



OMER BOBROWSKI

MATTHIAS SCHULTE

D. YOGESHWARAN

POISSON PROCESS
APPROXIMATION UNDER
STABILIZATION AND PALM
COUPLING

APPROXIMATION PAR PROCESSUS DE
POISSON SOUS STABILISATION ET
COUPLAGE DE PALM

ABSTRACT. — We present new Poisson process approximation results for stabilizing functionals of Poisson and binomial point processes. These functionals are allowed to have an unbounded range of interaction and encompass many examples in stochastic geometry. Our bounds are derived for the Kantorovich–Rubinstein distance using the generator approach to

Keywords: Functional limit theorems, Poisson process approximation, Kantorovich-Rubinstein distance, Point processes, Stein’s method, Glauber dynamics, Palm coupling, Stabilizing statistics, k -nearest neighbor balls, Morse critical points, Binomial point processes.

2020 *Mathematics Subject Classification:* 60D05, 60F17, 60G55, 60H07, 60J76.

DOI: <https://doi.org/10.5802/ahl.156>

(*) OB was supported in part by the Israel Science Foundation, Grant 1965/19. MS was supported by the Swiss National Science Foundation Grant No. 200021_175584. DY’s research was funded by DST-INSPIRE Faculty Award, SERB-MATRICES Grant and CPDA from the Indian Statistical Institute. DY wishes to thank Moritz Otto for explaining his results in advance. The authors are also grateful to the anonymous referees for thorough reviews and useful suggestions.

Stein’s method. We give different types of bounds for different point processes. While some of our bounds are given in terms of coupling of the point process with its Palm version, the others are in terms of the local dependence structure formalized via the notion of stabilization. We provide two supporting examples for our new framework – one is for Morse critical points of the distance function, and the other is for large k -nearest neighbor balls. Our bounds considerably extend the results in Barbour and Brown (1992), Decreasefond, Schulte and Thäle (2016) and Otto (2020).

RÉSUMÉ. — Nous présentons de nouveaux résultats d’approximation par des processus de Poisson pour les fonctionnelles stabilisantes de processus de Poisson et binomiaux. Ces fonctionnelles peuvent avoir des interactions non bornées, et recouvrent de nombreux exemples en géométrie stochastique. Nous obtenons des bornes pour la distance de Kantorovich–Rubinstein grâce par la méthode de Stein pour le générateur. Nous donnons différents types de bornes pour différents processus ponctuels. Alors que certaines de nos bornes soient données en termes de couplage du processus ponctuel avec sa version Palm, les autres sont exprimées en termes de la structure de dépendance locale, formalisée par la notion de stabilisation. Nous donnons deux exemples significatifs de ce nouveau cadre – l’un pour les points critiques Morse de la fonction distance, l’autre pour les grandes boules pour les k voisins les plus proches. Nos bornes étendent considérablement les résultats de Barbour et Brown (1992), Decreasefond, Schulte et Thäle (2016) et Otto (2020).

1. Introduction

Poisson processes arise in many fields of probability theory and are arguably the most prominent class of point processes. At the same time they are very convenient to work with since they exhibit complete spatial independence (i.e., point configurations within disjoint sets are independent). Therefore, approximating an involved point process by a Poisson process is an important problem. This is also the crucial step in establishing convergence in distribution to a Poisson process. This paper focuses on Poisson process approximation for a large class of point processes that often arise in stochastic geometry. The processes we study are functionals of Poisson (or binomial) point processes, which are themselves not Poisson, and in particular lack spatial independence. The results we present here significantly generalize recent ones [DST16, Ott20] by either considering a stronger approximation distance or more general functionals. Our approach is based on Stein’s method for Poisson process approximation; see e.g. [Bar88, BB92, BHJ92, CX04, DST16, Xia05]. Now we shall offer a quick preview of our setup, results and some background literature.

Let ν be either a Poisson process or a binomial point process (i.e., a collection of n i.i.d. points) in some space \mathbb{X} . We denote by $\mathbf{N}_{\mathbb{X}}$ the set of σ -finite point configurations on \mathbb{X} . For $k \in \mathbb{N}$, a second space \mathbb{Y} and measurable functions $f : \mathbb{X}^k \times \mathbf{N}_{\mathbb{X}} \rightarrow \mathbb{Y}$ and $g : \mathbb{X}^k \times \mathbf{N}_{\mathbb{X}} \rightarrow \{0, 1\}$ that are symmetric in the first k arguments, we consider the following point process on \mathbb{Y} ,

$$(1.1) \quad \xi := \xi[\nu] := \frac{1}{k!} \sum_{\mathbf{x} \in \nu_{\neq}^k} g(\mathbf{x}, \nu) \delta_{f(\mathbf{x}, \nu)},$$

where ν_{\neq}^k denotes the set of all k -tuples of distinct points of ν , and δ_y stands for the Dirac measure concentrated at $y \in \mathbb{Y}$. Point processes of the form (1.1) often arise

in stochastic geometry (cf. [Ott20, PS22]). For example, consider the case $\mathbb{X} = \mathbb{R}^d$, $\mathbb{Y} = \mathbb{R}$, $k = 1$, $g \equiv 1$ and $f(x, \nu)$ being the minimal distance from x to another point of ν . The point process ξ is then the collection of all nearest neighbor distances of ν .

Our main results deal with the comparison of ξ with a Poisson process ζ on \mathbb{Y} with a finite intensity measure. We consider the Kantorovich–Rubinstein distance between the distributions of the two processes ξ and ζ , defined as

$$(1.2) \quad \mathbf{d}_{\mathbf{KR}}(\xi, \zeta) := \sup_{h \in \text{LIP}(\mathbb{Y})} |\mathbb{E}\{h(\xi)\} - \mathbb{E}\{h(\zeta)\}|,$$

where $\text{LIP}(\mathbb{Y})$ is the class of measurable 1-Lipschitz functions with respect to the total variation distance on the space of finite point configurations on \mathbb{Y} (see (2.1)). Note that convergence in Kantorovich–Rubinstein distance implies convergence in distribution, and that the Kantorovich–Rubinstein distance dominates the total variation distance

$$(1.3) \quad \mathbf{d}_{\mathbf{TV}}(\xi, \zeta) := \sup_{A \in \mathcal{N}_{\mathbb{Y}}} |\mathbb{P}(\xi \in A) - \mathbb{P}(\zeta \in A)|,$$

where $\mathcal{N}_{\mathbb{Y}}$ is the σ -field on the space $\mathbf{N}_{\mathbb{Y}}$. The Kantorovich–Rubinstein distance, which is often studied in the context of optimal transport, is more sensitive to differences between the compared point processes than the total variation distance since the test functions can take larger values than only zero and one. Studying the Kantorovich–Rubinstein distance goes in the opposite direction to many other works on Poisson process approximation via Stein’s method, where weaker distances than the total variation distance were considered, see e.g. [BB92, BHJ92, CX04, Sch05, Sch09, Xia05].

In general, if the behaviour of ξ is close to that of a Poisson process, by taking $\mathbb{Y} = \mathbb{R}$ one gets an approximation for

$$\max \left\{ f(\mathbf{x}, \nu) : \mathbf{x} \in \nu_{\neq}^k, g(\mathbf{x}, \nu) = 1 \right\}.$$

Thus, approximation results for the point process ξ can be a crucial tool to compute the limiting distributions of some extreme values of ξ . In a similar fashion, we can treat functionals of ξ other than the maximum, such as order statistics, sums of the points or number of points in a certain region.

This paper continues the line of research initiated in [DST16]. There, processes of the form (1.1) were studied, but for a special case where g and f depend only on \mathbf{x} and not on ν , i.e., $f(\mathbf{x}, \nu) = f(\mathbf{x}), g(\mathbf{x}, \nu) = g(\mathbf{x})$. In other words, the results of [DST16] allow to deal with point processes that have the form of U -statistics. However, this formulation excludes point processes that arise naturally such as the nearest neighbor example mentioned above. This lacuna is addressed in the present paper. We summarize our main results as follows.

- We establish a general Poisson process approximation result (Theorem 3.1), which extends [BB92, Theorem 2.6] from the total variation distance to the Kantorovich–Rubinstein distance. Here we do not assume any structure on ξ and the bounds are given in terms of a coupling between ξ and its Palm measure.
- A similar result is derived for $\xi[\nu]$ under the assumption that ν is a Poisson process and that f in (1.1) depends only on \mathbf{x} (see Theorem 3.3). This result

- already extends the Poisson process approximation result of [DST16, Theorem 3.1] and may be considered as the process-level analogue of the Poisson approximation result in [Pen18, Theorem 3.1]. Theorem 3.1 and Theorem 3.3 already cover many existing applications in the literature (see [DST16, Ott20]).
- We establish a Poisson process approximation result for $\xi[\nu]$ when ν is a Poisson process and $f(\mathbf{x}, \nu), g(\mathbf{x}, \nu)$ depend on random neighborhoods around \mathbf{x} (see Theorem 4.1). This is an important concept in stochastic geometry known as stabilization. We elaborate on this in Remark 1.1. Here, our bounds are given more explicitly in terms of g as well as the tail probabilities of the random neighborhoods. This result in its full generality is one of the novel contributions of this work and an analogous result even in total variation distance is not available to the best of our knowledge. Under total variation distance, approximation results are proven only for $f(\mathbf{x}, \nu) = f(\mathbf{x})$ in [Ott20, Theorem 3.3] and the bounds therein are same as ours (barring a factor of 2) even though our bounds hold for more general functions f and in the stronger Kantorovich–Rubinstein distance.
 - For the binomial case, the lack of spatial independence leads to a more restricted statement and more involved bounds. In particular, for $k = 1$ we can show a Poisson process approximation result under stabilization of $f(x, \nu), g(x, \nu)$ (see Theorem 5.1). The restriction to $k = 1$ here is more for ease of stating the bounds, see Remark 5.2. We are not aware of such a general approximation result even under weaker distances for point processes driven by stabilizing functionals of binomial point processes. In fact, such results are not known even for functionals with finite range of stabilization, i.e., $f(x, \nu), g(x, \nu)$ depend on fixed compact neighborhoods around \mathbf{x} .

One point of difference between the four results is that the bounds in the first two results do not assume any local dependence structure on f and g , and the bounds are given in terms of a coupling. On the other hand, the latter two bounds are given explicitly in terms of the functional g alone. As will be illustrated in Section 6, our general bounds shall necessitate a few additional computations in applications compared to those for a Poisson (non-functional) approximation result. Since a marked point process with points in a space \mathbb{X}' and marks in a space \mathbb{M} can be seen as a point process on the larger space $\mathbb{X}' \times \mathbb{M}$, which one can choose as \mathbb{X} or \mathbb{Y} , our results also apply to marked point processes, but we shall not develop any applications to underlying marked point processes here.

Remark 1.1. — The notion of stabilization was introduced in the context of proving laws of large numbers and central limit theorems for functionals of point processes (see the surveys [Sch10, Yuk13]), and is meant to capture the spatial dependence of the functionals. Apart from its extensive use in proving laws of large numbers and central limit theorems, it has also been used to prove moderate deviations [ERS15] and normal approximation bounds [LRSY19]. Even though this notion is implicitly used in some of the Poisson approximation results, to the best of our knowledge there has not been a general Poisson process approximation result for stabilizing functionals until the recent work [Ott20]. In this regards, our article

makes an important contribution in furthering the usage of stabilization in Poisson process approximation. More detailed comparisons of our results with those in the literature (especially [BB92, DST16, Ott20]) are provided after the respective theorem statements.

Apart from extending total variation bounds in [Ott20] to Kantorovich–Rubinstein distance and the applications therein, we further envisage that the general Poisson process approximation result (Theorem 3.1) and the proof approach may be useful for deriving similar bounds when ν is not a Poisson process as well. Our latter results (Theorems 4.1 and 5.1), significantly enhance the scope of many applications and also allow us to consider more general point processes than those investigated in the literature so far.

In order to demonstrate the applicability and versatility of our results, we consider two examples – one from the field of random topology and the other from stochastic geometry:

- (1) Morse critical points: For a homogeneous Poisson process on a flat torus, we consider the distance function (from the points) and define critical points together with their Morse index. Grouping together all critical points of the same index, we have a point process for which we wish to prove Poisson convergence under suitable scaling. This convergence statement has a significant contribution to the analysis of the homological connectivity phenomenon studied in [Bob22]. In particular, it yields the asymptotic behaviour of the persistent homology in the critical window for homological connectivity (see Remark 6.3).
- (2) Large k -nearest neighbor balls: For an underlying Poisson or binomial point process we put around each point a ball whose radius is the distance to its k -nearest neighbor and establish that the scaled volumes (with respect to the intensity measure) of these balls can be approximated by a Poisson process. In both cases, our results yield explicit bounds for the rate of convergence (when the point processes are restricted to suitable sets). Here, by a more careful computation, we obtain better bounds than those in [Ott20, Theorem 4.2] for underlying Poisson processes and we also obtain a new approximation result when the underlying point process ν is a binomial point process. Comparisons between our applications and those existing in the literature as well as more potential applications are discussed in Section 6.

Before we end the introduction, we will say a quick word about our proofs. In order to control the Kantorovich–Rubinstein distance between ξ and ζ , we employ the same generator approach as in [DST16], see also [BB92]. To further bound the resulting expressions we use two classical approaches from Stein’s method, coupling and local dependence, which lead to slightly different results for the Poisson process approximation of ξ .

This paper is organized as follows. After introducing some preliminaries in Section 2 (including Stein’s method for Poisson process approximation), we present our coupling approach in Section 3. Local dependence (and stabilization) is used for underlying Poisson and binomial point processes in Sections 4 and 5, respectively.

Finally, we provide two applications of our main results to random topology and stochastic geometry in Section 6.

2. Preliminaries

2.1. Point processes

We shall adapt the notation from [DST16] and for more on point process theory, we refer the reader to [BBK20, LP17]. Let $(\mathbb{X}, \mathcal{X})$ be a locally compact second countable Hausdorff space (lcsch space). We denote by $\mathbf{N}_{\mathbb{X}}$ the space of σ -finite counting measures on \mathbb{X} , and by $\widehat{\mathbf{N}}_{\mathbb{X}}$ the space of all finite counting measures on \mathbb{X} . We equip the spaces $\mathbf{N}_{\mathbb{X}}$ and $\widehat{\mathbf{N}}_{\mathbb{X}}$ with the corresponding σ -algebras $\mathcal{N}_{\mathbb{X}}$ and $\widehat{\mathcal{N}}_{\mathbb{X}}$, which are induced by the maps $\omega \mapsto \omega(B)$ for all $B \in \mathcal{X}$.

By [LP17, Proposition 6.2], each $\omega \in \mathbf{N}_{\mathbb{X}}$ can be written as $\omega = \sum_{i \in I} \delta_{x_i}$, where $(x_i)_{i \in I}$ is a countable collection of points in \mathbb{X} and δ_x stands for the Dirac measure concentrated at $x \in \mathbb{X}$. Due to this representation we can think of counting measures as multisets and treat them as such by abusing notation. Thus, we write $\sum_{x \in \omega}$ for the sum over all $x_i, i \in I$, which is the same as the integral with respect to ω . More generally, we define for $k \in \mathbb{N}$,

$$\omega_{\neq}^k := \{(x_{i_1}, \dots, x_{i_k}) : i_1, \dots, i_k \in I, i_j \neq i_\ell, j \neq \ell\}$$

and denote by $\sum_{(x_1, \dots, x_k) \in \omega_{\neq}^k}$ the sum over all k -tuples of distinct points of ω , which is the same as the integral with respect to the k^{th} factorial measure of ω (see [LP17, Subsection 4.2]). By $\omega \cap S$ we mean the restriction of ω to $S \in \mathcal{X}$. Moreover, we write $\omega_1 \subset \omega_2$ for $\omega_1, \omega_2 \in \mathbf{N}_{\mathbb{X}}$ if $\omega_1(A) \leq \omega_2(A)$ for all $A \in \mathcal{X}$. In addition, for $\omega_1, \omega_2 \in \mathbf{N}_{\mathbb{X}}$ such that $\omega_1 = \sum_{i \in I} \delta_{x_i}$ and $\omega_2 = \sum_{j \in J} \delta_{y_j}$, we define

$$\omega_1 \setminus \omega_2 := \sum_{x \in \{x_i : i \in I\} \setminus \{y_j : j \in J\}} \delta_x.$$

The sets under the sum must be understood as multisets since ω_1, ω_2 and $\omega_1 \setminus \omega_2$ can have multiple points (i.e., points $x \in \mathbb{X}$ such that $\omega(\{x\}) > 1$). The symmetric difference $\omega_1 \triangle \omega_2$ is defined as $\omega_1 \triangle \omega_2 := \omega_1 \setminus \omega_2 + \omega_2 \setminus \omega_1$.

