LIMIT LAWS FOR LARGE kTH-NEAREST NEIGHBOR BALLS

A PREPRINT

Nicolas Chenavier
Université du Littoral Côte d'Opale
50 rue F. Buisson 62228 Calais
nicolas.chenavier@univ-littoral.fr

Norbert Henze
Institute of Stochastics
Karlsruhe Institute of Technology (KIT)
Englerstr. 2, D-76133 Karlsruhe
Norbert.Henze@kit.edu

Moritz Otto
Institute for Mathematical Stochastics
Otto von Guericke University Magdeburg
Universitätsplatz 2, D-39106 Magdeburg
Moritz.Otto@ovgu.de

April 30, 2021

ABSTRACT

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent random points in $\mathbb{R}^d$ with common Lebesgue density f. Under some conditions on f, we obtain a Poisson limit theorem, as $n \to \infty$, for the number of large probability kth-nearest neighbor balls of $X_1, \ldots, X_n$. Our result generalizes Theorem 2.2 of [10], which refers to the special case $k = 1$. Our proof is completely different since it employs the Chen-Stein method instead of the method of moments. Moreover, we obtain a rate of convergence for the Poisson approximation.

1 Introduction and main results

The starting point of this paper is the following result, see [10]. Let $X, X_1, \ldots, X_n, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random points in $\mathbb{R}^d$ that are defined on a common probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We assume that the distribution of X, which is denoted by μ, is absolutely continuous with respect to Lebesgue measure λ, and we denote the density of μ by f. Writing $\|\cdot\|$ for the Euclidean norm in $\mathbb{R}^d$, and putting $\mathcal{X}_n := \{X_1, \ldots, X_n\}$, let $R_{i,n} := \min_{j \neq i, j \leq n} \|X_i - X_j\|$ be the distance from X_i to its nearest neighbor in the set $\mathcal{X}_n \setminus \{X_i\}$. Moreover, let $\mathbb{1}\{A\}$ denote the indicator function of a set A, and write $B(x,r) = \{y \in \mathbb{R}^d : \|x - y\| \leq r\}$ for the closed ball centered at x with radius r. Finally, let

$$C_n := \sum_{i=1}^{n} \mathbb{1}\Big\{\mu\big(B(X_i, R_{i,n})\big) > \frac{t + \log n}{n}\Big\}$$

denote the number of exceedances of probability volumes of nearest neighbor balls that are larger than the threshold $(t + \log n)/n$. The main result of [10] is Theorem 2.2 of that paper, which states that, under a weak condition on the density f, for each fixed $t \in \mathbb{R}$, we have

$$C_n \xrightarrow{\mathcal{D}} \mathrm{Po}(\exp(-t)) \tag{1}$$

as $n \to \infty$, where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution, and $\mathrm{Po}(\xi)$ is the Poisson distribution with parameter $\xi > 0$.

Since the *maximum* probability content of these nearest balls, denoted by P_n, is at most $(t + \log n)/n$ if, and only if, $C_n = 0$, we immediately obtain a Gumbel limit $\lim_{n\to\infty} \mathbb{P}(nP_n - \log n \leq t) = \exp(-\exp(-t))$ for P_n.

MSC 2010 subject classifications. Primary 60F05 Secondary 60D05
Key words and phrases Binomial point process; large kth-nearest neighbor balls; Chen-Stein method; Poisson convergence; Gumbel distribution

To state a sufficient condition on f that guarantees (1), let supp$(\mu) := \{x \in \mathbb{R}^d : \mu(B(x,r)) > 0 \text{ for each } r > 0\}$ denote the support of μ. Theorem 2.2 of [10] requires that there are $\beta \in (0,1)$, $c_{max} < \infty$ and $\delta > 0$ such that, for any $r, s > 0$ and any $x, z \in$ supp(μ) with $\|x - z\| \geq \max\{r, s\}$ and $\mu(B(x,r)) = \mu(B(z,s)) \leq \delta$, one has

$$\frac{\mu(B(x,r) \cap B(z,s))}{\mu(B(z,s))} \leq \beta$$

and $\mu\big(B(z, 2s)\big) \leq c_{max}\mu(B(z,s))$.

These conditions hold if supp(f) is a compact set K (say), and there are $f_-, f_+ \in (0, \infty)$ such that

$$f_- \leq f(x) \leq f_+, \qquad x \in K. \tag{2}$$

Thus, the density f of X is bounded and bounded away from zero.

The purpose of this paper is to generalize (1) to kth-nearest neighbors, and to derive a rate of convergence for the Poisson approximation of the number of exceedances.

Before stating our main results, we give some more notation. For fixed $k \leq n-1$, we denote by $R_{i,n,k}$ the Euclidean distance of X_i to its kth-nearest neighbor among $\mathcal{X}_n \setminus \{X_i\}$, and we write $B(X_i, R_{i,n,k})$ for the kth-nearest neighbor ball centered at X_i with radius $R_{i,n,k}$. For fixed $t \in \mathbb{R}$, put

$$v_{n,k} := v_{n,k}(t) := \frac{t + \log n + (k-1)\log\log n - \log(k-1)!}{n}, \tag{3}$$

and let

$$C_{n,k} := \sum_{i=1}^{n} \mathbb{1}\big\{\mu\big(B(X_i, R_{i,n,k})\big) > v_{n,k}\big\} \tag{4}$$

denote the *number of exceedances* of probability contents of kth-nearest neighbor balls over the threshold $v_{n,k}$ defined in (3).

The threshold $v_{n,k}$ is in some sense *universal* in dealing with the number of exceedances of probability contents of kth-nearest neighbor balls. To this end, suppose that, in much more generality than considered so far, $X, X_1, X_2, \ldots$ are i.i.d. random elements taking values in a separable metric space (S, ρ). We retain the notations μ for the distribution of X and $B(x,r) := \{y \in S : \rho(x,y) \leq r\}$ for the closed ball with radius r centered at $x \in S$. Regarding the distribution μ, we assume that

$$\mu(\{y \in S : \rho(x,y) = r\}) = 0, \qquad x \in S,\ r \geq 0. \tag{5}$$

As a consequence, the distances $\rho(X_i, X_j)$, where $j \in \{1, \ldots, n\} \setminus \{i\}$, are different with probability one for each $i \in \{1, \ldots, n\}$. Thus, for fixed $k \leq n-1$, there is almost surely a unique kth-nearest neighbor of X_i, and we also retain the notations $R_{i,n,k}$ for the distance of X_i to its kth-nearest neighbor among $\mathcal{X}_n \setminus \{X_i\}$ and $B(X_i, R_{i,n,k})$ for the ball centered at X_i with radius $R_{i,n,k}$. Notice that the condition (5) excludes discrete metric spaces (see, e.g., Section 4 of [17]), but not function spaces like, e.g., the space $C[0,1]$ of continuous functions on $[0,1]$ with the supremum metric, and with Wiener measure μ.

