Electron. J. Probab. **26** (2021), article no. 149, 1–36.
ISSN: 1083-6489 https://doi.org/10.1214/21-EJP723

Poisson approximation with applications to stochastic geometry*

Federico Pianoforte† Matthias Schulte‡

Abstract

This article compares the distributions of integer-valued random variables and Poisson random variables. It considers the total variation and the Wasserstein distance and provides, in particular, explicit bounds on the pointwise difference between the cumulative distribution functions. Special attention is dedicated to estimating the difference when the cumulative distribution functions are evaluated at 0. This permits to approximate the minimum (or maximum) of a collection of random variables by a suitable random variable in the Kolmogorov distance. The main theoretical results are obtained by combining the Chen-Stein method with size-bias coupling and a generalization of size-bias coupling for integer-valued random variables developed herein. A wide variety of applications are then discussed with a focus on stochastic geometry. In particular, transforms of the minimal circumscribed radius and the maximal inradius of Poisson-Voronoi tessellations as well as the minimal inter-point distance of the points of a Poisson process are considered and bounds for their Kolmogorov distances to extreme value distributions are derived.

Keywords: Poisson approximation; Chen-Stein method; size-bias coupling; exponential approximation; stochastic geometry; U-statistics; Poisson-Voronoi tessellations; runs; extremes.
MSC2020 subject classifications: Primary 60F05, Secondary 60D05; 60G70; 60G55.
Submitted to EJP on April 7, 2021, final version accepted on November 12, 2021.

1 Introduction and main results

Let X be a random variable taking values in $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and let P_λ be a Poisson random variable with mean $\lambda > 0$. In this article we employ Stein's method, size-bias coupling and a generalization of size-bias coupling for integer-valued random variables

*This research was supported by the Swiss National Science Foundation (grant number 200021_175584).
†University of Bern, Switzerland.
E-mail: federico.pianoforte@stat.unibe.ch
‡Hamburg University of Technology, Germany.
E-mail: matthias.schulte@tuhh.de

developed herein to compare the distributions of X and P_λ. We derive upper bounds on the total variation distance

$$d_{TV}(X,P_\lambda)=\sup_{A\subset\mathbb{N}_0}|\mathbb{P}(X\in A)-\mathbb{P}(P_\lambda\in A)|$$

and the Wasserstein distance

$$d_W(X,P_\lambda)=\sup_{g\in\mathrm{Lip}(1)}|\mathbb{E}[g(X)]-\mathbb{E}[g(P_\lambda)]|$$

between X and P_λ, where $\mathrm{Lip}(1)$ denotes the set of all Lipschitz functions $g:\mathbb{N}_0\to\mathbb{R}$ with Lipschitz constant bounded by 1. In addition, we establish bounds on the pointwise differences

$$\big|\mathbb{P}(X\le v)-\mathbb{P}(P_\lambda\le v)\big|,\quad v\in\mathbb{N}_0,$$

between the cumulative distribution functions of X and P_λ, which are smaller than those for the total variation distance. Particular attention is paid to the case $v=0$. This permits to approximate the minimum (or maximum) of a collection of random variables by a suitable random variable in the Kolmogorov distance. For example, let λ_d denote the Lebesgue measure on $\mathbb{R}^d$, let k_d stand for the volume of the d-dimensional unit ball, and let η_t be a Poisson process on $\mathbb{R}^d$ with intensity measure $t\lambda_d$, $t>0$. From the aforementioned bounds for $v=0$ we deduce that the random variable Y_t given by

$$Y_t=\min_{(x,y)\in\eta_{t,\neq}^2\,:\,\frac{x+y}{2}\in[0,1]^d}2^{-1}t^2k_d\|x-y\|^d,$$

which is the rescaled minimum (Euclidean) distance between pairs of points of η_t with midpoint in $[0,1]^d$, satisfies

$$0\le\mathbb{P}(Y_t>u)-\mathbb{P}(E_1>u)\le\frac{80}{t}\tag{1.1}$$

for $u\ge0$ (see Theorem 3.9), where E_1 denotes an exponentially distributed random variable with mean 1. This is possible because $\mathbb{P}(Y_t>u)$ can be written as $\mathbb{P}(X_u=0)$ with

$$X_u=\frac12\sum_{(x,y)\in\eta_{t,\neq}^2}\mathbf{1}\Big\{\frac{x+y}{2}\in[0,1]^d,2^{-1}t^2k_d\|x-y\|^d\in[0,u]\Big\}$$

and $\mathbb{P}(E_1>u)=\mathbb{P}(P_u=0)$. By estimating $|\mathbb{P}(X_u=0)-\mathbb{P}(P_u=0)|$ uniformly for all $u\ge0$, one obtains (1.1), which provides a bound on the Kolmogorov distance

$$d_K(Y_t,E_1)=\sup_{u\in\mathbb{R}}|\mathbb{P}(Y_t>u)-\mathbb{P}(E_1>u)|$$

between Y_t and E_1.

Let us now give precise formulations of our main results. We use the shorthand notation $a\wedge b=\min\{a,b\}$ for $a,b\in\mathbb{R}$, and we indicate by W_+ and W_- the positive and negative part of a random variable W, respectively. Whenever we write $\alpha>0$, it is understood that $\alpha\in(0,\infty)$.

Theorem 1.1. *Let X be a random variable taking values in $\mathbb{N}_0$ and let P_λ be a Poisson random variable with mean $\lambda=\mathbb{E}[X]>0$. Assume there exists a random variable Z defined on the same probability space as X with values in $\mathbb{Z}$ such that*

$$i\mathbb{P}(X=i)=\lambda\mathbb{P}(X+Z=i-1),\quad i\in\mathbb{N},\tag{1.2}$$

is satisfied. Then,

$$d_{TV}(X,P_\lambda)\le(1\wedge\lambda)\mathbb{E}[|Z|]\quad\textit{and}\quad d_W(X,P_\lambda)\le(1.1437\sqrt{\lambda}\wedge\lambda)\mathbb{E}[|Z|]. \tag{1.3}$$

Furthermore for all $m\in\mathbb{N}_0$*,*

$$|\mathbb{P}(X=0)-\mathbb{P}(P_\lambda=0)|\le\frac{m!}{\lambda^m}\mathbb{E}[|Z|]+\sum_{k=0}^{m-1}\left(\frac{\lambda}{k+1}\wedge\frac{k!}{\lambda^k}\right)\mathbb{E}\big[|Z|\mathbf{1}\{X-Z_-=k\}\big] \tag{1.4}$$

and for all $v\in\mathbb{N}$*,*

$$|\mathbb{P}(X\le v)-\mathbb{P}(P_\lambda\le v)|\le\frac{(v+1)^2}{\lambda}\mathbb{E}[|Z|]+\mathbb{E}\big[|Z|\mathbf{1}\{X-Z_-\le v\}\big]. \tag{1.5}$$

Recall that for a random variable $Y\ge0$ with $\mu=\mathbb{E}[Y]>0$, a random variable Y^s on the same probability space as Y is a size-bias coupling of Y if it satisfies

$$\mathbb{E}[Yf(Y)]=\mu\mathbb{E}[f(Y^s)] \tag{1.6}$$

for all measurable f such that $\mathbb{E}[|Yf(Y)|]<\infty$. Thus, (1.2) implies that $X+Z+1$ is a size-bias coupling of X so that we can replace Z by X^s-X-1 with a size-bias coupling X^s of X in Theorem 1.1. In this form the bound for the total variation distance in (1.3) is a classical result (see [36, Theorem 4.13] and the discussion at the beginning of Section 5 in [3] for further references), whose proof is based on the Chen-Stein method and size-bias coupling. For the Chen-Stein method for Poisson approximation we refer the reader to e.g. [7, 12, 36], while [3] is a survey on size bias. Note that the bound on the Wasserstein distance in (1.3) can be derived by combining the proof of [36, Theorem 4.13] with [8, Theorem 1.1].

Remark 1.2. Let X be as in Theorem 1.1 and assume that (1.2) is satisfied.

(i) The last expressions on the right-hand sides of (1.4) and (1.5) can be further bounded using the inequalities

$$\begin{aligned}&\mathbb{E}[|Z|\mathbf{1}\{X-Z_-=k\}]\le\mathbb{E}[Z_-]+\mathbb{E}[Z_+\mathbf{1}\{X=k\}],\quad k\in\mathbb{N}_0,\\&\mathbb{E}\big[|Z|\mathbf{1}\{X-Z_-\le v\}\big]\le\mathbb{E}[Z_-]+\mathbb{E}[Z_+\mathbf{1}\{X\le v\}],\quad v\in\mathbb{N}.\end{aligned}$$

(ii) From (1.6) with $f(x)=x$ we obtain $\lambda\mathbb{E}[X^s]=\mathbb{E}[X^2]$ so that $Z=X^s-X-1$ yields

$$\mathbb{E}[Z]=\frac{1}{\lambda}\big\{\mathrm{Var}(X)-\lambda\big\}. \tag{1.7}$$

(iii) In this work, the random variable X is always assumed to be a sum of a possibly random number of indicator random variables. For sums of a fixed number of indicator random variables there is a standard approach to construct a size-bias coupling (see e.g. [36, Corollary 3.24]), which goes as follows: Assume that X is a sum of the form $X=\sum_{j=1}^n X_j$ with indicator random variables $X_1,\dots,X_n$. For $i\in\{1,\dots,n\}$ let $(X_j^{(i)})_{j\in\{1,\dots,n\}\setminus\{i\}}$ be random variables on the same probability space with the same distribution as $(X_j)_{j\in\{1,\dots,n\}\setminus\{i\}}$ conditioned on $X_i=1$. Moreover, let I be an independent random variable with $\mathbb{P}(I=i)=\mathbb{P}(X_i=1)/\sum_{j=1}^n\mathbb{P}(X_j=1)$ for $i\in\{1,\dots,n\}$. Then, $X^s=1+\sum_{j\in\{1,\dots,n\}\setminus\{I\}}X_j^{(I)}$ is a size-bias coupling of X.

The next result constitutes our main achievement and generalizes Theorem 1.1. Instead of assuming that Z satisfies (1.2) exactly, we allow error terms on the right-hand side of (1.2).

Theorem 1.3. *Let X be an integrable random variable with values in $\mathbb{N}_0$ and let P_λ be a Poisson random variable with mean $\lambda > 0$. Let Z be a random variable defined on the same probability space as X with values in $\mathbb{Z}$, and let $q_i, i \in \mathbb{N}_0$, be the sequence given by*

$$q_{i-1} = i\mathbb{P}(X = i) - \lambda\mathbb{P}(X + Z = i - 1), \quad i \in \mathbb{N}. \tag{1.8}$$

Then,

$$d_{TV}(X, P_\lambda) \le (1 \wedge \lambda)\mathbb{E}[|Z|] + \left(1 \wedge \frac{1}{\sqrt{\lambda}}\right) \sum_{i=0}^{\infty} |q_i| \tag{1.9}$$

and

$$d_W(X, P_\lambda) \le \lambda\mathbb{E}[|Z|] + \sum_{i=0}^{\infty} |q_i|. \tag{1.10}$$

Moreover, if $\mathbb{P}(X + Z \ge 0) = 1$, then

$$d_W(X, P_\lambda) \le (1.1437\sqrt{\lambda} \wedge \lambda)\mathbb{E}[|Z|] + \sum_{i=0}^{\infty} |q_i|, \tag{1.11}$$

for all $m \in \mathbb{N}_0$,

$$\begin{aligned} |\mathbb{P}(X = 0) - \mathbb{P}(P_\lambda = 0)| \le{}& \frac{m!}{\lambda^m}\mathbb{E}\big[|Z|\big] + \sum_{k=0}^{m-1} \left(\frac{\lambda}{k+1} \wedge \frac{k!}{\lambda^k}\right) \mathbb{E}\big[|Z|\mathbf{1}\{X - Z_- = k\}\big] \\ &+ \left(1 \wedge \frac{1}{\lambda}\right) |q_0| + \left(1 \wedge \frac{1}{\lambda^2}\right) \sum_{i=1}^{\infty} |q_i| \end{aligned} \tag{1.12}$$

and for all $v \in \mathbb{N}$,

$$\begin{aligned} |\mathbb{P}(X \le v) - \mathbb{P}(P_\lambda \le v)| \le{}& \frac{(v+1)^2}{\lambda}\mathbb{E}[|Z|] + \mathbb{E}\big[|Z|\mathbf{1}\{X - Z_- \le v\}\big] \\ &+ \left(1 \wedge \frac{1}{\sqrt{\lambda}}\right) \sum_{i=0}^{\infty} |q_i|. \end{aligned} \tag{1.13}$$

Note that Theorem 1.1 is a special case of Theorem 1.3. Indeed, if $q_i = 0$ for all $i \in \mathbb{N}_0$, (1.8) becomes (1.2) and the bounds in Theorem 1.3 simplify to those in Theorem 1.1. In this situation $X + Z + 1$ is a size-bias coupling of X. Thus, we can think of $X + Z + 1$ with Z satisfying (1.8) as a generalization of size-bias coupling. In order to have good bounds in Theorem 1.3, the error terms q_i, $i \in \mathbb{N}_0$, should be small. The important advantage of Theorem 1.3 compared to Theorem 1.1 is that one only needs to construct an *approximate* size-bias coupling instead of an exact size-bias coupling.

For our paper the so-called magic factors or Stein factors play a crucial role. These are bounds on the solutions of the Stein equation, which lead to the factors involving λ in our results. Since different classes of test functions have different magic factors, the upper bounds for the differences between $\mathbb{P}(X \le v)$ and $\mathbb{P}(P_\lambda \le v)$ for $v \in \mathbb{N}_0$ in Theorems 1.1 and 1.3 are of a better order in λ than those for the total variation distance or the Wasserstein distance. This observation is essential for obtaining approximation results in the Kolmogorov distance as (1.1) since it allows to bound the right-hand sides of (1.4), (1.5), (1.12) and (1.13) uniformly in λ. For a different Poisson approximation result where one has a better order in λ for the difference of the probabilities at zero than for the total variation distance we refer the reader to [1, Theorem 1].

To demonstrate the versatility of our general main results we apply them to several examples. In particular, we deduce bounds as (1.1), where we compare minima or maxima of collections of dependent random variables with random variables having an exponential, Weibull or Gumbel distribution.

As a first classical example, we use the standard size-bias coupling from Remark 1.2 (iii) and our general result, Theorem 1.1, to study the Poisson approximation of the number of non-overlapping k-runs in a sequence of n i.i.d. Bernoulli random variables (see Subsection 3.1). By a k-run one means at least k successes in a row. Here, we use Theorem 1.1 to bound the difference between the probability that among n trials there are no more than v non-overlapping k-runs and $\mathbb{P}(P_\alpha \leq v)$ for a certain Poisson random variable P_α; this bound is remarkable because it does not depend on k, i.e., the number of required successes in a row.

For Voronoi tessellations generated by a stationary Poisson process η_t on $\mathbb{R}^d$ of intensity $t > 0$, we consider statistics related to circumscribed radii and inradii (see Subsection 3.2 and 3.3). The circumscribed radius of a Voronoi cell with nucleus $x \in \eta_t$ is the radius of the smallest closed ball centered in x that contains the cell and is denoted by $C(x, \eta_t)$. The inradius of the cell corresponds to half of the Euclidean distance between x and its nearest neighbor in η_t. As a second example, we consider the transform

$$T_t = \min_{x \in \eta_t \cap W} \alpha_2 k_d t^{(d+2)/(d+1)} C(x, \eta_t)^d$$

of the minimal circumscribed radius for some observation window $W \subset \mathbb{R}^d$ with volume one, where k_d is the volume of the d-dimensional unit ball and α_2 is a dimension dependent constant (see (3.7)). Then, by the inequality (1.12) in Theorem 1.3, we show that

$$d_K(T_t, Y) \leq \frac{C_K}{t^{1/(d+1)}}$$

for $t \geq 1$, where Y denotes a Weibull random variable and C_K is some constant. For this example we use the full generality of Theorem 1.3 since we construct a coupling that satisfies (1.8), but which is not a size-bias coupling. As a third example, by applying the inequality (1.4) in Theorem 1.1 we approximate the transform

$$R_t = \max_{x \in \eta_t \cap W} \min_{y \in \eta_t \setminus \{x\}} t k_d \|x - y\|^d - \log(t)$$

of the maximal inradius over of the cells with nucleus in the observation window $W \subset \mathbb{R}^d$ by a Gumbel random variable G. We show that

$$d_K(R_t, G) \leq C \frac{\log(t)}{\sqrt{t}}$$

for $t \geq e^2$, where C is some dimension dependent constant.

Finally, we study the Poisson approximation of U-statistics constructed from an underlying Poisson or binomial point process (see Subsections 3.4 and 3.5). By applying Theorem 1.1, we obtain bounds which basically depend on integrals involving the kernel of the U-statistic. As application of our main finding on U-statistics with Poisson input, Theorem 3.8, we consider the minimum inter-point distance problem discussed at the beginning of the introduction and establish the bound (1.1) for the exponential approximation in Kolmogorov distance (see Subsection 3.6).

A crucial contribution of this paper to stochastic geometry is that we provide bounds with respect to the Kolmogorov distance for the distributional approximation of some minima and maxima. The limiting distributions of the minimal distance between the

points of a Poisson process and of large inradii and small circumscribed radii of Poisson-Voronoi tessellations have been studied before in e.g. [11, 13, 38, 39]. Some of these works provide quantitative bounds for the difference of the distribution functions at a fixed $u \in \mathbb{R}$, which depend on u. Thanks to our general Poisson approximation results Theorem 1.1 and Theorem 1.3, we are able to derive uniform bounds for all $u \in \mathbb{R}$. An alternative approach to deducing such results via Poisson approximation could be to apply directly Stein's method for the exponential, Weibull or Gumbel distribution; see e.g. [36] for more details on Stein's method for exponential approximation.

In our paper, an important and challenging part for the proofs of the applications is to construct the size-bias coupling or its approximate version defined by (1.8). As discussed in Remark 1.2 (iii), for the sum of a fixed number of indicator random variables this can be done by applying a standard technique. However, we mainly focus on sums over the points of Poisson processes, whence we have a random number of indicator random variables. In order to deal with this situation, we use the Mecke formula to obtain size-bias couplings similar to the classical construction.

In [31], a general result for the Poisson approximation of statistics of Poisson processes is derived by combining the Chen-Stein method and a kind of size-bias coupling and applied to study some statistics of inhomogeneous random graphs such as isolated vertices. Requiring some (stochastic) ordering assumptions between a random variable and its size-bias coupling leads to Poisson approximation results. In a similar spirit to our work, these ordering conditions were relaxed in [14]. For some recent Poisson process convergence results related to stochastic geometry we refer the reader to [28, 32].