The total variation distance between two measures μ_1, μ_2 on \mathbb{X} is defined as

$$d_{TV}(\mu_1, \mu_2) := \sup_{\substack{A \in \mathcal{X} \\ \mu_1(A), \mu_2(A) < \infty}} |\mu_1(A) - \mu_2(A)|.$$

We denote by $\text{LIP}(\mathbb{X})$ the class of measurable functionals $h : \widehat{\mathbf{N}}_{\mathbb{X}} \rightarrow \mathbb{R}$ such that for all $\omega_1, \omega_2 \in \widehat{\mathbf{N}}_{\mathbb{X}}$,

$$(2.1) \quad |h(\omega_1) - h(\omega_2)| \leq d_{TV}(\omega_1, \omega_2).$$

In other words, $\text{LIP}(\mathbb{X})$ is the class of measurable 1-Lipschitz functionals on $\widehat{\mathbf{N}}_{\mathbb{X}}$ with respect to the total variation distance between measures. Let $\|\cdot\|$ denote the *total variation* of a signed measure, i.e., $\|\mu\| = \mu_+(\mathbb{X}) + \mu_-(\mathbb{X})$ where μ_+, μ_- are the positive and negative parts of the signed measure μ . Note that if $\omega_1, \omega_2 \in \widehat{\mathbf{N}}_{\mathbb{X}}$, then $d_{TV}(\omega_1, \omega_2) = \max\{(\omega_1 \setminus \omega_2)(\mathbb{X}), (\omega_2 \setminus \omega_1)(\mathbb{X})\}$. Hence, we have the following

relation between the total variation distance and the total variation of the difference between counting measures

$$d_{TV}(\omega_1, \omega_2) \leq \|\omega_1 - \omega_2\| = (\omega_1 \Delta \omega_2)(\mathbb{X}) \leq 2d_{TV}(\omega_1, \omega_2).$$

The first inequality becomes an equality when $\omega_1 \subset \omega_2$. Thus, measuring distance between counting measures via total variation distance and total variation of measures differs at most by a factor of 2.

Recall the definitions of Kantorovich–Rubinstein (KR) distance $\mathbf{d}_{\mathbf{KR}}$ and the classical total variation distance $\mathbf{d}_{\mathbf{TV}}$ in (1.2) and (1.3) respectively. We refer to [DST16, Section 2.5] for a dual formulation as well as more details on the KR distance. Since, for $A \in \widehat{\mathcal{N}}_{\mathbb{X}}$, $\omega \mapsto \mathbb{1}\{\omega \in A\}$ is a function in $\text{LIP}(\mathbb{X})$, we have that $\mathbf{d}_{\mathbf{TV}}(\zeta_1, \zeta_2) \leq \mathbf{d}_{\mathbf{KR}}(\zeta_1, \zeta_2)$ for finite point processes ζ_1 and ζ_2 . For an example of a sequence of point processes that converges in $\mathbf{d}_{\mathbf{TV}}$ but not in $\mathbf{d}_{\mathbf{KR}}$, see [DST16, Example 2.2].

Let ζ be a point process with σ -finite intensity measure \mathbf{M} , i.e., $\mathbb{E}\{\zeta(A)\} = \mathbf{M}(A)$ for all $A \in \mathcal{X}$. The *reduced Palm expectation* of ζ at x denoted by $\mathbb{E}_x^!$ is defined via the Campbell–Mecke–Little formula [BBK20, Corollary 3.1.14], as follows. For a measurable function $f : \mathbb{X} \times \mathbf{N}_{\mathbb{X}} \rightarrow \mathbb{R}_+$, it holds that

$$(2.2) \quad \mathbb{E} \left\{ \sum_{x \in \zeta} f(x, \zeta) \right\} = \int_{\mathbb{X}} \mathbb{E}_x^! \{ f(x, \zeta + \delta_x) \} \mathbf{M}(dx).$$

The above formula can be extended to integrable real-valued functions by standard measure-theoretic arguments. Strictly speaking, $\mathbb{E}_x^!$ is defined only for \mathbf{M} -a.e. x , but this suffices for our purposes. The corresponding probability distribution is denoted by $\mathbb{P}_x^!$ and the point process with this probability distribution is denoted by $\zeta_x^!$.

An essential tool for studying Poisson processes is the multivariate Mecke equation (see [LP17, Theorem 4.4]). For a Poisson process η on \mathbb{X} with σ -finite intensity measure \mathbf{K} and a measurable function $f : \mathbb{X}^k \times \mathbf{N}_{\mathbb{X}} \rightarrow \mathbb{R}_+$, we have that

$$(2.3) \quad \mathbb{E} \left\{ \sum_{\mathbf{x} \in \eta_{\mathbb{X}^k}} f(\mathbf{x}, \eta) \right\} = \int_{\mathbb{X}^k} \mathbb{E} \{ f(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \} \mathbf{K}^k(d\mathbf{x}),$$

where $\mathbf{x} = (x_1, \dots, x_k)$ and $\delta_{\mathbf{x}} := \sum_{i=1}^k \delta_{x_i}$. The Mecke equation for $k = 1$ characterizes the Poisson process [LP17, Theorem 4.1] and by the definition of the reduced Palm version of a point process, we can also rephrase this characterization as $\eta_x^! \stackrel{d}{=} \eta$ for all $x \in \mathbb{X}$, where $\stackrel{d}{=}$ denotes equality in distribution of the two random elements.

For two Poisson processes ζ_1 and ζ_2 with finite intensity measures \mathbf{M} and \mathbf{L} , it is known (see [DST16, Remark 3.2(iv)]) that

$$(2.4) \quad \mathbf{d}_{\mathbf{KR}}(\zeta_1, \zeta_2) \leq d_{TV}(\mathbf{M}, \mathbf{L}).$$

Thus, it follows from the triangle inequality for $\mathbf{d}_{\mathbf{KR}}$ that, for any finite point process ξ ,

$$(2.5) \quad \mathbf{d}_{\mathbf{KR}}(\xi, \zeta_1) \leq d_{TV}(\mathbf{M}, \mathbf{L}) + \mathbf{d}_{\mathbf{KR}}(\xi, \zeta_2).$$

Due to this inequality it is often sufficient to compare a finite point process with a Poisson process with the same intensity measure.

2.2. Stein’s method for Poisson process approximation

Stein’s method is used to compare the distributions of two random objects (cf. [Ros11]). In our case we would like to compare a point process ξ to a given Poisson process ζ . We will use the generator approach with respect to the KR distance as in [DST16], using Glauber dynamics for Poisson process. See e.g. [BB92, BHJ92, CX04, Sch05, Sch09, Xia05] for more on Poisson process approximation under other metrics.

Let ζ be a Poisson process on $(\mathbb{X}, \mathcal{X})$, with a finite intensity measure \mathbf{M} . For $h : \widehat{\mathbf{N}}_{\mathbb{X}} \rightarrow \mathbb{R}$ and $\omega \in \widehat{\mathbf{N}}_{\mathbb{X}}$, we define

$$(2.6) \quad \mathcal{L}h(\omega) := \int_{\mathbb{X}} D_x h(\omega) \mathbf{M}(dx) - \int_{\mathbb{X}} D_x h(\omega - \delta_x) \omega(dx),$$

where $D_x h(\omega) := h(\omega + \delta_x) - h(\omega)$, and $\int_{\mathbb{X}} f(x) \omega(dx) = \sum_{x \in \omega} f(x)$. It can be shown (see [DVJ07, Proposition 10.4.VII] or [Pre75]) that the operator \mathcal{L} is the generator of a Markov process whose stationary distribution is the same as that of ζ . Further, \mathcal{L} satisfies the following equation (see [DST16, (4.3)])

$$(2.7) \quad \mathbb{E} \{h(\zeta)\} - h(\omega) = \int_0^\infty \mathcal{L}P_s h(\omega) ds,$$

where P_s denotes the Markov semigroup (also known as the Ornstein–Uhlenbeck semigroup) corresponding to the generator \mathcal{L} , and $P_0 h \equiv h$. The Markov process $(\zeta_s)_{s \geq 0}$ corresponding to the generator \mathcal{L} is called Glauber dynamics and is the spatial birth-death process in continuous time described as follows. If $\zeta_s = \omega$ at time s , each particle in ω dies at rate 1 and new particles are born at rate $\mathbf{M}(\mathbb{X})$ with their location chosen according to the probability measure $\mathbf{M}(\cdot)/\mathbf{M}(\mathbb{X})$. The new particles also have an exponentially distributed (with mean 1) lifetime. With this process representation, we can represent the semigroup P_s as follows. For $h : \widehat{\mathbf{N}}_{\mathbb{X}} \rightarrow \mathbb{R}$ we have

$$P_s h(\omega) = \mathbb{E} \{h(\zeta_s) \mid \zeta_0 = \omega\}.$$

An important property we will use later is that P_s satisfies the following Lipschitz property with respect to the total variation of measures [DST16, Lemma 5.2],

$$(2.8) \quad |P_s h(\omega_1) - P_s h(\omega_2)| \leq e^{-s} (\omega_1 \Delta \omega_2)(\mathbb{X}),$$

for $\omega_1, \omega_2 \in \widehat{\mathbf{N}}_{\mathbb{X}}$ and $h \in \text{LIP}(\mathbb{X})$. The property is deduced using the above representation of the semigroup and the exponential lifetimes of the particles.

3. Approximation for point processes via Palm coupling

We present two results in this section that provide Poisson process approximation bounds in the KR distance via suitable coupling between the original process and its Palm measure. While our first theorem holds for general point processes, the second theorem is specialized to the case of point processes driven by a Poisson process. These two results already significantly generalize many of the existing bounds in the literature. We shall discuss these connections in detail after the respective theorem statements.

THEOREM 3.1. — *Let ξ be a finite point process with intensity measure \mathbf{L} on \mathbb{X} , such that $\mathbf{L}(\mathbb{X}) < \infty$. Let ζ be a Poisson process with a finite intensity measure \mathbf{M} on \mathbb{X} . Suppose that for \mathbf{L} -a.e. $x \in \mathbb{X}$ we have coupled point processes $\xi^x, \tilde{\xi}^x$ such that $\xi^x \stackrel{d}{=} \xi$, and $\tilde{\xi}^x \stackrel{d}{=} \xi_x^!$, respectively ($\xi_x^!$ denotes the reduced Palm version of the point process ξ at x). We also assume that $x \mapsto \mathbb{E}\{(\xi^x \Delta \tilde{\xi}^x)(\mathbb{X})\}$ is measurable. Then*

$$\mathbf{d}_{\mathbf{KR}}(\xi, \zeta) \leq d_{TV}(\mathbf{L}, \mathbf{M}) + 2 \int_{\mathbb{X}} \mathbb{E}\{(\xi^x \Delta \tilde{\xi}^x)(\mathbb{X})\} \mathbf{L}(dx).$$

Remark 3.2. — Note that the couplings ξ^x and $\tilde{\xi}^x$ may be defined on different probability spaces for different x . However, the measurability condition on $x \mapsto \mathbb{E}\{(\xi^x \Delta \tilde{\xi}^x)(\mathbb{X})\}$ is satisfied if the couplings ξ^x and $\tilde{\xi}^x$ are defined on the same probability space for all x and the random function $x \mapsto (\xi^x, \tilde{\xi}^x)$ is measurable.

A version of the above theorem for the total variation distance was proven in [BB92, Theorem 2.6] without the factor of 2. The result of Barbour and Brown [BB92] was used in [Ott20, Theorem 3.3] to derive total variation distance bounds for Poisson approximation of stabilizing statistics of Euclidean Poisson processes. The above theorem implies that the bounds in [Ott20, Theorem 3.3] hold also under the stronger KR distance immediately, only with an additional factor of 2. We shall later state a much more general theorem for stabilizing statistics in Section 4. Our proof of Theorem 3.1 is similar in spirit to that of [BB92, Theorem 2.6], using Lipschitz functions rather than bounded functions of the point processes.

Proof of Theorem 3.1. — Due to (2.5) we can assume that $\mathbf{M} \equiv \mathbf{L}$ throughout the proof. Let $h \in \text{LIP}(\mathbb{X})$, and recall the definition of the generator \mathcal{L} of the Glauber dynamics in (2.6). Observe that $\mathcal{L}h(\omega) \leq \mathbf{L}(\mathbb{X}) + \omega(\mathbb{X})$ for $\omega \in \widehat{\mathbf{N}}_{\mathbb{X}}$. Using the finiteness of $\mathbf{L}(\mathbb{X})$ with Fubini–Tonelli Theorem, and the Campbell–Little–Mecke formula (2.2), we have that

$$\begin{aligned} \mathbb{E}\{\mathcal{L}h(\xi)\} &= \int_{\mathbb{X}} \mathbb{E}\{D_x h(\xi)\} \mathbf{L}(dx) - \int_{\mathbb{X}} \mathbb{E}_x^! \{D_x h(\xi)\} \mathbf{L}(dx) \\ &= \int_{\mathbb{X}} \left(\mathbb{E}\{D_x h(\xi)\} - \mathbb{E}\{D_x h(\xi_x^!)\} \right) \mathbf{L}(dx). \end{aligned}$$

Using our coupling assumption and the Lipschitz assumption on h , we can bound the absolute value of the last integrand by

$$\begin{aligned} \mathbb{E}\left\{\left|D_x h(\xi^x) - D_x h(\tilde{\xi}^x)\right|\right\} &\leq \mathbb{E}\left\{\left|h(\xi^x + \delta_x) - h(\tilde{\xi}^x + \delta_x)\right| + \left|h(\xi^x) - h(\tilde{\xi}^x)\right|\right\} \\ &\leq 2\mathbb{E}\left\{(\xi^x \Delta \tilde{\xi}^x)(\mathbb{X})\right\} \end{aligned}$$

for $x \in \mathbb{X}$ so that

$$|\mathbb{E}\{\mathcal{L}h(\xi)\}| \leq 2 \int_{\mathbb{X}} \mathbb{E}\left\{(\xi^x \Delta \tilde{\xi}^x)(\mathbb{X})\right\} \mathbf{L}(dx) =: R_{\xi}.$$

Similarly, replacing h by $P_s h$ and using (2.8), we have that

$$|\mathbb{E}\{\mathcal{L}P_s h(\xi)\}| \leq e^{-s} R_{\xi}.$$

Integrating over s yields

$$(3.1) \quad \int_0^{\infty} |\mathbb{E}\{\mathcal{L}P_s h(\xi)\}| ds \leq R_{\xi}.$$

Combining (2.6) with (2.8) leads to $|\mathcal{L}P_s h(\xi)| \leq e^{-s}(\mathbf{L}(\mathbb{X}) + \xi(\mathbb{X}))$. Using this and the finiteness of $\mathbf{L}(\mathbb{X})$, we can apply the Fubini–Tonelli Theorem in the Stein equation (2.7) to conclude that

$$(3.2) \quad \mathbb{E}\{h(\zeta)\} - \mathbb{E}\{h(\xi)\} = \mathbb{E}\left\{\int_0^\infty \mathcal{L}P_s h(\xi) ds\right\} = \int_0^\infty \mathbb{E}\{\mathcal{L}P_s h(\xi)\} ds.$$

Thus, from definition of KR distance (1.2), (3.2) and (3.1) we have that

$$\mathbf{d}_{\mathbf{KR}}(\xi, \zeta) = \sup_{h \in \text{LIP}(\mathbb{X})} \left| \int_0^\infty \mathbb{E}\{\mathcal{L}P_s h(\xi)\} ds \right| \leq \sup_{h \in \text{LIP}(\mathbb{X})} \int_0^\infty |\mathbb{E}\{\mathcal{L}P_s h(\xi)\}| ds \leq R_\xi,$$

as required. □

Next, we want to apply the proof strategy of Theorem 3.1 in the specific case when ξ is generated by statistics of a Poisson process. Let $(\mathbb{X}, \mathcal{X})$ and $(\mathbb{Y}, \mathcal{Y})$ be lscH spaces and let $f : \mathbb{X}^k \rightarrow \mathbb{Y}$ be a symmetric measurable function. Let $g : \mathbb{X}^k \times \mathbf{N}_{\mathbb{X}} \rightarrow \{0, 1\}$ be another measurable function that is also symmetric with respect to the \mathbf{x} coordinates. For $\omega \in \mathbf{N}_{\mathbb{X}}$, define

$$(3.3) \quad \xi[\omega] := \frac{1}{k!} \sum_{\mathbf{x} \in \omega_{\neq}^k} g(\mathbf{x}, \omega) \delta_{f(\mathbf{x})}.$$

In other words, the function g is used as a “flag” indicating whether or not to include a point at $f(\mathbf{x})$ in our point process $\xi[\omega]$. We will study the case $\xi := \xi[\eta]$, where η is a Poisson process on \mathbb{X} with a σ -finite intensity measure \mathbf{K} . A simple consequence of the Mecke formula (2.3) is that the intensity measure \mathbf{L} of ξ can be expressed as

$$(3.4) \quad \mathbf{L}(dy) = \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{1}\{f(\mathbf{x}) \in dy\} p(\mathbf{x}) \mathbf{K}^k(d\mathbf{x}),$$

where $p(\mathbf{x}) := \mathbb{E}\{g(\mathbf{x}, \eta + \delta_{\mathbf{x}})\}$.