In what follows, for sequences $(a_n)_{n\geq 0}$ and $(b_n)_{n \geq 0}$ of real numbers, write $a_n = O(b_n)$ if $|a_n| \leq C|b_n|$, $n \geq 1$, for some positive constant C.

Theorem 1.1. *If $X_1, X_2, \ldots$ are i.i.d. random elements of a metric space (S, ρ), and if* (5) *holds, then the sequence $(C_{n,k})$ satisfies*

$$\mathbb{E}[C_{n,k}] = \mathrm{e}^{-t} + O\left(\frac{\log\log n}{\log n}\right).$$

In particular, the mean number of exceedances $C_{n,k}$ converges to e^{-t} as n goes to infinity. By Markov's inequality, this result implies the tightness of the sequence $(C_{n,k})_{n\geq 1}$. Thus, at least a subsequence converges in distribution. The next result states convergence of $C_{n,k}$ to a Poisson distribution if $(S, \rho) = (\mathbb{R}^d, \|\cdot\|)$ and (2) holds. To this end, let $d_{TV}(Y, Z)$ be the total variation between two integer-valued random variables Y and Z, i.e.,

$$d_{TV}(Y,Z) = 2 \sup_{A \subset \mathbb{N}} |\mathbb{P}(Y \in A) - \mathbb{P}(Z \in A)|.$$

Theorem 1.2. *Let Z be a Poisson random variable with parameter e^{-t}. If $X, X_1, X_2, \ldots$ are i.i.d. in $\mathbb{R}^d$ with density f, and if the distribution μ of X has compact support $[0,1]^d$ and satisfies* (2)*, then, as $n \to \infty$,*

$$d_{TV}\left(C_{n,k}, Z\right) = O\left(\frac{\log\log n}{\log n}\right).$$

Theorem 1.2 is not only a generalization of Theorem 2.2 of [10] over all $k \geq 1$: it also provides a rate of convergence for the Poisson approximation of $C_{n,k}$. Our theorem is stated in the particular case that the support of μ is $[0,1]^d$ but we think that it can be extended to any measure μ whose support is a general convex body. For the sake of readibility of the manuscript, we did not deal with such a generalization.

Remark 1.3. The study of extremes of kth-nearest neighbor balls is classical in stochastic geometry, and it has various applications, see e.g. [16]. In Section 4 in [15], bounds for the total variation distance of the process of Poisson points with large kth-nearest neighbor ball (with respect to the intensity measure) and a Poisson process were obtained. Parallel to our work, these results have been extended by Bobrowski *et al.* to the Kantorovich-Rubinstein distance and generalized to the binomial process in Section 6.2 of a paper which has just been submitted [5]. While the results in [5] and [15] rely on Palm couplings of a thinned Poisson/binomial process and employ distances of point processes, we derive a bound on the total variation distance of the number of large kth-nearest-neighbor balls and a Poisson-distributed random variable. Our approach permits to build arguments on classical Poisson approximation theory [2] and an asymptotic independence property stated in Lemma 2.2 below, and it thus results in a considerably shorter and less technical proof.

Now, let $P_{n,k} = \max_{1 \leq i \leq n} \mu\big(B(X_i, R_{i,n,k})\big)$ be the maximum probability content of the kth-nearest neighbor balls. Since $C_{n,k} = 0$ if, and only if, $P_{n,k} \leq v_{n,k}$, we obtain the following corollary.

Corollary 1.4. *Under the conditions of Theorem 1.2, we have*

$$\lim_{n\to\infty} \mathbb{P}\big(nP_{n,k} - \log n - (k-1)\log\log n + \log(k-1)! \leq t\big) = \mathrm{G}(t), \quad t \in \mathbb{R},$$

where $\mathrm{G}(t) = \exp(-\exp(-t))$ is the distribution function of the Gumbel distribution.

Remark 1.5. If, in the Euclidean case, the density f is continuous, then $\mu(B(X_i, R_{i,n,k}))$ is approximately equal to $f(X_i)\kappa_d R_{i,n,k}^d$, where $\kappa_d = \pi^{d/2}/\Gamma(1+d/2)$ is the volume of the unit ball in $\mathbb{R}^d$. Under additional smoothness assumptions on f and (2), [11, 12] proved that

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{i=1,\ldots,n} f(X_i)\kappa_d \min\left(R_{i,n,k}^d, \|X_i - \partial K\|^d\right) \leq v_{n,k}\right) = \mathrm{G}(t), \tag{6}$$

where K is the support of μ. Here, the distance $\|X_i - \partial K\|$ of X_i to the boundary of K is important to overcome edge effects. These effects dominate the asymptotic behavior of the maximum of the kth-nearest neighbor distances if $k \geq d$, see [8, 9]. In fact, [11] proved convergence of the factorial moments of

$$\widetilde{C}_{n,k} := \sum_{i=1}^{n} \mathbb{1}\Big\{f(X_i)\kappa_d \min\left(R_{i,n,k}^d, \|X_i - \partial K\|^d\right) > v_{n,k}\Big\}$$

to the corresponding factorial moments of a random variable with the Poisson distribution $\mathrm{Po}(\mathrm{e}^{-t})$ and thus, by the method of moments, more than (6), namely $\widetilde{C}_{n,k} \xrightarrow{\mathcal{D}} \mathrm{Po}\big(\mathrm{e}^{-t}\big)$. However, our proof of Theorem 1.2 is completely different thereof, since it is based on the Chen-Stein method and provides a rate of convergence.

2 Proofs

2.1 Proof of Theorem 1.1

Proof. By symmetry, we have

$$\begin{aligned}
\mathbb{E}\big[C_{n,k}\big] &= n\,\mathbb{P}\left(\mu(B(X_1, R_{1,n,k})) > v_{n,k}\right) \\
&= n\,\mathbb{E}\left[\mathbb{P}\left(\mu(B(X_1, R_{1,n,k}))\right) > v_{n,k} | X_1\right].
\end{aligned}$$

For a fixed $x \in S$, let

$$H_x(r) := \mathbb{P}\left(\rho(x, X) \leq r\right), \quad r \geq 0,$$

be the distribution function of $\rho(x, X)$. Due to the condition (5), the function H_x is continuous, and by the probability integral transform (see e.g. [4], p. 8), the random variable

$$H_x(\rho(x, X)) = \mu\big(B(x, \rho(x, X))\big)$$

is uniformly distributed in the unit interval $[0, 1]$. Put $U_j := H_x(\rho(x, X_{j+1}))$, $j = 1, \ldots, n-1$. Then $U_1, \ldots, U_{n-1}$ are i.i.d. random variables with a uniform distribution in $(0, 1)$. Hence, conditionally on $X_1 = x$, the random variable $\mu(B(X_1, R_{1,n,k}))$ has the same distribution as $U_{k:n-1}$, where $U_{1:n-1} < \ldots < U_{n-1:n-1}$ are the order statistics of $U_1, \ldots, U_{n-1}$, and this distribution does not depend on x. Now, because of a well-known relation between the distribution of order statistics from the uniform distribution on $(0, 1)$ and the binomial distribution (see, e.g., [1], p. 16), we have

$$\mathbb{P}(U_{k:n-1} > s) = \sum_{j=0}^{k-1} \binom{n-1}{j} s^j (1-s)^{n-1-j}$$

and thus

$$\mathbb{E}\big[C_{n,k}\big] = n \sum_{j=0}^{k-1} \binom{n-1}{j} v_{n,k}^j (1-v_{n,k})^{n-1-j}. \tag{7}$$