Other noteworthy general results derived in this paper are lower and upper bounds on the probability that X equals 0, which are given in Proposition 2.6 and Corollary 2.7. Informally, they say that $\mathbb{P}(X = 0)$ can be bounded from above or below by $e^{-\lambda}$ for some $\lambda > 0$ if the random variable Z and the sequence $q_i, i \in \mathbb{N}_0$, in Theorem 1.1 and Theorem 1.3 satisfy certain conditions on their signs; for Z as in Theorem 1.1, it is understood that $q_i = 0$ for all $i \in \mathbb{N}_0$. These results sometimes allow us to remove the absolute values from the left-hand sides of (1.4) and (1.12).

The proof of Theorem 1.3 is based on the Chen-Stein method and the coupling in (1.8). Using the solution of the Stein equation for the Poisson distribution, we derive in Proposition 2.5 a new expression for the difference $|\mathbb{E}[g(P_\lambda)] - \mathbb{E}[g(X)]|$ for any $g \in \mathrm{Lip}(1)$. Taking in Proposition 2.5, the supremum over all functions in $\mathrm{Lip}(1)$ (or all indicator functions) establishes a different way to represent the Wasserstein distance (or the total variation distance). Moreover, choosing $g = \mathbf{1}\{\cdot \le v\}$ with $v \in \mathbb{N}$ gives a new expression for $|\mathbb{P}(X \le v) - \mathbb{P}(P_\lambda \le v)|$. These identities are then manipulated and combined with the magic factors and the coupling in (1.8) to prove Theorem 1.3.

Before we present our applications in Section 3, we prove our main results in the next section. Since the proofs for the applications to Poisson-Voronoi tessellations are rather long and technical, we postponed them to Section 4.

2 Proof of the main results

This section provides the proofs of Theorem 1.1 and Theorem 1.3. To this end, we first study the Stein equation for Poisson random variables. For any fixed $g \in \mathrm{Lip}(1)$, the solution of the Stein equation is a function $f_g : \mathbb{N}_0 \to \mathbb{R}$ with $f_g(0) = 0$ that satisfies

$$\lambda f_g(i+1) - i f_g(i) = g(i) - \mathbb{E}[g(P_\lambda)], \quad i \in \mathbb{N}_0. \tag{2.1}$$

The function f_g can be obtained by solving (2.1) recursively for $i = 0, 1, \ldots$. An explicit expression for this solution is given in [5, Lemma 1]. In particular, for $g = \mathbf{1}_A$ with $A \subset \mathbb{N}_0$, one has the following representation for f_g (see e.g. [36, Lemma 4.2]).

Lemma 2.1. *For any $\lambda > 0$ and $A \subset \mathbb{N}_0$ the unique solution f_A of*

$$\lambda f_A(i+1) - i f_A(i) = \mathbf{1}\{i \in A\} - \mathbb{P}(P_\lambda \in A), \quad i \in \mathbb{N}_0, \tag{2.2}$$

with $f_A(0) = 0$ is given by

$$f_A(i) = \frac{e^\lambda (i-1)!}{\lambda^i}\big[\mathbb{P}(P_\lambda \in A \cap \{0,1,\dots,i-1\}) - \mathbb{P}(P_\lambda \in A)\mathbb{P}(P_\lambda \le i-1)\big], \quad i \in \mathbb{N}.$$

From now on, we denote by f_A the solution of the Stein equation (2.1) for $g = \mathbf{1}_A$ with $A \subset \mathbb{N}_0$. Let X be a random variable with values in $\mathbb{N}_0$. The idea of the Chen-Stein method for the Poisson approximation of X is to plug X in (2.1) and to take the expectation, which yields

$$\mathbb{E}[\lambda f_g(X+1) - X f_g(X)] = \mathbb{E}[g(X)] - \mathbb{E}[g(P_\lambda)].$$

So we can control the difference between the expectations of $g(X)$ and $g(P_\lambda)$ on the right-hand side by estimating the term on the left-hand side. This requires some bounds on the solution of (2.1), which we give in the sequel. For a function $h : \mathbb{N}_0 \to \mathbb{R}$ we define $\Delta h : \mathbb{N}_0 \to \mathbb{R}$ by $\Delta h(i) = h(i+1) - h(i)$. The solution of the Stein equation (2.1) and its differences can be bounded by the following terms, which are called magic factors or Stein factors (see [8, Theorem 1.1]).

Lemma 2.2. *Let f_g be the solution of* (2.1)*. Then,*

$$\max_{i\in\mathbb{N}_0} |f_g(i)| \le 1 \quad \textit{and} \quad \max_{i\in\mathbb{N}} |\Delta f_g(i)| \le 1 \wedge \frac{8}{3\sqrt{2e\lambda}} \le 1 \wedge \frac{1.1437}{\sqrt{\lambda}}.$$

Since $f_g(0) = 0$, Lemma 2.2 implies that

$$\max_{i\in\mathbb{N}_0} |f_g(i)| \le 1 \quad \text{and} \quad \max_{i\in\mathbb{N}_0} |\Delta f_g(i)| \le 1. \tag{2.3}$$

Moreover, the solution of (2.2) for $A \subset \mathbb{N}_0$ has the following magic factors (see e.g. [36, Lemma 4.4]).

Lemma 2.3. *For f_A as in Lemma 2.1,*

$$\max_{i\in\mathbb{N}_0} |f_A(i)| \le 1 \wedge \frac{1}{\sqrt{\lambda}} \quad \textit{and} \quad \max_{i\in\mathbb{N}_0} |\Delta f_A(i)| \le 1 \wedge \frac{1}{\lambda}.$$

We now derive similar - potentially sharper - magic factors for the special cases $A = \{0,\dots,v\}, v \in \mathbb{N}_0$. Similar bounds for sets A that are singletons were deduced for the translated Poisson approximation in [35, Lemma 3.7].

Lemma 2.4. *Let $f_{\{0\}}$ be the unique solution of* (2.2) *for $A = \{0\}$. Then,*

$$|f_{\{0\}}(i)| \le \begin{cases} 1 \wedge \frac{1}{\lambda}, & \text{if } i = 1, \\ 1 \wedge \frac{1}{\lambda^2}, & \text{if } i \ge 2, \end{cases} \tag{2.4}$$

and for all $i \in \mathbb{N}$,

$$\Delta f_{\{0\}}(i) \le 0. \tag{2.5}$$

Furthermore for all $i, n \in \mathbb{N}$ with $i \ge n$,

$$|\Delta f_{\{0\}}(i)| \le \frac{1}{n} \wedge \frac{(n-1)!}{\lambda^n}. \tag{2.6}$$

Let $f_{\{0,\dots,v\}}$ be the unique solution of (2.2) *for $A = \{0,\dots,v\}$ with $v \in \mathbb{N}$ and $v \le \lambda$. Then for all $i \ge v+2$,*

$$\Delta f_{\{0,\dots,v\}}(i) \le 1 \wedge \frac{(v+1)^2}{\lambda^2}. \tag{2.7}$$

Proof. Obviously, the upper bound 1 in (2.4) follows from Lemma 2.3. Lemma 2.1 yields for $i \in \mathbb{N}$ that

$$f_{\{0\}}(i) = \frac{(i-1)!}{\lambda^i}(1-\mathbb{P}(P_\lambda \le i-1)) = \frac{(i-1)!}{\lambda^i}\sum_{m=i}^{\infty}\frac{\lambda^m}{m!}e^{-\lambda} = \sum_{\ell=0}^{\infty}\frac{\lambda^\ell}{(i+\ell)!}(i-1)!e^{-\lambda}. \quad (2.8)$$

This implies (2.4) for $i = 1, 2$, and yields for $i \ge 3$ that

$$f_{\{0\}}(i) = \sum_{\ell=0}^{\infty}\frac{\lambda^\ell}{(i+\ell)!}(i-1)!e^{-\lambda} = \frac{1}{\lambda^2}\sum_{\ell=0}^{\infty}\frac{\lambda^{\ell+2}}{(\ell+2)!}\frac{(i-1)!(\ell+2)!}{(i+\ell)!}e^{-\lambda}.$$

Thus, the elementary inequalities

$$\frac{(i-1)!(\ell+2)!}{(i+\ell)!} = \frac{(i-1)!}{(\ell+3)\cdot\ldots\cdot(\ell+i)} \le \frac{2(i-1)!}{i!} \le 1$$

establish (2.4) for $i \ge 3$. From (2.8) we also obtain for $n \in \mathbb{N}$,

$$\begin{aligned}
\Delta f_{\{0\}}(i) &= \sum_{\ell=0}^{\infty}\Big(\frac{\lambda^\ell}{(i+1+\ell)!}i! - \frac{\lambda^\ell}{(i+\ell)!}(i-1)!\Big)e^{-\lambda} \\
&= \sum_{\ell=0}^{\infty}\frac{\lambda^\ell}{(i+1+\ell)!}\big(i! - (i+1+\ell)(i-1)!\big)e^{-\lambda} \\
&= -\sum_{\ell=0}^{\infty}\frac{\lambda^\ell}{(i+1+\ell)!}(\ell+1)(i-1)!e^{-\lambda} \\
&= -\sum_{\ell=0}^{\infty}\frac{\lambda^\ell}{(n+\ell)!}\frac{(\ell+1)(n+\ell)!(i-1)!}{(i+1+\ell)!}e^{-\lambda},
\end{aligned}$$

which proves (2.5). For $i, n \in \mathbb{N}$ with $i \ge n$ the elementary inequalities

$$\frac{(\ell+1)(n+\ell)!(i-1)!}{(i+1+\ell)!} \le \frac{(n+\ell)!(i-1)!}{(i+\ell)!} \le (n-1)!$$

lead to

$$\big|\Delta f_{\{0\}}(i)\big| \le (n-1)!e^{-\lambda}\sum_{\ell=0}^{\infty}\frac{\lambda^\ell}{(n+\ell)!}.$$

Now the observations that

$$\sum_{\ell=0}^{\infty}\frac{\lambda^\ell}{(n+\ell)!} \le \frac{e^\lambda}{\lambda^n} \quad\text{and}\quad \sum_{\ell=0}^{\infty}\frac{\lambda^\ell}{(n+\ell)!} \le \frac{1}{n!}\sum_{\ell=0}^{\infty}\frac{\lambda^\ell}{\ell!}\frac{\ell!n!}{(n+\ell)!} \le \frac{1}{n!}\sum_{\ell=0}^{\infty}\frac{\lambda^\ell}{\ell!} \le \frac{e^\lambda}{n!}$$

show (2.6). Finally assume $\lambda \ge v$. By Lemma 2.1, we obtain for $i \ge v+2$,

$$\Delta f_{\{0,\ldots,v\}}(i) = e^\lambda \mathbb{P}(P_\lambda \in \{0,\ldots,v\})\Delta f_{\{0\}}(i).$$

Then (2.6) with $n = v+2$ implies that

$$|\Delta f_{\{0,\ldots,v\}}(i)| \le \frac{(v+1)!}{\lambda^{v+2}}\sum_{\ell=0}^{v}\frac{\lambda^\ell}{\ell!} = \frac{(v+1)!}{\lambda^2}\sum_{\ell=0}^{v}\frac{\lambda^{\ell-v}}{\ell!} \le \frac{(v+1)^2}{\lambda^2},$$

where we used the inequality $\lambda^{\ell-v}/\ell! \le 1/v!$ for $\ell = 0,\ldots,v$ and $\lambda \ge v$ in the last step. This and Lemma 2.3 establish (2.7). □

The next proposition compares the distributions of an integer-valued random variable and a Poisson distributed random variable.

Proposition 2.5. *Let X be an integrable random variable taking values in $\mathbb{N}_0$, let $\lambda \in (0,\infty)$, and define*

$$\mathcal{D}(i) = i\mathbb{P}(X=i) - \lambda\mathbb{P}(X=i-1), \quad i \in \mathbb{N}.$$

Then, for all $g \in \mathrm{Lip}(1)$,

$$\mathbb{E}[g(P_\lambda)] - \mathbb{E}[g(X)] = \sum_{i=1}^{\infty} f_g(i)\mathcal{D}(i),$$

where f_g is the solution of (2.1).

Proof. It follows from (2.1) and the definition of $\mathcal{D}(i)$, $i \in \mathbb{N}$, that

$$\begin{aligned}\mathbb{E}[g(P_\lambda)] - \mathbb{E}[g(X)] &= \mathbb{E}[Xf_g(X) - \lambda f_g(X+1)] = \sum_{i=0}^{\infty} \mathbb{P}(X=i)(if_g(i) - \lambda f_g(i+1)) \\ &= \sum_{i=1}^{\infty} \mathbb{P}(X=i) if_g(i) - \sum_{i=1}^{\infty} \mathbb{P}(X=i-1)\lambda f_g(i) = \sum_{i=1}^{\infty} f_g(i)\mathcal{D}(i),\end{aligned}$$

which gives the desired result. □

We are now in position to show Theorem 1.3.

Proof of Theorem 1.3. It follows from (1.8) that

$$\mathcal{D}(i) = i\mathbb{P}(X=i) - \lambda\mathbb{P}(X=i-1) = \lambda\mathbb{P}(X+Z=i-1) - \lambda\mathbb{P}(X=i-1) + q_{i-1}, \quad i \in \mathbb{N}.$$

Thus, Proposition 2.5 yields for $g \in \mathrm{Lip}(1)$ that

$$\begin{aligned}&\mathbb{E}[g(P_\lambda)] - \mathbb{E}[g(X)] \\ &= \lambda \sum_{i=1}^{\infty} f_g(i)\big(\mathbb{P}(X+Z=i-1) - \mathbb{P}(X=i-1)\big) + \sum_{i=1}^{\infty} f_g(i) q_{i-1} =: H_g + Q_g.\end{aligned} \tag{2.9}$$

With $f_g(0) = 0$ and the convention $f_g(i) = 0$ for $i < 0$, we obtain

$$H_g = \lambda\mathbb{E}[f_g(X+Z+1) - f_g(X+1)]. \tag{2.10}$$

Therefore the triangle inequality implies that

$$|H_g| \le \lambda \max_{i\in\mathbb{N}_0} |\Delta f_g(i)| \mathbb{E}[|Z|].$$

Furthermore, we have

$$|Q_g| \le \max_{i\in\mathbb{N}} |f_g(i)| \sum_{i=0}^{\infty} |q_i|. \tag{2.11}$$

Then combining (2.3) and the bounds on $|H_g|$ and $|Q_g|$ establishes (1.10). Moreover, from Lemma 2.3 and the bounds on $|H_g|$ and $|Q_g|$ with $g = \mathbf{1}_A$ for $A \subset \mathbb{N}_0$, we obtain (1.9).

Under the assumption $\mathbb{P}(X+Z \ge 0) = 1$, $X+Z+1$ and $X+1$ take only values in $\mathbb{N}$, whence we have even

$$|H_g| \le \lambda \max_{i\in\mathbb{N}} |\Delta f_g(i)| \mathbb{E}[|Z|].$$

Together with (2.11) and Lemma 2.2 this yields (1.11).

For $A \subset \mathbb{N}_0$ we obtain from (2.9) with $g = \mathbf{1}_A$,

$$\mathbb{P}(P_\lambda \in A) - \mathbb{P}(X \in A) = H_A + Q_A, \tag{2.12}$$

where $H_A = H_g$ and $Q_A = Q_g$ with $g = \mathbf{1}_A$. It follows from (2.10) and the fact that X takes only values in $\mathbb{N}_0$ that

$$H_A = \lambda \sum_{j\in\mathbb{Z}\setminus\{0\}} \sum_{i=1}^{\infty} \big(f_A(i+j) - f_A(i)\big)\mathbb{P}(X = i-1, Z = j). \tag{2.13}$$

The assumption $\mathbb{P}(X + Z \geq 0) = 1$ implies that $\mathbb{P}(X = i-1, Z = -j) = 0$ for all $i, j \in \mathbb{N}$ with $i \leq j$. Hence, we obtain

$$\begin{aligned}
&\lambda \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} \big(f_A(i-j) - f_A(i)\big)\mathbb{P}(X = i-1, Z = -j)\\
&= \lambda \sum_{j=1}^{\infty}\sum_{i=j+1}^{\infty} \big(f_A(i-j) - f_A(i)\big)\mathbb{P}(X = i-1, Z = -j)\\
&= \lambda \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} \big(f_A(i) - f_A(i+j)\big)\mathbb{P}(X = i+j-1, Z = -j).
\end{aligned} \tag{2.14}$$

Combining (2.13) and (2.14) leads to

$$\begin{aligned}
|H_A| &\leq \lambda \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} |f_A(i+j) - f_A(i)|\mathbb{P}(X = i-1, Z = j)\\
&\quad + \lambda \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} |f_A(i) - f_A(i+j)|\mathbb{P}(X = i+j-1, Z = -j) =: H_A^{(1)} + H_A^{(2)}.
\end{aligned}$$

For $A = \{0\}$, by (2.6) in Lemma 2.4 with $n = i$ for $i \leq m$ and $n = m+1$ for $i \geq m+1$, we have

$$\begin{aligned}
H_{\{0\}}^{(1)} &\leq \sum_{j=1}^{\infty}\sum_{i=1}^{m} \left(\frac{\lambda}{i} \wedge \frac{(i-1)!}{\lambda^{i-1}}\right) j\mathbb{P}(X = i-1, Z = j) + \sum_{j=1}^{\infty}\sum_{i=m+1}^{\infty} \frac{m!}{\lambda^m} j\mathbb{P}(X = i-1, Z = j)\\
&= \sum_{k=0}^{m-1} \left(\frac{\lambda}{k+1} \wedge \frac{k!}{\lambda^k}\right)\mathbb{E}[Z_+\mathbf{1}\{X = k\}] + \frac{m!}{\lambda^m}\mathbb{E}[Z_+\mathbf{1}\{X \geq m\}].
\end{aligned}$$

Again (2.6) in Lemma 2.4 with $n = i$ for $i \leq m$ and $n = m+1$ for $i \geq m+1$ leads to

$$\begin{aligned}
H_{\{0\}}^{(2)} &\leq \sum_{j=1}^{\infty}\sum_{i=1}^{m} \left(\frac{\lambda}{i} \wedge \frac{(i-1)!}{\lambda^{i-1}}\right) j\mathbb{P}(X = i+j-1, Z = -j)\\
&\quad + \sum_{j=1}^{\infty}\sum_{i=m+1}^{\infty} \frac{m!}{\lambda^m} j\mathbb{P}(X = i+j-1, Z = -j)\\
&= \sum_{i=1}^{m} \left(\frac{\lambda}{i} \wedge \frac{(i-1)!}{\lambda^{i-1}}\right)\mathbb{E}[Z_-\mathbf{1}\{X + Z = i-1\}] + \frac{m!}{\lambda^m}\mathbb{E}[Z_-\mathbf{1}\{X + Z \geq m\}]\\
&= \sum_{k=0}^{m-1} \left(\frac{\lambda}{k+1} \wedge \frac{k!}{\lambda^k}\right)\mathbb{E}[Z_-\mathbf{1}\{X + Z = k\}] + \frac{m!}{\lambda^m}\mathbb{E}[Z_-\mathbf{1}\{X + Z \geq m\}].
\end{aligned}$$

From (2.4) in Lemma 2.4 it follows that

$$|Q_{\{0\}}| \leq \left(1 \wedge \frac{1}{\lambda}\right)|q_0| + \left(1 \wedge \frac{1}{\lambda^2}\right)\sum_{i=1}^{\infty} |q_i|.$$

Combining (2.12) and the bounds on $|Q_{\{0\}}|, H_{\{0\}}^{(1)}$ and $H_{\{0\}}^{(2)}$ completes the proof of (1.12).