THEOREM 3.3. — *Let η be a Poisson process on \mathbb{X} with a σ -finite intensity measure \mathbf{K} . Let $\xi := \xi[\eta]$ be as defined in (3.3) and such that $\mathbf{L}(\mathbb{Y}) < \infty$. Suppose that for (almost) every $\mathbf{x} \in \mathbb{X}^k$ with $p(\mathbf{x}) > 0$, we have a pair of coupled point processes $\xi^{\mathbf{x}}$ and $\tilde{\xi}^{\mathbf{x}}$, such that $\xi^{\mathbf{x}} \stackrel{d}{=} \xi$, and $\tilde{\xi}^{\mathbf{x}} \stackrel{d}{=} (\xi[\eta + \delta_{\mathbf{x}}] - \delta_{f(\mathbf{x})}) | \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) = 1\}$ (i.e., the conditional distribution of $\xi[\eta + \delta_{\mathbf{x}}] - \delta_{f(\mathbf{x})}$ given $g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) = 1$). Assume also that $\mathbf{x} \mapsto \mathbb{E}\{(\xi^{\mathbf{x}} \Delta \tilde{\xi}^{\mathbf{x}})(\mathbb{Y})\}$ is measurable. Let ζ be a Poisson process on \mathbb{Y} with a finite intensity measure \mathbf{M} . Then,*

$$\mathbf{d}_{\mathbf{KR}}(\xi, \zeta) \leq d_{TV}(\mathbf{L}, \mathbf{M}) + \frac{2}{k!} \int_{\mathbb{X}^k} \mathbb{E}\{(\xi^{\mathbf{x}} \Delta \tilde{\xi}^{\mathbf{x}})(\mathbb{Y})\} p(\mathbf{x}) \mathbf{K}^k(d\mathbf{x}).$$

Remark 3.4. — Before proceeding to the proof, we show how the above theorem includes the results in [Pen18, Theorem 3.1] and [DST16, Theorem 3.1] as special cases.

- (1) If $g(\mathbf{x}, \omega) = g(\mathbf{x})$ we can set $\text{dom}(f) = \{\mathbf{x} : g(\mathbf{x}) = 1\}$ and then $p(\mathbf{x}) = \mathbb{1}\{\mathbf{x} \in \text{dom}(f)\}$. Further, there is a natural coupling where we take $\xi^{\mathbf{x}} = \xi$, $\tilde{\xi}^{\mathbf{x}} = \xi + \hat{\xi}[\mathbf{x}, \eta + \delta_{\mathbf{x}}]$, where $\hat{\xi}[\mathbf{x}, \eta + \delta_{\mathbf{x}}] := \xi[\eta + \delta_{\mathbf{x}}] - \delta_{f(\mathbf{x})} - \xi[\eta]$ (see [DST16,

(6.2)] for an explicit expression). By the assumptions on f and g we have that $\widehat{\xi}[\mathbf{x}, \eta + \delta_{\mathbf{x}}] \in \widehat{\mathbf{N}}_{\mathbb{Y}}$. In this case, we have that

$$(\xi^{\mathbf{x}} \Delta \widetilde{\xi}^{\mathbf{x}})(\mathbb{Y}) = \widehat{\xi}[\mathbf{x}, \eta + \delta_{\mathbf{x}}](\mathbb{Y}),$$

which is the same as the bound obtained for R_{η} in [DST16, page 2172]. Following the arguments there, we recover the Poisson case of [DST16, Theorem 3.1], which is one of the main theorems in that paper.

- (2) In many cases, the bound achieved using the coupling in Theorem 3.3 is very similar to the bound one often obtains when applying [Pen18, Theorem 3.1]. This bound is merely the Poisson approximation for $\xi(\mathbb{Y})$; see the proof of [IY20, Theorem 2.8], for example. In some of these examples, the bounds obtained for the approximation of the random *variable* $\xi(\mathbb{Y})$ can be extended to bounds for Poisson *process* approximation with a few additional calculations.
- (3) The above theorem suffices in order to extend the results of [Ott20] (see Theorem 3.3 therein) under the KR distance. However, in Theorem 4.1 we will present an extension to more general Poisson driven point processes.

Our proof combines ideas from the proof of Theorem 3.1 and those in the proof of [Pen18, Theorem 3.1], for the case of $\xi(\mathbb{Y})$. If we were to apply Theorem 3.1 directly, then we need a coupling of ξ and its Palm measure on the space \mathbb{Y} . This is not the coupling we have assumed in the above theorem.

Proof of Theorem 3.3. — Due to (2.5), it suffices to prove the statement for the case $\mathbf{L} \equiv \mathbf{M}$. Let $h \in \text{LIP}(\mathbb{Y})$. Recall that \mathcal{L} is the generator of the Glauber dynamics (2.6) as in the proof of Theorem 3.1. By (3.3), we obtain

$$\begin{aligned} \mathcal{L}h(\xi) &= \int_{\mathbb{Y}} D_y h(\xi) \mathbf{L}(dy) - \int_{\mathbb{Y}} D_y h(\xi - \delta_y) \xi(dy) \\ &= \int_{\mathbb{Y}} D_y h(\xi) \mathbf{L}(dy) - \frac{1}{k!} \sum_{\mathbf{x} \in \eta_{\neq}^k} g(\mathbf{x}, \eta) D_{f(\mathbf{x})} \left[h(\xi - \delta_{f(\mathbf{x})}) \right]. \end{aligned}$$

Using Fubini–Tonelli, the finiteness of $\mathbf{L}(\mathbb{Y})$, the multivariate Mecke formula (2.3) and the definition of \mathbf{L} in (3.4), we have that

$$\begin{aligned} (3.5) \quad \mathbb{E} \{ \mathcal{L}h(\xi) \} &= \int_{\mathbb{Y}} \mathbb{E} \{ D_y h(\xi) \} \mathbf{L}(dy) \\ &\quad - \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{E} \left\{ g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) D_{f(\mathbf{x})} \left[h(\xi[\eta + \delta_{\mathbf{x}}] - \delta_{f(\mathbf{x})}) \right] \right\} \mathbf{K}^k(d\mathbf{x}) \\ &= \frac{1}{k!} \int_{\mathbb{X}^k} \left(\mathbb{E} \{ D_{f(\mathbf{x})} [h(\xi)] \} \right. \\ &\quad \left. - \mathbb{E} \left\{ D_{f(\mathbf{x})} \left[h(\xi[\eta + \delta_{\mathbf{x}}] - \delta_{f(\mathbf{x})}) \right] \mid g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) = 1 \right\} \right) p(\mathbf{x}) \mathbf{K}^k(d\mathbf{x}). \end{aligned}$$

From the above identity, the existence of coupled point processes $\xi^{\mathbf{x}}, \widetilde{\xi}^{\mathbf{x}}$ and the Lipschitz property (2.1), we can derive that

$$\begin{aligned}
 |\mathbb{E}\{\mathcal{L}h(\xi)\}| &= \frac{1}{k!} \left| \int_{\mathbb{X}^k} \left(\mathbb{E}\{D_{f(\mathbf{x})}[h(\xi^{\mathbf{x}})]\} - \mathbb{E}\{D_{f(\mathbf{x})}[h(\tilde{\xi}^{\mathbf{x}})]\} \right) p(\mathbf{x}) \mathbf{K}^k(d\mathbf{x}) \right| \\
 &\leq \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{E}\{|h(\xi^{\mathbf{x}} + \delta_{f(\mathbf{x})}) - h(\tilde{\xi}^{\mathbf{x}} + \delta_{f(\mathbf{x})})|\} p(\mathbf{x}) \mathbf{K}^k(d\mathbf{x}) \\
 &\quad + \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{E}\{|h(\xi^{\mathbf{x}}) - h(\tilde{\xi}^{\mathbf{x}})|\} p(\mathbf{x}) \mathbf{K}^k(d\mathbf{x}) \\
 &\leq \frac{2}{k!} \int_{\mathbb{X}^k} \mathbb{E}\{(\xi^{\mathbf{x}} \Delta \tilde{\xi}^{\mathbf{x}})(\mathbb{Y})\} p(\mathbf{x}) \mathbf{K}^k(d\mathbf{x}).
 \end{aligned}$$

Now, we can follow the proof of Theorem 3.1 by substituting the above bound instead of R_ξ therein and complete the proof as in Theorem 3.1. \square

4. Stabilization for Poisson input

In this section, we consider the Poisson-driven point process $\xi = \xi[\eta]$, and extend the bound in Theorem 3.3 to the case where both f and g are locally dependent, i.e., may depend not only on \mathbf{x} but also on some (random) neighborhood around \mathbf{x} .

Recall that $(\mathbb{X}, \mathcal{X})$ and $(\mathbb{Y}, \mathcal{Y})$ are lscH spaces. Let $g : \mathbb{X}^k \times \mathbf{N}_{\mathbb{X}} \rightarrow \{0, 1\}$, $f : \mathbb{X}^k \times \mathbf{N}_{\mathbb{X}} \rightarrow \mathbb{Y}$ be measurable functions that are also symmetric in the \mathbf{x} coordinates. Let $\mathcal{F} := \mathcal{F}(\mathbb{X})$ be the space of all closed subsets of \mathbb{X} equipped with the smallest σ -algebra containing $\{F \in \mathcal{F} : F \cap K \neq \emptyset\}$ for all compact $K \subset \mathbb{X}$. This σ -algebra is the Borel σ -algebra generated by the Fell topology (see [BBK20, Chapter 9]). Let $\mathcal{S} : \mathbb{X}^k \times \mathbf{N}_{\mathbb{X}} \rightarrow \mathcal{F}$ be a measurable function. In this section we assume that f and g are *localized* to \mathcal{S} , i.e., for all $\omega \in \mathbf{N}_{\mathbb{X}}$, for all $\mathbf{x} \in \omega_{\neq}^k$ and for all $S \supset \mathcal{S}(\mathbf{x}, \omega)$, we have that

$$\begin{aligned}
 (4.1) \quad &g(\mathbf{x}, \omega) = g(\mathbf{x}, \omega \cap S), \quad \text{and} \\
 &f(\mathbf{x}, \omega) = f(\mathbf{x}, \omega \cap S), \quad \text{if } g(\mathbf{x}, \omega) = 1,
 \end{aligned}$$

where we use $\omega \cap S$ as a multiset. We also assume that $\mathcal{S}(\mathbf{x}, \omega)$ is a stopping set, i.e., for every compact set $S \subset \mathbb{X}$ we have that

$$(4.2) \quad \{\omega \in \mathbf{N}_{\mathbb{X}} : \mathcal{S}(\mathbf{x}, \omega) \subset S\} = \{\omega \in \mathbf{N}_{\mathbb{X}} : \mathcal{S}(\mathbf{x}, \omega \cap S) \subset S\}.$$

The construction of a measurable \mathcal{S} as above is not always obvious. Often, such an \mathcal{S} is constructed as a ball of random radius with the radius being called as *radius of stabilization*. The justification for the measurability of the radius of stabilization in the Euclidean case can be found in [Pen07, Definition 2.1 and below]. We also refer the reader to [Ott20, Section 2] for more information on stopping sets (see also [BL09, Appendix A] and [LPY21, Appendix A]). Finally, we point out that more general regions than balls can occur as “localizing regions” \mathcal{S} (see [BM22]).

Next, define

$$(4.3) \quad \xi[\omega] := \frac{1}{k!} \sum_{\mathbf{x} \in \omega_{\neq}^k} g(\mathbf{x}, \omega) \delta_{f(\mathbf{x}, \omega)},$$

and consider $\xi := \xi[\eta]$ where η is a Poisson process with a σ -finite intensity measure \mathbf{K} , as before. Using the multivariate Mecke formulae (2.3), we have that the intensity measure \mathbf{L} of ξ is given by

$$(4.4) \quad \mathbf{L}(dy) = \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{E} \{ \mathbb{1} \{ f(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \in dy \} g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \} \mathbf{K}^k(d\mathbf{x}).$$

For a set $A \subset \mathbb{X}$ and $\mathbf{x} = (x_1, \dots, x_k) \in \mathbb{X}^k$ we write $\mathbf{x} \subset A$ to denote $\{x_1, \dots, x_k\} \subset A$.

THEOREM 4.1. — *Let $\xi = \xi[\eta]$ be the point process defined as in (4.3) with $\mathbf{L}(\mathbb{Y}) < \infty$ and f, g and \mathcal{S} satisfying (4.1) and (4.2). Let ζ be a Poisson process with a finite intensity measure \mathbf{M} . Further, suppose that we are given a measurable mapping $\mathbf{x} \mapsto S_{\mathbf{x}}$ from \mathbb{X}^k to \mathcal{F} satisfying $\mathbf{x} \subset S_{\mathbf{x}}$. Define*

$$\tilde{g}(\mathbf{x}, \omega) := g(\mathbf{x}, \omega) \mathbb{1} \{ \mathcal{S}(\mathbf{x}, \omega) \subset S_{\mathbf{x}} \}$$

for $\omega \in \mathbf{N}_{\mathbb{X}}$, and

$$\xi_{\text{tr}} := \xi_{\text{tr}}[\eta] := \frac{1}{k!} \sum_{\mathbf{x} \in \eta_{\neq}^k} \tilde{g}(\mathbf{x}, \eta) \delta_{f(\mathbf{x}, \eta)}.$$

Then

$$\begin{aligned} \mathbf{d}_{\mathbf{KR}}(\xi, \zeta) &\leq d_{TV}(\mathbf{L}, \mathbf{M}) + 2(\text{Var}(\xi_{\text{tr}}(\mathbb{Y})) - \mathbb{E}\{\xi_{\text{tr}}(\mathbb{Y})\}) + E_1 + 2E_2 \\ &= d_{TV}(\mathbf{L}, \mathbf{M}) + E_1 + E_2 + E_3 + E_4 \end{aligned}$$

with

$$\begin{aligned} E_1 &:= \frac{2}{k!} \int_{\mathbb{X}^k} \mathbb{E} \{ g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \mathbb{1} \{ \mathcal{S}(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \not\subset S_{\mathbf{x}} \} \} \mathbf{K}^k(d\mathbf{x}), \\ E_2 &:= \frac{2}{(k!)^2} \int_{\mathbb{X}^k} \int_{\mathbb{X}^k} \mathbb{1} \{ S_{\mathbf{x}} \cap S_{\mathbf{z}} \neq \emptyset \} \mathbb{E} \{ \tilde{g}(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \} \mathbb{E} \{ \tilde{g}(\mathbf{z}, \eta + \delta_{\mathbf{z}}) \} \mathbf{K}^k(d\mathbf{z}) \mathbf{K}^k(d\mathbf{x}), \\ E_3 &:= \frac{2}{(k!)^2} \int_{\mathbb{X}^k} \int_{\mathbb{X}^k} \mathbb{1} \{ S_{\mathbf{x}} \cap S_{\mathbf{z}} \neq \emptyset \} \mathbb{E} \{ \tilde{g}(\mathbf{x}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}}) \tilde{g}(\mathbf{z}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}}) \} \\ &\quad \times \mathbf{K}^k(d\mathbf{z}) \mathbf{K}^k(d\mathbf{x}), \\ E_4 &:= \frac{2}{k!} \sum_{\emptyset \subsetneq I \subsetneq \{1, \dots, k\}} \frac{1}{(k - |I|)!} \int_{\mathbb{X}^k} \int_{\mathbb{X}^{k-|I|}} \mathbb{E} \{ \tilde{g}(\mathbf{x}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}}) \tilde{g}((\mathbf{x}_I, \mathbf{z}), \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}}) \} \\ &\quad \times \mathbf{K}^{k-|I|}(d\mathbf{z}) \mathbf{K}^k(d\mathbf{x}), \end{aligned}$$

where for $I = \{i_1, \dots, i_m\}$ we set $\mathbf{x}_I = (x_{i_1}, \dots, x_{i_m})$, and $(\mathbf{x}_I, \mathbf{z}) = (x_{i_1}, \dots, x_{i_m}, z_1, \dots, z_{k-m})$.