Here, the summand for $j = k-1$ equals

$$n\binom{n-1}{k-1} v_{n,k}^{k-1}(1-v_{n,k})^{n-k} = \frac{n}{(k-1)!}\,(nv_{n,k})^{k-1} \prod_{i=1}^{k-1} \frac{n-i}{n}\,(1-v_{n,k})^{n-k}.$$

Using Taylor expansions, (3) yields

$$nv_{n,k} = \log n + O\left(\log\log n\right), \quad \prod_{i=1}^{k-1} \frac{n-i}{n} = 1 + O\left(\frac{1}{n}\right)$$

and

$$(1-v_{n,k})^{n-k} = \frac{(k-1)!}{n} \exp\left(-t - (k-1)\log\log n + O\left(\frac{\log^2(n)}{n}\right)\right).$$

Straigthforward computations now give

$$n\binom{n-1}{k-1} v_{n,k}^{k-1}(1-v_{n,k})^{n-k} = \mathrm{e}^{-t} + O\left(\frac{\log\log n}{\log n}\right).$$

Regarding the remaining summands on the right hand side of (7), it is readily seen that

$$\sum_{j=0}^{k-2} \binom{n-1}{j} v_{n,k}^j (1-v_{n,k})^{n-1-j} = O\left(n\binom{n-1}{k-1} v_{n,k}^{k-1}(1-v_{n,k})^{n-k} \cdot \frac{1}{nv_{n,k}}\right),$$

with the convention that the sum equals 0 if $k = 1$. From the above computations and from (3), it follows that this sum equals $O\left(1/\log n\right)$, which concludes the proof of Theorem 1.1. $\square$

Remark 2.1. In the proof given above, we conditioned on the realizations x of X_1. Since the distribution of $H_x(\rho(x, X)) = \mu\big(B(x, \rho(x, X))\big)$ does not depend on X, we obtain as a by-product that

$$\mathbb{P}\left(\mu(B(X_1, R_{1,n,k})) > v_{n,k}\right) = \sum_{j=0}^{k-1} \binom{n-1}{j} v_{n,k}^j (1-v_{n,k})^{n-1-j} \sim \frac{\mathrm{e}^{-t}}{n},$$

if $X_1, X_2, \ldots, X_n$ are independent and $X_2, \ldots, X_n$ are i.i.d. according to μ. Here, X_1 may have an arbitrary distribution and $a_n \sim b_n$ means that $a_n/b_n \to 1$ as $n \to \infty$.

2.2 Proof of Theorem 1.2

The main idea to derive Theorem 1.2 is to discretize supp$(\mu) = [0, 1]^d$ into finitely many "small sets" and then to employ the Chen-Stein method. To apply this method, we will have to check an *asymptotic independence property* and a *local property* which ensures that, with high probability, two exceedances cannot appear in the same neighborhood. We introduce these properties below and recall a result due to Arratia *et al.* [2] on the Chen-Stein method.

The asymptotic independence property Fix $\varepsilon > 0$. Writing $\lfloor\cdot\rfloor$ for the floor function, we partition $[0,1]^d$ into a set $\mathcal{V}_n$ of N_n^d *subcubes* (i.e., subsets that are cubes) of equal size that can only have boundary points in common, where $N_n = \lfloor n/\log(n)^{1+\varepsilon}\rfloor$. The subcubes are indexed by the set $[1, N_n]^d = \{\mathbf{j} := (j_1,\ldots,j_d) : j_m \in \{1,\ldots,N_n\}$ for $m \in \{1,\ldots,d\}\}$. With a slight abuse of notation, we identify a cube with its index. Let

$$\mathcal{E}_n = \bigcap_{\mathbf{j}\in\mathcal{V}_n} \{\mathcal{X}_n \cap \mathbf{j} \neq \emptyset\}$$

be the event that each of the subcubes contains at least one of the points of $\mathcal{X}_n$. The event $\mathcal{E}_n$ is extensively used in stochastic geometry to derive central limit theorems or to deal with extremes [3, 6, 7], and it will play a crucial role throughout the rest of the paper. The following lemma, which captures the idea of "asymptotic independence" , is at the heart of our development.

Lemma 2.2. *For each $\alpha > 0$, we have $\mathbb{P}(\mathcal{E}_n^c) = o(n^{-\alpha})$ as $n \to \infty$.*

Proof. By subadditivity and independence, it follows that

$$\begin{aligned}\mathbb{P}(\mathcal{E}_n^c) &\leq \sum_{\mathbf{j}\in\mathcal{V}_n} \mathbb{P}(\mathcal{X}_n \cap \mathbf{j} = \emptyset)\\ &= \sum_{\mathbf{j}\in\mathcal{V}_n} \big(\mathbb{P}(X_1 \notin \mathbf{j})\big)^n\\ &= \sum_{\mathbf{j}\in\mathcal{V}_n} \big(1-\mu(\mathbf{j})\big)^n\\ &\leq \sum_{\mathbf{j}\in\mathcal{V}_n} \exp(-n\mu(\mathbf{j})).\end{aligned}$$

Here, the last inequality holds since $\log(1-x) \leq -x$ for each $x \in [0,1)$. Since $f \geq f_- > 0$ on K, we have $\mu(\mathbf{j}) = \int_{\mathbf{j}} f \mathrm{d}\lambda \geq f_-\lambda(\mathbf{j})$, whence – writing $\#M$ for the cardinality of a finite set M –

$$\begin{aligned}\mathbb{P}\big(\mathcal{E}_n^c\big) &\leq \sum_{\mathbf{j}\in\mathcal{V}_n} \exp\big(-nf_-\lambda(\mathbf{j})\big)\\ &\leq \#\mathcal{V}_n \exp\big(-f_-(\log n)^{1+\varepsilon}\big).\end{aligned}$$

Since $\#\mathcal{V}_n \leq n/(\log n)^{1+\varepsilon}$, it follows that $n^\alpha\,\mathbb{P}(\mathcal{E}_n^c) \to 0$ as $n \to \infty$. □