For $\lambda < v$, (1.13) follows directly from (1.9). By Lemma 2.3 for $i \le v+1$ and (2.7) in Lemma 2.4 for $i \ge v+2$, we obtain

$$H^{(1)}_{\{0,\dots,v\}} \le (1\wedge\lambda)\sum_{j=1}^{\infty}\sum_{i=1}^{v+1} j\mathbb{P}(X=i-1,Z=j) + \sum_{j=1}^{\infty}\sum_{i=v+2}^{\infty}\frac{(v+1)^2}{\lambda} j\mathbb{P}(X=i-1,Z=j)$$
$$= (1\wedge\lambda)\mathbb{E}[Z_+\mathbf{1}\{X\le v\}] + \frac{(v+1)^2}{\lambda}\mathbb{E}[Z_+\mathbf{1}\{X\ge v+1\}]$$

and

$$H^{(2)}_{\{0,\dots,v\}} \le (1\wedge\lambda)\sum_{j=1}^{\infty}\sum_{i=1}^{v+1} j\mathbb{P}(X=i+j-1,Z=-j)$$
$$+\sum_{j=1}^{\infty}\sum_{i=v+2}^{\infty}\frac{(v+1)^2}{\lambda} j\mathbb{P}(X=i+j-1,Z=-j)$$
$$= (1\wedge\lambda)\mathbb{E}[Z_-\mathbf{1}\{X+Z\le v\}] + \frac{(v+1)^2}{\lambda}\mathbb{E}[Z_-\mathbf{1}\{X+Z\ge v+1\}].$$

Moreover, Lemma 2.3 yields

$$|Q_{\{0,\dots,v\}}| \le \max_{i\in\mathbb{N}_0}|f_{\{0,\dots,v\}}(i)|\sum_{i=0}^{\infty}|q_i| \le \left(1\wedge\frac{1}{\sqrt{\lambda}}\right)\sum_{i=0}^{\infty}|q_i|.$$

Combining (2.12) with $A=\{0,\dots,v\}$ and the bounds on $|Q_{\{0,\dots,v\}}|, H^{(1)}_{\{0,\dots,v\}}$ and $H^{(2)}_{\{0,\dots,v\}}$ establishes (1.13). □

Next we derive Theorem 1.1 from Theorem 1.3.

Proof of Theorem 1.1. It follows from (1.2) that X and Z satisfy (1.8) with $\lambda=\mathbb{E}[X]$ and $q_i=0$ for $i\in\mathbb{N}_0$ and that

$$\lambda = \mathbb{E}[X] = \sum_{k=1}^{\infty} k\mathbb{P}(X=k) = \sum_{k=1}^{\infty}\lambda\mathbb{P}(X+Z=k-1) = \lambda\mathbb{P}(X+Z\ge 0),$$

whence $\mathbb{P}(X+Z\ge 0)=1$. This allows us to apply Theorem 1.3 which proves (1.3), (1.4) and (1.5). □

The next result provides some inequalities for the probability that a non-negative integer-valued random variable equals zero.

Proposition 2.6. *Let X be an integrable random variable with values in $\mathbb{N}_0$ and $\lambda>0$. Consider a random variable Z defined on the same probability space as X with values in $\mathbb{Z}$, and let $(q_i)_{i\in\mathbb{N}_0}$ be the sequence given by*

$$q_{i-1} = i\mathbb{P}(X=i) - \lambda\mathbb{P}(X+Z=i-1), \quad i\in\mathbb{N}.$$

a) If Z is non-negative and $q_i\le 0$ for $i\in\mathbb{N}_0$,

$$\mathbb{P}(X=0)\ge e^{-\lambda}.$$

b) If Z is non-positive, $\mathbb{P}(X+Z\ge 0)=1$ and $q_i\ge 0$ for $i\in\mathbb{N}_0$,

$$\mathbb{P}(X=0)\le e^{-\lambda}.$$

Proof. It follows from (2.12) and (2.13) for $A = \{0\}$ as well as $\mathbb{P}(P_\lambda = 0) = e^{-\lambda}$ that

$$\begin{aligned} &e^{-\lambda} - \mathbb{P}(X = 0) \\ &= \lambda \sum_{j \in \mathbb{Z}\setminus\{0\}} \sum_{i=1}^{\infty} \big(f_{\{0\}}(i+j) - f_{\{0\}}(i)\big)\mathbb{P}(X = i-1, Z = j) + \sum_{i=1}^{\infty} f_{\{0\}}(i) q_{i-1}. \end{aligned}$$

By the assumption that $Z \geq 0$ (resp. $Z \leq 0$ and $\mathbb{P}(X + Z \geq 0) = 1$) the first sum on the right-hand side runs only over $j \geq 1$ (resp. $j \leq -1$ and the inner sum runs over all $i \in \mathbb{N}$ with $i + j \geq 1$). Together with

$$f_{\{0\}}(i+j) - f_{\{0\}}(i) \leq 0 \text{ for } i, j \geq 1 \quad \text{and} \quad f_{\{0\}}(i+j) - f_{\{0\}}(i) \geq 0 \text{ for } j \leq -1,\ i + j \geq 1,$$

which follows from (2.5) in Lemma 2.4, $f_{\{0\}}(i) \geq 0$ for all $i \in \mathbb{N}$ and the assumptions on $(q_i)_{i \in \mathbb{N}_0}$, this leads to the desired results. □

Since (1.2) is a special case of (1.8) with $\mathbb{P}(X + Z \geq 0) = 1$ (see the proof of Theorem 1.1), the following corollary is a direct consequence of Proposition 2.6.

Corollary 2.7. *Let X be a random variable taking values in $\mathbb{N}_0$ and let $\lambda = \mathbb{E}[X] > 0$. Assume there exists a random variable Z such that* (1.2) *is satisfied.*

a) If Z is non-negative,

$$\mathbb{P}(X = 0) \geq e^{-\lambda}.$$

b) If Z is non-positive,

$$\mathbb{P}(X = 0) \leq e^{-\lambda}.$$

3 Applications

3.1 Long head runs

Consider n independent and identically distributed Bernoulli random variables. A k-head run is defined as an uninterrupted sequence of k successes, where k is a positive integer. For example, for $k = 1$, one simply studies the successes, while for $k = 2$, one considers the occurrence of two consecutive successes in a row. Several authors have investigated the number of k-head runs in a sequence of Bernoulli random variables; see e.g. the book [4]. In this subsection, we discuss the Poisson approximation of the number of non-overlapping k-runs among n i.i.d. Bernoulli random variables, denoted by $S_{n,k}$. In particular, we obtain an explicit bound on the pointwise difference between the cumulative distribution functions of $S_{n,k}$ and $P_{\mathbb{E}[S_{n,k}]}$ that is independent from the number k of required successes in a row.

Let $k \in \mathbb{N}$ and X_j, $j \in \mathbb{N}_0$, be a sequence of independent and Bernoulli distributed random variables with parameter $0 < p \leq 1/2$. We denote by $I^{(i)}$ with $i \in \mathbb{N}_0$ the random variable

$$I^{(i)} = \mathbf{1}\{X_{i-1} = 0, X_i = 1, \ldots, X_{i+k-1} = 1\},$$

where $X_{-1} = 0$. For $k \leq n$ the number $S_{n,k}$ of non-overlapping k-runs in $X_0, \ldots, X_{n-1}$ is given by

$$S_{n,k} = \sum_{i=0}^{n-k} I^{(i)}. \tag{3.1}$$

Theorem 3.1. *Let $S_{n,k}$ be the random variable given by* (3.1) *with $k, n \in \mathbb{N}, k \leq n$. Then,*

$$d_{TV}\big(S_{n,k}, P_{\mathbb{E}[S_{n,k}]}\big) \leq (2k+1)\big(1 \wedge \mathbb{E}[S_{n,k}]\big)p^k. \tag{3.2}$$

Moreover, for $v \in \mathbb{N}_0$ and $n \geq 2$,

$$\max_{1\leq k\leq n} |\mathbb{P}(S_{n,k} \leq v) - \mathbb{P}(P_{\mathbb{E}[S_{n,k}]} \leq v)| \leq 40(v+2)^2 \frac{\log n}{n}. \tag{3.3}$$

The bound (3.2) was shown in [27, Corollary 15] as a consequence of [1, Theorem 1]. The Poisson approximation for $S_{n,k}$ is also investigated in e.g. [2, 7, 16, 20]. The explicit bound in (3.3) on the pointwise difference between the cumulative distribution functions of $S_{n,k}$ and $P_{\mathbb{E}[S_{n,k}]}$ does not depend on the number k of required successes in a row. Hence, (3.3) improves [26, Corollary 3.23] and [27, Corollary 16] because we found an explicit bound. Since the proof of Theorem 3.1 is based on Theorem 1.1, by applying the second inequality of (1.3) in Theorem 1.1, it is possible to attain a bound on the Wasserstein distance between $S_{n,k}$ and $P_{\mathbb{E}[S_{n,k}]}$.

We believe that the inequalities in Theorem 3.1 can be extended in two directions. Firstly, by applying the full generality of our main result, Theorem 1.3, it might be possible to derive similar inequalities also for non-independent Bernoulli random variables if they satisfy some mixing and local conditions as in [32, Theorem 3.1]. Secondly, we may also establish a higher dimensional version of Theorem 3.1. Indeed, if we define a connected component in $\mathbb{Z}^d$ for $d \in \mathbb{N}$, as a family of points $A \subset \mathbb{Z}^d$ such that for any $x \in A$ the minimum Euclidean distance between x and $A \setminus \{x\}$ is 1, then $S_{n,k}$ counts the number of connected components of the set $\{j \in \{0, 1, \ldots, n-k\} : X_j = 1\}$ in $\mathbb{Z}$ with size at least k. Then, by taking a family of i.i.d. Bernoulli random variables $X_j, j \in \mathbb{Z}^d$, we may also also derive a Poisson approximation result for the number of connected components of the set of points $\{j \in [-n, n]^d \cap \mathbb{Z}^d : X_j = 1\}$ in $\mathbb{Z}^d$ of cardinality at least k.

For the proof of Theorem 3.1 we define

$$U_\ell = \sum_{i=0\vee(\ell-k)}^{(n-k)\wedge(\ell+k)} I^{(i)}, \quad \ell = 0, \ldots, n-k,$$

where $a \vee b = \max\{a, b\}$ for any $a, b \in \mathbb{R}$, and let Y be a random variable independent from $X_j, j \in \mathbb{N}_0$, and with distribution given by

$$\mathbb{P}(Y = \ell) = \frac{\mathbb{E}[I^{(\ell)}]}{\mathbb{E}[S_{n,k}]}, \quad \ell = 0, \ldots, n-k.$$

The next proposition follows directly from the construction of a size-bias coupling given in Remark 1.2 (iii). Nevertheless, we provide a complete proof, which is rather instructive.

Proposition 3.2. *Let $k, n \in \mathbb{N}$ with $k \leq n$. For any $m \in \mathbb{N}$,*

$$m\mathbb{P}(S_{n,k} = m) = \mathbb{E}[S_{n,k}]\mathbb{P}(S_{n,k} - U_Y = m-1).$$

Proof. Let $\ell \in \{0, \ldots, n-k\}$ and $m \in \mathbb{N}$ be fixed. Then, we have

$$\mathbb{E}[I^{(\ell)}\mathbf{1}\{S_{n,k} - I^{(\ell)} = m-1\}] = \mathbb{E}[I^{(\ell)}\mathbf{1}\{S_{n,k} - U_\ell = m-1\}],$$

where we used that $I^{(\ell)} = 1$ implies $U_\ell = 1$ because we consider non-overlapping k-runs. Since $I^{(\ell)}$ and $S_{n,k} - U_\ell$ are independent, it follows that

$$\begin{aligned} m\mathbb{P}(S_{n,k} = m) &= \sum_{\ell=0}^{n-k} \mathbb{E}[I^{(\ell)}\mathbf{1}\{S_{n,k} = m\}] = \sum_{\ell=0}^{n-k} \mathbb{E}[I^{(\ell)}\mathbf{1}\{S_{n,k} - I^{(\ell)} = m-1\}] \\ &= \sum_{\ell=0}^{n-k} \mathbb{E}[I^{(\ell)}]\mathbb{P}(S_{n,k} - U_\ell = m-1) = \mathbb{E}[S_{n,k}]\mathbb{P}(S_{n,k} - U_Y = m-1), \end{aligned}$$

which concludes the proof. □

Remark 3.3. Since $U_Y \geq 0$, from Corollary 2.7 b) it follows that $\mathbb{P}(S_{n,k}=0) \leq e^{-\mathbb{E}[S_{n,k}]}$. Thus, straightforward calculations imply that

$$\mathbb{P}(S_{n,k}=0) \leq \exp\big(-(n-k+1)p^k(1-p)\big).$$

Proof of Theorem 3.1. From (1.3) in Theorem 1.1 and Proposition 3.2, it follows that

$$d_{TV}(S_{n,k}, P_{\mathbb{E}[S_{n,k}]}) \leq (1 \wedge \mathbb{E}[S_{n,k}])\mathbb{E}[U_Y] \leq (2k+1)(1 \wedge \mathbb{E}[S_{n,k}])p^k,$$

where we used $\mathbb{E}[U_\ell] \leq (2k+1)p^k$ for $\ell = 0, \dots, n-k$ in the last step. This proves (3.2).

Let $n \geq 2$ be fixed. Since $(2k+1)p^k, k \geq 1$, is decreasing in k for any $p \leq 1/2$, by (3.2) we deduce for $k \geq 2\log n$ that

$$|\mathbb{P}(S_{n,k} \leq v) - \mathbb{P}(P_{\mathbb{E}[S_{n,k}]} \leq v)| \leq (2k+1)p^k \leq (4\log n + 1)2^{-2\log n} \leq \frac{4\log n + 1}{n}. \quad (3.4)$$

Let $k < 2\log n$. From (1.4) in Theorem 1.1 with $m = 1$ for $v = 0$ and (1.5) in Theorem 1.1 for $v \in \mathbb{N}$, it follows that

$$|\mathbb{P}(S_{n,k} \leq v) - \mathbb{P}(P_{\mathbb{E}[S_{n,k}]} \leq v)| \leq \frac{(v+1)^2\mathbb{E}[U_Y]}{\mathbb{E}[S_{n,k}]} + \mathbb{E}[U_Y \mathbf{1}\{S_{n,k} - U_Y \leq v\}]. \quad (3.5)$$

From $0 \leq U_\ell \leq 2$ for $\ell \in \{0, \dots, n-k\}$ and the definition of Y it follows that

$$\begin{aligned}\mathbb{E}[U_Y \mathbf{1}\{S_{n,k} - U_Y \leq v\}] &\leq \mathbb{E}[U_Y \mathbf{1}\{S_{n,k} \leq v+2\}] = \sum_{\ell=0}^{n-k} \frac{\mathbb{E}[I^{(\ell)}]}{\mathbb{E}[S_{n,k}]}\mathbb{E}[U_\ell \mathbf{1}\{S_{n,k} \leq v+2\}] \\ &\leq \frac{p^k}{\mathbb{E}[S_{n,k}]}\mathbb{E}\sum_{\ell=0}^{n-k}\sum_{i=0\vee(\ell-k)}^{(n-k)\wedge(\ell+k)} I^{(i)}\mathbf{1}\{S_{n,k} \leq v+2\}.\end{aligned}$$

Thus, by the inequality

$$\sum_{\ell=0}^{n-k}\sum_{i=0\vee(\ell-k)}^{(n-k)\wedge(\ell+k)} a_i \leq (2k+1)\sum_{m=0}^{n-k} a_m, \quad a_0, \dots, a_{n-k} \geq 0,$$

we obtain

$$\mathbb{E}[U_Y \mathbf{1}\{S_{n,k} - U_Y \leq v\}] \leq \frac{(2k+1)p^k}{\mathbb{E}[S_{n,k}]}\mathbb{E}[S_{n,k}\mathbf{1}\{S_{n,k} \leq v+2\}] \leq \frac{(2k+1)p^k(v+2)}{\mathbb{E}[S_{n,k}]}.$$

Together with (3.5) and the inequalities

$$\mathbb{E}[S_{n,k}] \geq (n-k+1)p^k/2 \quad \text{and} \quad \mathbb{E}\big[U_Y\big] \leq (2k+1)p^k,$$

this shows for $k < 2\log n$ and $n > 4\log n$ that

$$\begin{aligned}|\mathbb{P}(S_{n,k} \leq v) - \mathbb{P}(P_{\mathbb{E}[S_{n,k}]} \leq v)| &\leq \frac{2(v+1)^2(2k+1)}{n-k+1} + \frac{2(v+2)(2k+1)}{n-k+1} \\ &\leq \frac{4(v+2)^2(4\log n+1)}{n-2\log n} \leq \frac{40(v+2)^2\log n}{n},\end{aligned}$$

where we used the inequalities $4\log n + 1 \leq 5\log n$ and $n - 2\log n \geq n/2$ for $n > 4\log n$ in the last step. Combining this and (3.4) establishes (3.3) for $n > 4\log n$. In conclusion, note that $n > 4\log n$ for $n > 10$, and for $2 \leq n \leq 10$, the right-hand side of (3.3) is greater than 1. Thus, (3.3) holds for all $n \geq 2$. □

3.2 Minimal circumscribed radii of Poisson-Voronoi tessellations

In this subsection, we consider circumscribed radii of stationary Poisson-Voronoi tessellations. The aim is to continue the work started in [11] by proving that the Kolmogorov distance between a transform of the minimal circumscribed radius and a Weibull random variable converges to 0 at a rate of $1/t^{1/(d+1)}$ when the intensity t of the underlying Poisson process goes to infinity.

