{d(\gamma - 1)} + \left(\frac{n}{t}\right)^d \beta(\tau - 2(r_2 + Rb_n)).$$
Choosing $N$ as the largest integer satisfying $N^{-1} \geq \varepsilon \tau a_n b_n$ with a sufficiently small $\varepsilon > 0$ and putting

$$a_n = \ln^{1/2} n (n^{d} b_n)^{-1/2} + \ln^{\eta} n b_n^{-1} t^{d(\gamma-1)/(\gamma+1)} n^{-d(\gamma-1)/(\gamma+1)}$$

(6.3)

for $\eta > 1/(\gamma + 1)$ and $t_n = t$ from (5.15), we recognize that

$$\sum_{n \geq 1} P \left( \sup_{r_1 \leq r \leq r_2} |\tilde{g}_n(r) - E\tilde{g}_n(r)| \geq \tau a_n \right) < \infty.$$ 

Thus, by definition of $P$-a.s. convergence combined with the relation

$$\sup_{r_1 \leq r \leq r_2} |E\tilde{g}_n(r) - \lambda^2 g(r)| = O(b_n^p) \text{ as } n \to \infty$$

following from Prop. 3.1, we arrive at

$$\sup_{r_1 \leq r \leq r_2} |\tilde{g}_n(r) - \lambda^2 g(r)| = O(a_n + b_n^p) \text{ P-a.s. as } n \to \infty.$$ 

Finally, we specify the sequence $t_n$ from (5.15) in each of the Cases (i), (ii) and (iii) and determine the sequence $b_n$ such that $a_n + b_n^p$ becomes a minimum. This optimal selection of $b_n$ is achieved by equating the first and the second summand on the rhs of (6.3) with $b_n^p$ and then taking the maximum of the two possible solutions. This yields the bandwidths and rates of convergence stated in Theorem 3.2.

6.3. Proof of Theorem 3.3

Taking into account Remark 3 we prove (3.8) by verifying the assumptions of Theorem 3.2 for $d=1$, $p \geq 1$. Some slight changes in calculating mean and variance of $\hat{h}_n(x)$ are caused by the fact that the renewal process is defined on the positive real axis. First remember that a renewal process $\Psi = \sum_{i \geq 1} \delta_{\xi_i + \cdots + \xi_i}$ with $0 < m = \mathbb{E}\xi_2 < \infty$ possesses finite moments $\mathbb{E}(\Psi([0,1]))^k$ of any order $k \geq 1$ and the third – and fourth – order product densities are expressible by the renewal
density $h$ itself:

$$
\varrho^{(3)}(u, v) = h(u)h(v-u) \text{ and } \varrho^{(3)}(u, v, w) = h(u)h(v-u)h(w-v)
$$

for $0 \leq u < v < w < \infty$.

Therefore, provided $h$ is bounded, Condition $\mathcal{M}(2\gamma)$ and Condition $\mathcal{P}(K)$ are satisfied for any $\gamma > 2$ and any interval $K=[r_1,r_2]$, $0 < r_1 < r_2 < a$.

Estimates of the absolute regularity coefficient for stationary renewal processes were obtained in [8]. It was shown there that $\lim_{t \to \infty} t^q \beta_q(t) = 0$ holds if $\xi \gg 0$ and some convolution power of $F$ has an absolutely continuous component. Letting $\gamma \to \infty$ in Case (i) of Theorem 3.2 means that (3.8) holds whenever $\beta_q(t) = O(t^{-q})$ for $(q-1)/(q+1) > (1+p)/(1+2p)$, i.e. $q + 2 > (5p + 2)/p$.

It remains to verify that the $(p-1)$st-order derivative of $h$ is uniformly bounded and Lipschitz-continuous on $[0, a]$ if

$$
|f^{(p-1)}(x)| \leq b_{p-1} \text{ and } |f^{(p-1)}(x) - f^{(p-1)}(y)| \leq l_{p-1}|x - y| \quad (6.4)
$$

for $0 \leq x \leq a$ and $0 \leq b_{p-1}$, $l_{p-1} < \infty$.

From (3.6) we deduce that, for $k = 0, \ldots, p-1,$

$$
h^{(k)}(x) = f^{(k)}(x) + \sum_{i=0}^{k-1} f^{(i)}(0)h^{(k-i-1)}(x) + \int_0^x f^{(k)}(x-t)h(t)dt
$$

Now, by simple rearrangements, we may write

$$
|h^{(p-1)}(t) - h^{(p-1)}(s)| \leq |f^{(p-1)}(t) - f^{(p-1)}(s)|
$$

$$
+ \sum_{i=0}^{p-2} |f^{(i)}(0)||h^{(p-i-2)}(t) - h^{(p-i-2)}(s)|
$$

$$
+ \frac{1}{m} \int_t^s |f^{(p-1)}(x)|dx + \int_0^s |f^{(p-1)}(t-x) - f^{(p-1)}(s-x)||h(x) - \frac{1}{m}|dx
$$

$$
+ \int_s^t |f^{(p-1)}(t-x)||h(x) - \frac{1}{m}|dx.
$$
To accomplish the proof we make use of the fact that $\int_0^\infty |h(x) - \frac{1}{m}| \, dx < \infty$ which follows from our assumptions, see e.g. [8], and the estimates

$$|f^{(q)}(x)| \leq b_q \quad \text{and} \quad |f^{(q)}(x) - f^{(q)}(y)| \leq l_q |x - y|, \quad 0 \leq b_q, \quad l_q \leq \infty$$

for all $x, y \in [0, a]$ and $q = 0, \ldots, p - 1$. The uniform boundedness and Lipschitz-continuity of $f, f', \ldots, f^{(p-2)}$ can be derived from (6.3). This is quite obvious for $a < \infty$, but applies also for $a = \infty$.

Acknowledgements

The authors wish to thank the referee and the Editor for their useful comments on the earlier version of this article.

References