Note that the bound is similar to that in [Ott20, Theorem 3.3] while it holds for a more general class of point processes under the stronger KR distance and for a more general notion of “localizing region” than balls. The proof of [Ott20, Theorem 3.3] proceeds by constructing a coupling between ξ and $\xi_x^!$ and then using [BB92, Theorem 2.6] (the total variation distance analogue of our Theorem 3.1). The proof here is much simpler – we use the canonical truncation via $S_{\mathbf{x}}$ to reduce the proof to Poisson approximation bounds for truncated point processes with bounded range of dependence. We first prove the bound for the truncated point process and then

use this result to complete the proof of Theorem 4.1. Our proof techniques have been refined further in [Ott21, Theorem 8] and thereby leading to improved bounds. These improved bounds are crucial for some applications; see [Ott21].

Proof. — As before, we assume that $\mathbf{L} = \mathbf{M}$, and use (2.5) otherwise. The proof will proceed in two steps as described above.

Step 1. The case of bounded range of dependence. — We first assume that for all $\mathbf{x} \in \mathbb{X}^k$ and $\omega \in \mathbf{N}_{\mathbb{X}}$ we have that $\mathcal{S}(\mathbf{x}, \omega) \subset S_{\mathbf{x}}$. This implies that $\tilde{g} = g$, $\xi = \xi_{\text{tr}}$ and $E_1 = 0$. Moreover, (4.1) leads to

$$(4.5) \quad g(\mathbf{x}, \omega) = g(\mathbf{x}, \omega \cap S_{\mathbf{x}}) \quad \text{and} \quad \text{if} \quad g(\mathbf{x}, \omega) = 1, \quad f(\mathbf{x}, \omega) = f(\mathbf{x}, \omega \cap S_{\mathbf{x}}).$$

Fix $h \in \text{LIP}(\mathbb{Y})$. Similarly to the derivation in (3.5), we use (4.4) to obtain that

$$\begin{aligned} \mathbb{E} \{ \mathcal{L}h(\xi) \} &= \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{E} \left\{ g(\mathbf{x}, \eta' + \delta_{\mathbf{x}}) D_{f(\mathbf{x}, \eta' + \delta_{\mathbf{x}})} [h(\xi[\eta])] \right\} \mathbf{K}^k(d\mathbf{x}) \\ &\quad - \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{E} \left\{ g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) D_{f(\mathbf{x}, \eta + \delta_{\mathbf{x}})} \left[h \left(\xi[\eta + \delta_{\mathbf{x}}] - \delta_{f(\mathbf{x}, \eta + \delta_{\mathbf{x}})} \right) \right] \right\} \mathbf{K}^k(d\mathbf{x}), \end{aligned}$$

where η' is a Poisson process, independent of η and $\eta' \stackrel{d}{=} \eta$. Integrating over s as in (3.2), yields

$$(4.6) \quad \begin{aligned} &\mathbb{E} \{ h(\zeta) \} - \mathbb{E} \{ h(\xi) \} \\ &= \frac{1}{k!} \int_0^\infty \int_{\mathbb{X}^k} \left(\mathbb{E} \left\{ g(\mathbf{x}, \eta' + \delta_{\mathbf{x}}) D_{f(\mathbf{x}, \eta' + \delta_{\mathbf{x}})} [P_s h(\xi[\eta])] \right\} \right. \\ &\quad \left. - \mathbb{E} \left\{ g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) D_{f(\mathbf{x}, \eta + \delta_{\mathbf{x}})} \left[P_s h \left(\xi[\eta + \delta_{\mathbf{x}}] - \delta_{f(\mathbf{x}, \eta + \delta_{\mathbf{x}})} \right) \right] \right\} \right) \mathbf{K}^k(d\mathbf{x}) ds. \end{aligned}$$

For $\mathbf{x} \in \mathbb{X}^k$ define

$$(4.7) \quad \begin{aligned} \xi_{\mathbf{x}} &:= \xi_{\mathbf{x}}[\eta] \\ &= \frac{1}{k!} \sum_{\mathbf{z} \in \eta_{\neq}^k : S_{\mathbf{z}} \cap S_{\mathbf{x}} = \emptyset} g(\mathbf{z}, \eta) \delta_{f(\mathbf{z}, \eta)} \\ &= \frac{1}{k!} \sum_{\mathbf{z} \in \eta_{\neq}^k : S_{\mathbf{z}} \cap S_{\mathbf{x}} = \emptyset} g(\mathbf{z}, \eta + \delta_{\mathbf{x}}) \delta_{f(\mathbf{z}, \eta + \delta_{\mathbf{x}})} \\ &= \frac{1}{k!} \sum_{\mathbf{z} \in (\eta \cap S_{\mathbf{x}}^c)_{\neq}^k : S_{\mathbf{z}} \cap S_{\mathbf{x}} = \emptyset} g(\mathbf{z}, \eta \cap S_{\mathbf{x}}^c) \delta_{f(\mathbf{z}, \eta \cap S_{\mathbf{x}}^c)}, \end{aligned}$$

where A^c stands for the complement of A . The second and third equalities follow from the assumption (4.5) on f and g , and the fact that $\mathbf{z} \subset S_{\mathbf{z}} \subset S_{\mathbf{x}}^c$. Now, it follows from (2.8) that

$$\begin{aligned}
 (4.8) \quad & \left| \mathbb{E} \left\{ g(\mathbf{x}, \eta' + \delta_{\mathbf{x}}) D_{f(\mathbf{x}, \eta' + \delta_{\mathbf{x}})} [P_s h(\xi[\eta])] \right\} \right. \\
 & \quad \left. - \mathbb{E} \left\{ g(\mathbf{x}, \eta' + \delta_{\mathbf{x}}) D_{f(\mathbf{x}, \eta' + \delta_{\mathbf{x}})} [P_s h(\xi_{\mathbf{x}}[\eta])] \right\} \right| \\
 & \leq \frac{2}{k!} e^{-s} \mathbb{E} \left\{ g(\mathbf{x}, \eta' + \delta_{\mathbf{x}}) \sum_{\mathbf{z} \in \eta_{\neq}^k : S_{\mathbf{z}} \cap S_{\mathbf{x}} \neq \emptyset} g(\mathbf{z}, \eta) \right\} \\
 & \leq \frac{2}{k!} e^{-s} \int_{\mathbb{X}^k} \mathbb{1} \{S_{\mathbf{z}} \cap S_{\mathbf{x}} \neq \emptyset\} \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}})\} \mathbb{E} \{g(\mathbf{z}, \eta + \delta_{\mathbf{z}})\} \mathbf{K}^k(d\mathbf{z}).
 \end{aligned}$$

Observe that by definition (4.7) we have $\xi_{\mathbf{x}}[\eta] \subset \xi[\eta + \delta_{\mathbf{x}}] - g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \delta_{f(\mathbf{x}, \eta + \delta_{\mathbf{x}})}$ as multisets. Thus, we can define the point process

$$\widehat{\xi}_{\mathbf{x}} := \widehat{\xi}_{\mathbf{x}}[\eta] := \xi[\eta + \delta_{\mathbf{x}}] - g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \delta_{f(\mathbf{x}, \eta + \delta_{\mathbf{x}})} - \frac{1}{k!} \sum_{\mathbf{z} \in \eta_{\neq}^k} g(\mathbf{z}, \eta + \delta_{\mathbf{x}}) \delta_{f(\mathbf{z}, \eta + \delta_{\mathbf{x}})}.$$

In other words, $\widehat{\xi}_{\mathbf{x}}[\eta]$ contains the points of $\xi[\eta + \delta_{\mathbf{x}}]$ that are generated by nonempty strict subsets of \mathbf{x} . Using (2.8) and the Mecke formula, we derive that

$$\begin{aligned}
 (4.9) \quad & \left| \mathbb{E} \left\{ g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) D_{f(\mathbf{x}, \eta + \delta_{\mathbf{x}})} [P_s h(\xi[\eta + \delta_{\mathbf{x}}] - \delta_{f(\mathbf{x}, \eta + \delta_{\mathbf{x}})})] \right\} \right. \\
 & \quad \left. - \mathbb{E} \left\{ g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) D_{f(\mathbf{x}, \eta + \delta_{\mathbf{x}})} [P_s h(\xi_{\mathbf{x}}[\eta])] \right\} \right| \\
 & \leq 2e^{-s} \left(\mathbb{E} \left\{ g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \frac{1}{k!} \sum_{\mathbf{z} \in \eta_{\neq}^k : S_{\mathbf{z}} \cap S_{\mathbf{x}} \neq \emptyset} g(\mathbf{z}, \eta + \delta_{\mathbf{x}}) \right\} \right. \\
 & \quad \left. + \mathbb{E} \left\{ g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \widehat{\xi}_{\mathbf{x}}[\eta](\mathbb{Y}) \right\} \right) \\
 & \leq \frac{2}{k!} e^{-s} \int_{\mathbb{X}^k} \mathbb{1} \{S_{\mathbf{z}} \cap S_{\mathbf{x}} \neq \emptyset\} \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}}) g(\mathbf{z}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}})\} \mathbf{K}^k(d\mathbf{z}) \\
 & \quad + 2e^{-s} \mathbb{E} \left\{ g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \widehat{\xi}_{\mathbf{x}}[\eta](\mathbb{Y}) \right\}.
 \end{aligned}$$

By assumption (4.5), $f(\mathbf{x}, \eta + \delta_{\mathbf{x}})$ and $g(\mathbf{x}, \eta + \delta_{\mathbf{x}})$ depend only on $\eta \cap S_{\mathbf{x}}$, while $\xi_{\mathbf{x}}$ is a functional of $\eta \cap S_{\mathbf{x}}^c$ (see (4.7)). Together with the independence property of Poisson processes, we obtain that

$$\begin{aligned}
 (4.10) \quad & \mathbb{E} \left\{ g(\mathbf{x}, \eta' + \delta_{\mathbf{x}}) D_{f(\mathbf{x}, \eta' + \delta_{\mathbf{x}})} P_s h(\xi_{\mathbf{x}}[\eta]) \right\} \\
 & = \mathbb{E} \left\{ g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) D_{f(\mathbf{x}, \eta + \delta_{\mathbf{x}})} P_s h(\xi_{\mathbf{x}}[\eta]) \right\}.
 \end{aligned}$$

Substituting (4.8), (4.9) and (4.10) into (4.6), using the triangle inequality and integrating over s , yields

$$\begin{aligned}
 & \left| \mathbb{E} \{h(\zeta)\} - \mathbb{E} \{h(\xi)\} \right| \\
 & \leq \frac{2}{(k!)^2} \int_{\mathbb{X}^k} \int_{\mathbb{X}^k} \mathbb{1} \{S_{\mathbf{z}} \cap S_{\mathbf{x}} \neq \emptyset\} \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}})\} \mathbb{E} \{g(\mathbf{z}, \eta + \delta_{\mathbf{z}})\} \\
 & \quad \times \mathbf{K}^k(d\mathbf{z})\mathbf{K}^k(d\mathbf{x}) \\
 & + \frac{2}{(k!)^2} \int_{\mathbb{X}^k} \int_{\mathbb{X}^k} \mathbb{1} \{S_{\mathbf{z}} \cap S_{\mathbf{x}} \neq \emptyset\} \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}})g(\mathbf{z}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}})\} \\
 & \quad \times \mathbf{K}^k(d\mathbf{z})\mathbf{K}^k(d\mathbf{x}) \\
 & + \frac{2}{k!} \int_{\mathbb{X}^k} \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \widehat{\xi}_{\mathbf{x}}[\eta](\mathbb{Y})\} \mathbf{K}^k(d\mathbf{x}).
 \end{aligned}$$

Here, the first and the second term on the right-hand side are E_2 and E_3 , respectively. Notice that

$$\begin{aligned}
 & \frac{2}{k!} \int_{\mathbb{X}^k} \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \widehat{\xi}_{\mathbf{x}}[\eta](\mathbb{Y})\} \mathbf{K}^k(d\mathbf{x}) \\
 & = \frac{2}{k!} \sum_{\emptyset \subsetneq I \subsetneq \{1, \dots, k\}} \frac{1}{(k - |I|)!} \int_{\mathbb{X}^k} \int_{\mathbb{X}^{k-|I|}} \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}}) g((\mathbf{x}_I, \mathbf{z}), \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}})\} \\
 & \quad \times \mathbf{K}^{k-|I|}(d\mathbf{z})\mathbf{K}^k(d\mathbf{x}),
 \end{aligned}$$

which is equal to E_4 . Together with $E_1 = 0$, this proves the second bound in the theorem. A short computation shows that

$$\begin{aligned}
 E_4 & = 2\mathbb{E} \{\xi(\mathbb{Y})^2\} - 2\mathbb{E} \{\xi(\mathbb{Y})\} \\
 & \quad - 2 \left(\frac{1}{k!}\right)^2 \int_{\mathbb{X}^k} \int_{\mathbb{X}^k} \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}}) g(\mathbf{z}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}})\} \mathbf{K}^k(d\mathbf{z})\mathbf{K}^k(d\mathbf{x}).
 \end{aligned}$$

From (4.5) and the independence property of η it follows that

$$\begin{aligned}
 & 2 \left(\frac{1}{k!}\right)^2 \int_{\mathbb{X}^k} \int_{\mathbb{X}^k} \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}}) g(\mathbf{z}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}})\} \mathbf{K}^k(d\mathbf{z})\mathbf{K}^k(d\mathbf{x}) \\
 & = 2\mathbb{E} \{\xi(\mathbb{Y})\}^2 + 2 \left(\frac{1}{k!}\right)^2 \int_{\mathbb{X}^k} \int_{\mathbb{X}^k} \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}}) g(\mathbf{z}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}})\} \\
 & \quad - \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}})\} \mathbb{E} \{g(\mathbf{z}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}})\} \mathbf{K}^k(d\mathbf{z})\mathbf{K}^k(d\mathbf{x}) \\
 & = 2\mathbb{E} \{\xi(\mathbb{Y})\}^2 + 2 \left(\frac{1}{k!}\right)^2 \int_{\mathbb{X}^k} \int_{\mathbb{X}^k} \mathbb{1} \{S_{\mathbf{x}} \cap S_{\mathbf{y}} \neq \emptyset\} \\
 & \quad \times \left(\mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}}) g(\mathbf{z}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}})\} \right. \\
 & \quad \left. - \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}})\} \mathbb{E} \{g(\mathbf{z}, \eta + \delta_{\mathbf{x}} + \delta_{\mathbf{z}})\} \right) \mathbf{K}^k(d\mathbf{z})\mathbf{K}^k(d\mathbf{x}) \\
 & = 2\mathbb{E} \{\xi(\mathbb{Y})\}^2 + E_3 - E_2.
 \end{aligned}$$

Combining the previous identities yields that

$$E_4 = 2(\text{Var}(\xi(\mathbb{Y})) - \mathbb{E}\{\xi(\mathbb{Y})\}) + E_2 - E_3,$$

which proves the first bound in the Theorem 4.1.

Step 2. The general case. — We will use Step 1 to complete the proof of the general result in the Theorem 4.1.