For any locally finite counting measure ν on $\mathbb{R}^d$, we denote by $N(x,\nu)$ the Voronoi cell with nucleus $x \in \mathbb{R}^d$ generated by $\nu + \delta_x$, that is

$$N(x,\nu) = \left\{ y \in \mathbb{R}^d \,:\, \|y-x\| \leq \|y-x'\|, x \neq x' \in \nu \right\}.$$

Voronoi tessellations, i.e., tessellations consisting of Voronoi cells $N(x,\nu)$, $x \in \nu$, arise in different fields such as biology [33], astrophysics [34] and communication networks [9]. For more details on Poisson-Voronoi tessellations, i.e., Voronoi tessellations generated by an underlying Poisson process, we refer the reader to e.g. [10, 25, 37]. We denote by $\mathbf{B}(x,r)$ the open ball centered at $x \in \mathbb{R}^d$ with radius $r > 0$. The circumscribed radius of the Voronoi cell $N(x,\nu)$ is defined as

$$C(x,\nu) = \inf\left\{ R \geq 0 \,:\, \mathbf{B}(x,R) \supset N(x,\nu) \right\},$$

i.e., the circumscribed radius is the smallest radius for which the closed ball centered at the nucleus contains the cell.

Throughout this subsection we consider the stationary Poisson-Voronoi tessellation generated by a Poisson process η_t on $\mathbb{R}^d$ with intensity measure $t\lambda_d$, $t > 0$, where λ_d is the d-dimensional Lebesgue measure. Let $W \subset \mathbb{R}^d$ be a measurable set with $\lambda_d(W) = 1$. For any Voronoi cell $N(x,\eta_t)$ with $x \in \eta_t \cap W$, we take the circumscribed radius of the cell, and we define the point process ξ_t on the positive half line as

$$\xi_t = \sum_{x \in \eta_t \cap W} \delta_{\alpha_2 k_d t^{(d+2)/(d+1)} C(x,\eta_t)^d}. \tag{3.6}$$

Here δ_z is the Dirac measure at $z \in \mathbb{R}$, k_d denotes the volume of the d-dimensional unit ball, and the constant $\alpha_2 > 0$ is given by

$$\alpha_2 = \left(\frac{2^{d(d+1)}}{(d+1)!} p_{d+1} \right)^{1/(d+1)} \tag{3.7}$$

with

$$p_{d+1} := \mathbb{P}\Big(N\Big(0, \sum_{j=1}^{d+1} \delta_{Y_j}\Big) \subseteq \mathbf{B}(0,1) \Big), \tag{3.8}$$

where $Y_1, \ldots, Y_{d+1}$ are independent and uniformly distributed random points in $\mathbf{B}(0,2)$. We denote by T_t the first arrival time of ξ_t, i.e.,

$$T_t = \min_{x \in \eta_t \cap W} \alpha_2 k_d t^{(d+2)/(d+1)} C(x,\eta_t)^d, \tag{3.9}$$

which is - up to a rescaling - the d-th power of the minimal circumscribed radius of the cells with nucleus in W. Recall that a random variable Y has a Weibull distribution if its cumulative distribution function is given by $\mathbb{P}(Y \leq u) = 1 - e^{-(u/s)^k}$ for $u \geq 0$, and 0 otherwise; $k > 0$ is the shape parameter and $s > 0$ is the scale parameter.

Theorem 3.4. *Suppose $t \geq 1$. Let ξ_t and T_t be the point process and the random variable given by* (3.6) *and* (3.9)*, respectively. Let Y be a Weibull distributed random*

variable with shape parameter $d+1$ *and scale parameter* 1*. Then, there exist constants* $C_{\mathrm{TV}}, C_K > 0$ *only depending on* d *such that*

$$d_{TV}\big(\xi_t([0,u]), P_{u^{d+1}}\big) \leq C_{\mathrm{TV}} \frac{u^{d+2}}{t^{1/(d+1)}} \tag{3.10}$$

for $u > 0$*, and*

$$d_K(T_t, Y) \leq \frac{C_K}{t^{1/(d+1)}}. \tag{3.11}$$

Note that explicit formulas for the constants C_{TV} and C_K are given in the proof of Theorem 3.4, which can be found in Section 4.1. In [11, Theorem 1, Equation (2d)], the weak convergence of T_t to Y as $t \to \infty$ is shown. For an underlying inhomogeneous Poisson process, the weak convergence of ξ_t to a Poisson process and the weak convergence of T_t to Y are proven in [32, Section 3.3]. Although we only consider stationary Poisson processes, we believe that the arguments employed in this subsection may also establish similar results on the minimal circumscribed radius for more general Poisson processes with a different rate of convergence in t under some constraints on the density (e.g. Hölder continuity). To the best of our knowledge, the present paper is the first time the rates of convergence for the Poisson approximation of $\xi_t([0,u])$ and the Weibull approximation of T_t have been addressed. Since the proof of Theorem 3.4 is based on Theorem 1.3, together with (1.11) in Theorem 1.3, the same arguments used to show (3.10) may also lead to a bound on the Wasserstein distance between $\xi_t([0,u])$ and $P_{u^{d+1}}$.

3.3 Maximal inradii of Poisson-Voronoi tessellations

In this subsection, we consider the inradii of stationary Poisson-Voronoi tessellations. Recall that the inradius of a cell is the largest radius for which the ball centered at the nucleus is contained in the cell. The aim is to continue the work started in [11] by proving that the Kolmogorov distance between a transform of the largest inradius and a Gumbel random variable converges to 0 at a rate of $\log(t)/\sqrt{t}$ as the intensity t of the underlying Poisson process goes to infinity. More details on Poisson-Voronoi tessellations are given in Subsection 3.2.

Let $W \subset \mathbb{R}^d$ be a measurable set with Lebesgue measure $\lambda_d(W) = 1$. Let η_t be a Poisson process on $\mathbb{R}^d$ with intensity measure $t\lambda_d, t > 0$. For $x \in \mathbb{R}^d$ and a locally finite counting measure μ on $\mathbb{R}^d$ define

$$h_t(x, \mu) = \min_{y \in \mu \setminus \{x\}} t k_d \|x - y\|^d - \log(t),$$

where k_d is the volume of the d-dimensional unit ball. Note that for any $x \in \eta_t$, $\min\{\|x - y\| : y \in \eta_t \setminus \{x\}\}$ is twice the inradius of the Voronoi cell with nucleus x generated by η_t. Then, the random variable

$$R_t = \max_{x \in \eta_t \cap W} h_t(x, \eta_t) \tag{3.12}$$

is a transform of the maximal inradius over the cells with nucleus in W. We define the point process ξ_t as

$$\xi_t = \xi_t(\eta_t) = \sum_{x \in \eta_t \cap W} \delta_{h_t(x, \eta_t)}. \tag{3.13}$$

Recall that a random variable G has a standard Gumbel distribution if its cumulative distribution function is given by $\mathbb{P}(G \leq u) = e^{-e^{-u}}$ for $u \in \mathbb{R}$.

Theorem 3.5. *Suppose $t > e^2$. Let R_t and ξ_t be the random variable and the point process given by* (3.12) *and* (3.13)*, respectively. Let G be a random variable with a standard Gumbel distribution. Then,*

$$d_{TV}\big(\xi_t((u,\infty)), P_{e^{-u}}\big) \leq 2^d \frac{u+\log(t)}{e^{u/2}\sqrt{t}} + \frac{u+\log(t)}{e^u t} \tag{3.14}$$

for $u > -\log(t)$, and

$$d_K(R_t, G) \leq [2^{d+2}(4^d + 2^d + 2) + 1]\frac{\log(t)}{\sqrt{t}}. \tag{3.15}$$

The main achievement of Theorem 3.5 is the rate of convergence for the Kolmogorov distance in (3.15). In [11, Theorem 1, Equation (2a)], the weak convergence of R_t to a Gumbel random variable is proven. For $d = 2$ one obtains from the proof of [13, Proposition 8] that for any fixed $u \in \mathbb{R}$ the difference between $\mathbb{P}(R_t \leq u)$ and $\mathbb{P}(G \leq u)$ behaves like $O(\log(t)/\sqrt{t})$, where the constant hidden in the big-O-notation depends on u. However this result does not permit to bound the difference between $\mathbb{P}(R_t \leq u)$ and $\mathbb{P}(G \leq u)$ uniformly in $u \in \mathbb{R}$, whence it does not lead to a bound for the Kolmogorov distance. Note that [13, Proposition 8] concerns the maximal inradii of planar Gauss-Voronoi tessellations, which are generated by a Poisson cluster process and include planar Poisson-Voronoi tessellations as a special case. For this model it is shown that for any fixed $u \in \mathbb{R}$, $|\mathbb{P}(R_t \leq u) - \mathbb{P}(G \leq u)|$ behaves like $O(\log(t)^{-1/2})$, where the big-O-term depends on u.

For an underlying inhomogeneous Poisson process, the weak convergence of ξ_t to a Poisson process and the weak convergence of R_t to G are established in [32, Section 3.2], and for an underlying inhomogeneous binomial point process, the weak convergence of R_t to G is studied in [17, Theorem 1]. As for the results stated in Subsection 3.2 about the minimal circumscribed radius, we believe that similar arguments as in this subsection could lead to comparable results with a different rate of convergence in t for the maximal inradius of a Voronoi tessellation generated by an inhomogeneous Poisson processes under some constraints on the density.

Counting cells whose inradius is larger than a given value is equivalent to counting isolated vertices in random geometric graphs. The related problem of finding the longest edge of a k-nearest neighbor graph or a minimal spanning tree is studied, for example, in [30] or [29, Chapter 8] for underlying finite Poisson processes or binomial point processes, where one needs to take care of boundary effects.

The proof of Theorem 3.5 is given in Section 4.2. Since it is based on Theorem 1.1, together with the second inequality of (1.3) in Theorem 1.1, the same arguments used to show (3.14) may also lead to a bound on the Wasserstein distance between $\xi_t((u,\infty))$ and $P_{e^{-u}}$.

3.4 U-statistics of binomial point processes

Let $(\mathbb{X}, \mathcal{X})$ be a measurable space. A point process on $\mathbb{X}$ is a random element in the set of all σ-finite counting measures on $\mathbb{X}$, denoted by $\mathbf{N}_{\mathbb{X}}$, which is measurable with respect to the σ-field generated by the sets of the form

$$\{\mu \in \mathbf{N}_{\mathbb{X}} \,:\, \mu(B) = k\}, \quad k \in \mathbb{N}_0, B \in \mathcal{X}.$$

We consider a binomial point process β_n on $\mathbb{X}$ of $n \in \mathbb{N}$ independent points in $\mathbb{X}$ that are distributed according to a probability measure K. Let $\ell \in \mathbb{N}$ and let $h : \mathbb{X}^\ell \to \{0,1\}$ be a measurable symmetric function. In the following we study the U-statistic

$$S = \frac{1}{\ell!} \sum_{(x_1,\ldots,x_\ell) \in \beta^\ell_{n,\neq}} h(x_1, ..., x_\ell),$$

where $\beta^{\ell}_{n,\neq}$ denotes the set of all ℓ-tuples of distinct points of β_n. We refer to the monographs [19, 23] for more details on U-statistics and their applications in statistics. A straightforward computation shows that

$$\lambda := \mathbb{E}[S] = \frac{(n)_\ell}{\ell!} \int_{\mathbb{X}^\ell} h(x_1, \dots, x_\ell) dK^\ell(x_1, \dots, x_\ell),$$

where $(n)_\ell$ stands for the ℓ-th descending factorial.

In this subsection, we establish bounds on the Poisson approximation of S in the total variation and Wasserstein distances. We also provide bounds on the pointwise difference between the cumulative distribution functions of S and P_λ. To this end, we define

$$r = \max_{1\le i\le \ell-1} (n)_{2\ell-i} \int_{\mathbb{X}^i} \left(\int_{\mathbb{X}^{\ell-i}} h(x_1, \dots, x_\ell) dK^{\ell-i}(x_{i+1}, \dots, x_\ell) \right)^2 dK^i(x_1, \dots, x_i)$$

for $\ell \ge 2$, and put $r = 0$ for $\ell = 1$. Moreover for $n \ge 2\ell$, we define

$$\tilde{S} = \frac{1}{\ell!} \sum_{(x_1,\dots,x_\ell)\in\beta^{\ell}_{n-2\ell,\neq}} h(x_1, ..., x_\ell).$$

Theorem 3.6. *Let $n \ge 2\ell$ and let S, $\lambda > 0$, r and $\tilde{S}$ be as above. Then,*

$$d_{TV}(S, P_\lambda) \le (1 \wedge \lambda)\left(\frac{2^\ell r}{\ell!\lambda} + \frac{2\ell^2\lambda}{n}\right) \quad \textit{and} \quad d_W(S, P_\lambda) \le (1.1437\sqrt{\lambda} \wedge \lambda)\left(\frac{2^\ell r}{\ell!\lambda} + \frac{2\ell^2\lambda}{n}\right). \tag{3.16}$$

Moreover, for all $m \in \mathbb{N}$,

$$\left|\mathbb{P}(S = 0) - e^{-\lambda}\right| \le \left[\sum_{k=0}^{m-1} \left(\frac{\lambda}{k+1} \wedge \frac{k!}{\lambda^k}\right)\mathbb{P}(\tilde{S} \le k) + \frac{m!}{\lambda^m}\right]\left(\frac{2^\ell r}{\ell!\lambda} + \frac{2\ell^2\lambda}{n}\right) \tag{3.17}$$

and for all $v \in \mathbb{N}$,

$$|\mathbb{P}(S \le v) - \mathbb{P}(P_\lambda \le v)| \le \left[\frac{(v+1)^2}{\lambda} + \mathbb{P}(\tilde{S} \le v)\right]\left(\frac{2^\ell r}{\ell!\lambda} + \frac{2\ell^2\lambda}{n}\right). \tag{3.18}$$

For fixed ℓ and λ the right-hand sides in Theorem 3.6 converge to zero as $n \to \infty$ if and only if $r \to 0$ as $n \to \infty$. The term in the definition of r for $i \in \{1, \dots, \ell-1\}$ is the expected number of combinations of distinct points $x_1, \dots, x_i$, $y_1, \dots, y_{\ell-i}$ and $y'_1, \dots, y'_{\ell-i}$ of β_n such that $h(x_1, \dots, x_i, y_1, \dots, y_{\ell-i}) = 1$ and $h(x_1, \dots, x_i, y'_1, \dots, y'_{\ell-i}) = 1$, i.e., $x_1, \dots, x_i$ contribute to more than one non-zero summand in S. Thus, $r \to 0$ as $n \to \infty$ implies that the probability that there is a point of β_n that belongs to more than one k-tuple (up to permutations) with non-zero h goes to zero.

The bound on the Wasserstein distance in (3.16) slightly improves that in [15, Theorem 7.1] since it has a better order in λ. The bound for the total variation distance was also derived in [39, Proposition 2] by rewriting [6, Theorem 2]. By means of (3.17), one can study for some measurable symmetric function $g : \mathbb{X}^\ell \to \mathbb{R}$ the maximum (minimum) of $g(p)$ over all $p \in \beta^{\ell}_{n,\neq}$, which is called U-max-statistic (U-min-statistic). This is possible because for any $u \in \mathbb{R}$, the probability that $\max_{p\in\beta^{\ell}_{n,\neq}} g(p)$ is less than u can be written as the probability that $\sum_{p\in\beta^{\ell}_{n,\neq}} \mathbf{1}\{g(p) \ge u\}$ equals 0. Limit theorems for U-max-statistics were considered in [21], yet without providing approximation results with respect to any distance; see also [24]. In contrast to these works, (3.17) may lead to approximation results in the Kolmogorov distance; see Theorem 3.9 in Subsection 3.6 and the discussion below it. To the best of our knowledge, the last two inequalities presented in Theorem 3.6 have no analogues in the literature.

From now on assume that $n \geq \ell$. Let χ be a point process of ℓ random points $X'_1, \dots, X'_\ell$ in $\mathbb{X}$ that are independent of β_n and distributed such that

$$\mathbb{P}((X'_1, \dots, X'_\ell) \in A) = \frac{(n)_\ell}{\ell! \lambda} \int_{\mathbb{X}^\ell} \mathbf{1}\{(x_1, \dots, x_\ell) \in A\} h(x_1, \dots, x_\ell) dK^\ell(x_1, \dots, x_\ell)$$

for all A from the product σ-field $\mathcal{X}^\ell$. Now we define

$$S' = -h(X'_1, \dots, X'_\ell) + \frac{1}{\ell!} \sum_{(x_1, \dots, x_\ell) \in (\beta_{n-\ell} \cup \chi)^\ell_{\neq}} h(x_1, ..., x_\ell).$$

Although the next proposition can be derived by the construction of a size-bias coupling described in Remark 1.2 (iii), we provide a complete proof that relies on similar arguments as the proof for U-statistics of Poisson processes in the next subsection.