The local property We now define a metric d on $\mathcal{V}_n$ by putting $\mathrm{d}(\mathbf{j},\mathbf{j}') := \max_{1\leq s\leq d} |j_s - j_s'|$ for any two different subcubes $\mathbf{j}$ and $\mathbf{j}'$, and $\mathrm{d}(\mathbf{j},\mathbf{j}) := 0$, $\mathbf{j} \in \mathcal{V}_n$. Let $S(\mathbf{j}, r) = \{\mathbf{j}' \in \mathcal{V}_n : \mathrm{d}(\mathbf{j},\mathbf{j}') \leq r\}$ be the ball of subcubes of radius r centered at $\mathbf{j}$. For any $\mathbf{j} \in \mathcal{V}_n$, put

$$M_{\mathbf{j}} := \max_{i\leq n, X_i\in\mathbf{j}} \mu(B(X_i, R_{i,n,k})),$$

with the convention $M_{\mathbf{j}} = 0$ if $\mathcal{X}_n \cap \mathbf{j} = \emptyset$. Conditionally on the event $\mathcal{E}_n$, and provided that $\mathrm{d}(\mathbf{j},\mathbf{j}') \geq 2k+1$, the random variables $M_{\mathbf{j}}$ and $M_{\mathbf{j}'}$ are independent. Lemma 2.2 is referred to as the *asymptotic independence property*: conditionally on the event $\mathcal{E}_n$, which occurs with high probability, the extremes $M_{\mathbf{j}}$ and $M_{\mathbf{j}}'$ attained on two subcubes which are sufficiently distant from each other are independent.

The following lemma claims that, with high probability, two exceedances cannot occur in the same neighborhood.

Lemma 2.3. *With the notation $a \wedge b := \min(a,b)$ for $a, b \in \mathbb{R}$, let*

$$R(n) = \sup_{\mathbf{j}\in\mathcal{V}_n} \sum_{i\neq i'\leq n} \mathbb{P}\big(X_i, X_{i'} \in S(\mathbf{j}, 2k); \mu(B(X_i, R_{i,n,k})) \wedge \mu(B(X_{i'}, R_{i',n,k})) > v_{n,k}\big).$$

Then $R(n) = O(n^{-1}(\log n)^{2-d+\varepsilon})$ as $n \to \infty$.

Here, with a slight abuse of notation, we have identified the family of subcubes $S(\mathbf{j}, 2k) = \{\mathbf{j}' \in \mathcal{V}_n : \mathrm{d}(\mathbf{j},\mathbf{j}') \leq 2k\}$ with the set $\bigcup\{\mathbf{j}' : \mathbf{j}' \in \mathcal{V}_n \text{ and } \mathrm{d}(\mathbf{j},\mathbf{j}') \leq 2k\}$.

We prepare the proof of Lemma 2.3 with the following result that gives the volume of two d-dimensional balls.

Lemma 2.4. *If* $x \in B(0,2)$ *then*

$$\lambda(B(0,1) \cup B(x,1)) = 2\left(\kappa_d\left(1 - \frac{\arccos(\|x\|/2)}{\pi}\right) + \frac{\|x\|\kappa_{d-1}}{2d}\left(\sqrt{1-(\|x\|/2)^2}\right)^{d-1}\right).$$

Proof. We calculate the volume of $\lambda(B(0,1) \cup B(x,1))$ as the sum of the volumes of the following two congruent sets. The first one, say B, is given by the set of all points in $B(0,1) \cup B(x,1)$ that are closer to 0 than to x and for the second one we change the roles of 0 and x. The set B is the union of a cone C with radius $\sqrt{1-(\|x\|/2)^2}$, height $\|x\|/2$ and apex at the origin and a set $D := B(0,1) \setminus S$, where S is a simplicial cone with external angle $\arccos(\|x\|/2)$. From elementary geometry, we obtain that the volumes of C and D are given by

$$\lambda(C) = \frac{\|x\|\kappa_{d-1}}{2d}\left(\sqrt{1-(\|x\|/2)^2}\right)^{d-1}, \quad \lambda(D) = \kappa_d\left(1 - \frac{\arccos(\|x\|/2)}{\pi}\right).$$

This finishes the proof of the lemma. ☐

Proof of Lemma 2.3. For $z \in [0,1]^d$, let

$$r_{n,k}(z) := \inf\{r > 0 : \mu(B(z,r)) > v_{n,k}\}.$$

Writing $\#\mathcal{Y}(A)$ for the number of points of a finite set $\mathcal{Y}$ of random points in $\mathbb{R}^d$ that fall into a Borel set A, we have

$$\mu(B(z, R_{n,k}(z))) > v_{n,k} \quad \Longleftrightarrow \quad \#\mathcal{X}_n(B(z, r_{n,k}(z))) \le k-1.$$

In the following, we assume that $r_{n,k}(X_{i'}) \le r_{n,k}(X_i)$ (which is at the cost of a factor 2) and distinguish the two cases $X_{i'} \in B(X_i, r_{n,k}(X_i))$ and $X_{i'} \in S(\mathbf{j}, 2k) \setminus B(X_i, r_{n,k}(X_i))$. This distinction of cases gives

$$\begin{aligned}
&\mathbb{P}\big(X_i, X_{i'} \in S(\mathbf{j},2k); \mu(B(X_i, R_{i,n,k})) \wedge \mu(B(X_{i'}, R_{i',n,k})) > v_{n,k}\big) \\
&\quad\le 2\mathbb{P}\big(X_i, X_{i'} \in S(\mathbf{j},2k); r_{n,k}(X_{i'}) \le r_{n,k}(X_i); \mu(B(X_i, R_{i,n,k})) \wedge \mu(B(X_{i'}, R_{i',n,k})) > v_{n,k}\big) \\
&\quad\le 2\mathbb{P}\big(X_i \in S(\mathbf{j},2k); X_{i'} \in B(X_i, r_{n,k}(X_i)); r_{n,k}(X_{i'}) \le r_{n,k}(X_i); \\
&\qquad\qquad \mu(B(X_i, R_{i,n,k})) \wedge \mu(B(X_{i'}, R_{i',n,k})) > v_{n,k}\big) &(8)\\
&\qquad + 2\mathbb{P}\big(X_i \in S(\mathbf{j},2k), X_{i'} \in S(\mathbf{j},2k) \setminus B(X_i, r_{n,k}(X_i)); r_{n,k}(X_{i'}) \le r_{n,k}(X_i); \\
&\qquad\qquad \mu(B(X_i, R_{i,n,k})) \wedge \mu(B(X_{i'}, R_{i',n,k})) > v_{n,k}\big). &(9)
\end{aligned}$$

We bound the summands (8) and (9) separately. Since X_i and $X_{i'}$ are independent, (8) takes the form

$$\begin{aligned}
&2\int_{S(\mathbf{j},2k)}\int_{B(x,r_{n,k}(x))} \mathbb{1}\{r_{n,k}(y) \le r_{n,k}(x)\}\, \mathbb{P}(\#(\mathcal{X}_n \setminus \{X_i, X_{i'}\} \cup \{x\})(B(y, r_{n,k}(y))) \le k-1; \\
&\qquad \#(\mathcal{X}_n \setminus \{X_i, X_{i'}\} \cup \{y\})(B(x, r_{n,k}(x))) \le k-1)\, \mu(\mathrm{d}y)\, \mu(\mathrm{d}x). &(10)
\end{aligned}$$

