Universität Stuttgart

The limit distribution of the maximum probability nearest neighbor ball

László Györfi
Norbert Henze
Harro Walk

Stuttgarter Mathematische Berichte

2019-002

mathematik

Fachbereich Mathematik
Fakultät Mathematik und Physik
Universität Stuttgart
Pfaffenwaldring 57
D-70 569 Stuttgart

E-Mail: preprints@mathematik.uni-stuttgart.de
WWW: http://www.mathematik.uni-stuttgart.de/preprints

ISSN **1613-8309**

The limit distribution of the maximum probability nearest neighbor ball

László Györfi* Norbert Henze† Harro Walk‡

November 17, 2018

Abstract

Let $X_1, \dots, X_n$ be independent random points drawn from an absolutely continuous probability measure with density f in $\mathbb{R}^d$. Under mild conditions on f, we derive a Poisson limit theorem for the number of large probability nearest neighbor balls. Denoting by P_n the maximum probability measure of nearest neighbor balls, this limit theorem implies a Gumbel extreme value distribution for $nP_n - \ln n$ as $n \to \infty$. Moreover, we derive a tight upper bound on the upper tail of the distribution of $nP_n - \ln n$, which does not depend on f.

Keywords: Nearest neighbors; Gumbel extreme value distribution; Poisson limit theorem; exchangeable events

2010 AMS subject classifications: Primary 60F05; Secondary 60G09, 60G70.

*Budapest University of Technology and Economics, gyorfi@cs.bme.hu
†Karlsruhe Institute of Technology, henze@kit.edu
‡Universität Stuttgart, walk@mathematik.uni-stuttgart.de

1 Introduction

Let $X, X_1, \ldots, X_n, \ldots$ be independent, identically distributed (i.i.d.) random vectors taking values in $\mathbb{R}^d$. We assume throughout the paper that the distribution of X, which is denoted by μ, has a density f with respect to Lebesgue measure λ.

Writing $\|\cdot\|$ for the Euclidean norm on $\mathbb{R}^d$, put

$$R_{i,n} := \min_{j \neq i, j \leq n} \|X_i - X_j\|,$$

and let

$$P_n := \max_{1 \leq i \leq n} \mu\{S(X_i, R_{i,n})\}$$

denote the maximum probability of the nearest neighbor (NN) balls, where $S(x, r) := \{y \in \mathbb{R}^d : \|y - x\| \leq r\}$ stands for the closed ball with center x and radius r. This paper deals with both the finite-sample and the asymptotic distribution of

$$nP_n - \ln n,$$

as $n \to \infty$.

There is a huge related literature for Poisson sample size. Let N be a random variable that is independent of $X_1, X_2, \ldots$ and has a Poisson distribution with $\mathbb{E}(N) = n$. Then

$$X_1, \ldots, X_N \tag{1}$$

is a non-homogeneous Poisson process with intensity function nf. For the nucleuses $X_1, \ldots, X_N$, $\widetilde{A}_n(X_j)$ denotes the Voronoi cell around X_j, and $\widehat{r}_j$ and $\widehat{R}_j$ stand for the inscribed and circumscribed radii of $\widetilde{A}_n(X_j)$, respectively, i.e., we have

$$\widehat{r}_j = \sup\{r > 0 : S(X_j, r) \subset \widetilde{A}_n(X_j)\}$$

and

$$\widehat{R}_j = \inf\{r > 0 : \widetilde{A}_n(X_j) \subset S(X_j, r)\}.$$

If $X_1, X_2, \ldots$ are i.i.d. uniformly distributed on the unit cube $[0, 1]^d$, then (2a) and (2c) of Theorem 1 in Calka and Chenavier [3] read

$$\lim_{n\to\infty} \mathbb{P}\left(2^d n\lambda\left\{S\left(0, \max_{1\leq j\leq N} \widehat{r}_j\right)\right\} - \ln n \leq y\right) = G(y)$$

and

$$\lim_{n\to\infty} \mathbb{P}\left(n\lambda\left\{S\left(0, \max_{1\le j\le N} \widehat{R}_j)\right)\right\} - \ln\left(\alpha_d n\,(\ln n)^{d-1}\right) \le y\right) = G(y),$$

$y \in \mathbb{R}$. Here, $\alpha_d > 0$ is a universal constant, and

$$G(y) = \exp(-\exp(-y))$$

denotes the distribution function of the Gumbel extreme value distribution.

The paper is organized as follows. In Section 2 we study the distribution of $nP_n - \ln n$. Theorem 1 is on a universal and tight bound on the upper tail of $nP_n - \ln n$. Under mild conditions on the density, Theorem 2 shows that the number of exceedances of nearest neighbor ball probabilities over a certain sequence of thresholds has an asymptotic Poisson distribution as $n \to \infty$. As a consequence, the limit distribution of $nP_n - \ln n$ is the Gumbel extreme value distribution. Theorem 3 in Section 3 is the extension of Theorem 1 for Poisson sample size. All proofs are presented in Section 4. The main tool for proving Theorem 2 is a novel Poisson limit theorem for sums of indicators of exchangeable events, which is formulated as Proposition 1. The final section sheds some light on a technical condition on f that is used in the proof of the main result.

Although there is a weak dependence between the probabilities of nearest neighbor balls, a main message of this paper is that one can neglect this dependence when looking for the limit distribution of the maximum probability.

2 The maximum nearest neighbor ball

Under the assumption that the density f is sufficiently smooth and bounded away from zero, Henze [7] and [8] derived the limit distribution of the maximum *approximate* probability measure

$$\max_{1\le i\le n} f(X_i) R_{i,n}^d v_d \tag{2}$$

of NN-balls. Here, $v_d = \pi^{d/2}/\Gamma(1 + d/2)$ stands for the volume of the unit ball in $\mathbb{R}^d$.

In the following, we consider the number of points among $X_1, \ldots, X_n$ for which the probability content of the nearest neighbor ball exceeds some (large) threshold. To be more specific, we fix $y \in \mathbb{R}$ and consider the random variable

$$C_n := \sum_{i=1}^{n} \mathbb{I}\big\{n\mu\{S(X_i, R_{i,n})\} > y + \ln n\big\},$$

where $\mathbb{I}\{\cdot\}$ denotes the indicator function. Writing "$\xrightarrow{\mathcal{D}}$" for convergence in distribution, we will show that, under some conditions on the density f,

$$C_n \xrightarrow{\mathcal{D}} Z \quad \text{as } n \to \infty,$$

where Z is a random variable with the Poisson distribution $\mathrm{Po}(\exp(-y))$. Now, $C_n = 0$ if, and only if, $nP_n - \ln n \le y$, and it follows that

$$\lim_{n\to\infty} \mathbb{P}\left(nP_n - \ln n \le y\right) = \mathbb{P}(Z = 0) = G(y), \quad y \in \mathbb{R}. \tag{3}$$

Since $1 - G(y) \le \exp(-y)$ if $y \ge 0$, (3) implies

$$\limsup_{n\to\infty} \mathbb{P}\left(nP_n - \ln n \ge y\right) \le \mathrm{e}^{-y}, \quad y \ge 0. \tag{4}$$

Our first result is a non-asymptotic upper bound on the upper tail of the distribution of $nP_n - \ln n$. This bound holds without any condition on the density and thus entails (4) universally.