Proposition 3.7. *For all $n \geq \ell$ and $k \in \mathbb{N}$,*

$$k\mathbb{P}(S = k) = \lambda \mathbb{P}(S' = k - 1).$$

Proof. We have that

$$k\mathbb{P}(S = k) = \mathbb{E}[k\mathbf{1}\{S = k\}] = \frac{1}{\ell!} \mathbb{E} \sum_{(x_1, \dots, x_\ell) \in \beta^\ell_{n, \neq}} h(x_1, \dots, x_\ell) \mathbf{1}\{S = k\}.$$

Using the fact that for any measurable map $g : \mathbb{X}^u \times \mathbf{N}_\mathbb{X} \to [0, \infty)$ with $u \in \mathbb{N}$,

$$\mathbb{E} \sum_{(x_1, \dots, x_u) \in \beta^u_{n, \neq}} g(x_1, \dots, x_u, \beta_n) = (n)_u \int_{\mathbb{X}^u} \mathbb{E}[g(x_1, \dots, x_u, \beta_{n-u} + \sum_{i=1}^u \delta_{x_i})] dK^u(x_1, \dots, x_u),$$

we obtain

$$\begin{aligned} k\mathbb{P}(S = k) &= \frac{(n)_\ell}{\ell!} \int_{\mathbb{X}^\ell} h(x_1, \dots, x_\ell) \mathbb{P}\bigg(\frac{1}{\ell!} \sum_{(y_1, \dots, y_\ell) \in (\beta_{n-\ell} \cup \{x_1, \dots, x_\ell\})^\ell_{\neq}} h(y_1, \dots, y_\ell) = k \bigg) \\ &\qquad \times dK^\ell(x_1, \dots, x_\ell) \\ &= \lambda \mathbb{P}(S' + h(X'_1, \dots, X'_\ell) = k) = \lambda \mathbb{P}(S' = k - 1), \end{aligned}$$

where we used $h(X'_1, \dots, X'_\ell) = 1$ a.s. in the last step. This concludes the proof. □

Proof of Theorem 3.6. Suppose $n \geq 2\ell$. Our goal is to apply Theorem 1.1 with $Z = S' - S$, which satisfies the assumption (1.2) by Proposition 3.7. We define $s : \mathbf{N}_\mathbb{X} \to \mathbb{R}$ by

$$s(\nu) = \frac{1}{\ell!} \sum_{(x_1, \dots, x_\ell) \in \nu^\ell_{\neq}} h(x_1, \dots, x_\ell)$$

so that $S = s(\beta_n)$ and $S' = s(\beta_{n-\ell} + \chi) - h(X'_1, \dots, X'_\ell)$. By the monotonicity of s, we have

$$\begin{aligned} |Z| = |S' - S| &= |s(\beta_{n-\ell} + \chi) - h(X'_1, \dots, X'_\ell) - s(\beta_{n-\ell}) - (s(\beta_n) - s(\beta_{n-\ell}))| \\ &\leq (s(\beta_{n-\ell} + \chi) - h(X'_1, \dots, X'_\ell) - s(\beta_{n-\ell})) + s(\beta_n) - s(\beta_{n-\ell}). \end{aligned} \tag{3.19}$$

Together with

$$\begin{aligned} &s(\beta_{n-\ell} + \chi) - h(X'_1, \dots, X'_\ell) - s(\beta_{n-\ell}) + s(\beta_n) - s(\beta_{n-\ell}) \\ &= s(\beta_{n-\ell} + \chi) - h(X'_1, \dots, X'_\ell) - s(\beta_n) + 2(s(\beta_n) - s(\beta_{n-\ell})) = Z + 2(s(\beta_n) - s(\beta_{n-\ell})) \end{aligned} \tag{3.20}$$

this implies

$$\mathbb{E}[|Z|] \leq \mathbb{E}[Z] + 2\mathbb{E}[s(\beta_n) - s(\beta_{n-\ell})].$$

From (1.7) in Remark 1.2 we know that

$$\mathbb{E}[Z] = \frac{1}{\lambda}(\mathrm{Var}(S) - \lambda) = \frac{1}{\lambda}(\mathbb{E}[S^2] - \lambda^2 - \lambda).$$

Thus, it follows from [15, Lemma 6.1] and the definition of r that

$$\mathbb{E}[Z] \leq \frac{2^\ell r}{\ell!\lambda}.$$

A straightforward computation shows that

$$\mathbb{E}[s(\beta_n) - s(\beta_{n-\ell})] = \left(1 - \frac{(n-\ell)_\ell}{(n)_\ell}\right)\lambda = \frac{(n)_\ell - (n-\ell)_\ell}{(n)_\ell}\lambda \leq \frac{\ell^2(n-1)_{\ell-1}}{(n)_\ell}\lambda = \frac{\ell^2\lambda}{n}.$$

Combining the previous estimates yields

$$\mathbb{E}[|Z|] \leq \frac{2^\ell r}{\ell!\lambda} + \frac{2\ell^2\lambda}{n}$$

so that (3.16) follows from (1.3).

Let $k \in \mathbb{N}$ be fixed. Note that $S \geq s(\beta_{n-\ell})$ and $S' \geq s(\beta_{n-\ell})$. If $Z \geq 0$, this implies

$$\mathbf{1}\{S - Z_- \leq k\} = \mathbf{1}\{S \leq k\} \leq \mathbf{1}\{s(\beta_{n-\ell}) \leq k\}.$$

For $Z \leq 0$ we obtain

$$\mathbf{1}\{S - Z_- \leq k\} = \mathbf{1}\{S + Z \leq k\} = \mathbf{1}\{S' \leq k\} \leq \mathbf{1}\{s(\beta_{n-\ell}) \leq k\}.$$

Combing the two cases leads to

$$\mathbf{1}\{S - Z_- = k\} \leq \mathbf{1}\{S - Z_- \leq k\} \leq \mathbf{1}\{s(\beta_{n-\ell}) \leq k\}.$$

Together with (3.19) we obtain

$$\begin{aligned}&\mathbb{E}[|Z|\mathbf{1}\{S - Z_- = k\}]\\ &\leq \mathbb{E}[\mathbf{1}\{s(\beta_{n-\ell}) \leq k\}(s(\beta_{n-\ell} + \chi) - h(X'_1, \dots, X'_\ell) - s(\beta_{n-\ell}) + s(\beta_n) - s(\beta_{n-\ell}))].\end{aligned} \tag{3.21}$$

For $u \in \{1, \dots, \ell - 1\}$ and $g : \mathbb{X}^u \to [0, \infty)$, we have

$$\begin{aligned}&\mathbb{E}[\mathbf{1}\{s(\beta_{n-\ell}) \leq k\} \sum_{(x_1,\dots,x_u)\in\beta^u_{n-\ell,\neq}} g(x_1, \dots, x_u)]\\ &= (n-\ell)_u \int_{\mathbb{X}^u} \mathbb{P}(s(\beta_{n-\ell-u} + \sum_{i=1}^u \delta_{x_i}) \leq k) g(x_1, \dots, x_u) dK^u(x_1, \dots, x_u)\\ &\leq \mathbb{P}(s(\beta_{n-2\ell}) \leq k)(n-\ell)_u \int_{\mathbb{X}^u} g(x_1, \dots, x_u) dK^u(x_1, \dots, x_u)\\ &= \mathbb{P}(s(\beta_{n-2\ell}) \leq k)\mathbb{E} \sum_{(x_1,\dots,x_u)\in\beta^u_{n-\ell,\neq}} g(x_1, \dots, x_u),\end{aligned} \tag{3.22}$$

where the inequality follows from the monotonicity of s. Because of

$$s(\beta_{n-\ell} + \chi) - h(X'_1, \dots, X'_\ell) - s(\beta_{n-\ell}) = \sum_{u=1}^{\ell-1} \sum_{(x_1,\dots,x_u)\in\beta^u_{n-\ell,\neq}} \tilde{h}_u(x_1, \dots, x_u; \chi)$$

and

$$s(\beta_n) - s(\beta_{n-\ell}) = \sum_{u=1}^{\ell-1} \sum_{(x_1,\dots,x_u)\in\beta^u_{n-\ell,\neq}} \overline{h}_u(x_1,\dots,x_u;\beta_n \setminus \beta_{n-\ell})$$

with suitable functions $\tilde{h}_u$ and $\overline{h}_u$, $u \in \{1,\dots,\ell-1\}$, we can rewrite the second factor on the right-hand side of (3.21) as sum of U-statistics with respect to $\beta_{n-\ell}$. Now an application of (3.22) and (3.20) yield

$$\begin{aligned}
&\mathbb{E}[|Z|\mathbf{1}\{S - Z_- = k\}] \\
&\le \mathbb{P}(s(\beta_{n-2\ell}) \le k)\mathbb{E}[s(\beta_{n-\ell}+\chi) - h(X'_1,\dots,X'_\ell) - s(\beta_{n-\ell}) + s(\beta_n) - s(\beta_{n-\ell})] \\
&= \mathbb{P}(s(\beta_{n-2\ell}) \le k)\big(\mathbb{E}[Z] + 2\mathbb{E}[s(\beta_n) - s(\beta_{n-\ell})]\big).
\end{aligned}$$

Bounding the second factor on the right-hand side as above leads to

$$\mathbb{E}[|Z|\mathbf{1}\{S - Z_- = k\}] \le \mathbb{P}(s(\beta_{n-2\ell}) \le k)\left(\frac{2^\ell r}{\ell!\lambda} + \frac{2\ell^2\lambda}{n}\right).$$

Thus, (3.17) and (3.18) are immediate consequences of (1.4) and (1.5). □

3.5 U-statistics of Poisson processes

In this subsection, we study the Poisson approximation of U-statistics, where one sums over all ℓ-tuples of distinct points of a Poisson process instead of those of a binomial point process as in the previous subsection. In this case, the summation can run over infinitely many ℓ-tuples. As the results for U-statistics with binomial input in Subsection 3.4, the theory developed herein permits to study extreme value problems arising in stochastic geometry. For example, in the next subsection, we employ our main result for U-statistics with Poisson input to investigate the limiting behavior of the minimum inter-point distance between the points of a Poisson process in $\mathbb{R}^d$.

Let $(\mathbb{X},\mathcal{X})$ be a measurable space and let η be a Poisson process with a σ-finite intensity measure L on $\mathbb{X}$. For a fixed $\ell \in \mathbb{N}$ and a symmetric measurable function $h : \mathbb{X}^\ell \to \{0,1\}$ that is integrable with respect to L^ℓ we consider the U-statistic

$$S = \frac{1}{\ell!} \sum_{(x_1,\dots,x_\ell)\in\eta^\ell_{\neq}} h(x_1,\dots,x_\ell),$$

where $\eta^\ell_{\neq}$ denotes the set of all ℓ-tuples of distinct points of η. It follows from the multivariate Mecke formula that

$$\lambda := \mathbb{E}[S] = \frac{1}{\ell!}\int_{\mathbb{X}^\ell} h(x_1,\dots,x_\ell)dL^\ell(x_1,\dots,x_\ell).$$

We define

$$r = \max_{1\le i\le \ell-1} \int_{\mathbb{X}^i}\left(\int_{\mathbb{X}^{\ell-i}} h(x_1,\dots,x_\ell)dL^{\ell-i}(x_{i+1},\dots,x_\ell)\right)^2 dL^i(x_1,\dots,x_i)$$

for $\ell \ge 2$, and put $r = 0$ for $\ell = 1$. The expression r is used to quantify the accuracy of the Poisson approximation for S and it is the analogue of r given in Subsection 3.4 for binomial U-statistics.

Theorem 3.8. *Let S, $\lambda > 0$ and r be as above. Then,*

$$d_{TV}(S,P_\lambda) \le \left(1 \wedge \frac{1}{\lambda}\right)\frac{2^\ell r}{\ell!} \quad \textit{and} \quad d_W(S,P_\lambda) \le \left(1 \wedge \frac{1.1437}{\sqrt{\lambda}}\right)\frac{2^\ell r}{\ell!}. \tag{3.23}$$

Moreover, for all $m \in \mathbb{N}$,

$$0 \le \mathbb{P}(S=0) - e^{-\lambda} \le \left[\sum_{k=0}^{m-1} \left(\frac{1}{k+1} \wedge \frac{k!}{\lambda^{k+1}} \right) \mathbb{P}\big(S \le k\big) + \frac{m!}{\lambda^{m+1}} \right] \frac{2^\ell r}{\ell!} \tag{3.24}$$

and for all $v \in \mathbb{N}$,

$$|\mathbb{P}(S \le v) - \mathbb{P}(P_\lambda \le v)| \le \left[\frac{(v+1)^2}{\lambda} + \mathbb{P}(S \le v) \right] \frac{2^\ell r}{\ell! \lambda}. \tag{3.25}$$

Similarly as discussed next to Theorem 3.6, the right-hand sides in Theorem 3.8 become small for fixed ℓ and λ if r is small, which implies that the probability that there exists a point of η that belongs to more than one k-tuple (up to permutations) with non-zero h is small.

The result for the total variation distance in (3.23) was shown in [39, Proposition 1], which improved [38, Proposition 4.1], and in [31, Section 8]. The bound for the Wasserstein distance in (3.23) was also derived in [31, Section 8] and has a slightly better order in λ than that in [15, Theorem 7.1]. To the best of our knowledge, the other inequalities presented in Theorem 3.8 have no analogues in the literature.

Proof of Theorem 3.8. We follow a similar approach as in the proof of Theorem 3.6. For $\ell = 1$, Theorem 3.8 is a direct consequence of [22, Theorem 5.1], whence we assume $\ell \ge 2$ from now on.

Let χ be a point process of ℓ random points $X_1', \dots, X_\ell'$ that are independent of η and distributed according to

$$\mathbb{P}((X_1', \dots, X_\ell') \in A) = \frac{1}{\ell! \lambda} \int_{\mathbb{X}^\ell} \mathbf{1}\{(x_1, \dots, x_\ell) \in A\} h(x_1, \dots, x_\ell) dL^\ell(x_1, \dots, x_\ell)$$

for $A \in \mathcal{X}^\ell$. We define

$$S' = -h(X_1', \dots, X_\ell') + \frac{1}{\ell!} \sum_{(x_1, \dots, x_\ell) \in (\eta \cup \chi)^k_{\neq}} h(x_1, \dots, x_\ell).$$

For $k \in \mathbb{N}$ the multivariate Mecke formula implies that

$$\begin{aligned}
&k\mathbb{P}(S=k) = \mathbb{E}[S\mathbf{1}\{S=k\}] \\
&= \frac{1}{\ell!} \mathbb{E} \sum_{(x_1, \dots, x_\ell) \in \eta^\ell_{\neq}} h(x_1, \dots, x_\ell) \mathbf{1}\left\{ \frac{1}{\ell!} \sum_{(y_1, \dots, y_\ell) \in \eta^\ell_{\neq}} h(y_1, \dots, y_\ell) = k \right\} \\
&= \frac{1}{\ell!} \int_{\mathbb{X}^\ell} h(x_1, \dots, x_\ell) \mathbb{P}\left(\frac{1}{\ell!} \sum_{(y_1, \dots, y_\ell) \in (\eta \cup \{x_1, \dots, x_\ell\})^\ell_{\neq}} h(y_1, \dots, y_\ell) = k \right) dL^\ell(x_1, \dots, x_\ell) \\
&= \lambda \mathbb{P}(S' + h(X_1', \dots, X_\ell') = k) = \lambda \mathbb{P}(S' = k-1),
\end{aligned}$$

where we used $h(X_1', \dots, X_\ell') = 1$ a.s. in the last step. Thus, we see that S satisfies the hypothesis of Theorem 1.1 with $Z = S' - S \ge 0$.

Next we compute the expressions on the right-hand sides of the bounds in Theorem 1.1. Let $k \in \mathbb{N}$ be fixed. Define $s(\nu) = \frac{1}{\ell!} \sum_{(x_1, \dots, x_\ell) \in \nu^\ell_{\neq}} h(x_1, \dots, x_\ell)$ for $\nu \in \mathbf{N}_\mathbb{X}$ and note that $S = s(\eta)$. Since

$$s(\nu + \chi + \delta_x) - s(\nu + \delta_x) \ge s(\nu + \chi) - s(\nu) \quad \text{and} \quad \mathbf{1}\{s(\nu + \delta_x) \le k\} \le \mathbf{1}\{s(\nu) \le k\}$$

for all $\nu \in \mathbf{N}_\mathbb{X}$ and $x \in \mathbb{X}$, by [22, Theorem 20.4] we obtain

$$\mathbb{E}\big[Z\mathbf{1}\{S \le k\}\big] \le \mathbb{E}[Z]\mathbb{P}(S \le k).$$

Together with $Z \geq 0$, we have

$$\mathbb{E}[|Z|\mathbf{1}\{S - Z_- = k\}] = \mathbb{E}[Z\mathbf{1}\{S = k\}] \leq \mathbb{E}[Z\mathbf{1}\{S \leq k\}] \leq \mathbb{E}[Z]\mathbb{P}(S \leq k). \tag{3.26}$$

Furthermore, from Remark 1.2 (ii) it follows that

$$\mathbb{E}[Z] = \frac{1}{\lambda}\big\{\mathrm{Var}(S) - \lambda\big\} = \frac{1}{\lambda}\big\{\mathbb{E}[S^2] - \lambda^2 - \lambda\big\}. \tag{3.27}$$

Then, from $Z \geq 0$ and [15, Lemma 6.1] we deduce

$$\mathbb{E}[|Z|] = \mathbb{E}[Z] \leq \sum_{i=1}^{\ell-1} \frac{1}{\ell!\lambda}\binom{\ell}{i} r \leq \frac{2^\ell}{\ell!\lambda} r. \tag{3.28}$$

Finally, combining this bound with (1.3) shows (3.23), while (1.4) and (1.5) together with (3.26) and (3.28) lead to (3.24), where the first inequality is a consequence of Corollary 2.7 a), and (3.25). □

3.6 The distances between the points of a Poisson process

We consider random points in $\mathbb{R}^d$ distributed according to a Poisson process. For any pair of these points with the midpoint in a bounded measurable set $W \subset \mathbb{R}^d$, we take a transform of the Euclidean distance, and we study the Poisson approximation for the number of times that these quantities belong to a certain range of values. More importantly, we consider the exponential approximation for a transform of the minimal distance between pairs of points with midpoint in W.

Let η_t be a Poisson process on $\mathbb{R}^d$ with intensity measure $t\lambda_d$, $t > 0$, where we denote by λ_d the d-dimensional Lebesgue measure. For convenience, we assume $\lambda_d(W) = 1$; nonetheless, the following arguments are valid for every W with a positive and finite volume. Define

$$\xi_t = \frac{1}{2} \sum_{(x,y)\in\eta_{t,\neq}^2} \mathbf{1}\Big\{\frac{x+y}{2} \in W\Big\}\delta_{2^{-1}t^2k_d\|x-y\|^d}, \qquad t > 0,$$

$$Y_t = \min_{(x,y)\in\eta_{t,\neq}^2 : \frac{x+y}{2}\in W} 2^{-1}t^2k_d\|x-y\|^d, \quad t > 0,$$

where δ_z stands for the Dirac measure at $z \in \mathbb{R}$, $\|\cdot\|$ is the Euclidean norm and k_d denotes the volume of the d-dimensional unit ball.

Theorem 3.9. *Let ξ_t and Y_t be as above for $t > 0$. Let γ be a Poisson process on $[0,\infty)$ with the restriction of the Lebesgue measure to $[0,\infty)$ as intensity measure. Then for all $u \geq 0$ and all measurable $B \subset [0,u]$,*

$$d_{TV}(\xi_t(B), \gamma(B)) \leq (1 \wedge u)\frac{8u}{t} \tag{3.29}$$

and

$$0 \leq \mathbb{P}(Y_t > u) - e^{-u} \leq \frac{80}{t}. \tag{3.30}$$

The minimal distance between the points of a Poisson process was also considered in [11, 15, 38, 39], sometimes formulated as minimal edge length of the random geometric graph or the minimal inradius of a Poisson-Voronoi tessellation. The important achievement of Theorem 3.9 is that a rate of convergence for the Kolmogorov distance is provided in (3.30). So far it was only possible to prove bounds on the difference between $\mathbb{P}(Y_t > u)$ and e^{-u} that depend on $u > 0$ (see e.g. [38, Theorem 2.4] or [39, Corollary 3]).