We denote by \mathbf{L}_{tr} the intensity measure of the truncated point process ξ_{tr} , which was defined in the statement of the theorem. It follows from the multivariate Mecke equation (2.3) that

$$\begin{aligned} & \mathbf{L}_{\text{tr}}(\mathrm{d}y) \\ &= \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \mathbb{1} \{ \mathcal{S}(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \subset S_{\mathbf{x}} \} \mathbb{1} \{ f(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \in \mathrm{d}y \} \} \mathbf{K}^k(\mathrm{d}\mathbf{x}). \end{aligned}$$

Let ζ_{tr} be a Poisson process with intensity measure \mathbf{L}_{tr} . From the triangle inequality for the KR distance along with (2.1) and (2.4) we obtain

$$\begin{aligned} \mathbf{d}_{\mathbf{KR}}(\xi, \zeta) &\leq \mathbf{d}_{\mathbf{KR}}(\xi, \xi_{\text{tr}}) + \mathbf{d}_{\mathbf{KR}}(\xi_{\text{tr}}, \zeta_{\text{tr}}) + \mathbf{d}_{\mathbf{KR}}(\zeta, \zeta_{\text{tr}}) \\ &\leq \mathbb{E} \{d_{TV}(\xi, \xi_{\text{tr}})\} + \mathbf{d}_{\mathbf{KR}}(\xi_{\text{tr}}, \zeta_{\text{tr}}) + d_{TV}(\mathbf{L}, \mathbf{L}_{\text{tr}}). \end{aligned}$$

Since $\mathbf{L}_{\text{tr}}(A) \leq \mathbf{L}(A)$ and $\xi_{\text{tr}}(A) \leq \xi(A)$ for all $A \in \mathcal{Y}$, and using (4.4), we have

$$\begin{aligned} d_{TV}(\mathbf{L}, \mathbf{L}_{\text{tr}}) &= \mathbf{L}(\mathbb{Y}) - \mathbf{L}_{\text{tr}}(\mathbb{Y}) \\ &= \frac{1}{k!} \int_{\mathbb{X}^k} \mathbb{E} \{g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \mathbb{1} \{ \mathcal{S}(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \not\subset S_{\mathbf{x}} \} \} \mathbf{K}^k(\mathrm{d}\mathbf{x}) \\ &= \frac{E_1}{2}, \end{aligned}$$

and

$$\mathbb{E} \{d_{TV}(\xi, \xi_{\text{tr}})\} = \mathbb{E} \{ \xi(\mathbb{Y}) - \xi_{\text{tr}}(\mathbb{Y}) \} = \mathbf{L}(\mathbb{Y}) - \mathbf{L}_{\text{tr}}(\mathbb{Y}) = \frac{E_1}{2}.$$

Thus, we have shown that

$$(4.11) \quad \mathbf{d}_{\mathbf{KR}}(\xi, \zeta) \leq \mathbf{d}_{\mathbf{KR}}(\xi_{\text{tr}}, \zeta_{\text{tr}}) + E_1.$$

Recall that $\tilde{g}(\mathbf{x}, \eta + \delta_{\mathbf{x}}) := g(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \mathbb{1} \{ \mathcal{S}(\mathbf{x}, \eta + \delta_{\mathbf{x}}) \subset S_{\mathbf{x}} \}$, and observe that by the stopping set property of \mathcal{S} (4.2), both f and \tilde{g} satisfy the local dependence assumptions in (4.5). Hence, using Step 1 of the proof, we can derive the necessary bounds for $\mathbf{d}_{\mathbf{KR}}(\xi_{\text{tr}}, \zeta_{\text{tr}})$ in (4.11) and complete the proof. \square

5. Stabilization for binomial input

In this section we focus on the binomial point process β_n (replacing the Poisson process), i.e., a point process consisting of n independent points, distributed according to some probability measure \mathbf{Q} . In this case, we have that for any measurable function $h : \mathbb{X}^k \times \mathbf{N}_{\mathbb{X}} \rightarrow \mathbb{R}_+$ and $n \geq k$,

$$(5.1) \quad \mathbb{E} \left\{ \sum_{\mathbf{x} \in \beta_{n, \neq}^k} h(\mathbf{x}, \beta_n) \right\} = (n)_k \int_{\mathbb{X}^k} \mathbb{E} \{ h(\mathbf{x}, \beta_{n-k} + \delta_{\mathbf{x}}) \} \mathbf{Q}^k(\mathrm{d}\mathbf{x}),$$

with $(n)_k := n \cdot \dots \cdot (n - k + 1)$. Note that (5.1) is the analogue for binomial point processes of the multivariate Mecke formula (2.3).

Let β_n be a binomial point process, and let

$$\xi := \xi[\beta_n] := \sum_{x \in \beta_n} g(x, \beta_n) \delta_{f(x, \beta_n)}.$$

Here, g and f are as in Section 4 for an underlying Poisson process, but we allow only $k = 1$ in the sequel. By (5.1), the point process ξ has an intensity measure

$$(5.2) \quad \mathbf{L}(dy) = n \int_{\mathbb{X}} \mathbb{E} \{g(x, \beta_{n-1} + \delta_x) \mathbb{1} \{f(x, \beta_{n-1} + \delta_x) \in dy\}\} \mathbf{Q}(dx).$$

THEOREM 5.1. — *Let ξ and \mathbf{L} be as above with f, g and a measurable function $\mathcal{S} : \mathbb{X} \times \mathbf{N}_{\mathbb{X}} \rightarrow \mathcal{F}$ satisfying (4.1) and (4.2) and let ζ be a Poisson process with a finite intensity measure \mathbf{M} . Further, suppose that we are given a measurable mapping $x \mapsto S_x$ from \mathbb{X} to \mathcal{F} satisfying $x \in S_x$ and $\mathbf{Q}(S_x) < 1$ for all $x \in \mathbb{X}$. For $\omega \in \mathbf{N}_{\mathbb{X}}$ and $x \in \omega$, define*

$$\tilde{g}(x, \omega) := g(x, \omega) \mathbb{1} \{\mathcal{S}(x, \omega) \subset S_x\}.$$

Then

$$\mathbf{d}_{\mathbf{KR}}(\xi, \zeta) \leq d_{TV}(\mathbf{L}, \mathbf{M}) + E_1 + E_2 + E_3 + E_4 + E_5 + E_6,$$

where

$$\begin{aligned} E_1 &:= 2n \int_{\mathbb{X}} \mathbb{E} \{g(x, \beta_{n-1} + \delta_x) \mathbb{1} \{\mathcal{S}(x, \beta_{n-1} + \delta_x) \not\subset S_x\}\} \mathbf{Q}(dx), \\ E_2 &:= 2n^2 \int_{\mathbb{X}^2} \mathbb{1} \{S_x \cap S_y \neq \emptyset\} \mathbb{E} \{\tilde{g}(x, \beta_{n-1} + \delta_x)\} \mathbb{E} \{\tilde{g}(y, \beta_{n-1} + \delta_y)\} \mathbf{Q}^2(d(x, y)), \\ E_3 &:= 2n^2 \int_{\mathbb{X}^2} \mathbb{1} \{S_x \cap S_y \neq \emptyset\} \\ &\quad \times \mathbb{E} \{\tilde{g}(x, \beta_{n-2} + \delta_x + \delta_y) \tilde{g}(y, \beta_{n-2} + \delta_x + \delta_y)\} \mathbf{Q}^2(d(x, y)), \\ E_4 &:= 2n \int_{\mathbb{X}} \left((1 + n\mathbf{Q}(S_x)) \mathbb{E} \{\tilde{g}(x, \beta_{n-1} + \delta_x)\} \right. \\ &\quad \left. + n \int_{S_x} \mathbb{E} \{\tilde{g}(x, \beta_{n-2} + \delta_x + \delta_z)\} \mathbf{Q}(dz) \right) \\ &\quad \times \left(\int_{\mathbb{X}} \mathbb{1} \{S_x \cap S_y = \emptyset\} \mathbb{E} \{\tilde{g}(y, \tilde{\beta}_{x, n-1} + \delta_y)\} \mathbf{Q}_x(dy) \right. \\ &\quad \left. + n \int_{\mathbb{X}^2} \mathbb{1} \{S_x \cap S_{y_1} = \emptyset, y_2 \in S_{y_1}\} \right. \\ &\quad \left. \times \mathbb{E} \{\tilde{g}(y_1, \tilde{\beta}_{x, n-2} + \delta_{y_1} + \delta_{y_2})\} \mathbf{Q}_x^2(d(y_1, y_2)) \right) \mathbf{Q}(dx), \\ E_5 &:= 2n^3 \int_{\mathbb{X}^2} \mathbb{1} \{S_x \cap S_y = \emptyset\} \mathbb{E} \{\tilde{g}(x, \beta_{n-1} + \delta_x)\} \mathbb{E} \{\tilde{g}(y, \beta_{n-2} + \delta_y)\} \\ &\quad \times \mathbf{Q}(S_x) \mathbf{Q}_x(S_y) \mathbf{Q}^2(d(x, y)), \end{aligned}$$

$$\begin{aligned}
 E_6 &:= 2n^2 \int_{\mathbb{X}^2} \mathbb{1} \{S_x \cap S_y = \emptyset\} \mathbb{E} \{ \tilde{g}(x, \beta_{n-2} + \delta_x) \tilde{g}(y, \beta_{n-2} + \delta_y) \} \\
 &\quad \times \mathbf{Q}_x(S_y) \mathbf{Q}^2(d(x, y)) \\
 &\quad + 2n^3 \int_{\mathbb{X}^2} \int_{S_x} \mathbb{1} \{S_x \cap S_y = \emptyset\} \mathbb{E} \{ \tilde{g}(x, \beta_{n-3} + \delta_x + \delta_z) \tilde{g}(y, \beta_{n-3} + \delta_y) \} \\
 &\quad \times \mathbf{Q}_x(S_y) \mathbf{Q}(dz) \mathbf{Q}^2(d(x, y)),
 \end{aligned}$$

and where $\tilde{\beta}_{x,m}$ is a binomial point process of m independent points distributed according to $\mathbf{Q}_x(\cdot) := \mathbf{Q}(\cdot \cap S_x^c) / \mathbf{Q}(S_x^c)$.

Remark 5.2. — A few remarks are in place regarding the above theorem in relation to the approximation results in the previous sections.

- (1) The binomial point process does not possess the same spatial independence as the Poisson process and hence the bounds are more complicated here. However, we expect that in many applications, the computation of the bounds should not be more difficult than in the Poisson case.
- (2) Since the bounds are already quite involved, we have restricted ourselves to only the case $k = 1$, i.e., to functions f, g that take $x \in \mathbb{X}$ as an input instead of $\mathbf{x} \in \mathbb{X}^k$.
- (3) Our proof shall follow the same strategy as in Theorem 4.1 but accounting for the added complication due to the lack of spatial independence in the binomial point process. In the special case where $f(x, \omega) = f(x)$ (as in Theorem 3.3) we can use a coupling approach as in the proof of Theorem 3.3, but this shall not significantly simplify the proof or the bounds in the above theorem.

Proof. — By (2.5), we can assume as before that $\mathbf{L} = \mathbf{M}$. Further, we shall prove the theorem under the assumption that $\tilde{g} \equiv g$ and follow the arguments as in Step 2 in the proof of Theorem 4.1, to complete the proof of the general case. Observe that in the case of $\tilde{g} \equiv g$ we have $E_1 = 0$.

Assume from here onwards that $\tilde{g} \equiv g$, i.e., $\mathcal{S}(x, \omega) \subset S_x$ for all $x \in \mathbb{X}$ and $\omega \in \mathbf{N}_{\mathbb{X}}$. For a Lipschitz function h , we have by (2.6) and (2.7) that

$$\begin{aligned}
 &\mathbb{E} \{h(\zeta)\} - \mathbb{E} \{h(\xi)\} \\
 &= \int_0^\infty \int_{\mathbb{Y}} \mathbb{E} \{D_y P_s h(\xi)\} \mathbf{L}(dy) ds - \int_0^\infty \mathbb{E} \left\{ \sum_{y \in \xi} D_y P_s h(\xi - \delta_y) \right\} ds.
 \end{aligned}$$

Using (5.1) we have

$$\begin{aligned}
 &\int_0^\infty \mathbb{E} \left\{ \sum_{y \in \xi} D_y P_s h(\xi - \delta_y) \right\} ds \\
 &= \int_0^\infty \mathbb{E} \left\{ \sum_{x \in \beta_n} g(x, \beta_n) D_{f(x, \beta_n)} P_s h(\xi - \delta_{f(x, \beta_n)}) \right\} ds \\
 &= n \int_0^\infty \int_{\mathbb{X}} \mathbb{E} \left\{ g(x, \beta_{n-1} + \delta_x) D_{f(x, \beta_{n-1} + \delta_x)} P_s h(\xi [\beta_{n-1} + \delta_x] - \delta_{f(x, \beta_{n-1} + \delta_x)}) \right\} \\
 &\quad \times \mathbf{Q}(dx) ds.
 \end{aligned}$$

Denoting by β'_m an independent copy of β_m , we obtain from (5.2) that

$$\begin{aligned} & \int_0^\infty \int_{\mathbb{Y}} \mathbb{E} \{ D_y P_s h(\xi) \} \mathbf{L}(dy) ds \\ &= n \int_0^\infty \int_{\mathbb{X}} \mathbb{E} \left\{ g \left(x, \beta'_{n-1} + \delta_x \right) D_{f(x, \beta'_{n-1} + \delta_x)} P_s h \left(\xi [\beta_n] \right) \right\} \mathbf{Q}(dx) ds. \end{aligned}$$

Combining the three previous identities leads to

$$\begin{aligned} & \mathbb{E} \{ h(\zeta) \} - \mathbb{E} \{ h(\xi) \} \\ &= n \int_0^\infty \int_{\mathbb{X}} \mathbb{E} \left\{ g \left(x, \beta'_{n-1} + \delta_x \right) D_{f(x, \beta'_{n-1} + \delta_x)} P_s h \left(\xi [\beta_n] \right) \right\} \\ & \quad - \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) D_{f(x, \beta_{n-1} + \delta_x)} P_s h \left(\xi [\beta_{n-1} + \delta_x] - \delta_{f(x, \beta_{n-1} + \delta_x)} \right) \right\} \mathbf{Q}(dx) ds. \end{aligned}$$

For $x \in \mathbb{X}$ and $\omega \in \mathbf{N}$ we define

$$\xi_x[\omega] = \sum_{y \in \omega, S_x \cap S_y = \emptyset} g(y, \omega) \delta_{f(y, \omega)}.$$

For $x \in \mathbb{X}$ and $s \geq 0$ let

$$\begin{aligned} T_{1,x,s} &:= \left| \mathbb{E} \left\{ g \left(x, \beta'_{n-1} + \delta_x \right) D_{f(x, \beta'_{n-1} + \delta_x)} P_s h \left(\xi [\beta_n] \right) \right\} \right. \\ & \quad \left. - \mathbb{E} \left\{ g \left(x, \beta'_{n-1} + \delta_x \right) D_{f(x, \beta'_{n-1} + \delta_x)} P_s h \left(\xi_x [\beta_n] \right) \right\} \right|, \\ T_{2,x,s} &:= \left| \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) D_{f(x, \beta_{n-1} + \delta_x)} P_s h \left(\xi [\beta_{n-1} + \delta_x] - \delta_{f(x, \beta_{n-1} + \delta_x)} \right) \right\} \right. \\ & \quad \left. - \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) D_{f(x, \beta_{n-1} + \delta_x)} P_s h \left(\xi_x [\beta_{n-1}] \right) \right\} \right|, \\ T_{3,x,s} &= \left| \mathbb{E} \left\{ g \left(x, \beta'_{n-1} + \delta_x \right) D_{f(x, \beta'_{n-1} + \delta_x)} P_s h \left(\xi_x [\beta_n] \right) \right\} \right. \\ & \quad \left. - \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) D_{f(x, \beta_{n-1} + \delta_x)} P_s h \left(\xi_x [\beta_{n-1}] \right) \right\} \right| \end{aligned}$$

so that by the triangle inequality

$$(5.3) \quad \left| \mathbb{E} \{ h(\zeta) \} - \mathbb{E} \{ h(\xi) \} \right| \leq n \int_0^\infty \int_{\mathbb{X}} (T_{1,x,s} + T_{2,x,s} + T_{3,x,s}) \mathbf{Q}(dx) ds.$$

Fix $x \in \mathbb{X}$ and $s \geq 0$. It follows from (2.8) that

$$\begin{aligned} T_{1,x,s} &\leq 2e^{-s} \mathbb{E} \left\{ g \left(x, \beta'_{n-1} + \delta_x \right) d_{TV} \left(\xi [\beta_n], \xi_x [\beta_n] \right) \right\} \\ &= 2e^{-s} \mathbb{E} \left\{ g \left(x, \beta'_{n-1} + \delta_x \right) \right\} \mathbb{E} \left\{ d_{TV} \left(\xi [\beta_n], \xi_x [\beta_n] \right) \right\}, \end{aligned}$$

and

$$T_{2,x,s} \leq 2e^{-s} \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) d_{TV} \left(\xi [\beta_{n-1} + \delta_x] - \delta_{f(x, \beta_{n-1} + \delta_x)}, \xi_x [\beta_{n-1}] \right) \right\}.$$