Theorem 1 *Without any restriction on the density f, we have*

$$\mathbb{P}\left(nP_n - \ln n \ge y\right) \le \exp\left(-\frac{n-1}{n}y + \frac{\ln n}{n}\right)\mathbb{I}\{y \le n - \ln n\}, \quad y \in \mathbb{R}. \tag{5}$$

Theorem 1 implies a non-asymptotic upper bound on the mean of $nP_n - \ln n$, since

$$\begin{aligned}
\mathbb{E}\left[nP_n - \ln n\right] &\le \mathbb{E}\left[(nP_n - \ln n)^+\right] \\
&= \int_0^\infty \mathbb{P}\left(nP_n - \ln n \ge y\right) \mathrm{d}y \\
&\le \int_0^\infty \exp\left(-\frac{n-1}{n}y + \frac{\ln n}{n}\right) \mathrm{d}y \\
&= \frac{n}{n-1}\exp\left(\frac{\ln n}{n}\right).
\end{aligned}$$

Notice that this upper bound approaches 1 for large n, and that the mean of the standard Gumbel distribution is the Euler-Mascheroni constant, which is $-\int_0^\infty \mathrm{e}^{-y} \ln y \, \mathrm{d}y = 0.5772\ldots$

Recall that the support of μ is defined by

$$\mathrm{supp}(\mu) := \{x \in \mathbb{R}^d : \mu\{S(x,r)\} > 0 \text{ for each } r > 0\},$$

i.e., the support of μ is the smallest closed set in $\mathbb{R}^d$ having μ-measure one.

Theorem 2 *Assume there are $\beta \in (0,1)$, $c_{max} < \infty$ and $\delta > 0$ such that, for any $r, s > 0$ and any $x, z \in supp(\mu)$ with $\|x - z\| \geq \max\{r, s\}$ and $\mu(S(x,r)) = \mu(S(z,s)) \leq \delta$, one has*

$$\frac{\mu(S(x,r) \cap S(z,s))}{\mu(S(z,s))} \leq \beta \tag{6}$$

and

$$\mu(S(z,2s)) \leq c_{max}\mu(S(z,s)). \tag{7}$$

Then

$$\sum_{i=1}^{n} \mathbb{I}\{n\mu\{S(X_i, R_{i,n})\} > y + \ln n\} \xrightarrow{\mathcal{D}} \mathrm{Po}(\exp(-y)), \quad y \in \mathbb{R}, \tag{8}$$

and hence

$$\lim_{n\to\infty} \mathbb{P}(nP_n - \ln n \leq y) = G(y), \quad y \in \mathbb{R}. \tag{9}$$

Remark 1 *It is easy to see that (6) and (7) hold if the density is both bounded from above by f_{max} and bounded away from zero by $f_{min} > 0$. Indeed, putting*

$$\beta := 1 - \frac{1}{2} \cdot \frac{f_{min}}{f_{max}}, \quad c_{max} := 2^d \cdot \frac{f_{max}}{f_{min}},$$

we have

$$\begin{aligned}\frac{\mu(S(x,r) \cap S(z,s))}{\mu(S(z,s))} &= 1 - \frac{\mu(S(z,s) \setminus S(x,r))}{\mu(S(z,s))} \\ &\leq 1 - \frac{f_{min}\,\lambda(S(z,s) \setminus S(x,r))}{f_{max}\,\lambda(S(z,s))} \\ &\leq \beta\end{aligned}$$

and

$$\begin{aligned}\mu\left(S(z,2s)\right) &\leq f_{max}\,\lambda\left(S(z,2s)\right)\\ &= f_{max}\,2^d\lambda\left(S(z,s)\right)\\ &\leq c_{max}\,\mu\left(S(z,s)\right).\end{aligned}$$

A challenging problem left is to weaken the conditions of Theorem 2 or to prove that (8) and (9) hold without any conditions on the density. We believe that such universal limit results are possible, because the summands in (8) are identically distributed, and their distribution does not depend on the actual density. More discussion on condition (6) is given in Section 5.

3 The maximum nearest neighbor ball for a non-homogeneous Poisson process

In this section we consider the non-homogeneous Poisson process $X_1,\ldots,X_N$ defined by (1). Putting

$$\widetilde{R}_{i,n} := \min_{j\neq i, j\leq N} \|X_i - X_j\|$$

and

$$\widetilde{P}_n = \max_{1\leq i\leq N} \mu\{S(X_i,\widetilde{R}_{i,n})\},$$

the following result is the Poisson-analogue to Theorem 1.

Theorem 3 *Without any restriction on the density f we have*

$$\mathbb{P}\left(n\widetilde{P}_n - \ln n \geq y\right) \leq \mathrm{e}^{-y}\exp\left(\frac{(y+\ln n)^2}{n}\right), \quad y\in\mathbb{R}.$$

4 Proofs

Proof of Theorem 1. Since the right hand side of (5) is larger than 1 if $y<0$, we take $y\geq 0$ in what follows. Moreover, in view of $P_n\leq 1$ the left hand side of (5) vanishes if $y>n-\ln n$. We therefore assume without loss of generality that

$$\frac{y+\ln n}{n}\leq 1. \tag{10}$$

For a fixed $x \in \mathbb{R}^d$, let

$$H_x(r) := \mathbb{P}\left(\|x - X\| \le r\right), \quad r \ge 0, \tag{11}$$

be the distribution function of $\|x-X\|$. By the probability integral transform (cf. Biau and Devroye [1], p. 8), the random variable

$$H_x(\|x - X\|) = \mu\{S(x, \|x - X\|)\}$$

is uniformly distributed on $[0, 1]$. We thus have

$$\mu\left\{S(x, H_x^{-1}(p))\right\} = p, \quad 0 < p < 1, \tag{12}$$

where $H_x^{-1}(p) = \inf\{r : H_x(r) \ge p\}$. It follows that

$$\begin{aligned}
\mathbb{P}\left(nP_n - \ln n \ge y\right) &= \mathbb{P}\left(n \max_{1\le i\le n} \mu\{S(X_i, R_{i,n})\} - \ln n \ge y\right) \\
&\le n\,\mathbb{P}\left(n\mu\{S(X_1, R_{1,n})\} - \ln n \ge y\right) \\
&= n\,\mathbb{P}\left(\mu\{S(X_1, R_{1,n})\} \ge \frac{y + \ln n}{n}\right) \\
&= n\,\mathbb{P}\left(\min_{2\le j\le n} \mu\{S(X_1, \|X_1 - X_j\|)\} \ge \frac{y + \ln n}{n}\right).
\end{aligned}$$

Now, (10) implies

$$\begin{aligned}
\mathbb{P}\left(nP_n - \ln n \ge y\right) &\le n\,\mathbb{E}\left[\mathbb{P}\left(\min_{2\le j\le n} \mu\{S(X_1, \|X_1 - X_j\|)\} \ge \frac{y + \ln n}{n}\Big| X_1\right)\right] \\
&= n\left(1 - \frac{y + \ln n}{n}\right)^{n-1} \\
&\le n \exp\left(-\frac{(y + \ln n)(n-1)}{n}\right) \\
&= \exp\left(-\frac{n-1}{n}y + \frac{\ln n}{n}\right).
\end{aligned}$$

□

Proof of Theorem 3. We again assume (10) in what follows. By conditioning on N, we have

$$\mathbb{P}\left(n\widetilde{P}_n - \ln n \geq y\right) = \sum_{k=1}^{\infty} \mathbb{P}\left(n\widetilde{P}_n - \ln n \geq y \mid N = k\right) \mathbb{P}(N = k)$$
$$= \sum_{k=1}^{\infty} \mathbb{P}(nP_k - \ln n \geq y) \, \mathbb{P}(N = k) \,.$$

Putting $y_n := (y + \ln n)/n$, we obtain

$$\mathbb{P}(nP_k - \ln n \geq y) = \mathbb{P}(kP_k - \ln k \geq ky_n - \ln k) \,,$$

and Theorem 1 implies

$$\mathbb{P}(kP_k - \ln k \geq ky_n - \ln k) \leq \exp\left(-\frac{k-1}{k}(ky_n - \ln k) + \frac{\ln k}{k}\right)$$
$$= \exp(-(k-1)y_n + \ln k) \,.$$

It follows that

$$\begin{aligned}
\mathbb{P}\left(n\widetilde{P}_n - \ln n \geq y\right) &\leq \sum_{k=1}^{\infty} \exp(-(k-1)y_n + \ln k) \, \mathbb{P}(N = k) \\
&= \mathrm{e}^{y_n - n} \sum_{k=1}^{\infty} k \left(\mathrm{e}^{-y_n}\right)^k \frac{n^k}{k!} \\
&= \mathrm{e}^{y_n - n} \sum_{k=1}^{\infty} \frac{(n\mathrm{e}^{-y_n})^k}{(k-1)!} \\
&= n\mathrm{e}^{y_n - n - y_n} \exp\left(n\mathrm{e}^{-y_n}\right) \\
&= n \exp(-n(1 - \mathrm{e}^{-y_n})).
\end{aligned}$$

Since $z \geq 0$ entails $\mathrm{e}^{-z} \leq 1 - z + z^2$, we finally obtain

$$\mathbb{P}\left(n\widetilde{P}_n - \ln n \geq y\right) \leq n \exp\left(-n(y_n - y_n^2)\right) = \mathrm{e}^{-y} \exp\left(\frac{(y + \ln n)^2}{n}\right).$$

□

The main tool in the proof Theorem 2 is the following result.