In the works mentioned above all pairs of points are considered such that one or both points belong to W. Our approach, where we only require that the midpoint of the points is in W, can be extended to this different way of counting, but one might get additional terms in the bounds since $\mathbb{E}[\xi_t([0,u])]$ is not necessarily u due to boundary effects.

In [38, 39], beside Poisson approximation results for the number of inter-point distances below a given threshold it was shown that the point process of rescaled inter-point distances converges weakly to a Poisson process. By (3.29) and [18, Theorem 16.16], we can also deduce that ξ_t converges weakly to γ as $t \to \infty$.

The related problem of small distances between the points of a binomial point process was first studied in [40]. Because of the similarity to Theorem 3.8, we believe that by applying Theorem 3.6 it is possible to prove a similar result to Theorem 3.9 for an underlying binomial point process.

By using in the proof of Theorem 3.9 the corresponding bound of Theorem 3.8 for the Wasserstein distance, one can obtain the counterpart of (3.29) for the Wasserstein distance with a different power in u and the same rate of convergence in t.

Proof of Theorem 3.9. First, we show that the intensity measure of the point process ξ_t is the restriction of the Lebesgue measure to $[0,\infty)$. Let $v_t = \left(\frac{2u}{k_d t^2}\right)^{1/d}$. The change of variable $z = \frac{x+y}{2}$ yields

$$\begin{aligned}\mathbb{E}[\xi_t([0,u])] &= \frac{t^2}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathbf{1}\Big\{\frac{x+y}{2} \in W\Big\}\mathbf{1}\{\|x-y\| \le v_t\}\,dydx \\ &= 2^{d-1}t^2 \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathbf{1}\big\{z \in W\big\}\mathbf{1}\{2\|x-z\| \le v_t\}\,dzdx \\ &= 2^{d-1}t^2 \int_{W}\int_{\mathbb{R}^d} \mathbf{1}\{2\|x-z\| \le v_t\}\,dxdz = u.\end{aligned}$$

For measurable $B \subset [0,u]$ with $u > 0$ define

$$r_t(B) = t\int_{\mathbb{R}^d}\left(t\int_{\mathbb{R}^d} \mathbf{1}\Big\{\frac{x+y}{2} \in W\Big\}\mathbf{1}\{2^{-1}t^2k_d\|x-y\|^d \in B\}\,dy\right)^2 dx.$$

Again from the change of variable $z = \frac{x+y}{2}$, it follows that

$$\begin{aligned}r_t(B) \le r_t([0,u]) &= 2^{2d}t^3\int_{\mathbb{R}^d}\left(\int_W \mathbf{1}\{2\|x-z\| \le v_t\}\,dz\right)^2 dx \\ &\le 2^{2d}t^3\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} \mathbf{1}\{2\|x-z\| \le v_t\}\,dz\int_W \mathbf{1}\{2\|x-\tilde{z}\| \le v_t\}\,d\tilde{z}\right)dx \\ &= 2^{d+1}ut\int_{\mathbb{R}^d}\int_W \mathbf{1}\{2\|x-\tilde{z}\| \le v_t\}\,d\tilde{z}dx \\ &= 2^{d+1}ut\int_{W}\int_{\mathbb{R}^d} \mathbf{1}\{2\|x-\tilde{z}\| \le v_t\}\,dxd\tilde{z} = \frac{4u^2}{t}.\end{aligned}$$

Therefore (3.23) in Theorem 3.8 with $h(x,y) = \mathbf{1}\big\{\frac{x+y}{2} \in W\big\}\mathbf{1}\{2^{-1}t^2k_d\|x-y\|^d \in B\}$ yields for measurable $B \subset [0,u]$ that

$$d_{TV}(\xi_t(B),\gamma(B)) \le \left(1 \wedge \frac{1}{u}\right)2r_t(B) \le (1\wedge u)\frac{8u}{t}.$$

From (3.27) and (3.28) in the proof of Theorem 3.8 with $S = \xi_t([0,u])$, $r = r_t([0,u])$ and h as above, we know that

$$\mathrm{Var}(\xi_t([0,u])) \le \mathbb{E}[\xi_t([0,u])] + 2r_t([0,u]) = u + \frac{8u^2}{t}.$$

Thus it follows from the Chebyshev inequality that

$$\mathbb{P}(\xi_t([0,u])=0) \le \mathbb{P}(|\xi_t([0,u])-u| \ge u) \le \frac{\mathrm{Var}(\xi_t([0,u]))}{u^2} = \frac{1}{u}+\frac{8}{t}. \tag{3.31}$$

Together with (3.24) in Theorem 3.8 with $m=1$ and straightforward arguments, this leads to

$$\begin{aligned} 0 \le \mathbb{P}(\xi_t([0,u])=0) - e^{-u} = \mathbb{P}(Y_t>u) - e^{-u} &\le \left[\frac{1}{u}\mathbb{P}(\xi_t([0,u])=0)+\frac{1}{u^2}\right]\frac{8u^2}{t} \\ &\le \left(\frac{1}{u^2}+\frac{8}{ut}+\frac{1}{u^2}\right)\frac{8u^2}{t} = \frac{16}{t}+\frac{64u}{t^2} \end{aligned} \tag{3.32}$$

so that

$$\sup_{u\in[0,t]} |\mathbb{P}(Y_t>u)-e^{-u}| \le \frac{80}{t}.$$

Moreover, from (3.31) we obtain

$$\mathbb{P}(Y_t>t) = \mathbb{P}(\xi_t([0,t])=0) \le \frac{9}{t}.$$

Therefore we have

$$\sup_{u\in[0,\infty)} |\mathbb{P}(Y_t>u)-e^{-u}| \le \max\Big\{ \sup_{u\in[0,t]} |\mathbb{P}(Y_t>u)-e^{-u}|, \mathbb{P}(Y_t>t), e^{-t}\Big\} \le \frac{80}{t},$$

which combined with the left-hand side of (3.32) completes the proof. □

4 Proofs for Poisson-Voronoi tessellations

4.1 Proof of Theorem 3.4

The proof of Theorem 3.4 requires several preparations. We set

$$s_t = \alpha_2 k_d t^{(d+2)/(d+1)}.$$

Let M_t denote the intensity measure of ξ_t, and define

$$\begin{aligned} \widehat{M}_t([0,u]) &= t\int_W \mathbb{E}\big[\mathbf{1}\{s_t C(x,\eta_t+\delta_x)^d \le u\}\mathbf{1}\{\eta_t(\mathbf{B}(x,4(u/s_t)^{1/d}))=d+1\}\big]dx, \\ \theta_t([0,u]) &= t\int_W \mathbb{E}\big[\mathbf{1}\{s_t C(x,\eta_t+\delta_x)^d \le u\}\mathbf{1}\{\eta_t(\mathbf{B}(x,4(u/s_t)^{1/d}))>d+1\}\big]dx \end{aligned}$$

for $u\ge 0$. For $x\in W$ and $u\ge 0$ we have

$$\eta_t(\mathbf{B}(x,2(u/s_t)^{1/d}))\ge d+1 \quad \text{whenever} \quad s_t C(x,\eta_t+\delta_x)^d \le u. \tag{4.1}$$

This is the case since $s_t C(x,\eta_t+\delta_x)^d \le u$ implies that the nuclei of the neighboring cells of x are in $\mathbf{B}(x,2(u/s_t)^{1/d})$ and each Voronoi cell has at least $d+1$ neighboring cells. From the Mecke formula and (4.1) it follows that

$$M_t([0,u]) = \widehat{M}_t([0,u]) + \theta_t([0,u]), \quad u\ge 0.$$

Lemma 4.1. *For all $u>0$ and $t>0$,*

$$\widehat{M}_t([0,u]) = u^{d+1}\exp\Big(-\frac{4^d u}{\alpha_2 t^{1/(d+1)}}\Big),\ \theta_t([0,u]) \le \frac{2^{d(d+3)}}{\alpha_2 p_{d+1}}\frac{u^{d+2}}{t^{1/(d+1)}} \text{ and } M_t([0,u]) \le \frac{u^{d+1}}{p_{d+1}}.$$

Proof. First we compute $\widehat{M}_t([0,u])$. From (4.1) and the definition of p_{d+1} in (3.8) we derive

$$\widehat{M}_t([0,u]) = t\int_W e^{-2^d k_d tu/s_t}\frac{\big(2^d k_d tu/s_t\big)^{d+1}}{(d+1)!}p_{d+1} \times \mathbb{P}\big(\eta_t\big(\mathbf{B}\big(x,4(u/s_t)^{1/d}\big)\setminus\mathbf{B}\big(x,2(u/s_t)^{1/d}\big)\big)=0\big)dx.$$

Substituting $s_t = \alpha_2 k_d t^{(d+2)/(d+1)}$ and $\alpha_2 = \big(\frac{2^{d(d+1)}}{(d+1)!}p_{d+1}\big)^{1/(d+1)}$ into the previous equation implies that the right-hand side equals

$$u^{d+1}\int_W \exp\Big(-\frac{2^d u}{\alpha_2 t^{1/(d+1)}} - t\lambda_d\big(\mathbf{B}\big(x,4(u/s_t)^{1/d}\big)\setminus\mathbf{B}\big(x,2(u/s_t)^{1/d}\big)\big)\Big)dx$$
$$= u^{d+1}\exp\Big(-\frac{2^d u}{\alpha_2 t^{1/(d+1)}} - \frac{2^d u}{\alpha_2 t^{1/(d+1)}}(2^d-1)\Big) = u^{d+1}\exp\Big(-\frac{4^d u}{\alpha_2 t^{1/(d+1)}}\Big),$$

which completes the first part of the proof.

For $u>0$, we have

$$\theta_t([0,u]) \le t\int_W \mathbb{E}\big[\mathbf{1}\big\{\eta_t\big(\mathbf{B}\big(x,4(u/s_t)^{1/d}\big)\big)>d+1\big\}\big]dx = t\sum_{k=d+2}^{\infty} e^{-\beta_t}\frac{\beta_t^k}{k!}$$

with $\beta_t = 4^d k_d tu/s_t$. Elementary calculations imply that

$$t\sum_{k=d+2}^{\infty} e^{-\beta_t}\frac{\beta_t^k}{k!} = t\beta_t^{d+2}\sum_{k=d+2}^{\infty} e^{-\beta_t}\frac{\beta_t^{k-d-2}}{k!} = t\beta_t^{d+2}\sum_{\ell=0}^{\infty} e^{-\beta_t}\frac{\beta_t^{\ell}}{(\ell+d+2)!}$$
$$\le \frac{t\beta_t^{d+2}}{(d+2)!} = \frac{t\big(4^d k_d tu/s_t\big)^{d+2}}{(d+2)!}.$$

Substituting $s_t = \alpha_2 k_d t^{(d+2)/(d+1)}$ and $\alpha_2 = \big(\frac{2^{d(d+1)}}{(d+1)!}p_{d+1}\big)^{1/(d+1)}$ into the latter term yields

$$\theta_t([0,u]) \le \frac{2^{d(d+3)}}{\alpha_2 p_{d+1}}\frac{u^{d+2}}{t^{1/(d+1)}},$$

which is the desired result.

From the Mecke formula, (4.1) and the same arguments as above, we obtain

$$M_t([0,u]) \le t\int_W \mathbb{P}(\eta_t(\mathbf{B}(x,2(u/s_t)^{1/d}))\ge d+1)dx = t\sum_{k=d+1}^{\infty}\frac{(2^d k_d tu/s_t)^k}{k!}e^{-2^d k_d tu/s_t}$$
$$\le \frac{t(2^d k_d tu/s_t)^{d+1}}{(d+1)!} = \frac{2^{d(d+1)}k_d^{d+1}t^{d+2}u^{d+1}}{k_d^{d+1}\frac{2^{d(d+1)}p_{d+1}}{(d+1)!}(d+1)!t^{d+2}} = \frac{u^{d+1}}{p_{d+1}},$$

which concludes the proof. □

We now provide a statement from [32, Lemma 3.14], which will be employed in the proof of the subsequent proposition.

Lemma 4.2. *Let $x_0,\dots,x_{d+1}\in\mathbb{R}^d$ be in general position (i.e., no k-dimensional affine subspace of $\mathbb{R}^d$ with $k\in\{0,\dots,d-1\}$ contains more than $k+1$ of the points) and assume that $N(x_0,\sum_{j=0}^{d+1}\delta_{x_i})$ is bounded. Then $N(x_i,\sum_{j=0}^{d+1}\delta_{x_i})$ is unbounded for any $i\in\{1,\dots,d+1\}$.*

Next we construct a random variable that satisfies (1.8) for $\xi_t([0,u])$ with remainder terms $q_i, i\in\mathbb{N}_0$, which vanish as $t\to\infty$. By B^c we denote the complement of $B\subset\mathbb{R}^d$ and by $\eta_t|_B$ the restriction of η_t to B.

Proposition 4.3. *Let X be uniformly distributed in W and independent of η_t. Then for $u > 0$,*

$$k\mathbb{P}(\xi_t([0,u]) = k) = \widehat{M}_t([0,u])\mathbb{P}(\xi_t([0,u]) + Z_{t,u} = k-1) + q_{k-1}(t,u), \quad k \in \mathbb{N},$$

with

$$Z_{t,u} = \xi_t\big(\eta_t|_{\mathbf{B}(X,4(u/s_t)^{1/d})^c}\big)([0,u]) - \xi_t([0,u])$$

and

$$q_i(t,u) = t\int_W \mathbb{E}\Big[\mathbf{1}\big\{s_tC(x,\eta_t+\delta_x)^d \le u\big\}\mathbf{1}\big\{\eta_t\big(\mathbf{B}\big(x,4(u/s_t)^{1/d}\big)\big) > d+1\big\} \times \mathbf{1}\Big\{\sum_{y\in\eta_t\cap W}\mathbf{1}\big\{s_tC(y,\eta_t+\delta_x)^d \le u\big\} = i\Big\}\Big]dx$$

for $i \in \mathbb{N}_0$.

Proof. The Mecke equation implies for $k \in \mathbb{N}$ that

$$k\mathbb{P}(\xi_t([0,u]) = k) = t\int_W \mathbb{E}\big[\mathbf{1}\big\{s_tC(x,\eta_t+\delta_x)^d \le u\big\}\mathbf{1}\big\{\xi_t(\eta_t+\delta_x)([0,u]) = k\big\}\big]dx$$
$$= t\int_W \mathbb{E}\Big[\mathbf{1}\big\{s_tC(x,\eta_t+\delta_x)^d \le u\big\}\mathbf{1}\Big\{\sum_{y\in\eta_t\cap W}\mathbf{1}\big\{s_tC(y,\eta_t+\delta_x)^d \le u\big\} = k-1\Big\}\Big]dx.$$

Now we divide the integral in

$$A_k + q_{k-1}(t,u) := t\int_W \mathbb{E}\Big[\mathbf{1}\big\{s_tC(x,\eta_t+\delta_x)^d \le u\big\}\mathbf{1}\big\{\eta_t\big(\mathbf{B}\big(x,4(u/s_t)^{1/d}\big)\big) = d+1\big\} \times \mathbf{1}\Big\{\sum_{y\in\eta_t\cap W}\mathbf{1}\big\{s_tC(y,\eta_t+\delta_x)^d \le u\big\} = k-1\Big\}\Big]dx$$
$$+ t\int_W \mathbb{E}\Big[\mathbf{1}\big\{s_tC(x,\eta_t+\delta_x)^d \le u\big\}\mathbf{1}\big\{\eta_t\big(\mathbf{B}\big(x,4(u/s_t)^{1/d}\big)\big) > d+1\big\} \times \mathbf{1}\Big\{\sum_{y\in\eta_t\cap W}\mathbf{1}\big\{s_tC(y,\eta_t+\delta_x)^d \le u\big\} = k-1\Big\}\Big]dx.$$

Then, it is enough to show that $A_k = \widehat{M}_t([0,u])\mathbb{P}(\xi_t([0,u]) + Z_{t,u} = k-1)$. In order to simplify the notation throughout this proof, we write

$$\mathbf{B}_2(x) := \mathbf{B}\big(x,2(u/s_t)^{1/d}\big) \quad \text{and} \quad \mathbf{B}_4(x) := \mathbf{B}\big(x,4(u/s_t)^{1/d}\big), \quad x \in \mathbb{R}^d.$$

In case there are only $d+1$ points of η_t in $\mathbf{B}_4(x)$, we have by (4.1) that $s_tC(x,\eta_t+\delta_x)^d \le u$ only if the $d+1$ elements of η_t belong to $\mathbf{B}_2(x)$. Therefore we obtain

$$A_k = t\int_W \mathbb{E}\Big[\mathbf{1}\big\{s_tC(x,\eta_t+\delta_x)^d \le u\big\}\mathbf{1}\big\{\eta_t(\mathbf{B}_4(x)\setminus\mathbf{B}_2(x)) = 0, \eta_t(\mathbf{B}_2(x)) = d+1\big\} \times \mathbf{1}\Big\{\sum_{y\in\eta_t\cap W}\mathbf{1}\{s_tC(y,\eta_t+\delta_x)^d \le u\} = k-1\Big\}\Big]dx. \tag{4.2}$$

The observation that

$$s_tC(y,\eta_t+\delta_x)^d \le u \quad \text{if and only if} \quad s_tC\big(y,(\eta_t+\delta_x)|_{\mathbf{B}_2(y)}\big)^d \le u \tag{4.3}$$

for $y \in \eta_t$ establishes that

$$A_k = t\int_W \mathbb{E}\Big[\mathbf{1}\big\{s_tC(x,\eta_t+\delta_x)^d \le u\big\}\mathbf{1}\big\{\eta_t(\mathbf{B}_4(x)\setminus\mathbf{B}_2(x)) = 0, \eta_t(\mathbf{B}_2(x)) = d+1\big\} \times \mathbf{1}\Big\{\xi_t(\eta_t|_{\mathbf{B}_4(x)^c})([0,u]) + \sum_{y\in\eta_t\cap\mathbf{B}_2(x)\cap W}\mathbf{1}\big\{s_tC(y,\eta_t+\delta_x)^d \le u\big\} = k-1\Big\}\Big]dx.$$