For $y \in B(x, r_{n,k}(x))$, the probability in the integrand figuring above is bounded from above by

$$\begin{aligned}
&\mathbb{P}(\#(\mathcal{X}_n \setminus \{X_i, X_{i'}\})(B(y, r_{n,k}(y))) \le k-1; \\
&\quad \#(\mathcal{X}_n \setminus \{X_i, X_{i'}\})(B(x, r_{n,k}(x))) \le k-2) \\
&\quad\le \mathbb{P}(\#(\mathcal{X}_n \setminus \{X_i, X_{i'}\})(B(y, r_{n,k}(y))) \le k-1; \\
&\qquad \#(\mathcal{X}_n \setminus \{X_i, X_{i'}\})(B(x, r_{n,k}(x)) \setminus B(y, r_{n,k}(y)) \le k-2). &(11)
\end{aligned}$$

Since the random vector

$$(\#(\mathcal{X}_n \setminus \{X_i, X_{i'}\})(B(y, r_{n,k}(y))), \#(\mathcal{X}_n \setminus \{X_i, X_{i'}\})(B(x, r_{n,k}(x)) \setminus B(y, r_{n,k}(y))))$$

is negatively quadrant dependent (see [13, Section 3.1]), Equation (11) has the upper bound

$$\begin{aligned}
&\mathbb{P}(\#(\mathcal{X}_n \setminus \{X_i, X_{i'}\})(B(y, r_{n,k}(y))) \le k-1) \\
&\quad \times \mathbb{P}(\#(\mathcal{X}_n \setminus \{X_i, X_{i'}\})(B(x, r_{n,k}(x)) \setminus B(y, r_{n,k}(y))) \le k-2) \\
&\le \mathbb{P}(\#(\mathcal{X}_n \setminus \{X_i, X_{i'}\})(B(y, r_{n,k}(y))) \le k-1) \\
&\quad \times \mathbb{P}(\#(\mathcal{X}_n \setminus \{X_i, X_{i'}\})(B(x, r_{n,k}(x)) \setminus B(y, r_{n,k}(x))) \le k-2), &(12)
\end{aligned}$$

where the last inequality holds since $r_{n,k}(y) \le r_{n,k}(x)$. Analogously to Remark 2.1, the first probability is

$$\mathbb{P}(\#(\mathcal{X}_n \setminus \{X_i, X_{i'}\})(B(y, r_{n,k}(y))) \le k-1) = \sum_{j=0}^{k-1}\binom{n-2}{j} v_{n,k}^j (1-v_{n,k})^{n-2-j} \sim \frac{\mathrm{e}^{-t}}{n}.$$

The latter probability in (12) is given by

$$\sum_{\ell=0}^{k-2}\binom{n-2}{\ell}\mu\big(B(x,r_{n,k}(x))\setminus B(y,r_{n,k}(x))\big)^{\ell}\big(1-\mu(B(x,r_{n,k}(x))\setminus B(y,r_{n,k}(x)))\big)^{n-2-\ell}. \tag{13}$$

In a next step, we estimate $\mu\big(B(x,r_{n,k}(x))\setminus B(y,r_{n,k}(x))\big)$. Since $f(x)\geq f_->0$, $x\in[0,1]^d$, and by the homogeneity of d-dimensional Lebesgue measure λ, we obtain

$$\begin{aligned}\mu\big(B(x,r_{n,k}(x))\setminus B(y,r_{n,k}(x))\big)&\geq f_-\lambda(B(x,r_{n,k}(x))\setminus B(y,r_{n,k}(x)))\\&=f_-r_{n,k}(x)^d\lambda(B(0,1)\setminus B(r_{n,k}(x)^{-1}(y-x),1))\\&=f_-r_{n,k}(x)^d\left(\lambda(B(0,1)\cup B(r_{n,k}(x)^{-1}(y-x),1))-\kappa_d\right).\end{aligned}$$

For $y\in B(x,r_{n,k}(x))$, Lemma 2.4 yields

$$\begin{aligned}&\mu\big(B(x,r_{n,k}(x))\setminus B(y,r_{n,k}(x))\big)\\&\quad\geq f_-r_{n,k}(x)^d\left(\kappa_d\left(1-\frac{2\arccos(\|x-y\|/2r_{n,k}(x))}{\pi}\right)+\frac{\|x-y\|\kappa_{d-1}}{2dr_{n,k}(x)}\left(\sqrt{1-(\|x-y\|/2r_{n,k}(x))^2}\right)^{d-1}\right).\end{aligned}$$

Since $\inf_{s>0}s^{-1}(1-2\arccos(s)/\pi)>0$, there is $c_0>0$ such that

$$\mu\big(B(x,r_{n,k}(x))\setminus B(y,r_{n,k}(x))\big)\geq c_0\|x-y\|r_{n,k}(x)^{d-1},\quad x\in S(\mathbf{j},2k),\ y\in B(x,r_{n,k}(x)).$$

Equation (13) and the bound $f(x)\leq f_+,\ x\in[0,1]^d$, give

$$\begin{aligned}&\int_{B(x,r_{n,k}(x))}\mathbb{1}\{r_{n,k}(y)\leq r_{n,k}(x)\}\mathbb{P}\big(\#(\mathcal{X}_n\setminus\{X_i,X_{i'}\})(B(x,r_{n,k}(x))\setminus B(y,r_{n,k}(x)))\leq k-1\big)\,\mu(\mathrm{d}y)\\&\quad\leq f_+\sum_{\ell=0}^{k-2}\binom{n-2}{\ell}\int_{B(x,r_{n,k}(x))}\Big(c_0\|x-y\|r_{n,k}(x)^{d-1}\Big)^{\ell}\Big(1-c_0\|x-y\|r_{n,k}(x)^{d-1}\Big)^{n-2-\ell}\lambda(\mathrm{d}y).\end{aligned}$$

We now introduce spherical coordinates and obtain

$$\begin{aligned}&f_+d\kappa_d\sum_{\ell=0}^{k-2}\binom{n-2}{\ell}\int_0^{r_{n,k}(x)}\Big(c_0tr_{n,k}(x)^{d-1}\Big)^{\ell}\Big(1-c_0tr_{n,k}(x)^{d-1}\Big)^{n-2-\ell}t^{d-1}\mathrm{d}t\\&\quad=f_+d\kappa_d\sum_{\ell=0}^{k-2}\binom{n-2}{\ell}\int_0^{r_{n,k}(x)}\Big(c_0tr_{n,k}(x)^{d-1}\Big)^{\ell}\exp\Big((n-2-\ell)\log(1-c_0tr_{n,k}(x)^{d-1})\Big)t^{d-1}\mathrm{d}t\\&\quad\leq f_+d\kappa_d\sum_{\ell=0}^{k-2}\binom{n-2}{\ell}\int_0^{r_{n,k}(x)}\Big(c_0tr_{n,k}(x)^{d-1}\Big)^{\ell}\exp\Big(-c_0(n-2-\ell)tr_{n,k}(x)^{d-1}\Big)t^{d-1}\mathrm{d}t.\end{aligned}$$