By (5.1) we have

$$\begin{aligned} \mathbb{E} \left\{ d_{TV} \left(\xi [\beta_n], \xi_x [\beta_n] \right) \right\} &\leq \mathbb{E} \left\{ \sum_{y \in \beta_n, S_x \cap S_y \neq \emptyset} g(y, \beta_n) \right\} \\ &= n \int_{\mathbb{X}} \mathbf{1} \{ S_x \cap S_y \neq \emptyset \} \mathbb{E} \left\{ g(y, \beta_{n-1} + \delta_y) \right\} \mathbf{Q}(dy), \end{aligned}$$

so that

$$T_{1,x,s} \leq 2e^{-s} \mathbb{E} \left\{ g \left(x, \beta'_{n-1} + \delta_x \right) \right\} n \int_{\mathbb{X}} \mathbb{1} \{ S_x \cap S_y \neq \emptyset \} \mathbb{E} \{ g \left(y, \beta_{n-1} + \delta_y \right) \} \mathbf{Q}(dy).$$

The inequality

$$\begin{aligned} g \left(x, \beta_{n-1} + \delta_x \right) d_{TV} \left(\xi \left[\beta_{n-1} + \delta_x \right] - \delta_{f \left(x, \beta_{n-1} + \delta_x \right)}, \xi_x \left[\beta_{n-1} \right] \right) \\ \leq g \left(x, \beta_{n-1} + \delta_x \right) \sum_{y \in \beta_{n-1}, S_x \cap S_y \neq \emptyset} g \left(y, \beta_{n-1} + \delta_x \right), \end{aligned}$$

where we used that by our assumptions $g \left(y, \beta_{n-1} + \delta_x \right) = g \left(y, \beta_{n-1} \right)$ if $S_x \cap S_y = \emptyset$, and (5.1) lead to

$$\begin{aligned} T_{2,x,s} \\ \leq 2e^{-s} n \int_{\mathbb{X}} \mathbb{1} \{ S_x \cap S_y \neq \emptyset \} \mathbb{E} \{ g \left(x, \beta_{n-2} + \delta_x + \delta_y \right) g \left(y, \beta_{n-2} + \delta_x + \delta_y \right) \} \mathbf{Q}(dy). \end{aligned}$$

Thus, we have shown that

$$(5.4) \quad n \int_0^\infty \int_{\mathbb{X}} T_{1,x,s} \mathbf{Q}(dx) ds \leq E_2 \quad \text{and} \quad n \int_0^\infty \int_{\mathbb{X}} T_{2,x,s} \mathbf{Q}(dx) ds \leq E_3.$$

Next, we define $N_m(x) = \beta_m(S_x)$ and $N'_m(x) = \beta'_m(S_x)$. Recall from the statement of the theorem that $\tilde{\beta}_{x,m}$ is a binomial point process of m independent points distributed according to $\mathbf{Q}_x(\cdot) = \mathbf{Q}(\cdot \cap S_x^c) / \mathbf{Q}(S_x^c)$, which we assume to be independent from $(\beta_\ell)_{\ell \in \mathbb{N}}$ and $(\beta'_\ell)_{\ell \in \mathbb{N}}$. From (2.8) we obtain

$$\begin{aligned} T_{3,x,s} &= \left| \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) D_{f \left(x, \beta_{n-1} + \delta_x \right)} P_s h \left(\xi_x \left[\beta'_n \right] \right) \right\} \right. \\ &\quad \left. - \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) D_{f \left(x, \beta_{n-1} + \delta_x \right)} P_s h \left(\xi_x \left[\beta_{n-1} \right] \right) \right\} \right| \\ &= \left| \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) D_{f \left(x, \beta_{n-1} + \delta_x \right)} P_s h \left(\xi_x \left[\tilde{\beta}_{x,n-N'_n(x)} \right] \right) \right\} \right. \\ &\quad \left. - \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) D_{f \left(x, \beta_{n-1} + \delta_x \right)} P_s h \left(\xi_x \left[\tilde{\beta}_{x,n-1-N_{n-1}(x)} \right] \right) \right\} \right| \\ &\leq 2e^{-s} \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) \left(\xi_x \left[\tilde{\beta}_{x,n-N'_n(x)} \right] \Delta \xi_x \left[\tilde{\beta}_{x,n-1-N_{n-1}(x)} \right] \right) (\mathbb{Y}) \right\}. \end{aligned}$$

We denote the points of $\tilde{\beta}_{x,m}$, by $\tilde{X}_1, \dots, \tilde{X}_m$. For $\ell \leq n$ we have

$$\begin{aligned} (5.5) \quad &\left(\xi_x \left[\tilde{\beta}_{x,n-\ell} \right] \Delta \xi_x \left[\tilde{\beta}_{x,n} \right] \right) (\mathbb{Y}) \\ &\leq \sum_{i=n-\ell+1}^n \mathbb{1} \{ S_x \cap S_{\tilde{X}_i} = \emptyset \} g \left(\tilde{X}_i, \tilde{\beta}_{x,n} \right) \\ &\quad + \sum_{i=n-\ell+1}^n \sum_{j=1}^{n-\ell} \mathbb{1} \{ S_x \cap S_{\tilde{X}_j} = \emptyset \} g \left(\tilde{X}_j, \tilde{\beta}_{x,n-\ell} \right) \mathbb{1} \{ \tilde{X}_i \in S_{\tilde{X}_j} \} \\ &\quad + \sum_{i=n-\ell+1}^n \sum_{j=1}^{n-\ell} \mathbb{1} \{ S_x \cap S_{\tilde{X}_j} = \emptyset \} g \left(\tilde{X}_j, \tilde{\beta}_{x,n} \right) \mathbb{1} \{ \tilde{X}_i \in S_{\tilde{X}_j} \}. \end{aligned}$$

Here, the first sum on the right-hand side counts the points of $\xi_x[\tilde{\beta}_{n,x}]$ associated with points from $\tilde{\beta}_{x,n} \setminus \tilde{\beta}_{x,n-\ell}$, while the second and the third term bound the number

of points that are present in $\xi_x[\tilde{\beta}_{x,n-\ell}]$ or $\xi_x[\tilde{\beta}_{x,n}]$ and not present or different in the other point process. We will use the triangle inequality

$$\begin{aligned} & \left(\xi_x \left[\tilde{\beta}_{x,n-N'_n(x)} \right] \Delta \xi_x \left[\tilde{\beta}_{x,n-1-N_{n-1}(x)} \right] \right) (\mathbb{Y}) \\ & \leq \left(\xi_x \left[\tilde{\beta}_{x,n-N'_n(x)} \right] \Delta \xi_x \left[\tilde{\beta}_{x,n} \right] \right) (\mathbb{Y}) + \left(\xi_x \left[\tilde{\beta}_{x,n} \right] \Delta \xi_x \left[\tilde{\beta}_{x,n-1-N_{n-1}(x)} \right] \right) (\mathbb{Y}) \end{aligned}$$

in the following. It follows from (5.5) with $\ell = N'_n(x)$ that

$$\begin{aligned} & \mathbb{E} \left\{ \left(\xi_x \left[\tilde{\beta}_{x,n-N'_n(x)} \right] \Delta \xi_x \left[\tilde{\beta}_{x,n} \right] \right) (\mathbb{Y}) \right\} \\ & \leq \mathbb{E} \left\{ \sum_{i=n-N'_n(x)+1}^n \mathbb{1} \{ S_x \cap S_{\tilde{X}_i} = \emptyset \} g \left(\tilde{X}_i, \tilde{\beta}_{x,n} \right) \right\} \\ & \quad + \mathbb{E} \left\{ \sum_{i=n-N'_n(x)+1}^n \sum_{j=1}^{n-N'_n(x)} \mathbb{1} \{ S_x \cap S_{\tilde{X}_j} = \emptyset \} g \left(\tilde{X}_j, \tilde{\beta}_{x,n-N'_n(x)} \right) \mathbb{1} \{ \tilde{X}_i \in S_{\tilde{X}_j} \} \right\} \\ & \quad + \mathbb{E} \left\{ \sum_{i=n-N'_n(x)+1}^n \sum_{j=1}^{n-N'_n(x)} \mathbb{1} \{ S_x \cap S_{\tilde{X}_j} = \emptyset \} g \left(\tilde{X}_j, \tilde{\beta}_{x,n} \right) \mathbb{1} \{ \tilde{X}_i \in S_{\tilde{X}_j} \} \right\} \\ & =: U'_{1,x} + U'_{2,x} + U'_{3,x}. \end{aligned}$$

We define $U_{i,x} := \mathbb{E} \{ g(x, \beta_{n-1} + \delta_x) \} U'_{i,x}$ for $i \in \{1, 2, 3\}$. Applying once more (5.5) with $\ell = N_{n-1}(x) + 1$ leads to

$$\begin{aligned} & \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) \left(\xi_x \left[\tilde{\beta}_{x,n} \right] \Delta \xi_x \left[\tilde{\beta}_{x,n-1-N_{n-1}(x)} \right] \right) (\mathbb{Y}) \right\} \\ & \leq \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) \sum_{i=n-N_{n-1}(x)}^n \mathbb{1} \{ S_x \cap S_{\tilde{X}_i} = \emptyset \} g \left(\tilde{X}_i, \tilde{\beta}_{x,n} \right) \right\} \\ & \quad + \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) \right. \\ & \quad \times \left. \sum_{i=n-N_{n-1}(x)}^n \sum_{j=1}^{n-1-N_{n-1}(x)} \mathbb{1} \{ S_x \cap S_{\tilde{X}_j} = \emptyset \} g \left(\tilde{X}_j, \tilde{\beta}_{x,n-1-N_{n-1}(x)} \right) \mathbb{1} \{ \tilde{X}_i \in S_{\tilde{X}_j} \} \right\} \\ & \quad + \mathbb{E} \left\{ g \left(x, \beta_{n-1} + \delta_x \right) \sum_{i=n-N_{n-1}(x)}^n \sum_{j=1}^{n-1-N_{n-1}(x)} \mathbb{1} \{ S_x \cap S_{\tilde{X}_j} = \emptyset \} \right. \\ & \quad \times \left. g \left(\tilde{X}_j, \tilde{\beta}_{x,n} \right) \mathbb{1} \{ \tilde{X}_i \in S_{\tilde{X}_j} \} \right\} \\ & =: U_{4,x} + U_{5,x} + U_{6,x}. \end{aligned}$$

From (5.1), which can be adapted to the situation, where one sums only over a fraction of the points, and the independence of $N'_n(x)$ and $\tilde{\beta}_{n,x}$ it follows that

$$U'_{1,x} = \mathbb{E} \{ N'_n(x) \} \int_{\mathbb{X}} \mathbb{1} \{ S_x \cap S_y = \emptyset \} \mathbb{E} \left\{ g \left(y, \tilde{\beta}_{x,n-1} + \delta_y \right) \right\} \mathbf{Q}_x(dy),$$

while the independence of β_{n-1} and $\tilde{\beta}_{n,x}$ and (5.1) lead to

$$\begin{aligned} U_{4,x} &= \mathbb{E} \left\{ g(x, \beta_{n-1} + \delta_x) (1 + N_{n-1}(x)) \int_{\mathbb{X}} \mathbf{1} \{S_x \cap S_y = \emptyset\} \right. \\ &\quad \left. \times \mathbb{E} \left\{ g(y, \tilde{\beta}_{x,n-1} + \delta_y) \right\} \mathbf{Q}_x(dy) \right\} \\ &= \left(\mathbb{E} \{g(x, \beta_{n-1} + \delta_x)\} + (n-1) \int_{S_x} \mathbb{E} \{g(x, \beta_{n-2} + \delta_x + \delta_z)\} \mathbf{Q}(dz) \right) \\ &\quad \times \int_{\mathbb{X}} \mathbf{1} \{S_x \cap S_y = \emptyset\} \mathbb{E} \left\{ g(y, \tilde{\beta}_{x,n-1} + \delta_y) \right\} \mathbf{Q}_x(dy). \end{aligned}$$

Together with $\mathbb{E} \{N'_n(x)\} = n\mathbf{Q}(S_x)$ and $U_{1,x} = \mathbb{E} \{g(x, \beta_{n-1} + \delta_x)\} U'_{1,x}$ we obtain

$$\begin{aligned} (5.6) \quad U_{1,x} + U_{4,x} &= \\ &\left((1 + n\mathbf{Q}(S_x)) \mathbb{E} \{g(x, \beta_{n-1} + \delta_x)\} + (n-1) \int_{S_x} \mathbb{E} \{g(x, \beta_{n-2} + \delta_x + \delta_z)\} \mathbf{Q}(dz) \right) \\ &\quad \times \int_{\mathbb{X}} \mathbf{1} \{S_x \cap S_y = \emptyset\} \mathbb{E} \left\{ g(y, \tilde{\beta}_{x,n-1} + \delta_y) \right\} \mathbf{Q}_x(dy). \end{aligned}$$

By similar independence arguments as above and (5.1), we have

$$\begin{aligned} U'_{3,x} &\leq \mathbb{E} \left\{ \sum_{i=n-N'_n(x)+1}^n \sum_{j=1}^n \mathbf{1} \{S_x \cap S_{\tilde{X}_j} = \emptyset\} g(\tilde{X}_j, \tilde{\beta}_{x,n}) \mathbf{1} \{\tilde{X}_i \in S_{\tilde{X}_j}\} \right\} \\ &\leq n\mathbb{E} \{N'_n(x)\} \int_{\mathbb{X}^2} \mathbf{1} \{S_x \cap S_{y_1} = \emptyset\} \mathbf{1} \{y_2 \in S_{y_1}\} \\ &\quad \times \mathbb{E} \left\{ g(y_1, \tilde{\beta}_{x,n-2} + \delta_{y_1} + \delta_{y_2}) \right\} \mathbf{Q}_x^2(d(y_1, y_2)) \end{aligned}$$

and

$$\begin{aligned} U_{6,x} &\leq n\mathbb{E} \left\{ g(x, \beta_{n-1} + \delta_x) (1 + N_{n-1}(x)) \right\} \\ &\quad \times \int_{\mathbb{X}^2} \mathbf{1} \{S_x \cap S_{y_1} = \emptyset\} \mathbf{1} \{y_2 \in S_{y_1}\} \\ &\quad \times \mathbb{E} \left\{ g(y_1, \tilde{\beta}_{x,n-2} + \delta_{y_1} + \delta_{y_2}) \right\} \mathbf{Q}_x^2(d(y_1, y_2)) \\ &\leq n \left(\mathbb{E} \{g(x, \beta_{n-1} + \delta_x)\} + n \int_{S_x} \mathbb{E} \{g(x, \beta_{n-2} + \delta_x + \delta_z)\} \mathbf{Q}(dz) \right) \\ &\quad \times \int_{\mathbb{X}^2} \mathbf{1} \{S_x \cap S_{y_1} = \emptyset\} \mathbf{1} \{y_2 \in S_{y_1}\} \\ &\quad \times \mathbb{E} \left\{ g(y_1, \tilde{\beta}_{x,n-2} + \delta_{y_1} + \delta_{y_2}) \right\} \mathbf{Q}_x^2(d(y_1, y_2)). \end{aligned}$$

From the definition of $U_{3,x}$ and $\mathbb{E} \{N'_n(x)\} = n\mathbf{Q}(S_x)$ we derive

$$\begin{aligned} (5.7) \quad U_{3,x} + U_{6,x} &\leq n \left((1 + n\mathbf{Q}(S_x)) \mathbb{E} \{g(x, \beta_{n-1} + \delta_x)\} + n \int_{S_x} \mathbb{E} \{g(x, \beta_{n-2} + \delta_x + \delta_z)\} \mathbf{Q}(dz) \right) \\ &\quad \times \int_{\mathbb{X}^2} \mathbf{1} \{S_x \cap S_{y_1} = \emptyset\} \mathbf{1} \{y_2 \in S_{y_1}\} \mathbb{E} \left\{ g(y_1, \tilde{\beta}_{x,n-2} + \delta_{y_1} + \delta_{y_2}) \right\} \mathbf{Q}_x^2(d(y_1, y_2)). \end{aligned}$$

Combining (5.6) and (5.7) yields

$$(5.8) \quad 2n \int_{\mathbb{X}} (U_{1,x} + U_{3,x} + U_{4,x} + U_{6,x}) \mathbf{Q}(dx) \leq E_4.$$