Proposition 1 *For each $n \geq 2$, let $A_{n,1}, \ldots, A_{n,n}$ be exchangeable events, and let*

$$Y_n := \sum_{j=1}^{n} \mathbb{I}\{A_{n,j}\}.$$

If, for some $\nu \in (0, \infty)$,

$$\lim_{n\to\infty} n^k \mathbb{P}\left(A_{n,1} \cap \ldots \cap A_{n,k}\right) = \nu^k \quad \textit{for each} \;\; k \geq 1, \tag{13}$$

then

$$Y_n \xrightarrow{\mathcal{D}} Y \qquad \textit{as} \;\; n \to \infty,$$

where Y has the Poisson distribution $\mathrm{Po}(\nu)$.

Proof. The proof uses the method of moments, see, e.g., [2], Section 30. Putting

$$S_{n,k} = \sum_{1\leq i_1 < \cdots < i_k \leq n} \mathbb{P}\left(A_{n,i_1} \cap \ldots \cap A_{n,i_k}\right), \quad k \in \{1, \ldots, n\},$$

and writing $Z^{(k)} = Z(Z-1)\cdots(Z-k+1)$ for the kth descending factorial of a random variable Z, we have

$$\mathbb{E}\left[Y_n^{(k)}\right] = k!\, S_{n,k}.$$

Since $A_{n,1}, \ldots, A_{n,n}$ are exchangeable, (13) implies

$$\lim_{n\to\infty} \mathbb{E}\left[Y_n^{(k)}\right] = \nu^k, \quad k \geq 1.$$

Now, $\nu^k = \mathbb{E}\left[Y^{(k)}\right]$, where Y has the Poisson distribution $\mathrm{Po}(\nu)$. We thus have

$$\lim_{n\to\infty} \mathbb{E}\left[Y_n^{(k)}\right] = \mathbb{E}\left[Y^{(k)}\right], \quad k \geq 1. \tag{14}$$

Since

$$Y_n^k = \sum_{j=0}^{k} \left\{ {k \atop j} \right\} Y_n^{(j)},$$

where $\left\{ {k \atop 0} \right\}, \ldots, \left\{ {k \atop k} \right\}$ denote Stirling numbers of the second kind (see, e.g., [5], p. 262), (14) entails $\lim_{n\to\infty} \mathbb{E}[Y_n^k] = \mathbb{E}[Y^k]$ for each $k \geq 1$. Since the distribution of Y is uniquely determined by the sequence of moments $(\mathbb{E}[Y^k])$, $k \geq 1$, the assertion follows. □

Proof of Theorem 2. Fix $y \in \mathbb{R}$. In what follows, we will verify (13) for

$$A_{n,i} := \big\{n\mu\{S(X_i, R_{i,n})\} \geq y + \ln n\big\}, \quad i \in \{1, \ldots, n\},$$

and $\nu = \exp(-y)$. Throughout the proof we tacitly assume

$$0 < y_n := \frac{y + \ln n}{n} < 1.$$

This assumption entails no loss of generality since n tends to infinity. With $H_x(\cdot)$ given in (11), we put

$$R^*_{i,n} := H^{-1}_{X_i}((y + \ln n)/n), \quad i \in \{1, \ldots, n\}.$$

For the special case $k = 1$, conditioning on X_1 and (12) yield

$$\begin{aligned}
n\,\mathbb{P}\left(A_{n,1}\right) &= n\,\mathbb{P}\left(\mu\left(S(X_1, R_{1,n})\right) \geq y_n\right)\\
&= n\,\mathbb{E}\Big[\mathbb{P}\left(\mu\left(S(X_1, R_{1,n})\right) \geq y_n \mid X_1\right)\Big]\\
&= n\,\mathbb{E}\Big[\left(1 - \mu\left(S(X_1, H^{-1}_{X_1}(y_n))\right)\right)^{n-1}\Big]\\
&= n\left(1 - \frac{y + \ln n}{n}\right)^{n-1}.
\end{aligned}$$

Using the inequalities $1 - 1/t \leq \ln t \leq t - 1$ gives $\lim_{n\to\infty} n\,\mathbb{P}(A_{n,1}) = \mathrm{e}^{-y}$. Thus (13) is proved for $k = 1$, remarkably without any condition on the underlying density f. We now assume $k \geq 2$ and put

$$\widetilde{R}_{i,k,n} := \min_{k+1\leq j\leq n} \|X_i - X_j\|, \quad r_{i,k} := \min_{j\neq i, j\leq k} \|X_i - X_j\|.$$

Then

$$R_{i,n} = \min\{\widetilde{R}_{i,k,n}, r_{i,k}\},$$

and because of $\widetilde{R}_{i,k,n} \to 0$ $\mathbb{P}$-almost surely as $n \to \infty$, it follows that, on a set of probability 1,

$$R_{i,n} = \widetilde{R}_{i,k,n} \quad \text{for each } i \in \{1, \ldots, k\}$$

if n is large enough. Conditioning on $X_1, \ldots, X_k$ we have

$$
\begin{aligned}
&\mathbb{P}\left(\cap_{i=1}^k A_{n,i}\right) \\
&\quad = \mathbb{P}\left(\cap_{i=1}^k \{\mu(S(X_i, \min\{\widetilde{R}_{i,k,n}, r_{i,k}\})) \geq y_n\}\right) \\
&\quad = \mathbb{P}\left(\cap_{i=1}^k \{\mu(S(X_i, \widetilde{R}_{i,k,n})) \geq y_n, \mu(S(X_i, r_{i,k})) \geq y_n\}\right) \\
&\quad = \mathbb{E}\left[\mathbb{P}\left(\cap_{i=1}^k \{\mu(S(X_i, \widetilde{R}_{i,k,n})) \geq y_n, \mu(S(X_i, r_{i,k})) \geq y_n\} \mid X_1, \ldots, X_k\right)\right] \\
&\quad = \mathbb{E}\left[\mathbb{P}\left(\cap_{i=1}^k \{\mu(S(X_i, \widetilde{R}_{i,k,n})) \geq y_n\} \mid X_1, \ldots, X_k\right) \prod_{i=1}^k \mathbb{I}\{\mu(S(X_i, r_{i,k})) \geq y_n\}\right].
\end{aligned}
$$

Furthermore, we obtain

$$
\begin{aligned}
&\mathbb{P}\left(\cap_{i=1}^k \{\mu(S(X_i, \widetilde{R}_{i,k,n})) \geq y_n\} \mid X_1, \ldots, X_k\right) \\
&\quad = \mathbb{P}\left(\cap_{i=1}^k \{\widetilde{R}_{i,k,n} \geq H_{X_i}^{-1}(y_n)\} \mid X_1, \ldots, X_k\right) \\
&\quad = \mathbb{P}\left(\cap_{i=1}^k \{\widetilde{R}_{i,k,n} \geq R_{i,n}^*\} \mid X_1, \ldots, X_k\right) \\
&\quad = \mathbb{P}\left(X_{k+1}, \ldots, X_n \notin \cup_{i=1}^k S(X_i, R_{i,n}^*) \mid X_1, \ldots, X_k\right) \\
&\quad = \left(1 - \mu\left(\cup_{i=1}^k S(X_i, R_{i,n}^*)\right)\right)^{n-k}.
\end{aligned}
$$