Here, the last line follows from the inequality $\log s\leq s-1,\ s>0$. Next, we apply the change of variables

$$t:=(c_0(n-2-\ell))^{-1}r_{n,k}(x)^{1-d}s\quad\Big(\text{i.e., }s=c_0(n-2-\ell)tr_{n,k}(x)^{d-1}\Big),$$

which shows that the last upper bound takes the form

$$f_+d\kappa_dc_0^{-d}r_{n,k}(x)^{d(1-d)}\sum_{\ell=0}^{k-2}\binom{n-2}{\ell}(n-2-\ell)^{-d-\ell}\int_0^{c_0(n-2-\ell)r_{n,k}(x)^d}s^{\ell+d-1}e^{-s}\,\mathrm{d}s. \tag{14}$$

We now use the bounds $f_-\kappa_dr_{n,k}(x)^d\leq v_{n,k}$, $\binom{n-2}{\ell}\leq n^{\ell}/\ell!$, and the fact that the integral figuring in (14) converges as $n\to\infty$. Hence, the expression in (14) is bounded from above by $c_1n^{-1}(\log n)^{1-d}$, where c_1 is some positive constant. Consequently, (8) is bounded from above by

$$\begin{aligned}&c_1n^{-1}(\log n)^{1-d}\lambda(S(\mathbf{j},2k))\sup_{y\in S(\mathbf{j},2k)}\mathbb{P}(\#(\mathcal{X}_n\setminus\{X_i,X_{i'}\})(B(y,r_{n,k}(y)))\leq k-1)\\&\quad\sim c_2n^{-3}(\log n)^{2-d+\varepsilon}\end{aligned} \tag{15}$$

for some $c_2>0$.

By analogy with the reasoning above, (9) is given by the integral

$$2\int_{S(\mathbf{j},2k)}\int_{S(\mathbf{j},2k)\setminus B(x,r_{n,k}(x))} \mathbb{1}\{r_{n,k}(y)\le r_{n,k}(x)\}\,\mathbb{P}\big(\#(\mathcal{X}_n\setminus\{X_i,X_{i'}\}\cup\{x\})(B(y,r_{n,k}(y)))\le k-1\big) \\ \times \mathbb{P}(\#(\mathcal{X}_n\setminus\{X_i,X_{i'}\})(B(x,r_{n,k}(x))\setminus B(y,r_{n,k}(y)))\le k-1)\,\mu(\mathrm{d}y)\,\mu(\mathrm{d}x). \tag{16}$$

If $y\notin B(x,r_{n,k}(x))$ and $r_{n,k}(x)\ge r_{n,k}(y)$, we have the lower bound

$$\lambda(B(x,r_{n,k}(x))\setminus B(y,r_{n,k}(y)))\ge\frac{\lambda(B(x,r_{n,k}(x)))}{2}.$$

Since $f_+\kappa_d r_{n,k}(x)^d\ge v_{n,k}$, we find a constant $c_3>0$ such that

$$\lambda(B(x,r_{n,k}(x))\setminus B(y,r_{n,k}(y)))\ge c_3 v_{n,k},$$

whence

$$\begin{aligned}&\mathbb{P}\big(\#(\mathcal{X}_n\setminus\{X_i,X_{i'}\})(B(x,r_{n,k}(x))\setminus B(y,r_{n,k}(y)))\le k-1\big)\\ &\le\sum_{\ell=0}^{k-1}\binom{n-2}{\ell}\big(c_3v_{n,k}\big)^{\ell}\big(1-c_3v_{n,k}\big)^{n-2-\ell}\\ &\sim\frac{c_3^{k-1}}{(k-1)!}\big(\log n\big)^{k-1}\exp\big(n\log(1-c_3v_{n,k})\big)\end{aligned}$$

as $n\to\infty$. Since $\log s\le s-1$ for $s>0$, (16) is bounded from above by

$$\begin{aligned}&c_4n^{-c_3}\lambda(S(\mathbf{j},2k))^2\sup_{y\in S(\mathbf{j},2k)}\mathbb{P}(\#(\mathcal{X}_n\setminus\{X_i,X_{i'}\})(B(y,r_{n,k}(y)))\le k-1)\\ &\sim c_5(4k+1)^{2d}\frac{(\log n)^{2+2\varepsilon}}{n^{3+c_3}},\end{aligned} \tag{17}$$

where c_4 and c_5 are positive constants. Summing over all $i\neq i'\le n$, it follows from (15) and (17) that $R(n)=O(n^{-1}(\log n)^{2-d+\varepsilon})$ as $n\to\infty$, which finishes the proof of Lemma 2.3. □

A Poisson approximation result based on the Chen-Stein method In this paragraph, we recall a Poisson approximation result due to Arratia *et al.* [2], which is based on the Chen-Stein method. To this end, we consider a finite or countable collection $(Y_\alpha)_{\alpha\in I}$ of $\{0,1\}$-valued random variables and we let $p_\alpha=\mathbb{P}(Y_\alpha=1)>0$, $p_{\alpha\beta}=\mathbb{P}(Y_\alpha=1,Y_\beta=1)$. Moreover, suppose that for each $\alpha\in I$, there is a set $B_\alpha\subset I$ that contains α. The set B_α is regarded as a neighborhood of α that consists of the set of indices β such that Y_α and Y_β are *not* independent. Finally, put

$$b_1=\sum_{\alpha\in I}\sum_{\beta\in B_\alpha}p_\alpha p_\beta,\quad b_2=\sum_{\alpha\in I}\sum_{\alpha\neq\beta\in B_\alpha}p_{\alpha\beta},\quad b_3=\sum_{\alpha\in I}\mathbb{E}\big[\,|\mathbb{E}[Y_\alpha-p_\alpha|\sigma(Y_\beta:\beta\notin B_\alpha)]|\,\big]. \tag{18}$$

Theorem 2.5. (*Theorem 1 of [2]*) *Let $W=\sum_{\alpha\in I}Y_\alpha$, and assume $\lambda:=\mathbb{E}(W)\in(0,\infty)$. Then*

$$d_{TV}(W,\mathrm{Po}(\lambda))\le 2(b_1+b_2+b_3).$$

Proof of Theorem 1.2 Recall $v_{n,k}$ from (3) and $C_{n,k}$ from (4). Put

$$\widehat{C}_{n,k}:=\sum_{\mathbf{j}\in\mathcal{V}_n}\mathbb{1}\big\{M_{\mathbf{j}}>v_{n,k}\big\}.$$

The following lemma claims that the number $C_{n,k}$ of exceedances is close to the number of subcubes for which there exists at least one exceedance, i.e. $\widehat{C}_{n,k}$, and that $\widehat{C}_{n,k}$ can be approximated by a Poisson random variable.