Suppose that $s_t C(x, \eta_t + \delta_x)^d \le u$ and that there are exactly $d+1$ points $y_1, \dots, y_{d+1}$ of η_t in $\mathbf{B}_2(x)$ and $\eta_t \cap \mathbf{B}_4(x) \cap \mathbf{B}_2(x)^c = \emptyset$. From Lemma 4.2 it follows that the Voronoi cells $N(y_i, \eta_t|_{\mathbf{B}_4(x)} + \delta_x), i = 1, \dots, d+1$, are unbounded. In particular, we have

$$C(y_i, \eta_t + \delta_x) > (u/s_t)^{1/d}, \quad i = 1, \dots, d+1.$$

Together with the same arguments used to show (4.2) and independence, this implies that

$$\begin{aligned} A_k &= t \int_W \mathbb{E}\big[\mathbf{1}\{s_t C(x, \eta_t + \delta_x)^d \le u\}\mathbf{1}\{\eta_t(\mathbf{B}_4(x) \setminus \mathbf{B}_2(x)) = 0, \eta_t(\mathbf{B}_2(x)) = d+1\} \\ &\qquad \times \mathbf{1}\{\xi_t(\eta_t|_{\mathbf{B}_4(x)^c})([0,u]) = k-1\}\big] dx \\ &= t \int_W \mathbb{E}\big[\mathbf{1}\{s_t C(x, \eta_t + \delta_x)^d \le u\}\mathbf{1}\{\eta_t(\mathbf{B}_4(x)) = d+1\}\big] \\ &\qquad \times \mathbb{P}\big(\xi_t(\eta_t|_{\mathbf{B}_4(x)^c})([0,u]) = k-1\big) dx. \end{aligned}$$

Then, because the expectation in the latter equation does not depend on the choice of $x \in W$, we have that

$$\begin{aligned} A_k &= \widehat{M}_t([0,u]) \int_W \mathbb{P}\big(\xi_t\big(\eta_t|_{\mathbf{B}_4(x)^c}\big)([0,u]) = k-1\big) dx \\ &= \widehat{M}_t([0,u]) \mathbb{P}(\xi_t([0,u]) + Z_{t,u} = k-1) \end{aligned}$$

with

$$Z_{t,u} = \xi_t\big(\eta_t|_{\mathbf{B}_4(X)^c}\big)([0,u]) - \xi_t([0,u]).$$

This and $\mathbf{B}_4(X) = \mathbf{B}\big(X, 4(u/s_t)^{1/d}\big)$ give the desired conclusion. □

Lemma 4.4. *For $u > 0$, $t > 0$ and $Z_{t,u}$ as in Proposition 4.3,*

$$\mathbb{E}[|Z_{t,u}|] \le \frac{6^d}{\alpha_2 p_{d+1}} \frac{u^{d+2}}{t^{(d+2)/(d+1)}}.$$

Proof. For $x \in W$ it follows from the observation in (4.3) that

$$0 \le \xi_t([0,u]) - \xi_t\big(\eta_t|_{\mathbf{B}(x, 4(u/s_t)^{1/d})^c}\big)([0,u]) \le \sum_{y \in \eta_t \cap W \cap \mathbf{B}(x, 6(u/s_t)^{1/d})} \mathbf{1}\{s_t C(y, \eta_t)^d \le u\}.$$

By the Mecke formula and the stationarity of η_t, we obtain

$$\begin{aligned} \mathbb{E} \sum_{y \in \eta_t \cap W \cap \mathbf{B}(x, 6(u/s_t)^{1/d})} \mathbf{1}\{s_t C(y, \eta_t)^d \le u\} &\le t \lambda_d(W \cap \mathbf{B}(x, 6(u/s_t)^{1/d})) \mathbb{P}(s_t C(0, \eta_t + \delta_0)^d \le u) \\ &\le \frac{6^d u}{\alpha_2 t^{(d+2)/(d+1)}} t \mathbb{P}(s_t C(0, \eta_t + \delta_0)^d \le u). \end{aligned}$$

From Lemma 4.1 we deduce

$$t \mathbb{P}(s_t C(0, \eta_t + \delta_0)^d \le u) = M_t([0,u]) \le \frac{u^{d+1}}{p_{d+1}},$$

which proves the assertion. □

Lemma 4.5. *For $u > 0$ and $t > 0$,*

$$\mathbb{P}(T_t > u) = \mathbb{P}(\xi_t([0,u]) = 0) \le e^{-\widehat{M}_t([0,u])}.$$

Proof. The first identity is obvious. Let $Z_{t,u}$ be the random variable defined in Proposition 4.3. Since $Z_{t,u} \le 0$, $\mathbb{P}(\xi_t([0,u]) + Z_{t,u} \ge 0) = 1$ and $q_i(t,u) \ge 0$ for all $i \in \mathbb{N}_0$, the inequality follows from Proposition 2.6 b). □

In the next lemma, we combine the results obtained above and Theorem 1.3 to derive intermediate bounds on the quantities considered in Theorem 3.4.

Lemma 4.6. *For $u > 0$ and $t > 0$,*

$$d_{TV}\big(\xi_t([0,u]), P_{\widehat{M}_t([0,u])}\big) \le \frac{6^d}{\alpha_2 p_{d+1}} \frac{u^{d+2}}{t^{(d+2)/(d+1)}} + \theta_t([0,u]) \tag{4.4}$$

and

$$0 \le e^{-\widehat{M}_t([0,u])} - \mathbb{P}(T_t > u) \le \left(1 + \frac{1}{\widehat{M}_t([0,u])}\right) \frac{6^d}{\alpha_2 p_{d+1}} \frac{u^{d+2}}{t^{(d+2)/(d+1)}} + \frac{2\theta_t([0,u])}{\widehat{M}_t([0,u])^2}. \tag{4.5}$$

Proof. From Proposition 4.3 it follows that the assumptions of Theorem 1.3 are satisfied. Then, (1.9) in Theorem 1.3 and $\sum_{i=0}^{\infty} |q_i(t,u)| = \theta_t([0,u])$ yield

$$d_{TV}(\xi_t([0,u]), P_{\widehat{M}_t([0,u])}) \le (1 \wedge \widehat{M}_t([0,u]))\mathbb{E}[|Z_{t,u}|] + \big(1 \wedge \widehat{M}_t([0,u])^{-1/2}\big)\theta_t([0,u])$$

so that (4.4) follows from Lemma 4.4.

Let us now prove (4.5). From Lemma 4.5, (1.12) in Theorem 1.3 with $m = 1$ and $\sum_{i=1}^{\infty} q_i(t,u) \le \theta_t([0,u])$ we obtain

$$0 \le e^{-\widehat{M}_t([0,u])} - \mathbb{P}(T_t > u) \le \frac{\mathbb{E}\big[|Z_{t,u}|\big]}{\widehat{M}_t([0,u])} + \mathbb{E}\big[|Z_{t,u}|\big] + \frac{q_0(t,u)}{\widehat{M}_t([0,u])} + \frac{\theta_t([0,u])}{\widehat{M}_t([0,u])^2}.$$

The first two terms on the right-hand side can be bounded by Lemma 4.4. Recall that

$$q_0(t,u) = t \int_W \mathbb{E}\Big[\mathbf{1}\big\{s_t C(x, \eta_t + \delta_x)^d \le u\big\}\mathbf{1}\big\{\eta_t\big(\mathbf{B}\big(x, 4(u/s_t)^{1/d}\big)\big) > d+1\big\} \times \mathbf{1}\Big\{\sum_{y \in \eta_t \cap W} \mathbf{1}\{s_t C(y, \eta_t + \delta_x)^d \le u\} = 0\Big\}\Big] dx.$$

Since the product of the first two indicator functions is increasing with respect to additional points, while the third indicator function is decreasing, it follows from [22, Theorem 20.4] that

$$q_0(t,u) \le t \int_W \mathbb{E}\Big[\mathbf{1}\big\{s_t C(x, \eta_t + \delta_x)^d \le u\big\}\mathbf{1}\big\{\eta_t\big(\mathbf{B}\big(x, 4(u/s_t)^{1/d}\big)\big) > d+1\big\}\Big] \times \mathbb{P}\Big(\sum_{y \in \eta_t \cap W} \mathbf{1}\{s_t C(y, \eta_t + \delta_x)^d \le u\} = 0\Big) dx.$$

Now Lemma 4.5 and the elementary inequality $ve^{-v} \le 1$ for $v \ge 0$ lead to

$$q_0(t,u) \le \theta_t([0,u])\mathbb{P}(\xi_t([0,u]) = 0) \le \theta_t([0,u])e^{-\widehat{M}_t([0,u])} \le \frac{\theta_t([0,u])}{\widehat{M}_t([0,u])},$$

which concludes the proof. □

Proof of Theorem 3.4. Let $u > 0$ be fixed. From (4.4) in Lemma 4.6, Lemma 4.1 and $t \ge 1$ it follows that

$$d_{TV}\big(\xi_t([0,u]), P_{\widehat{M}_t([0,u])}\big) \le \frac{6^d}{\alpha_2 p_{d+1}} \frac{u^{d+2}}{t^{(d+2)/(d+1)}} + \theta_t([0,u]) \le \frac{6^d + 2^{d(d+3)}}{\alpha_2 p_{d+1}} \frac{u^{d+2}}{t^{1/(d+1)}}. \tag{4.6}$$

Using a well-known bound for the total variation distance between two Poisson distributed random variables, Lemma 4.1 and the inequality $1-e^{-v}\leq v$ for $v\geq 0$, we obtain

$$d_{TV}\big(P_{u^{d+1}},P_{\widehat{M}_t([0,u])}\big)\leq u^{d+1}-\widehat{M}_t([0,u])=u^{d+1}\bigg(1-\exp\bigg(-\frac{4^d u}{\alpha_2 t^{1/(d+1)}}\bigg)\bigg)\leq\frac{4^d u^{d+2}}{\alpha_2 t^{1/(d+1)}}.$$

Now the triangle inequality yields

$$d_{TV}\big(\xi_t([0,u]),P_{u^{d+1}}\big)\leq\frac{3\cdot 2^{d(d+3)}}{\alpha_2 p_{d+1}}\frac{u^{d+2}}{t^{1/(d+1)}},$$

which proves (3.10).

Let us now show (3.11). From (4.6) and Lemma 4.5 we have that, for $u\in[0,1]$,

$$0\leq e^{-\widehat{M}_t([0,u])}-\mathbb{P}(T_t>u)\leq\frac{2^{d(d+3)+1}}{\alpha_2 p_{d+1}}\frac{1}{t^{1/(d+1)}}.$$

In the following we consider the case $1\leq u\leq t^{1/(2d+2)}\tau$ with $\tau=\alpha_2/4^d$. From Lemma 4.1 and $t\geq 1$ we obtain

$$u^{d+1}\geq\widehat{M}_t([0,u])\geq\frac{u^{d+1}}{e}. \tag{4.7}$$

Together with Lemma 4.5, (4.5) in Lemma 4.6, Lemma 4.1 and $u\geq 1$ we obtain

$$\begin{aligned}0\leq e^{-\widehat{M}_t([0,u])}-\mathbb{P}(T_t>u)&\leq\bigg(1+\frac{1}{\widehat{M}_t([0,u])}\bigg)\frac{6^d}{\alpha_2 p_{d+1}}\frac{u^{d+2}}{t^{(d+2)/(d+1)}}+\frac{2\theta_t([0,u])}{\widehat{M}_t([0,u])^2}\\&\leq(1+e)\frac{6^d}{\alpha_2 p_{d+1}}\frac{u^{d+2}}{t^{(d+2)/(d+1)}}+2e^2\frac{1}{u^{2d+2}}\frac{2^{d(d+3)}}{\alpha_2 p_{d+1}}\frac{u^{d+2}}{t^{1/(d+1)}}.\end{aligned}$$

Using $1\leq u^{d+2}\leq t^{(d+2)/(2d+2)}\alpha_2^{d+2}/4^{d(d+2)}$, $t\geq 1$ and the definition of α_2 in (3.7), we deduce

$$\begin{aligned}0\leq e^{-\widehat{M}_t([0,u])}-\mathbb{P}(T_t>u)&\leq\frac{(1+e)6^d}{4^{d(d+2)}}\frac{\alpha_2^{d+1}}{p_{d+1}}\frac{1}{t^{1/(d+1)}}+2e^2\frac{1}{u^{2d+2}}\frac{2^{d(d+3)}}{\alpha_2 p_{d+1}}\frac{u^{d+2}}{t^{1/(d+1)}}\\&\leq\frac{1}{t^{1/(d+1)}}+\frac{2^{d(d+3)+4}}{\alpha_2 p_{d+1}}\frac{1}{t^{1/(d+1)}}\end{aligned}$$

so that

$$\sup_{u\in[0,t^{1/(2d+2)}\tau]}|e^{-\widehat{M}_t([0,u])}-\mathbb{P}(T_t>u)|\leq\bigg[1+\frac{2^{d(d+3)+4}}{\alpha_2 p_{d+1}}\bigg]\frac{1}{t^{1/(d+1)}}.$$

Moreover, by Lemma 4.1, (4.7) and elementary arguments we obtain for $0\leq u\leq t^{1/(2d+2)}\tau$ that

$$\begin{aligned}0\leq e^{-\widehat{M}_t([0,u])}-e^{-u^{d+1}}&\leq\Big[u^{d+1}-\widehat{M}_t([0,u])\Big]e^{-\widehat{M}_t([0,u])}\\&\leq\frac{4^d u^{d+2}}{\alpha_2 t^{1/(d+1)}}e^{-u^{d+1}e^{-1}}\leq\frac{4^d e^{\frac{d+2}{d+1}}}{\alpha_2 t^{1/(d+1)}}\leq\frac{2^{2d+3}}{\alpha_2 t^{1/(d+1)}},\end{aligned}$$

where we used the inequalities $1-e^{-x}\leq x$ and $e^{-x^{d+1}}x^{d+2}\leq 1$ for $x\geq 0$. This implies that

$$\sup_{u\in[0,t^{1/(2d+2)}\tau]}|e^{-u^{d+1}}-\mathbb{P}(T_t>u)|\leq\bigg[1+\frac{2^{d(d+3)+4}}{\alpha_2 p_{d+1}}+\frac{2^{2d+3}}{\alpha_2}\bigg]\frac{1}{t^{1/(d+1)}}.$$

On the other hand, $x^2e^{-x^{d+1}}\leq 1$ for $x\geq 0$ leads to

$$\exp\Big(-(t^{1/(2d+2)}\tau)^{d+1}\Big)\leq\frac{1}{(t^{1/(2d+2)}\tau)^2}\leq\frac{16^d}{\alpha_2^2}\frac{1}{t^{1/(d+1)}}.$$

Combining the two previous inequalities gives a bound for $\mathbb{P}\left(T_t > t^{1/(2d+2)}\tau\right)$ and it implies

$$\begin{aligned}
&\sup_{u\in[0,\infty)} |e^{-u^{d+1}} - \mathbb{P}(T_t > u)| \\
&\leq \max\Big\{ \sup_{u\in[0,t^{1/(2d+2)}\tau]} |e^{-u^{d+1}} - \mathbb{P}(T_t > u)|, \mathbb{P}\left(T_t > t^{1/(2d+2)}\tau\right), \exp\left(-(t^{1/(2d+2)}\tau)^{d+1}\right) \Big\} \\
&\leq \left[1 + \frac{2^{d(d+3)+4}}{\alpha_2 p_{d+1}} + \frac{2^{2d+3}}{\alpha_2} + \frac{16^d}{\alpha_2^2}\right] \frac{1}{t^{1/(d+1)}}.
\end{aligned}$$

Now the identity $\mathbb{P}(T_t > 0) = 1$ concludes the proof. □

4.2 Proof of Theorem 3.5

For the proof of Theorem 3.5 we introduce some notation. By M_t we denote the intensity measure of ξ_t. For $u > -\log(t)$, set

$$v_t = v_t(u) = \left(\frac{u + \log(t)}{t k_d}\right)^{1/d}. \tag{4.8}$$

Then, for $u > -\log(t)$ we have

$$\begin{aligned}
M_t((u,\infty)) &= t\int_W \mathbb{E}\big[\mathbf{1}\{h_t(x, \eta_t + \delta_x) > u\}\big]dx = t\int_W \mathbb{P}\big(\eta_t(\mathbf{B}(x, v_t)) = 0\big)dx \\
&= t\int_W e^{-t v_t^d k_d} dx = t e^{-u - \log(t)} = e^{-u}.
\end{aligned}$$

Let X be a uniformly distributed random vector in W independent of η_t. In the next proposition we show that for each $u > -\log(t)$, and for an opportune choice of a random ball $\mathbf{B}$ centered at X, the random variable $\xi_t(\eta_t|_{\mathbf{B}^c})((u,\infty)) - \xi_t((u,\infty))$ satisfies (1.2) for $\xi_t((u,\infty))$.

Proposition 4.7. *For any $t > e$ and $u > -\log(t)$,*

$$k\mathbb{P}(\xi_t((u,\infty)) = k) = M_t((u,\infty))\mathbb{P}(\xi_t((u,\infty)) + Z_t(u) = k-1), \quad k \in \mathbb{N},$$

where the random variable $Z_t(u)$ is defined as

$$Z_t(u) = \xi_t(\eta_t|_{\mathbf{B}(X, v_t)^c})((u,\infty)) - \xi_t((u,\infty))$$

with $v_t = v_t(u)$ given by (4.8).