Notice that we have the obvious lower bound

$$
\begin{aligned}
&n^k \left(1 - \mu\left(\cup_{i=1}^k S(X_i, R_{i,n}^*)\right)\right)^{n-k} \prod_{i=1}^k \mathbb{I}\{\mu(S(X_i, r_{i,k})) \geq y_n\} \\
&\geq n^k \left(1 - \sum_{i=1}^k \mu\left(S(X_i, R_{i,n}^*)\right)\right)^{n-k} \prod_{i=1}^k \mathbb{I}\{\mu(S(X_i, r_{i,k}) \geq y_n\} \\
&= n^k \left(1 - k\,\frac{y + \ln n}{n}\right)^{n-k} \prod_{i=1}^k \mathbb{I}\{\mu(S(X_i, r_{i,k})) \geq y_n\}.
\end{aligned}
$$

Since the latter converges almost surely to e^{-ky} as $n \to \infty$, Fatou's lemma implies

$$\begin{aligned}
&\liminf_{n\to\infty} n^k \mathbb{P}\left(\cap_{i=1}^k A_{n,i}\right) \\
&= \liminf_{n\to\infty} \mathbb{E}\Big[n^k \left(1-\mu\left(\cup_{i=1}^k S_{X_i,R^*_{i,n}}\right)\right)^{n-k} \prod_{i=1}^k \mathbb{I}\{\mu(S(X_i,r_{i,k}))\geq y_n\}\Big] \\
&\geq \mathbb{E}\Big[\liminf_{n\to\infty} n^k \left(1-\mu\left(\cup_{i=1}^k S_{X_i,R^*_{i,n}}\right)\right)^{n-k} \prod_{i=1}^k \mathbb{I}\{\mu(S(X_i,r_{i,k}))\geq y_n\}\Big] \\
&= \mathrm{e}^{-ky}.
\end{aligned}$$

It thus remains to show

$$\limsup_{n\to\infty} n^k \mathbb{P}\left(\cap_{i=1}^k A_{n,i}\right) \leq \mathrm{e}^{-ky}. \tag{15}$$

Let D_n be the event that the balls $S(X_i, R^*_{i,n})$, $i = 1, \ldots, k$, are pairwise disjoint. Putting

$$\mathbb{I}_{n,k} := \prod_{i=1}^k \mathbb{I}\{\mu(S(X_i,r_{i,k}))\geq y_n\},$$

we have

$$\begin{aligned}
&\limsup_{n\to\infty} n^k \mathbb{P}(A_{n,1}\cap\ldots\cap A_{n,k}) \\
&= \limsup_{n\to\infty} n^k\, \mathbb{E}\Big[\left(1-\mu\left(\cup_{i=1}^k S(X_i,R^*_{i,n})\right)\right)^{n-k} \mathbb{I}_{n,k}\Big] \\
&\leq \limsup_{n\to\infty} n^k\, \mathbb{E}\Big[\exp\left(-(n-k)\mu\left(\cup_{i=1}^k S(X_i,R^*_{i,n})\right)\right) \mathbb{I}_{n,k}\Big] \\
&\leq \limsup_{n\to\infty} n^k\, \mathbb{E}\Big[\exp\left(-(n-k)k\frac{y+\ln n}{n}\right) \mathbb{I}\{D_n\}\Big] \\
&+ \limsup_{n\to\infty} n^k\, \mathbb{E}\Big[\exp\left(-(n-k)\mu\left(\cup_{i=1}^k S(X_i,R^*_{i,n})\right)\right) \mathbb{I}\{D_n^c\}\mathbb{I}_{n,k}\Big] \\
&\leq \mathrm{e}^{-ky} + \limsup_{n\to\infty} n^k\, \mathbb{E}\Big[\exp\left(-(n-k)\mu\left(\cup_{i=1}^k S(X_i,R^*_{i,n})\right)\right) \mathbb{I}\{D_n^c\}\mathbb{I}_{n,k}\Big].
\end{aligned}$$

It thus remains to show

$$\lim_{n\to\infty} n^k\, \mathbb{E}\big[\exp\left(-(n-k)\mu\left(\cup_{i=1}^k S(X_i,R^*_{i,n})\right)\right) \mathbb{I}\{D_n^c\}\mathbb{I}_{n,k}\big] = 0. \tag{16}$$

Under some additional smoothness conditions on the density, Henze [7] verified (16) for the related problem of finding the limit distribution of the random variable figuring in (2). By analogy with his way of proof, we introduce an equivalence relation on the set $\{1,\ldots,k\}$ as follows: An equivalence class consists of a singleton $\{i\}$ if

$$S(X_i, R_{i,n}^*) \cap S(X_j, R_{j,n}^*) = \emptyset$$

for each $j \neq i$. Otherwise, i and j are called equivalent if there is a subset $\{i_1,\ldots,i_\ell\}$ of $\{1,\ldots,k\}$ such that $i = i_1$ and $j = i_\ell$ and

$$S(X_{i_m}, R_{i_m,n}^*) \cap S(X_{i_{m+1}}, R_{i_{m+1},n}^*) \neq \emptyset$$

for each $m \in \{1,\ldots,\ell-1\}$. Let $\mathcal{P} = \{Q_1,\ldots,Q_q\}$ be a partition of $\{1,\ldots,k\}$, and denote by E_u the event that Q_u forms an equivalence class. For the event D_n, the partition $\mathcal{P}_0 := \{\{1\},\ldots,\{k\}\}$ is the trivial one, while on the complement D_n^c any partition $\mathcal{P}$ is non-trivial, which means that $q < k$. In order to prove (16), we have to show that

$$\limsup_{n\to\infty} n^k \mathbb{E}\left[\exp\left(-(n-k)\mu\left(\cup_{i=1}^k S(X_i, R_{i,n}^*)\right)\right) \mathbb{I}_{n,k} \prod_{u=1}^q \mathbb{I}\{E_u\}\right] = 0 \quad (17)$$

for each non-trivial partition $\mathcal{P}$. Since balls that belong to different equivalence classes are disjoint, we have

$$\begin{aligned}\mu\left(\cup_{i=1}^k S(X_i, R_{i,n}^*)\right) \prod_{u=1}^q \mathbb{I}\{E_u\} &= \mu\left(\cup_{u=1}^q \cup_{i\in Q_u} S(X_i, R_{i,n}^*)\right) \prod_{u=1}^q \mathbb{I}\{E_u\} \\ &= \sum_{u=1}^q \mu\left(\cup_{i\in Q_u} S(X_i, R_{i,n}^*)\right) \prod_{u=1}^q \mathbb{I}\{E_u\}.\end{aligned}$$

Writing $|B|$ for the number of elements of a finite set B, it follows that

$$\begin{aligned}&n^k \exp\left(-(n-k)\mu\left(\cup_{i=1}^k S(X_i, R_{i,n}^*)\right)\right) \prod_{i=1}^k \mathbb{I}\{\mu(S(X_i, r_{i,k})) \geq y_n\} \prod_{u=1}^q \mathbb{I}\{E_u\} \\ &\leq \mathrm{e}^k \prod_{u=1}^q n^{|Q_u|} \prod_{u=1}^q \mathrm{e}^{-n\mu\left(\cup_{i\in Q_u} S(X_i, R_{i,n}^*)\right)} \prod_{u=1}^q \prod_{i\in Q_u} \mathbb{I}\{\mu(S(X_i, r_{i,k})) \geq y_n\} \prod_{u=1}^q \mathbb{I}\{E_u\} \\ &= \mathrm{e}^k \prod_{u=1}^q \left(n^{|Q_u|} \mathrm{e}^{-n\mu\left(\cup_{i\in Q_u} S(X_i, R_{i,n}^*)\right)} \prod_{i\in Q_u} \mathbb{I}\{\mu(S(X_i, r_{i,k})) \geq y_n\} \mathbb{I}\{E_u\}\right).\end{aligned}$$