Lemma 2.6. *We have*

a) $\mathbb{P}(C_{n,k}\neq\widehat{C}_{n,k})=O\Big((\log n)^{1-d}\Big)$,

b) $d_{TV}(\widehat{C}_{n,k},\mathrm{Po}(\mathbb{E}[\widehat{C}_{n,k}]))\ =O\Big((\log n)^{1-d}\Big)$,

c) $\mathbb{E}[\widehat{C}_{n,k}] = \mathrm{e}^{-t} + O\left(\frac{\log\log n}{\log n}\right)$.

Proof. Assertion a) is a direct consequence of Lemma 2.3 and of the inequalities

$$\begin{aligned}\mathbb{P}(C_{n,k} \neq \widehat{C}_{n,k}) &= \mathbb{P}\big(\exists \mathbf{j} \in \mathcal{V}_n, \exists i, \ell \text{ s.t. } X_i, X_\ell \in \mathbf{j};\ \mu(B(X_i, R_{i,n,k})) \wedge \mu(B(X_\ell, R_{\ell,n,k})) > v_{n,k}\big)\\ &\leq \sum_{\mathbf{j}\in\mathcal{V}_n} \sum_{i\neq \ell \leq n} \mathbb{P}\left(\, X_i, X_\ell \in \mathbf{j};\ \mu(B(X_i, R_{i,n,k})) \wedge \mu(B(X_\ell, R_{\ell,n,k})) > v_{n,k}\,\right)\\ &\leq \frac{n}{(\log n)^{1+\varepsilon}} \times R(n).\end{aligned}$$

To prove b), we apply Theorem 2.5 to the collection $(Y_\alpha)_{\alpha\in I} = (M_\mathbf{j})_{\mathbf{j}\in\mathcal{V}_n}$. Recall that, conditionally on the event $\mathcal{E}_n$, the random variables $M_\mathbf{j}$ and $M_{\mathbf{j}'}$ are independent provided that $\mathrm{d}(\mathbf{j}, \mathbf{j}') \geq 2k+1$. With a slight abuse of notation, we omit to condition on $\mathcal{E}_n$ since this event occurs with probability tending to 1 as $n \to \infty$ (Lemma 2.2) at a rate which is at least polynomial. The first two terms in (18) are

$$b_1 = \sum_{\mathbf{j}\in\mathcal{V}_n} \sum_{\mathbf{j}'\in S(\mathbf{j},2k)} p_\mathbf{j} p_{\mathbf{j}'}, \quad b_2 = \sum_{\mathbf{j}\in\mathcal{V}_n} \sum_{\mathbf{j}\neq\mathbf{j}'\in S(\mathbf{j},2k)} p_{\mathbf{j}\mathbf{j}'},$$

where

$$p_\mathbf{j} = \mathbb{P}(M_\mathbf{j} > v_{n,k}), \quad p_{\mathbf{j}\mathbf{j}'} = \mathbb{P}(M_\mathbf{j} > v_{n,k}, M_{\mathbf{j}'} > v_{n,k}).$$

The term b_3 figuring in (18) equals 0 since, conditionally on $\mathcal{E}_n$, the random variable $M_\mathbf{j}$ is independent of the σ-field $\sigma(M_{\mathbf{j}'} : \mathbf{j}' \notin S(\mathbf{j}, 2k))$. Thus, according to Theorem 2.5, we have

$$d_{TV}(\widehat{C}_{n,k}, \mathrm{Po}(\mathbb{E}[\widehat{C}_{n,k}])) \leq 2(b_1 + b_2).$$

Fist, we deal with b_1. As for the first assertion, notice that for each $\mathbf{j} \in \mathcal{V}_n$, using symmetry, we obtain

$$\begin{aligned}p_\mathbf{j} &= \mathbb{P}\bigg(\bigcup_{i\leq n} \{X_i \in \mathbf{j}, \mu(B(X_i, R_{i,n,k})) > v_{n,k}\}\bigg)\\ &\leq n \cdot \mathbb{P}\big(X_1 \in \mathbf{j}, \mu(B(X_1, R_{1,n,k})) > v_{n,k}\big)\\ &= n \cdot \int_\mathbf{j} \mathbb{P}(\mu(B(x, R_{1,n,k})) > v_{n,k} | X_1 = x) f(x)\,\mathrm{d}x\\ &\leq n f^+ \lambda(\mathbf{j}) \int_\mathbf{j} \mathbb{P}(\mu(B(x, R_{1,n,k})) > v_{n,k} | X_1 = x)\, \frac{1}{\lambda(\mathbf{j})}\,\mathrm{d}x\\ &= n f^+ \lambda(\mathbf{j}) \mathbb{P}(\mu(B(\widetilde{X}_1, R_{1,n,k})) > v_{n,k}),\end{aligned}$$

where $\widetilde{X}_1$ is independent of $X_2, \ldots, X_n$ and has a uniform distribution over $\mathbf{j}$. Invoking Remark 2.1, the probability figuring in the last line is asymptotically equal to e^{-t}/n as $n \to \infty$. Since $\lambda(\mathbf{j}) = O\left((\log n)^{1+\varepsilon}/n\right)$, we thus have

$$p_\mathbf{j} \leq C \cdot \frac{(\log n)^{1+\varepsilon}}{n},$$

where C is a constant that does not depend on $\mathbf{j}$. Since $\#\mathcal{V}_n \leq \frac{n}{(\log n)^{1+\varepsilon}}$ and $\#S(\mathbf{j}, 2k) \leq (4k+1)^d$, summing over $\mathbf{j}, \mathbf{j}'$ gives

$$b_1 \leq C^2 \sum_{\mathbf{j}\in\mathcal{V}_n} \sum_{\mathbf{j}'\in S(\mathbf{j},2k)} \left(\frac{(\log n)^{1+\varepsilon}}{n}\right)^2 = O\left(\frac{(\log n)^{1+\varepsilon}}{n}\right).$$

To deal with b_2, notice that, for each $\mathbf{j}, \mathbf{j}' \in \mathcal{V}_n$ and $\mathbf{j}' \in S(\mathbf{j}, 2k)$, we have

$$\begin{aligned}p_{\mathbf{j}\mathbf{j}'} &= \mathbb{P}\Big(\bigcup_{i\neq i'\leq n} \{X_i \in \mathbf{j}, X_{i'} \in S(\mathbf{j}, 2k), \mu(B(X_i, R_{i,n,k})) \wedge \mu(B(X_{i'}, R_{i',n,k})) > v_{n,k}\}\Big)\\ &\leq \mathbb{P}\Big(\bigcup_{i\neq i'\leq n} \{X_i, X_{i'} \in S(\mathbf{j}, 2k); \mu(B(X_i, R_{i,n,k})) \wedge \mu(B(X_{i'}, R_{i',n,k})) > v_{n,k}\}\Big).\end{aligned}$$

Using subadditivity, and taking the supremum, we obtain

$$\begin{aligned} b_2 &\le \sum_{\mathbf{j}\in\mathcal{V}_n} \sum_{\mathbf{j}'\in S(\mathbf{j},2k)} \sup_{\mathbf{j}\in\mathcal{V}_n} \sum_{i\neq i'\le n} \mathbb{P}\big(X_i, X_{i'} \in S(\mathbf{j},2k), \mu(B(X_i,R_{i,n,k})) \wedge \mu(B(X_{i'},R_{i',n,k})) > v_{n,k}\big) \\ &\le \frac{n}{(\log n)^{1+\varepsilon}} \times (4k+1)^d \times R(n). \end{aligned}$$

According to Lemma 2.3, the last term equals $O\left((\log n)^{1-d}\right)$, which concludes the proof of b).