By (4.1) and the assumption $\mathcal{S}(z, \omega) \subset S_z$ for all $z \in \mathbb{X}, \omega \in \mathbf{N}_{\mathbb{X}}$, we obtain

$$(5.9) \quad \mathbb{1} \{z_1 \notin S_{z_2}\} g(z_2, \omega + \delta_{z_1}) = \mathbb{1} \{z_1 \notin S_{z_2}\} g(z_2, \omega)$$

for $z_1, z_2 \in \mathbb{X}$ and $\omega \in \mathbf{N}_{\mathbb{X}}$ with $z_2 \in \omega$. For $x \in \mathbb{X}$ we deduce from (5.1) and (5.9) that

$$\begin{aligned} U'_{2,x} &= \mathbb{E} \left\{ N'_n(x) \sum_{j=1}^{n-N'_n(x)} \mathbb{1} \{S_x \cap S_{\tilde{X}_j} = \emptyset\} g(\tilde{X}_j, \tilde{\beta}_{x, n-N'_n(x)}) \mathbf{Q}_x(S_{\tilde{X}_j}) \right\} \\ &= \mathbb{E} \left\{ \sum_{(x_1, x_2) \in \beta_{n, \neq}^2} \mathbb{1} \{x_1 \in S_x\} \mathbb{1} \{S_x \cap S_{x_2} = \emptyset\} g(x_2, \beta_n) \mathbf{Q}_x(S_{x_2}) \right\} \\ &= n(n-1) \mathbf{Q}(S_x) \int_{\mathbb{X}} \mathbb{1} \{S_x \cap S_y = \emptyset\} \mathbb{E} \{g(y, \beta_{n-2} + \delta_y)\} \mathbf{Q}_x(S_y) \mathbf{Q}(dy) \end{aligned}$$

so that, with $U_{2,x} = \mathbb{E} \{g(x, \beta_{n-1} + \delta_x)\} U'_{2,x}$,

$$(5.10) \quad 2n \int_{\mathbb{X}} U_{2,x} \mathbf{Q}(dx) \leq E_5.$$

Finally, applying again (5.1) and (5.9) leads to

$$\begin{aligned} U_{5,x} &= \mathbb{E} \left\{ g(x, \beta_{n-1} + \delta_x) (1 + N_{n-1}(x)) \sum_{j=1}^{n-1-N_{n-1}(x)} \right. \\ &\quad \left. \times \mathbb{1} \{S_x \cap S_{\tilde{X}_j} = \emptyset\} g(\tilde{X}_j, \tilde{\beta}_{x, n-1-N_{n-1}(x)}) \mathbf{Q}_x(S_{\tilde{X}_j}) \right\} \\ &= \mathbb{E} \left\{ g(x, \beta_{n-1} + \delta_x) \sum_{y \in \beta_{n-1}} \mathbb{1} \{S_x \cap S_y = \emptyset\} g(y, \beta_{n-1}) \mathbf{Q}_x(S_y) \right\} \\ &\quad + \mathbb{E} \left\{ g(x, \beta_{n-1} + \delta_x) \sum_{(y_1, y_2) \in \beta_{n-1, \neq}^2} \mathbb{1} \{y_1 \in S_x\} \mathbb{1} \{S_x \cap S_{y_2} = \emptyset\} \right. \\ &\quad \left. \times g(y_2, \beta_{n-1}) \mathbf{Q}_x(S_{y_2}) \right\} \\ &\leq n \int_{\mathbb{X}} \mathbb{1} \{S_x \cap S_y = \emptyset\} \mathbb{E} \{g(x, \beta_{n-2} + \delta_x) g(y, \beta_{n-2} + \delta_y)\} \mathbf{Q}_x(S_y) \mathbf{Q}(dy) \\ &\quad + n^2 \int_{\mathbb{X}} \int_{S_x} \mathbb{1} \{S_x \cap S_y = \emptyset\} \mathbb{E} \{g(x, \beta_{n-3} + \delta_x + \delta_z) g(y, \beta_{n-3} + \delta_y)\} \\ &\quad \times \mathbf{Q}_x(S_y) \mathbf{Q}(dz) \mathbf{Q}(dy) \end{aligned}$$

for $x \in \mathbb{X}$, which implies

$$(5.11) \quad 2n \int_{\mathbb{X}} U_{5,x} \mathbf{Q}(dx) \leq E_6.$$

Combining (5.8), (5.10) and (5.11) provides a bound for the integral over $T_{3,x,s}$, which together with (5.3) and (5.4) completes the proof of Theorem 5.1. \square

6. Applications

In this section we present applications of our results to (a) the critical points of a random distance function and (b) the volume (with respect to the intensity measure) of large k -nearest neighbor balls. Though some Poisson approximation results are known in both of these models, our results extend these in two ways. Firstly, we consider more general point processes than those considered before and secondly, we provide rates of convergence under a stronger metric. More details on comparison with the existing literature for these specific applications will be given in the respective subsections.

We wish to point out that these applications are to illustrate our generic results and it is conceivable that many more such applications of our results would be possible. For example, one may consider extremes of circum-radii and in-radii of Poisson–Voronoi tessellations as in [CC14, Che14, PS22] and also other stabilizing statistics as in [Ott20]. For example, we can deduce immediately from Theorem 4.1 and the proofs in [Ott20] that [Ott20, Theorems 4.2, 5.5 and 6.5] hold under the stronger KR distance. Using Theorems 3.3 or 4.1, we may extend many Poisson approximation results to Poisson *process* approximation results. For example, a coupling similar to that in Theorem 3.3 was used in [IY20, Theorem 2.8] to prove Poisson convergence for the number of isolated faces in a Vietoris–Rips complex. We expect that other Poisson approximation results in [Pen18] could also be extended using our framework (see Remark 3.4). Also, we believe that Poisson convergence results (for example, see [OA17, Theorem 2.1]) based on dependency graph method [AGG89] could be extended to a Poisson process approximation using our Theorem 4.1. Another point to emphasize is that most of the afore-mentioned articles consider only approximation for point processes induced by Poisson point processes but our bounds will help extending them to point processes induced by binomial point processes as well. For example, see Theorem 6.5.

6.1. Critical points for the random distance function

Let ω be a point configuration in some metric space. In this section we are interested in the space generated by the union of balls around ω ,

$$B_r(\omega) := \bigcup_{p \in \omega} B_r(p),$$

where $B_r(p)$ denotes the closed ball of radius r centred at p . In [Kah11] the theoretical study of *homology* of $B_r(\cdot)$ taken over random point processes was initiated. Briefly, homology is an algebraic-topological structure representing information about cycles in various dimensions, where 0-dimensional cycles correspond to connected components, 1-dimensional cycles correspond to loops surrounding “holes”, 2-dimensional

cycles correspond to surfaces enclosing “air pockets”, etc. (cf. [Hat02]). Increasing the radius r , *Morse Theory* (cf. [GR97, Mil63]) states that changes in the homology of $B_r(\omega)$ occur at critical levels of the distance function

$$\rho(x; \omega) := \min_{p \in \omega} \text{dist}(x, p),$$

where $\text{dist}(\cdot, \cdot)$ is the distance under the corresponding metric. For every critical point, Morse theory also assigns an *index* which, roughly speaking, counts the number of independent directions around the critical point, along which the distance function is decreasing. Critical points of index k can then affect the homology of $B_r(\eta)$ either in dimension k (creating new k -dimensional cycles) or in dimension $k - 1$ (terminating a $(k - 1)$ -dimensional cycle). Thus, the work in [BA14, Bob22] focused on the critical points and their indexes, as a proxy to the homology. We shall define critical points more formally below. Note that we will not elaborate more on the connection between critical points and homology as this will involve introducing algebraic topology basics. For more on the homology of $B_r(\omega)$, we refer the reader to the survey [BK18].

The work in [Bob22] focused on a homogeneous Poisson process defined on a flat torus $\mathbb{X} = \mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$ (which can be thought of as the unit box $[0, 1]^d$ with a periodic boundary). Specifically, it considered a Poisson process $\eta = \eta_n$ on \mathbb{T}^d with a fixed n . The main objective there was to characterize the phase transition for *homological connectivity*, i.e., where the k^{th} homology of the random union $B_r(\eta_n)$ converges to the k^{th} homology of the underlying torus \mathbb{T}^d . The main part of the proof there, shows that the very last obstructions to the k^{th} homological connectivity are in one-to-one correspondence with critical points of index k . In fact, it is shown in [Bob22, Section 8] that in terms of *persistent homology* these obstructions form infinitesimally small persistence intervals that are the very last ones to appear in the persistence barcode. The conditions for points to be critical, as presented in [Bob22, Section 2.3] are localized in the sense of Section 4. Hence, we are interested in providing a limit theorem for this process of obstructions.

Let r_n be defined via

$$(6.1) \quad a_n := \omega_d n r_n^d = \log n + (k - 1) \log \log n + \alpha_0,$$

where ω_d is the volume of a d -dimensional unit ball, and $\alpha_0 \in \mathbb{R}$ is fixed. We call the values of r_n satisfying (6.1) the “ k -homological connectivity” regime. Let R_n be any value satisfying

$$(6.2) \quad \lim_{n \rightarrow \infty} R_n = 0, \quad \text{and} \quad \lim_{n \rightarrow \infty} r_n / R_n = 0.$$

The analysis in [Bob22] focused on the number of critical points $c \in \mathbb{T}^d$ with Morse index k , such that $\rho(c; \eta_n) \in (r_n, R_n]$. Defining by $C_{k,n}$ the number of such critical points, in [Bob22, Proposition 4.1] the following was proved.

PROPOSITION 6.1. — *Let $\alpha_0 \in \mathbb{R}$, and let r_n, R_n satisfy (6.1) and (6.2), respectively. Then*

$$\lim_{n \rightarrow \infty} \mathbb{E} \{C_{k,n}\} = \lim_{n \rightarrow \infty} \text{Var} (C_{k,n}) = D_k e^{-\alpha_0},$$

where D_k is a known constant.

Note that the limiting results are independent of the choice of R_n (provided that (6.2) is satisfied). To prove Proposition 6.1, [Bob22] used the fact that critical points of index k are generated by subsets $\mathbf{x} \in (\eta_n)_{\neq}^{k+1}$. Defining

$$(6.3) \quad g(\mathbf{x}, \eta_n) := \mathbb{1} \{ \mathbf{x} \subset \eta_n \text{ generates a critical point of index } k, \text{ with } \rho \in (r_n, R_n] \},$$

then we can write

$$C_{k,n} := \frac{1}{(k+1)!} \sum_{\mathbf{x} \in (\eta_n)_{\neq}^{k+1}} g(\mathbf{x}, \eta_n).$$

Our goal here is not only to prove a Poisson limit for the random variable $C_{k,n}$ but to provide an elaborate point-process limit for the actual configuration of critical points that appear in $(r_n, R_n]$. To this end, we require a few more definitions.

We will consider subsets $\mathbf{x} \in (\mathbb{T}^d)^{k+1}$ that are (a) in general position, and (b) contained in a ball of radius R_n . For such subsets we can define $c(\mathbf{x})$ and $\rho(\mathbf{x})$ as the center and radius of the unique $(k-1)$ -sphere containing \mathbf{x} . In [Bob22, Lemma 2.3] it was shown that every critical point of $\rho(\cdot; \eta_n)$ with Morse index k , is of the form $c(\mathbf{x})$ for some $\mathbf{x} \in (\eta_n)_{\neq}^{k+1}$, and the corresponding critical value satisfies $\rho(c(\mathbf{x}); \eta_n) = \rho(\mathbf{x})$. Next, following (6.1) we define

$$\alpha(\mathbf{x}) := \omega_d n \rho(\mathbf{x})^d - \log n - (k-1) \log \log n.$$

Our goal is to define a point process on $\mathbb{T}^d \times \mathbb{R}$ representing pairs of the form $(c(\mathbf{x}), \alpha(\mathbf{x}))$ for those critical points in $(r_n, R_n]$, by defining

$$\xi_k = \xi_k[\eta_n] := \frac{1}{(k+1)!} \sum_{\mathbf{x} \in (\eta_n)_{\neq}^{k+1}} g(\mathbf{x}, \eta_n) \delta_{f(\mathbf{x})},$$

where $f(\mathbf{x}) = (c(\mathbf{x}), \alpha(\mathbf{x}))$, and r_n satisfies (6.1). In other words, the domain space here is $\mathbb{X} = \mathbb{T}^d$ and the image space is $\mathbb{Y} := \mathbb{T}^d \times \mathbb{R}_0$, where $\mathbb{R}_0 := [\alpha_0, \infty)$. From Proposition 6.1 and noting that $C_{k,n} = \xi_k(\mathbb{Y})$, we obtain that

$$(6.4) \quad \lim_{n \rightarrow \infty} \mathbb{E} \{ \xi_k(\mathbb{Y}) \} = \lim_{n \rightarrow \infty} \text{Var} (\xi_k(\mathbb{Y})) = D_k e^{-\alpha_0}.$$

We will prove that ξ_k converges to a Poisson process on \mathbb{Y} .

THEOREM 6.2. — *Let $\alpha_0 \in \mathbb{R}$, and ξ_k as defined above. Then for $n \geq 3$ and $R_n = \sqrt{r_n}$, we have*

$$\mathbf{d}_{\mathbf{KR}}(\xi_k, \zeta_k) \leq C_{\alpha_0} (\log \log n)^d (\log n)^{-\frac{d-k}{d+1}},$$

for some $C_{\alpha_0} > 0$, and where ζ_k is a Poisson process on $\mathbb{Y} = \mathbb{T}^d \times \mathbb{R}_0$, with intensity $D_k e^{-\alpha} d\alpha dc$. In particular, this implies that $\xi_k \xrightarrow{\mathbf{KR}} \zeta_k$ as $n \rightarrow \infty$.

In other words, the process of all pairs (c, α) representing critical points and critical values, has a limit of a Poisson process which is homogeneous in c and has an exponentially decaying intensity in α .

Remark 6.3. —

- (1) [Bob22, Theorem 8.1] presents a weaker statement than Theorem 6.2 here. The limiting point process in [Bob22] was only for the critical radii, while here we show that a Poisson limit also extends to the combined location+radius point process. Regarding the proofs, the calculations in [Bob22] provide all the moment estimates needed to invoke Theorem 4.1. However, the key ingredient needed to prove the Poisson limit in [Bob22] is in fact Theorem 4.1 which appears here for the first time.
- (2) We can also conclude from Theorem 6.2 the convergence of the entire point process of critical points and distances (i.e., on $\mathbb{T}^d \times \mathbb{R}$). More precisely, consider

$$\xi'_k[\eta_n] := \frac{1}{(k+1)!} \sum_{\mathbf{x} \in (\eta_n)^{k+1}_{\neq}} \mathbb{1}\{\mathbf{x} \subset \eta_n \text{ generates a critical point of index } k \text{ with } \rho(\mathbf{x}) \leq R_n\} \delta_{f(\mathbf{x})},$$

where $f(\mathbf{x}), \rho(\mathbf{x})$ are as defined above. Since Theorem 6.2 holds for all $\alpha_0 \in \mathbb{R}$, by the characterization of convergence in distribution of point processes in [Kal02, Theorem 16.16], we have that $\xi'_k[\eta_n] \xrightarrow{d} \zeta'_k$, where ζ'_k is a Poisson process on $\mathbb{Y} = \mathbb{T}^d \times \mathbb{R}$, with intensity $D_k e^{-\alpha} d\alpha dc$.

- (3) The choice of $R_n = \sqrt{r_n}$ was required in order to get the bound presented in Theorem 6.2. In terms of the homological connectivity phenomenon, this choice has no practical consequence. In fact it can be shown that even at radius $r = 2r_n$ with high probability $B_r(\eta_n)$ covers the torus \mathbb{T}^d completely. This implies that homology exhibits no further changes, and the only changes we expect to see are at radii smaller than $2r_n \ll \sqrt{r_n}$. Choosing a larger R_n leads to process convergence for a larger point process which is more desirable.

Proof of Theorem 6.2. — Following the notation above, we denote by \mathbf{L} and \mathbf{M} the intensity measures of ξ_k and ζ_k , respectively. Denote $\mathbb{X} := \mathbb{T}^d$, and note that since η_n is a homogeneous Poisson process with rate n on \mathbb{X} , we have that $\mathbf{K}(dx) = ndx$. Also, setting $S_{\mathbf{x}} = B_{\rho(\mathbf{x})}(c(\mathbf{x}))$ then the definitions for the critical points in [Bob22] imply that $g(\mathbf{x}, \eta_n)$ is indeed localized to $S_{\mathbf{x}}$, i.e., condition (4.1) holds with $\mathcal{S}(\mathbf{x}, \omega) \equiv S_{\mathbf{x}}$.