In view of independence, we have

$$n^k \mathbb{E}\left[\mathrm{e}^{-n\mu\left(\cup_{i=1}^k S(X_i, R_{i,n}^*)\right)} \prod_{i=1}^k \mathbb{I}\{\mu(S(X_i, r_{i,k})) \geq y_n\} \prod_{u=1}^q \mathbb{I}\{E_u\} \right]$$
$$= \prod_{u=1}^q \mathbb{E}\left[n^{|Q_u|} \mathrm{e}^{-n\mu\left(\cup_{i\in Q_u} S(X_i, R_{i,n}^*)\right)} \prod_{i\in Q_u} \mathbb{I}\{\mu(S(X_i, r_{i,k})) \geq y_n\} \mathbb{I}\{E_u\} \right].$$

Thus, (17) is proved if we can show

$$\lim_{n\to\infty} \mathbb{E}\left[n^{|Q_u|} \mathrm{e}^{-n\mu\left(\cup_{i\in Q_u} S(X_i, R_{i,n}^*)\right)} \prod_{i\in Q_u} \mathbb{I}\{\mu(S(X_i, r_{i,k})) \geq y_n\} \mathbb{I}\{E_u\} \right] = 0$$

for each u with $2 \leq |Q_u| < k$. Without loss of generality assume

$$Q_u = \{1, \ldots, |Q_u|\}.$$

Then

$$\begin{aligned}
&\cap_{i=1}^{|Q_u|} \{\mu(S(X_i, r_{i,k})) \geq y_n\} \\
&= \cap_{i=1}^{|Q_u|} \left\{ \mu(S(X_i, \min_{j\neq i, j\leq |Q_u|} \|X_i - X_j\|)) \geq y_n \right\} \\
&= \cap_{i=1}^{|Q_u|} \left\{ \min_{j\neq i, j\leq |Q_u|} \mu(S(X_i, \|X_i - X_j\|)) \geq y_n \right\} \\
&= \cap_{i=1}^{|Q_u|} \cap_{j\neq i, j\leq |Q_u|} \{\mu(S(X_i, \|X_i - X_j\|)) \geq y_n\} \\
&= \cap_{i=1}^{|Q_u|} \cap_{j\neq i, j\leq |Q_u|} \{\|X_i - X_j\| \geq H_{X_i}^{-1}(y_n)\} \\
&= \cap_{i,j\leq |Q_u|, i\neq j} \{\|X_i - X_j\| \geq \max(R_{i,n}^*, R_{j,n}^*)\},
\end{aligned}$$

and we obtain

$$\begin{aligned}
&n^{|Q_u|} \mathrm{e}^{-n\mu\left(\cup_{i\in Q_u} S(X_i, R_{i,n}^*)\right)} \prod_{i\in Q_u} \mathbb{I}\{\mu(S_{X_i, r_{i,k}}) \geq y_n\} \mathbb{I}\{E_u\} \\
&= n^{|Q_u|} \mathrm{e}^{-n\mu\left(\cup_{i=1}^{|Q_u|} S(X_i, R_{i,n}^*)\right)} \mathbb{I}\{\cap_{i,j\leq |Q_u|, i\neq j} \{\|X_i - X_j\| \geq \max\{R_{i,n}^*, R_{j,n}^*\}\}\} \mathbb{I}\{E_u\} \\
&\leq n^{|Q_u|} \mathrm{e}^{-n\mu\left(\cup_{i=1}^{2} S(X_i, R_{i,n}^*)\right)} \mathbb{I}\{\cap_{i,j\leq |Q_u|, i\neq j} \{\|X_i - X_j\| \geq \max\{R_{i,n}^*, R_{j,n}^*\}\}\} \mathbb{I}\{E_u\}.
\end{aligned}$$

Now, condition (6) implies

$$
\begin{aligned}
&n\mu\left(\cup_{i=1}^{2} S(X_i, R_{i,n}^*)\right) \\
&= n\mu\left(S(X_1, R_{1,n}^*)\right) + n\mu\left(S(X_2, R_{2,n}^*)\right) - n\mu\left(S(X_2, R_{2,n}^*) \cap S(X_1, R_{1,n}^*)\right) \\
&= n\,\frac{y+\ln n}{n}\left(2 - \frac{\mu\left(S(X_2, R_{2,n}^*) \cap S(X_1, R_{1,n}^*)\right)}{\mu\left(S(X_2, R_{2,n}^*)\right)}\right) \\
&\geq (y+\ln n)(2-\beta) \\
&=: (y+\ln n)(1+\varepsilon)
\end{aligned}
$$

(say). Notice that $\varepsilon > 0$ since $0 < \beta < 1$. Thus,

$$
\begin{aligned}
&n^{|Q_u|}\mathbb{E}\left[\mathrm{e}^{-n\mu\left(\cup_{i\in Q_u} S(X_i, R_{i,n}^*)\right)} \prod_{i\in Q_u} \mathbb{I}\{\mu(S(X_i, r_{i,k})) \geq y_n\}\mathbb{I}\{E_u\}\right] \\
&\leq n^{|Q_u|}\mathrm{e}^{-(y+\ln n)(1+\varepsilon)}\mathbb{E}\left[\mathbb{I}\{\cap_{i,j\leq |Q_u|, i\neq j}\{\|X_i - X_j\| \geq \max\{R_{i,n}^*, R_{j,n}^*\}\}\}\mathbb{I}\{E_u\}\right] \\
&= \mathrm{O}\left(n^{|Q_u|-1-\varepsilon}\right)\mathbb{P}\left(E_u\right).
\end{aligned}
$$

In order to bound $\mathbb{P}(E_u)$ we need the following lemma:

Lemma 1 *On E_u there is a random integer $L \in \{1, \ldots, |Q_u|\}$ depending on $X_1, \ldots, X_{|Q_u|}$ such that $Q_u \setminus \{L\}$ forms an equivalence class.*

Proof. Let $m := |Q_u|$. Regard $X_1, \ldots, X_m$ as vertices of a graph in which any two vertices X_i and X_j are connected by a node if $S(X_i, R_{i,n}^*) \cap S(X_j, R_{j,n}^*) \neq \emptyset$. Since $Q_u = \{1, \ldots, m\}$ is an equivalence class, this graph is connected. If there is at least one vertex X_j (say) with degree 1, put $L := j$. Otherwise, the degree of each vertex is at least two, and we have $m \geq 3$. If $m = 3$, the graph is a triangle, and we can choose L arbitrarily. Now suppose the lemma is true for any graph having $m \geq 3$ vertices, in which each vertex degree is at least 2. If we have an additional $(m+1)$th vertex X_{m+1}, this is connected to at least two other vertices X_i and X_j (say). Of the graph with vertices $X_1, \ldots, X_m$ we can delete one vertex, and the remaining graph is connected. But X_{m+1} is then connected to either X_i or X_j, and we may choose $L = i$ or $L = j$. Notice that for $d = 1$ the proof is trivial since $\cup_{i\in Q_u} S(X_i, R_{i,n}^*)$ is an interval, and we can take either $L = 1$ or $L = m$. □