To prove c), observe that

$$\left|\mathbb{E}[\widehat{C}_{n,k}] - \mathrm{e}^{-t}\right| \le \left|\mathbb{E}[\widehat{C}_{n,k}] - \mathbb{E}[C_{n,k}]\right| + \left|\mathbb{E}[C_{n,k}] - \mathrm{e}^{-t}\right|.$$

By Theorem 1.1, the last summand is $O\left(\frac{\log\log n}{\log n}\right)$. Since $C_{n,k} \ge \widehat{C}_{n,k}$, we further have

$$\begin{aligned} \left|\mathbb{E}[\widehat{C}_{n,k}] - \mathbb{E}[C_{n,k}]\right| &= \mathbb{E}[C_{n,k} - \widehat{C}_{n,k}] \\ &= \mathbb{E}\Big(\sum_{i\le n} \mathbb{1}\{\mu(B(X_i,R_{i,n,k})) > v_{n,k}\} - \sum_{\mathbf{j}\in\mathcal{V}_n} \mathbb{1}\{M_{\mathbf{j}} > v_{n,k}\}\Big) \\ &= \sum_{\mathbf{j}\in\mathcal{V}_n} \mathbb{E}\Big[\Big(\sum_{i\le n} \mathbb{1}\{X_i\in\mathbf{j}\}\mathbb{1}\{\mu(B(X_i,R_{i,n,k})) > v_{n,k}\} - 1\Big)\mathbb{1}\{M_{\mathbf{j}} > v_{n,k}\}\Big] \\ &\le \sum_{\mathbf{j}\in\mathcal{V}_n} \sum_{i\neq i'\le n} \mathbb{P}\big(X_i, X_{i'}\in\mathbf{j}, \mu(B(X_i,R_{i,n,k})), \mu(B(X_{i'},R_{i',n,k})) > v_{n,k}\big) \\ &\le \#\mathcal{V}_n \times R(n). \end{aligned}$$

According to Lemma 2.3, the last term equals $O\left((\log n)^{1-d}\right)$. This concludes the proof of Lemma 2.6 and thus of Theorem 1.2.

□

3 Concluding remarks

When dealing with limit laws for large kth-nearest neighbor *distances* of a sequence of i.i.d. random points in $\mathbb{R}^d$ with density f, which take values in a bounded region K, the modification of the kth-nearest neighbor distances made in (6) (by introducing the "boundary distances" $\|X_i - \partial K\|$) and the condition that f is bounded away from zero, which have been adopted in [11] and [12], seem to be crucial, since boundary effects play a decisive role ([8, 9]). Regarding kth-nearest neighbor balls with *large probability volume*, there is no need to introduce $\|X_i - \partial K\|$. It is an open problem, however, whether Theorem 1.2 continues to hold for densities that are not bounded away from zero.

A second open problem refers to Theorem 1.1, which states convergence of expectations of $C_{n,k}$ in a setting beyond the finite-dimensional case. Since $C_{n,k}$ is non-negative, the sequence $(C_{n,k})_k$ is tight by Markov's inequality. Can one find conditions on the underlying distribution that ensure convergence in distribution to some random element of the metric space?

References

[1] Ahsanullah, M., Nevzzorov, V.B., and Shakil, M. (2013). An Introduction to Order Statistics. *Atlantis Press. Amsterdam, Paris, Beijing.*

[2] Arratia, R., Goldstein, L., and Gordon, L., (1990). Poisson approximation and the Chen-Stein method. *Statist. Sci.* 5: 403–434.

[3] Avram, F. and Bertsimas, D., (1993). On central limit theorem in geometrical probability. *Ann. Appl. Probab.* 3(4):1033–1046, 1993.

[4] Biau, G. and Devroye, L. (2015). Lectures on the Nearest Neighbor Method. *Springer, New York.*

[5] Bobrowski, O., Schulte, M., and Yogeshwaran, D. (2021). Poisson process approximation under stabilization and Palm coupling. *Available at https://arxiv.org/abs/2104.13261.*

[6] Bonnet, G. and Chenavier, N., (2020). The maximal degree in a Poisson-Delaunay graph. *Bernoulli* 26: 948–979.

[7] Chenavier, N. and Robert, C. Y. (2018). Cluster size distributions of extreme values for the Poisson-Voronoi tessellation. *Ann. Appl. Probab.* 28(6): 3291–3323.

[8] Dette, H. and Henze, N. (1989). The limit distributionof the largest nearest-neighbour link in the unit d-cube. *J. Appl. Probab.* 26:67–80.

[9] Dette, H. and Henze, N. (1990). Some peculiar boundary phenomena for extremes of rth nearest neighbor links. *Statist. & Prob. Lett.* 10:381–390.

[10] Györfi, L., Henze, N., and Walk, H. (2019). The limit distribution of the maximum probability nearest neighbor ball. *J. Appl. Probab.* 56: 574–589.

[11] Henze, N. (1982). The limit distribution for maxima of "weighted" rth-nearest-neighbour distances. *J. Appl. Probab.* 19:344–354.

[12] Henze, N. (1983). Ein asymptotischer Satz über den maximalen Minimalabstand von unabhängigen Zufallsvektoren mit Anwendung auf einen Anpassungstest im $\mathbb{R}^p$ und auf der Kugel. [An asymptotic theorem on the maximum minimum distance of independent random vectors, with application to a goodness-of-fit test in $\mathbb{R}^p$ and on the sphere]. *Metrika* 30:245–259 (in German).

[13] Joag-Dev, K. and Proschan, F. (1983). Negative association of random variables with applications. *Ann. Stat.* 11:286–295.

[14] Last, G. and Penrose, M. D. (2017). Lectures on the Poisson Process. Cambridge University Press (IMS Textbook).

[15] Otto, M. (2020). Poisson approximation of Poisson-driven point processes and extreme values in stochastic geometry. *Available at https://arxiv.org/pdf/2005.10116.pdf.*

[16] Penrose, M. D. (1997). The longest edge of the random minimal spanning tree. *Ann. Appl. Probab.* 7(2): 340-361.

[17] Zubkov, A. N. and Orlov, O. P. (2018). Limit distributions of extremal distances to the nearest neighbor. *Discrete Math. Appl.* 28:189–199.