Proof. Let $B = (u,\infty)$ with $u > -\log(t)$. The Mecke equation yields for $k \in \mathbb{N}$ that

$$k\mathbb{P}(\xi_t(B) = k) = t\int_W \mathbb{E}\big[\mathbf{1}\{h_t(x, \eta_t + \delta_x) > u\}\mathbf{1}\{\xi_t(\eta_t + \delta_x)(B) = k\}\big]dx.$$

Since $h_t(x, \eta_t + \delta_x) > u$ if and only if $\eta_t(\mathbf{B}(x, v_t)) = 0$, the right-hand side equals

$$\begin{aligned}
&t\int_W \mathbb{E}\big[\mathbf{1}\{\eta_t(\mathbf{B}(x, v_t)) = 0\}\mathbf{1}\{\xi_t(\eta_t|_{\mathbf{B}(x,v_t)^c})(B) = k-1\}\big]dx \\
&= t\int_W \mathbb{P}\big(\eta_t(\mathbf{B}(x, v_t)) = 0\big)\mathbb{E}\big[\mathbf{1}\{\xi_t(\eta_t|_{\mathbf{B}(x,v_t)^c})(B) = k-1\}\big]dx \\
&= e^{-u}\int_W \mathbb{P}\big(\xi_t(\eta_t|_{\mathbf{B}(x,v_t)^c})(B) = k-1\big)dx.
\end{aligned}$$

Hence, elementary arguments lead to

$$\begin{aligned} k\mathbb{P}(\xi_t(B)=k) &= M_t(B)\mathbb{P}\big(\xi_t(\eta_t|_{\mathbf{B}(X,v_t)^c})(B)=k-1\big)\\ &= M_t(B)\mathbb{P}\big(\xi_t(B)+Z_t(u)=k-1\big), \end{aligned}$$

which is the desired conclusion. □

Proof of Theorem 3.5. Suppose $u > -\log(t)$ and let $Z_t(u)$ be as in Proposition 4.7. We can rewrite $Z_t(u)$ as

$$\begin{aligned} Z_t(u) &= \xi_t(\eta_t|_{\mathbf{B}(X,v_t)^c})((u,\infty)) - \xi_t((u,\infty))\\ &= \sum_{z\in\eta_t\cap W\cap\mathbf{B}(X,2v_t)\cap\mathbf{B}(X,v_t)^c} \mathbf{1}\{h_t(z,\eta_t|_{\mathbf{B}(X,v_t)^c})>u\} - \mathbf{1}\{h_t(z,\eta_t)>u\}\\ &\quad - \sum_{z\in\eta_t\cap\mathbf{B}(X,v_t)\cap W} \mathbf{1}\{h_t(z,\eta_t)>u\}\\ &=: Z'_{t,X}(u) - Z''_{t,X}(u), \end{aligned}$$

where $Z'_{t,X}(u)$ and $Z''_{t,X}(u)$ are non-negative. For a fixed $x\in W$, the Mecke formula and short computations yield

$$\begin{aligned} \mathbb{E}[Z'_{t,X}(u)] &\le \mathbb{E}\Big[\sum_{z\in\eta_t\cap\mathbf{B}(x,2v_t)\cap\mathbf{B}(x,v_t)^c} \mathbf{1}\big\{h_t(z,\eta_t|_{\mathbf{B}(x,v_t)^c})>u\big\}\Big]\\ &= t\int_{\mathbf{B}(x,2v_t)\cap\mathbf{B}(x,v_t)^c} \mathbb{P}\big(\eta_t(\mathbf{B}(z,v_t)\cap\mathbf{B}(x,v_t)^c)=0\big)dz\\ &\le t\int_{\mathbf{B}(x,2v_t)\cap\mathbf{B}(x,v_t)^c} e^{-tv_t^dk_d/2}dz \le 2^d(u+\log(t))e^{-(u+\log(t))/2} = 2^d\frac{u+\log(t)}{e^{u/2}\sqrt{t}} \end{aligned} \tag{4.9}$$

and, similarly,

$$\begin{aligned} \mathbb{E}[Z''_{t,X}(u)] &\le \mathbb{E}\Big[\sum_{z\in\eta_t\cap\mathbf{B}(x,v_t)} \mathbf{1}\big\{h_t(z,\eta_t)>u\big\}\Big] = t\int_{\mathbf{B}(x,v_t)} \mathbb{P}\big(\eta_t(\mathbf{B}(z,v_t))=0\big)dz\\ &\le t\int_{\mathbf{B}(x,v_t)} e^{-tv_t^dk_d}dz \le (u+\log(t))e^{-u-\log(t)} = \frac{u+\log(t)}{e^ut}. \end{aligned} \tag{4.10}$$

It follows from the triangle inequality that

$$\mathbb{E}[|Z_t(u)|] \le 2^d\frac{u+\log(t)}{e^{u/2}\sqrt{t}} + \frac{u+\log(t)}{e^ut}.$$

Then, by the first inequality of (1.3) in Theorem 1.1, we obtain (3.14).

Let us now show (3.15). We consider the cases $u\ge 0$, $u\in[-\log(\log(t)),0)$ and $u<-\log(\log(t))$ separately. Because of $ue^{-u}\le 1$ and $ue^{-u/2}\le 1$ for $u\ge 0$ and $\log(t)\ge 1$, by (3.14) we have

$$d_{TV}\big(\xi_t((u,\infty)),P_{e^{-u}}\big) \le (2^{d+1}+2)\frac{\log(t)}{\sqrt{t}}$$

for $u\ge 0$, which proves (3.15) for $u\ge 0$.

In the following let $u\in[-\log(\log(t)),0)$ be fixed. Since $Z_t(u)=Z'_{t,X}(u)-Z''_{t,X}(u)$ and the terms on the right-hand side are both non-negative, we obtain that

$$Z_t(u)_+ \le Z'_{t,X}(u) \quad\text{and}\quad Z_t(u)_- \le Z''_{t,X}(u).$$

Combining these inequalities and (1.4) in Theorem 1.1 with $m = 1$ establishes

$$\begin{aligned}&|\mathbb{P}(R_t \le u) - \mathbb{P}(P_{e^{-u}} = 0)| = |\mathbb{P}(\xi_t((u,\infty)) = 0) - \mathbb{P}(P_{e^{-u}} = 0)|\\ &\le e^u \mathbb{E}[|Z_t(u)|] + \mathbb{E}[|Z_t(u)|\mathbf{1}\{\xi_t((u,\infty)) - Z_t(u)_- = 0\}]\\ &\le e^u \mathbb{E}[|Z_t(u)|] + \mathbb{E}[Z'_{t,X}(u)\mathbf{1}\{\xi_t((u,\infty)) = 0\}] + \mathbb{E}[Z''_{t,X}(u)].\end{aligned}$$

Moreover, by (4.9) and (4.10) we have

$$\mathbb{E}[Z'_{t,X}(u)] \le 2^d \frac{u + \log(t)}{e^{u/2}\sqrt{t}} \le 2^d \frac{\log(t)}{e^{u/2}\sqrt{t}} \text{ and } \mathbb{E}[Z''_{t,X}(u)] \le \frac{u + \log(t)}{e^u t} \le \frac{(\log(t))^2}{t} \le \frac{\log(t)}{\sqrt{t}}.$$

Thus the identity $Z_t(u) = Z'_{t,X}(u) - Z''_{t,X}(u)$ with $Z'_{t,X}(u), Z''_{t,X}(u) \ge 0$ implies that

$$|\mathbb{P}(R_t \le u) - \mathbb{P}(P_{e^{-u}} = 0)| \le (2^d + 2)\frac{\log(t)}{\sqrt{t}} + \mathbb{E}[Z'_{t,X}(u)\mathbf{1}\{\xi_t((u,\infty)) = 0\}]. \tag{4.11}$$

For $x \in W$ we define

$$\xi_{t,x}((u,\infty)) = \sum_{z \in \eta_t \cap W \cap \mathbf{B}(x,4v_t)^c} \mathbf{1}\{h_t(z,\eta_t) > u\}.$$

Since, for $x \in W$, $\mathbf{1}\{\xi_t((u,\infty)) = 0\} \le \mathbf{1}\{\xi_{t,x}((u,\infty)) = 0\}$ and $Z'_{t,x}(u)$ and $\mathbf{1}\{\xi_{t,x}((u,\infty)) = 0\}$ are independent, we have

$$\begin{aligned}\mathbb{E}[Z'_{t,X}(u)\mathbf{1}\{\xi_t((u,\infty)) = 0\}] &\le \int_W \mathbb{E}[Z'_{t,x}(u)\mathbf{1}\{\xi_{t,x}((u,\infty)) = 0\}]dx\\ &= \int_W \mathbb{E}[Z'_{t,x}(u)]\mathbb{P}(\xi_{t,x}((u,\infty)) = 0)dx.\end{aligned} \tag{4.12}$$

For $x \in W$, the Markov and the triangle inequalities, (3.14) and $e^{u/2}\sqrt{t} \ge 1$ imply that

$$\begin{aligned}\mathbb{P}(\xi_{t,x}((u,\infty)) = 0) &\le \mathbb{P}(\xi_t((u,\infty)) = 0) + \mathbb{P}\Big(\sum_{z \in \eta_t \cap \mathbf{B}(x,4v_t)} \mathbf{1}\{h_t(z,\eta_t) > u\} > 0\Big)\\ &\le 2^d \frac{\log(t)}{e^{u/2}\sqrt{t}} + \frac{\log(t)}{e^u t} + e^{-e^{-u}} + \mathbb{E}\Big[\sum_{z \in \eta_t \cap \mathbf{B}(x,4v_t)} \mathbf{1}\big\{h_t(z,\eta_t) > u\big\}\Big]\\ &\le (2^d + 1)\frac{\log(t)}{e^{u/2}\sqrt{t}} + e^{-e^{-u}} + \mathbb{E}\Big[\sum_{z \in \eta_t \cap \mathbf{B}(x,4v_t)} \mathbf{1}\big\{h_t(z,\eta_t) > u\big\}\Big].\end{aligned}$$

Similar arguments as used in (4.10) and $e^{u/2}\sqrt{t} \ge 1$ lead to

$$\mathbb{E}\Big[\sum_{z \in \eta_t \cap \mathbf{B}(x,4v_t)} \mathbf{1}\big\{h_t(z,\eta_t) > u\big\}\Big] \le \frac{4^d(u + \log(t))}{e^u t} \le \frac{4^d \log(t)}{e^{u/2}\sqrt{t}}.$$

Since $\log(t)e^u \ge 1$ and $\frac{\log(t)^2}{\sqrt{t}} \le 4$ for $t > e^2$, we obtain

$$\frac{\log(t)}{e^{u/2}\sqrt{t}} \le \frac{\log(t)^2 e^u}{e^{u/2}\sqrt{t}} \le \frac{\log(t)^2}{\sqrt{t}} e^{u/2} \le 4e^{u/2}.$$

Together with $\exp(-e^{-u} - u/2) \le 1$, which follows from $u < 0$, we have shown

$$\mathbb{P}(\xi_{t,x}((u,\infty)) = 0) \le 4(4^d + 2^d + 1)e^{u/2} + e^{u/2} \le (4(4^d + 2^d + 1) + 1)e^{u/2}$$

so that, by (4.9) and (4.12),

$$\begin{aligned}\mathbb{E}[Z'_{t,X}(u)\mathbf{1}\{\xi_t((u,\infty))=0\}] &\le (4(4^d+2^d+1)+1)e^{u/2}\frac{2^d\log(t)}{e^{u/2}\sqrt{t}}\\ &= (2^{d+2}(4^d+2^d+1)+2^d)\frac{\log(t)}{\sqrt{t}}.\end{aligned}$$

Combining this with (4.11) leads to

$$\begin{aligned}\left|\mathbb{P}(R_t\le u)-e^{-e^{-u}}\right| &\le (2^{d+2}(4^d+2^d+1)+2^d+2^d+2)\frac{\log(t)}{\sqrt{t}}\\ &\le 2^{d+2}(4^d+2^d+2)\frac{\log(t)}{\sqrt{t}},\end{aligned} \tag{4.13}$$

which establishes (3.15) for $u\in[-\log(\log(t)),0)$.

Finally for $u<-\log(\log(t))$ we have

$$\mathbb{P}(R_t\le u)\le\mathbb{P}(R_t\le-\log(\log(t))),$$

which by (4.13) and the triangle inequality is bounded by

$$2^{d+2}(4^d+2^d+2)\frac{\log(t)}{\sqrt{t}}+\frac{1}{t}.$$

Therefore elementary arguments lead to

$$\sup_{u<-\log(\log(t))}\left|\mathbb{P}(R_t\le u)-e^{-e^{-u}}\right|\le[2^{d+2}(4^d+2^d+2)+1]\frac{\log(t)}{\sqrt{t}},$$

which concludes the proof of (3.15). □

Remark 4.8. Note that the integral in the middle of (4.9) cannot be bounded with a better exponent for t. Indeed, using substitution, we can rewrite the integral as

$$\frac{u+\log(t)}{k_d}\int_{\mathbf{B}(0,2)\cap\mathbf{B}(0,1)^c}e^{-(u+\log(t))\frac{\lambda_d(\mathbf{B}(y,1)\cap\mathbf{B}(0,1)^c)}{k_d}}dy.$$

For any sufficiently small $\varepsilon>0$ there exists a set $A\subset\mathbf{B}(0,2)\cap\mathbf{B}(0,1)^c$ with $\lambda_d(A)>0$ such that the ratio in the exponent is at least $(1+\varepsilon)/2$ for all $y\in A$. This provides a lower bound of the order $\log(t)t^{-(1+\varepsilon)/2}$.

References

[1] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *Ann. Probab.*, 17(1):9–25, 1989.

[2] R. Arratia, L. Goldstein, and L. Gordon. Poisson approximation and the Chen-Stein method. *Statist. Sci.*, 5(4):403–434, 1990.

[3] R. Arratia, L. Goldstein, and F. Kochman. Size bias for one and all. *Probab. Surv.*, 16:1–61, 2019.

[4] N. Balakrishnan and M. V. Koutras. *Runs and scans with applications*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], New York, 2002.

[5] A. D. Barbour. Stein's method and Poisson process convergence. *J. Appl. Probab.*, (25A):175–184, 1988.

[6] A. D. Barbour and G. K. Eagleson. Poisson convergence for dissociated statistics. *J. Roy. Statist. Soc. Ser. B*, 46(3):397–402, 1984.

[7] A. D. Barbour, L. Holst, and S. Janson. *Poisson approximation*, volume 2 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York, 1992.

[8] A. D. Barbour and A. Xia. On Stein's factors for Poisson approximation in Wasserstein distance. *Bernoulli*, 12(6):943–954, 2006.

[9] B. Błaszczyszyn and R. Schott. Approximations of functionals of some modulated-Poisson Voronoi tessellations with applications to modeling of communication networks. *Japan J. Indust. Appl. Math.*, 22(2):179–204, 2005.

[10] P. Calka. Tessellations. In *New perspectives in stochastic geometry*, pages 145–169. Oxford Univ. Press, Oxford, 2010.

[11] P. Calka and N. Chenavier. Extreme values for characteristic radii of a Poisson-Voronoi tessellation. *Extremes*, 17(3):359–385, 2014.

[12] L. H. Y. Chen. Stein's method: some perspectives with applications. In *Probability towards 2000 (New York, 1995)*, volume 128 of *Lect. Notes Stat.*, pages 97–122. Springer, New York, 1998.

[13] N. Chenavier. A general study of extremes of stationary tessellations with examples. *Stochastic Processes and their Applications*, 124(9):2917–2953, 2014.

[14] F. Daly and O. Johnson. Relaxation of monotone coupling conditions: Poisson approximation and beyond. *J. Appl. Probab.*, 55(3):742–759, 2018.

[15] L. Decreusefond, M. Schulte, and C. Thäle. Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry. *Ann. Probab.*, 44(3):2147–2197, 2016.

[16] A. P. Godbole. Poisson approximations for runs and patterns of rare events. *Adv. in Appl. Probab.*, 23(4):851–865, 1991.

[17] N. Henze. Ein asymptotischer Satz über den maximalen Minimalabstand von unabhängigen Zufallsvektoren mit Anwendung auf einen Anpassungstest im $\mathbf{R}^p$ und auf der Kugel. *Metrika*, 30(4):245–259, 1983.

[18] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[19] V. S. Koroljuk and Y. V. Borovskich. *Theory of U-statistics*, volume 273 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1994. Translated from the 1989 Russian original by P. V. Malyshev and D. V. Malyshev and revised by the authors.

[20] M. V. Koutras, G. K. Papadopoulos, and S. G. Papastavridis. Runs on a circle. *J. Appl. Probab.*, 32(2):396–404, 1995.

[21] W. Lao and M. Mayer. U-max-statistics. *J. Multivariate Anal.*, 99(9):2039–2052, 2008.

[22] G. Last and M. Penrose. *Lectures on the Poisson process*, volume 7 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, Cambridge, 2018.

[23] A. J. Lee. *U-statistics*, volume 110 of *Statistics: Textbooks and Monographs*. Marcel Dekker, Inc., New York, 1990.

[24] M. Mayer and I. Molchanov. Limit theorems for the diameter of a random sample in the unit ball. *Extremes*, 10(3):129–150, 2007.

[25] J. Møller. *Lectures on random Voronoĭ tessellations*, volume 87 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1994.

[26] S. Y. Novak. *Extreme value methods with applications to finance*, volume 122 of *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL, 2012.

[27] S. Y. Novak. Poisson approximation. *Probab. Surv.*, 16:228–276, 2019.

[28] M. Otto. Poisson approximation of Poisson-driven point processes and extreme values in stochastic geometry. arXiv:2005.10116, 2020.

[29] M. Penrose. *Random geometric graphs*, volume 5 of *Oxford Studies in Probability*. Oxford University Press, Oxford, 2003.

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

[31] M. D. Penrose. Inhomogeneous random graphs, isolated vertices, and Poisson approximation. *J. Appl. Probab.*, 55(1):112–136, 2018.

[32] F. Pianoforte and M. Schulte. Criteria for Poisson process convergence with applications to inhomogeneous Poisson-Voronoi tessellations. arXiv:2101.07739, 2021.

[33] A. Poupon. Voronoi and Voronoi-related tessellations in studies of protein structure and interaction. *Current Opinion in Structural Biology*, 14(2):233–241, 2004.

[34] M. Ramella, W. Boschin, D. Fadda, and M. Nonino. Finding galaxy clusters using Voronoi tessellations. *A&A*, 368(3):776–786, 2001.

[35] A. Röllin. Translated Poisson approximation using exchangeable pair couplings. *Ann. Appl. Probab.*, 17(5-6):1596–1614, 2007.

[36] N. Ross. Fundamentals of Stein's method. *Probab. Surv.*, 8:210–293, 2011.

[37] R. Schneider and W. Weil. *Stochastic and integral geometry*. Probability and its Applications (New York). Springer-Verlag, Berlin, 2008.

[38] M. Schulte and C. Thäle. The scaling limit of Poisson-driven order statistics with applications in geometric probability. *Stochastic Process. Appl.*, 122(12):4096–4120, 2012.

[39] M. Schulte and C. Thäle. Poisson point process convergence and extreme values in stochastic geometry. In *Stochastic analysis for Poisson point processes*, volume 7 of *Bocconi Springer Ser.*, pages 255–294. Bocconi Univ. Press, 2016.

[40] B. Silverman and T. Brown. Short distances, flat triangles and Poisson limits. *J. Appl. Probab.*, 15(4):815–825, 1978.

Acknowledgments. We would like to thank Fraser Daly for some valuable comments. We are thankful to an anonymous referee and an associate editor for their attentive reading and their helpful suggestions.