By induction, we now show that

$$
\mathbb{P}(E_u) = \mathrm{O}\left((\ln n/n)^{|Q_u|-1}\right) \tag{18}
$$

as $n \to \infty$ for each $m := |Q_u| \in \{2, \ldots, k-1\}$. We start with the base case $m = 2$. Notice that $\mathbb{P}(E_u) \le \mathbb{P}(\|X_2 - X_1\| \le R^*_{2,n} + R^*_{1,n})$ and

$$\begin{aligned}
&\mathbb{P}(\|X_2 - X_1\| \le R^*_{2,n} + R^*_{1,n} \mid X_1) \\
&= \mathbb{P}(\|X_2 - X_1\| \le R^*_{2,n} + R^*_{1,n}, R^*_{2,n} \le R^*_{1,n} \mid X_1) \\
&\quad + \mathbb{P}(\|X_2 - X_1\| \le R^*_{2,n} + R^*_{1,n}, R^*_{2,n} > R^*_{1,n} \mid X_1) \\
&\le \mathbb{P}(\|X_2 - X_1\| \le 2R^*_{1,n} \mid X_1) + \mathbb{P}(\|X_2 - X_1\| \le 2R^*_{2,n} \mid X_1).
\end{aligned}$$

Now, condition (7) entails

$$\begin{aligned}
\mathbb{P}(\|X_2 - X_1\| \le 2R^*_{1,n} \mid X_1) = \mu(S(X_1, 2R^*_{1,n})) &\le c_{max}\, \mu(S(X_1, R^*_{1,n})) \\
&= c_{max}\, \frac{y + \ln n}{n}.
\end{aligned}$$

Putting $\widetilde{R}_{2,n} := H^{-1}_{X_2}(c_{max}(y + \ln n)/n)$, a second appeal to (7) yields

$$\mu(S(X_2, 2R^*_{2,n})) \le c_{max}\, \mu(S(X_2, R^*_{2,n})) = c_{max}\, \frac{y + \ln n}{n}$$

and thus $2R^*_{2,n} \le \widetilde{R}_{2,n}$. Consequently,

$$\mathbb{P}(\|X_2 - X_1\| \le 2R^*_{2,n} \mid X_1) \le \mathbb{P}(\|X_2 - X_1\| \le \widetilde{R}_{2,n} \mid X_1).$$

Let γ_d be the minimum number of cones of angle $\pi/3$ centered at 0 such that their union covers $\mathbb{R}^d$. Then the cone covering lemma (cf. Lemma 10.1 in Devroye and Györfi [4], and Lemma 6.2 in Györfi et al. [6]) says that, for any $0 \le a \le 1$ and any x_1, we have

$$\mu(\{x_2 \in \mathbb{R}^d : \mu(S(x_2, \|x_2 - x_1\|)) \le a\}) \le \gamma_d\, a. \tag{19}$$

Now, (19) implies

$$\mu(\{x_2 \in \mathbb{R}^d : \|x_2 - x_1\| \le H^{-1}_{x_2}(a)\})\gamma_d\, a,$$

whence

$$\mathbb{P}(\|X_2 - X_1\| \le \widetilde{R}_{2,n} \mid X_1) \le \gamma_d c_{max}\, \frac{y + \ln n}{n}.$$

We thus obtain

$$\mathbb{P}(\|X_2 - X_1\| \le R^*_{2,n} + R^*_{1,n} \mid X_1) = \mathrm{O}\left(\frac{\ln n}{n}\right), \tag{20}$$

and so (18) is proved for $m = 2$. For the induction step, assume (18) holds for $|Q_u| = m \in \{2, \dots, k-2\}$. If Q_u with $|Q_u| = m+1$ is an equivalence class, then by Lemma 1 there are random integers L_1 and L_2 less than $m+2$, such that $Q_u \setminus \{L_1\}$ forms an equivalence class, and

$$\|X_{L_1} - X_{L_2}\| \leq R^*_{L_1,n} + R^*_{L_2,n}.$$

It follows that

$$\begin{aligned}
\mathbb{P}(E_u) &\leq (m+1)m\mathbb{P}\left(E_u \cap \{L_1 = m+1, L_2 = 1\}\right) \\
&\leq k(k-1)\mathbb{P}\big(\{Q_u \setminus \{m+1\} \text{ forms an equivalence class }\} \\
&\quad \cap \{\|X_{m+1} - X_1\| \leq R^*_{m+1,n} + R^*_{1,n}\}\big) \\
&= k(k-1)\mathbb{E}\big[\mathbb{I}\{Q_u \setminus \{m+1\} \text{ forms an equivalence class }\} \\
&\quad \cdot \mathbb{P}\left(\|X_{m+1} - X_1\| \leq R^*_{m+1,n} + R^*_{1,n} \mid X_1, \dots, X_m\right)\big] \\
&= k(k-1)\mathbb{E}\big[\mathbb{I}\{Q_u \setminus \{m+1\} \text{ forms an equivalence class }\} \\
&\quad \cdot \mathbb{P}\left(\|X_{m+1} - X_1\| \leq R^*_{m+1,n} + R^*_{1,n} \mid X_1\right)\big] \\
&\leq \mathrm{O}\left(\frac{\ln n}{n}\right)\mathbb{P}\left(Q_u \setminus \{m+1\} \text{ forms an equivalence class}\right) \\
&= \mathrm{O}\left(\frac{\ln n}{n}\right)\mathrm{O}\left((\ln n/n)^{m-1}\right) \\
&= \mathrm{O}\left((\ln n/n)^{m}\right).
\end{aligned}$$

Notice that the penultimate equation follows from the induction hypothesis, and the last "$\leq$" is a consequence of (20). Notice further that these limit relations imply (18), whence

$$\begin{aligned}
&n^{|Q_u|}\mathbb{E}\left[\mathrm{e}^{-n\mu\left(\cup_{i\in Q_u} S(X_i, R^*_{i,n})\right)} \prod_{i\in Q_u} \mathbb{I}\{\mu(S(X_i, r_{i,k})) \geq y_n\}\mathbb{I}\{E_u\}\right] \\
&= \mathrm{O}\left(n^{|Q_u|-1-\varepsilon}\right)\mathbb{P}(E_u) \\
&= \mathrm{O}\left(n^{|Q_u|-1-\varepsilon}\right)\mathrm{O}\left((\ln n/n)^{|Q_u|-1}\right) \\
&= \mathrm{o}(1).
\end{aligned}$$

Summarizing, we have shown (17) and thus (15). Hence (13) is verified with $\nu = \exp(-y)$, and the theorem is proved. □

5 Discussion on condition (6)

In this final section we comment on condition (6). For $d = 1$, we verify (6) if on $S(x,r) \cup S(z,s)$ the distribution function F of μ is either convex or concave. If $\|x - z\| \geq r + s$, then $S(x,r)$ and $S(z,s)$ are disjoint, therefore suppose $r + s \geq \|x - z\| \geq \max(r,s)$. Assume that F is convex, the proof for concave F is similar. If $x < z$, the convexity of F and

$$\mu\left(S(z,s)\right) = F(z+s) - F(z-s) =: p$$

(say) imply $F(z) - F(z-s) \leq p/2$. Thus

$$\begin{aligned}\mu\left(S(x,r) \cap S(z,s)\right) &= \mu([z-s, x+r]) \\ &\leq \min\{\mu([z-s,z]), \mu([x,x+r])\} \\ &= \min\{F(z) - F(z-s), F(x+r) - F(x)\} \\ &\leq F(z) - F(z-s) \\ &\leq p/2\end{aligned}$$

and hence

$$\frac{\mu\left(S(x,r) \cap S(z,s)\right)}{\mu\left(S(z,s)\right)} \leq \frac{1}{2}.$$

Thus (6) is satisfied with $\beta = 1/2$.

For $d > 1$, the problem is more involved. Again, suppose $r+s \geq \|x-z\| \geq \max(r,s)$. Writing $\langle \cdot, \cdot \rangle$ for the inner product in $\mathbb{R}^d$, introduce the half spaces

$$H_1 := \{u \in \mathbb{R}^d : \langle u - x, z - x \rangle \geq 0\}, \quad H_2 := \{u \in \mathbb{R}^d : \langle u - z, x - z \rangle \geq 0\}.$$

Then

$$\begin{aligned}\mu\left(S(x,r) \cap S(z,s)\right) &= \mu\left((S(z,s) \cap H_2) \cap (S(x,r) \cap H_1)\right) \\ &\leq \frac{\mu\left(S(z,s) \cap H_2\right) + \mu\left(S(x,r) \cap H_1\right)}{2}.\end{aligned}$$

We introduce another implicit condition as follows: Assume there are $\alpha \in (1,2)$ and $\delta > 0$ such that, for any $r, s > 0$ and any $x, z \in \text{supp}(\mu)$ with $r + s \geq \|x - z\| \geq \max(r,s)$ and $\mu\left(S(x,r)\right) = \mu\left(S(z,s)\right) \leq \delta$, one has either

$$\mu\left(S(z,s) \cap H_2\right) \leq \alpha\, \mu\left(S(x,r) \cap H_1^c\right) \tag{21}$$

or

$$\mu\left(S(x,r)\cap H_1\right) \le \alpha\mu\left(S(z,s)\cap H_2^c\right). \tag{22}$$

In case of (21) we have

$$\begin{aligned}
\frac{\mu\left(S(z,s)\cap H_2\right)+\mu\left(S(x,r)\cap H_1\right)}{2} &\le \frac{\alpha\,\mu\left(S(x,r)\cap H_1^c\right)+\mu\left(S(x,r)\cap H_1\right)}{2}\\
&\le \alpha\,\frac{\mu\left(S(x,r)\cap H_1^c\right)+\mu\left(S(x,r)\cap H_1\right)}{2}\\
&= \frac{\alpha}{2}\,\mu\left(S(x,r)\right),
\end{aligned}$$

and (6) is verified with $\beta = \alpha/2$. The case of (22) is similar. For the univariate case and for $x < z$, (21) and (22) mean

$$F(z) - F(z-s) \le \alpha\left(F(x) - F(x-r)\right) \tag{23}$$

and

$$F(x+r) - F(x) \le \alpha\left(F(z+s) - F(z)\right). \tag{24}$$

For convex F and small δ, (24) is approximately satisfied with $\alpha \approx 1$. Vice versa, (23) holds for concave F.

References

[1] Biau, G. and Devroye, L.: *Lectures on the Nearest Neighbor Method*, Springer-Verlag, New York, 2015.

[2] Billingsley, P.: *Probability and Measure*. John Wiley & Sons, New York 1986.

[3] Calka P. and Chenavier, N.: Extreme values for characteristic radii of a Poisson-Voronoi tessellation. *Journal of Applied Probability*, 17:359-385, 2014.

[4] Devroye, L. and Györfi, L.: *Nonparametric Density Estimation: the L_1 View*. Wiley, 1985.

[5] Graham, R.L., Knuth, D.E., Patashnik, O.: *Concrete Mathematics*. Addison-Wesley, Reading, Massachusetts, 1994.

[6] Györfi, L., Kohler, M., Krzyżak, A. and Walk, H.: *A Distribution-Free Theory of Nonparametric Regression*. Springer-Verlag, New York, 2002.

[7] Henze, N.: Ein asymptotischer Satz über den maximalen Minimalabstand von unabhängigen Zufallsvektoren mit Anwendungen auf einen Anpassungstest im $\mathbb{R}^p$ und auf der Kugel. (German) [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, 1983.

[8] Henze, N.: The limit distribution for maxima of weighted rth-nearest-neighbour distances. *J. Applied Probability*, 19:344–354, 1982.

László Györfi
Budapest University of Technology and Economics
E-Mail: gyorfi@cs.bme.hu

Norbert Henze
Karlsruhe Institute of Technology
E-Mail: henze@kit.edu

Harro Walk
Fachbereich Mathematik, Universität Stuttgart, Stuttgart, Germany
E-Mail: walk@mathematik.uni-stuttgart.de

Erschienene Preprints ab Nummer 2012-001

Komplette Liste: `http://www.mathematik.uni-stuttgart.de/preprints`

2019-003 *Braun, A.; Kohler, M.; Walk, H.:* On the rate of convergence of a neural network regression estimate learned by gradient descent

2019-002 *Györfi, L.; Henze, N.; Walk, H.:* The limit distribution of the maximum probability nearest neighbor ball

2019-001 *Gaspoz, F.D.; Heine, C.-J.; Siebert, K.G.:* An Alternative Proof of H^1-Stability of the L_2-Projection on Graded Meshes

2018-003 *Kollross, A.:* Octonions, triality, the exceptional Lie algebra F4, and polar actions on the Cayley hyperbolic plane

2018-002 *Diaź-Ramos, J.C.; Domínguez-Vázquez, M.; Kollross, A.:* On homogeneous manifolds whose isotropy actions are polar

2018-001 *Grundhöfer, T.; Stroppel, M.; Van Maldeghem, H.:* Embeddings of hermitian unitals into pappian projective planes

2017-011 *Hansmann, M.; Kohler, M.; Walk, H.:* On the strong universal consistency of local averaging regression estimates

2017-010 *Devroye, L.; Györfi, L.; Lugosi, G.; Walk, H.:* A nearest neighbor estimate of a regression functional

2017-009 *Steinke, G.; Stroppel, M.:* On elation Laguerre planes with a two-transitive orbit on the set of generators

2017-008 *Steinke, G.; Stroppel, M.:* Laguerre planes and shift planes

2017-007 *Blunck, A.; Knarr, N.; Stroppel, B.; Stroppel, M.:* Transitive groups of similitudes generated by octonions

2017-006 *Blunck, A.; Knarr, N.; Stroppel, B.; Stroppel, M.:* Clifford parallelisms defined by octonions

2017-005 *Knarr, N.; Stroppel, M.:* Subforms of Norm Forms of Octonion Fields

2017-004 *Apprich, C.; Dieterich, A.; Höllig, K.; Nava-Yazdani, E.:* Cubic Spline Approximation of a Circle with Maximal Smoothness and Accuracy

2017-003 *Fischer, S.; Steinwart, I.:* Sobolev Norm Learning Rates for Regularized Least-Squares Algorithm

2017-002 *Farooq, M.; Steinwart, I.:* Learning Rates for Kernel-Based Expectile Regression

2017-001 *Bauer, B.; Devroye, L; Kohler, M.; Krzyzak, A.; Walk, H.:* Nonparametric Estimation of a Function From Noiseless Observations at Random Points

2016-006 *Devroye, L.; Györfi, L.; Lugosi, G.; Walk, H.:* On the measure of Voronoi cells

2016-005 *Kohls, C.; Kreuzer, C.; Rösch, A.; Siebert, K.G.:* Convergence of Adaptive Finite Elements for Optimal Control Problems with Control Constraints

2016-004 *Blaschzyk, I.; Steinwart, I.:* Improved Classification Rates under Refined Margin Conditions

2016-003 *Feistauer, M.; Roskovec, F.; Sändig, AM.:* Discontinuous Galerkin Method for an Elliptic Problem with Nonlinear Newton Boundary Conditions in a Polygon

2016-002 *Steinwart, I.:* A Short Note on the Comparison of Interpolation Widths, Entropy Numbers, and Kolmogorov Widths

2016-001 *Köster, I.:* Sylow Numbers in Spectral Tables

2015-016 *Hang, H.; Steinwart, I.:* A Bernstein-type Inequality for Some Mixing Processes and Dynamical Systems with an Application to Learning

2015-015 *Steinwart, I.:* Representation of Quasi-Monotone Functionals by Families of Separating Hyperplanes

2015-014 *Muhammad, F.; Steinwart, I.:* An SVM-like Approach for Expectile Regression

2015-013 *Nava-Yazdani, E.:* Splines and geometric mean for data in geodesic spaces

2015-012 *Kimmerle, W.; Köster, I.:* Sylow Numbers from Character Tables and Group Rings

2015-011 *Györfi, L.; Walk, H.:* On the asymptotic normality of an estimate of a regression functional

2015-010 *Gorodski, C, Kollross, A.:* Some remarks on polar actions

2015-009 *Apprich, C.; Höllig, K.; Hörner, J.; Reif, U.:* Collocation with WEB–Splines

2015-008 *Kabil, B.; Rodrigues, M.:* Spectral Validation of the Whitham Equations for Periodic Waves of Lattice Dynamical Systems

2015-007 *Kollross, A.:* Hyperpolar actions on reducible symmetric spaces

2015-006 *Schmid, J.; Griesemer, M.:* Well-posedness of Non-autonomous Linear Evolution Equations in Uniformly Convex Spaces

2015-005 *Hinrichs, A.; Markhasin, L.; Oettershagen, J.; Ullrich, T.:* Optimal quasi-Monte Carlo rules on higher order digital nets for the numerical integration of multivariate periodic functions

2015-004 *Kutter, M.; Rohde, C.; Sändig, A.-M.:* Well-Posedness of a Two Scale Model for Liquid Phase Epitaxy with Elasticity

2015-003 *Rossi, E.; Schleper, V.:* Convergence of a numerical scheme for a mixed hyperbolic-parabolic system in two space dimensions

2015-002 *Döring, M.; Györfi, L.; Walk, H.:* Exact rate of convergence of kernel-based classification rule

2015-001 *Kohler, M.; Müller, F.; Walk, H.:* Estimation of a regression function corresponding to latent variables

2014-021 *Neusser, J.; Rohde, C.; Schleper, V.:* Relaxed Navier-Stokes-Korteweg Equations for Compressible Two-Phase Flow with Phase Transition

2014-020 *Kabil, B.; Rohde, C.:* Persistence of undercompressive phase boundaries for isothermal Euler equations including configurational forces and surface tension

2014-019 *Bilyk, D.; Markhasin, L.:* BMO and exponential Orlicz space estimates of the discrepancy function in arbitrary dimension

2014-018 *Schmid, J.:* Well-posedness of non-autonomous linear evolution equations for generators whose commutators are scalar

2014-017 *Margolis, L.:* A Sylow theorem for the integral group ring of $PSL(2, q)$

2014-016 *Rybak, I.; Magiera, J.; Helmig, R.; Rohde, C.:* Multirate time integration for coupled saturated/unsaturated porous medium and free flow systems

2014-015 *Gaspoz, F.D.; Heine, C.-J.; Siebert, K.G.:* Optimal Grading of the Newest Vertex Bisection and H^1-Stability of the L_2-Projection

2014-014 *Kohler, M.; Krzyżak, A.; Walk, H.:* Nonparametric recursive quantile estimation

2014-013 *Kohler, M.; Krzyżak, A.; Tent, R.; Walk, H.:* Nonparametric quantile estimation using importance sampling

2014-012 *Györfi, L.; Ottucsák, G.; Walk, H.:* The growth optimal investment strategy is secure, too.

2014-011 *Györfi, L.; Walk, H.:* Strongly consistent detection for nonparametric hypotheses

2014-010 *Köster, I.:* Finite Groups with Sylow numbers $\{q^x, a, b\}$

2014-009 *Kahnert, D.:* Hausdorff Dimension of Rings

2014-008 *Steinwart, I.:* Measuring the Capacity of Sets of Functions in the Analysis of ERM

2014-007 *Steinwart, I.:* Convergence Types and Rates in Generic Karhunen-Loève Expansions with Applications to Sample Path Properties

2014-006 *Steinwart, I.; Pasin, C.; Williamson, R.; Zhang, S.:* Elicitation and Identification of Properties

2014-005 *Schmid, J.; Griesemer, M.:* Integration of Non-Autonomous Linear Evolution Equations

2014-004 *Markhasin, L.:* L_2- and $S^r_{p,q}B$-discrepancy of (order 2) digital nets

2014-003 *Markhasin, L.:* Discrepancy and integration in function spaces with dominating mixed smoothness

2014-002 *Eberts, M.; Steinwart, I.:* Optimal Learning Rates for Localized SVMs

2014-001 *Giesselmann, J.:* A relative entropy approach to convergence of a low order approximation to a nonlinear elasticity model with viscosity and capillarity

2013-016 *Steinwart, I.:* Fully Adaptive Density-Based Clustering

2013-015 *Steinwart, I.:* Some Remarks on the Statistical Analysis of SVMs and Related Methods

2013-014 *Rohde, C.; Zeiler, C.:* A Relaxation Riemann Solver for Compressible Two-Phase Flow with Phase Transition and Surface Tension

2013-013 *Moroianu, A.; Semmelmann, U.:* Generalized Killling spinors on Einstein manifolds

2013-012 *Moroianu, A.; Semmelmann, U.:* Generalized Killing Spinors on Spheres

2013-011 *Kohls, K; Rösch, A.; Siebert, K.G.:* Convergence of Adaptive Finite Elements for Control Constrained Optimal Control Problems

2013-010 *Corli, A.; Rohde, C.; Schleper, V.:* Parabolic Approximations of Diffusive-Dispersive Equations

2013-009 *Nava-Yazdani, E.; Polthier, K.:* De Casteljau's Algorithm on Manifolds

2013-008 *Bächle, A.; Margolis, L.:* Rational conjugacy of torsion units in integral group rings of non-solvable groups

2013-007 *Knarr, N.; Stroppel, M.J.:* Heisenberg groups over composition algebras

2013-006 *Knarr, N.; Stroppel, M.J.:* Heisenberg groups, semifields, and translation planes

2013-005 *Eck, C.; Kutter, M.; Sändig, A.-M.; Rohde, C.:* A Two Scale Model for Liquid Phase Epitaxy with Elasticity: An Iterative Procedure

2013-004 *Griesemer, M.; Wellig, D.:* The Strong-Coupling Polaron in Electromagnetic Fields

2013-003 *Kabil, B.; Rohde, C.:* The Influence of Surface Tension and Configurational Forces on the Stability of Liquid-Vapor Interfaces

2013-002 *Devroye, L.; Ferrario, P.G.; Györfi, L.; Walk, H.:* Strong universal consistent estimate of the minimum mean squared error

2013-001 *Kohls, K.; Rösch, A.; Siebert, K.G.:* A Posteriori Error Analysis of Optimal Control Problems with Control Constraints

2012-013 *Diaz Ramos, J.C.; Dominguez Vázquez, M.; Kollross, A.:* Polar actions on complex hyperbolic spaces

2012-012 *Moroianu; A.; Semmelmann, U.:* Weakly complex homogeneous spaces

2012-011 *Moroianu; A.; Semmelmann, U.:* Invariant four-forms and symmetric pairs

2012-010 *Hamilton, M.J.D.:* The closure of the symplectic cone of elliptic surfaces

2012-009 *Hamilton, M.J.D.:* Iterated fibre sums of algebraic Lefschetz fibrations

2012-008 *Hamilton, M.J.D.:* The minimal genus problem for elliptic surfaces

2012-007 *Ferrario, P.:* Partitioning estimation of local variance based on nearest neighbors under censoring

2012-006 *Stroppel, M.:* Buttons, Holes and Loops of String: Lacing the Doily

2012-005 *Hantsch, F.:* Existence of Minimizers in Restricted Hartree-Fock Theory

2012-004 *Grundhöfer, T.; Stroppel, M.; Van Maldeghem, H.:* Unitals admitting all translations

2012-003 *Hamilton, M.J.D.:* Representing homology classes by symplectic surfaces

2012-002 *Hamilton, M.J.D.:* On certain exotic 4-manifolds of Akhmedov and Park

2012-001 *Jentsch, T.:* Parallel submanifolds of the real 2-Grassmannian