Using Theorem 4.1 we have,

$$\begin{aligned} (6.5) \quad \mathbf{d}_{\mathbf{KR}}(\xi_k, \zeta_k) &\leq d_{TV}(\mathbf{L}, \mathbf{M}) \\ &\quad + 2 \{ \text{Var}(\xi_k(\mathbb{Y})) - \mathbb{E}\{\xi_k(\mathbb{Y})\} \} \\ &\quad + \left(\frac{2}{(k+1)!} \right)^2 \int_{\mathbb{X}^{k+1}} \int_{\mathbb{X}^{k+1}} \mathbb{1}\{S_{\mathbf{z}} \cap S_{\mathbf{x}} \neq \emptyset\} \\ &\quad \times \mathbb{E}\{g(\mathbf{x}, \eta_n + \delta_{\mathbf{x}})\} \mathbb{E}\{g(\mathbf{z}, \eta_n + \delta_{\mathbf{z}})\} \mathbf{K}^{k+1}(d\mathbf{z}) \mathbf{K}^{k+1}(d\mathbf{x}). \end{aligned}$$

First, recall that ζ_k is a Poisson process on $\mathbb{Y} = \mathbb{T}^d \times \mathbb{R}_0$, with intensity $D_k e^{-\alpha} d\alpha dc$. Then for any measurable $A \subset \mathbb{Y}$,

$$\mathbf{M}(A) = \mathbb{E}\{\zeta_k(A)\} = D_k \int_A e^{-\alpha} d\alpha dc.$$

Next, from Lemma A.1 we have

$$\mathbf{L}(A) = \mathbb{E} \{ \xi_k(A) \} = D_k n \int_{\hat{A}} s^{k-1} e^{-s} \mathbb{1} \{ s \leq \omega_d n R_n^d \} ds dc,$$

where

$$\hat{A} := \{ (c, s) : (c, \alpha) \in A, s = \log n + (k - 1) \log \log n + \alpha \}.$$

Therefore, defining $I_n(\alpha) := \mathbb{1} \{ \log n + (k - 1) \log \log n + \alpha \leq \omega_d n R_n^d \}$, we have

$$\begin{aligned} (6.6) \quad \mathbf{L}(A) &= D_k n \int_A (\log n + (k - 1) \log \log n + \alpha)^{k-1} e^{-(\log n + (k-1) \log \log n + \alpha)} I_n(\alpha) d\alpha dc \\ &= D_k \int_A \left(1 + \frac{(k - 1) \log \log n + \alpha}{\log n} \right)^{k-1} e^{-\alpha} I_n(\alpha) d\alpha dc. \end{aligned}$$

If n is large enough, then $\alpha_0 > -\log \log n$, and we have

$$\begin{aligned} (6.7) \quad |\mathbf{L}(A) - \mathbf{M}(A)| &= D_k \left| \int_A \left(I_n(\alpha) \left(1 + \frac{(k - 1) \log \log n + \alpha}{\log n} \right)^{k-1} - 1 \right) e^{-\alpha} d\alpha dc \right| \\ &\leq D_k \sum_{i=1}^{k-1} \binom{k-1}{i} \int_A \left(\frac{(k-1) \log \log n + \alpha}{\log n} \right)^i e^{-\alpha} d\alpha dc + D_k \int_{\mathbb{R}} (1 - I_n(\alpha)) e^{-\alpha} d\alpha \\ &= D_k \sum_{i=1}^{k-1} \sum_{j=0}^i \binom{k-1}{i} \binom{i}{j} \frac{((k-1) \log \log n)^{i-j}}{(\log n)^i} \int_A \alpha^j e^{-\alpha} d\alpha dc \\ &\quad + D_k e^{-\omega_d n R_n^d + \log n + (k-1) \log \log n} \\ &= O \left(\frac{\log \log n}{\log n} \right), \end{aligned}$$

where we used the fact that $\int_{\mathbb{R}_0} \alpha^j e^{-\alpha} d\alpha < \infty$ and that $R_n = \sqrt{r_n}$.

The fact that $\text{Var}(\xi_k(\mathbb{Y})) - \mathbb{E} \{ \xi_k(\mathbb{Y}) \} \rightarrow 0$ is given by (6.4). Further, from Lemma A.2, we have

$$(6.8) \quad \text{Var}(\xi_k(\mathbb{Y})) - \mathbb{E} \{ \xi_k(\mathbb{Y}) \} = O \left((\log \log n)^d (\log n)^{-\frac{d-k}{d+1}} \right).$$

We are left now with bounding the last term in (6.5).

Fix \mathbf{x} and note that if $S_{\mathbf{z}} \cap S_{\mathbf{x}} \neq \emptyset$, then $|c(\mathbf{z}) - c(\mathbf{x})| \leq \rho(\mathbf{z}) + \rho(\mathbf{x}) \leq 2R_n$. Therefore, using (6.6), we derive that

$$\begin{aligned} &\frac{1}{(k+1)!} \int_{\mathbb{X}^{k+1}} \mathbb{1} \{ S_{\mathbf{z}} \cap S_{\mathbf{x}} \neq \emptyset \} \mathbb{E} \{ g(\mathbf{z}, \eta_n + \delta_{\mathbf{z}}) \} \mathbf{K}^{k+1}(d\mathbf{z}) \\ &\leq \frac{1}{(k+1)!} \int_{\mathbb{X}^{k+1}} \mathbb{E} \{ \mathbb{1} \{ c(\mathbf{z}) \in B_{2R_n}(c(\mathbf{x})) \} g(\mathbf{z}, \eta_n + \delta_{\mathbf{z}}) \} \mathbf{K}^{k+1}(d\mathbf{z}) \\ &= \mathbf{L}(B_{2R_n}(c(\mathbf{x})) \times \mathbb{R}_0) = \text{Vol}(B_{2R_n}(c(\mathbf{x}))) \cdot \mathbf{L}(\mathbb{T}^d \times \mathbb{R}_0) = \omega_d (2R_n)^d D_k e^{-\alpha_0}, \end{aligned}$$

where we used the fact that $\text{Vol}(\mathbb{T}^d) = 1$. Finally, we can conclude that

$$\begin{aligned}
 (6.9) \quad & \frac{1}{((k+1)!)^2} \int_{\mathbb{X}^{k+1}} \int_{\mathbb{X}^{k+1}} \mathbb{1}\{S_{\mathbf{z}} \cap S_{\mathbf{x}} \neq \emptyset\} \mathbb{E}\{g(\mathbf{x}, \eta_n + \delta_{\mathbf{x}})\} \\
 & \times \mathbb{E}\{g(\mathbf{z}, \eta_n + \delta_{\mathbf{z}})\} \mathbf{K}^{k+1}(d\mathbf{z}) \mathbf{K}^{k+1}(d\mathbf{x}) \\
 & \leq \omega_d (2R_n)^d D_k e^{-\alpha_0} \frac{1}{(k+1)!} \int_{\mathbb{X}^{k+1}} \mathbb{E}\{g(\mathbf{z}, \eta_n + \delta_{\mathbf{z}})\} \mathbf{K}^{k+1}(d\mathbf{x}) = O(R_n^d).
 \end{aligned}$$

Since $R_n = \sqrt{r_n}$, we have $R_n^d = O(\sqrt{\log n/n})$. Combining the bounds in (6.7)-(6.9) completes the proof. \square

6.2. Large k -nearest neighbor balls

We look at the point process of the scaled volumes of k -nearest neighbor balls. We shall consider the setup as in [Ott20, Section 4]. This is a well-studied statistic in computational and stochastic geometry with varied applications (see [Pen97, Pg. 342]). It is also closely related to another important object, the minimal spanning tree. We remark on this briefly in our discussion after Theorem 6.4.

Let $\mathbb{X} = \mathbb{T}^d$, $d \geq 2$ as in the previous section and identify it with $[0, 1]^d$, and let \mathbf{K} be a finite measure on $(\mathbb{X}, \mathcal{X})$. Define for $k \geq 1$ and $\omega \in \widehat{\mathbf{N}}_{\mathbb{X}}$,

$$\begin{aligned}
 R_k(x, \omega) & := \inf\{r > 0 : \omega(B_r(x) \setminus \{x\}) \geq k\}, \\
 \mathbf{K}_k(x, \omega) & := \mathbf{K}(B_{R_k(x, \omega)}(x)),
 \end{aligned}$$

where $B_r(x)$ is the closed ball of radius r around x . In other words, R_k is the k -nearest neighbor distance of x , and \mathbf{K}_k is the measure of the ball of radius R_k around x . Let $\mathbb{Y} = \mathbb{X} \times \mathbb{R}$, and define

$$\xi[\omega] := \sum_{x \in \omega} \delta_{(x, n\mathbf{K}_k(x, \omega) - a_n)},$$

where

$$(6.10) \quad a_n := \log n + (k-1) \log \log n - \log(k-1)!.$$

In other words, for an underlying point process ν , $\xi[\nu]$ is a point process of pairs, each pair consisting of a point and the scaled volume of its k -nearest neighbor ball. The scaling is chosen such that we count only the extremal (maximal) ones.

THEOREM 6.4. — *Let $d \geq 2$ and $k \geq 1$. Let η_n be a Poisson process with intensity measure $n\mathbf{K}$ and assume that \mathbf{K} is a probability measure with a density $\lambda : \mathbb{X} \rightarrow (0, \infty)$ such that $0 < \lambda_- \leq \lambda(x) \leq \lambda_+ < \infty$ for all $x \in \mathbb{X}$. Let ζ be a Poisson process with intensity measure $\mathbf{M}(dx, dt) = \lambda(x) dx e^{-t} dt$ on $\mathbb{X} \times \mathbb{R}$. Then, for any $b_0 \in \mathbb{R}$, there exists a constant $C \in (0, \infty)$ only depending on b_0, k, d, λ_+ and λ_- such that*

$$(6.11) \quad \mathbf{d}_{\mathbf{KR}}(\xi[\eta_n] \cap (\mathbb{X} \times (b_0, \infty)), \zeta \cap (\mathbb{X} \times (b_0, \infty))) \leq C \frac{\log \log n}{\log n}$$

for all $n \geq 3$. Moreover, $\xi[\eta_n] \xrightarrow{d} \zeta$ as $n \rightarrow \infty$.

We have assumed $\lambda_- > 0$ and $\lambda_+ < \infty$ for technical reasons to simplify some estimates involving \mathbf{K} .

The constant C in the theorem can be deduced from the bounds in (6.12), (6.13), (6.14), (6.15) and (6.16). The convergence of $\xi[\eta_n]$ was obtained in [Pen97, Theorem 2] in the case of homogeneous Poisson processes. Our above theorem considers inhomogeneous Poisson processes as well. Further [Pen97, Theorem 2] is used to prove Poisson process convergence of extremal edge lengths of the minimal spanning tree built on the complete graph on η_n with Euclidean distances as the edge-weights (see [Pen97, Theorem 3]). The proof technique of [Pen97, Theorem 2] is to discretize and use the Chen–Stein bound of [AGG89].

[Ott20, Theorem 4.2] gives bounds under total variation distance for the projection of $\xi[\eta_n] \cap (\mathbb{X} \times (0, \infty))$ to \mathbb{X} . We will now show that using our Theorem 4.1 with some additional computations compared to [Ott20, Theorem 4.2], we obtain improved rates of convergence for the more general point process $\xi[\eta_n] \cap (\mathbb{X} \times (b_0, \infty))$ for any $b_0 \in \mathbb{R}$.

Proof. — Let $b_0 \in \mathbb{R}$ be fixed. Set $g(x, \omega) := \mathbb{1} \{n\mathbf{K}_k(x, \omega) - a_n > b_0\}$ and $f(x, \omega) := (x, n\mathbf{K}_k(x, \omega) - a_n)$. Define

$$\xi_0[\omega] := \sum_{x \in \omega} g(x, \omega) \delta_{f(x, \omega)}, \quad \omega \in \widehat{\mathbf{N}}_{\mathbb{X}},$$

and observe that $\xi_0[\omega] = \xi[\omega] \cap (\mathbb{X} \times (b_0, \infty))$. We put

$$r_n(x, u) := \inf \{r : n\mathbf{K}(B_r(x)) \geq a_n + u\}.$$

We will apply Theorem 4.1 with f and g as above to compare $\xi_0[\eta_n]$ with the restriction of ζ to $\mathbb{X} \times (b_0, \infty)$. Denoting $\mathcal{S}(x, \omega) = B_{R_k(x, \omega)}(x)$, observe that f and g are both localized to \mathcal{S} , which is a stopping set. Further, we set $S_x = B_{r_n(x, b)}(x)$ for a fixed $b \in (\max\{0, b_0\}, \infty)$, which will be chosen later as a function of n .

We shall compute one-by-one the terms in the KR distance bound provided by Theorem 4.1. Observe that since we are summing over points, the last term E_4 is absent. Throughout this proof all inequalities are supposed to hold for all $n \in \mathbb{N}$ such that $a_n + b_0 > 0$.

Note that for $x \in \mathbb{X}$, we have $\{n\mathbf{K}_k(x, \eta_n + \delta_x) - a_n > u\} = \{\eta_n(B_{r_n(x, u)}(x)) < k\}$ (a.s.). Consider the first term in the bound of Theorem 4.1. Define $\varrho_n := n\lambda_- \omega_d / 2^d - a_n$, where ω_d is the volume of the unit ball in \mathbb{R}^d . Denote the intensity measure of ξ_0 by \mathbf{L} . Then for $B \in \mathcal{X}$ and $u \in (b_0, \varrho_n)$ we have by the Mecke formula that

$$\begin{aligned} \mathbf{L}(B \times (u, \infty)) &= n \int_B \mathbb{P}(n\mathbf{K}_k(x, \eta_n + \delta_x) - a_n > u) \lambda(x) dx \\ &= n \int_B \mathbb{P}(\eta_n(B_{r_n(x, u)}(x)) \leq k - 1) \lambda(x) dx \\ &= n\mathbf{K}(B) \sum_{i=0}^{k-1} e^{-(a_n+u)} \frac{(a_n+u)^i}{i!} = e^{-u} \mathbf{K}(B) \sum_{i=0}^{k-1} \frac{(k-1)!(a_n+u)^i}{i!(\log n)^{k-1}}, \end{aligned}$$

where we used (6.10) in the last equality. Thus, the density ℓ of \mathbf{L} satisfies

$$\ell(x, u) = \lambda(x) e^{-u} \frac{(a_n+u)^{k-1}}{(\log n)^{k-1}}$$

for $x \in \mathbb{X}$ and $u \in (b_0, \varrho_n)$. Let \mathbf{M}_0 be the intensity measure of $\zeta_0 := \zeta \cap (\mathbb{X} \times (b_0, \infty))$. Now we obtain

$$\begin{aligned}
 (6.12) \quad d_{TV}(\mathbf{L}, \mathbf{M}_0) &\leq \int_{b_0}^{\infty} \left| e^{-u} \frac{(a_n + u)^{k-1}}{(\log n)^{k-1}} - e^{-u} \right| du + \mathbf{L}(\mathbb{X} \times (\varrho_n, \infty)) + \mathbf{M}_0(\mathbb{X} \times (\varrho_n, \infty)) \\
 &\leq C_0 \frac{\log \log n}{\log n}
 \end{aligned}$$

with a constant $C_0 \in (0, \infty)$ only depending on b_0, k, d and λ_- . Next, we compute the integral in the second term of the bound in Theorem 4.1. Similar to the above calculation, we can derive that

$$\begin{aligned}
 (6.13) \quad E_1 &\leq 2n \int_{\mathbb{X}} \mathbb{P}(\eta_n(B_{r_n(x,b)}(x)) \leq k-1) \lambda(x) dx \\
 &\leq C_1 e^{-b} \left(1 + \frac{(k-1) \log \log n + b}{\log n} \right)^{k-1}
 \end{aligned}$$

for $b < \varrho_n$ with $C_1 := 2k!$.

Observing that $\{\mathcal{S}(x, \omega) \subset S_x\} = \{R_k(x, \omega) \leq r_n(x, b)\}$ and using the definition of g , the integral in the third term of the bound in Theorem 4.1 can be bounded from above by

$$\begin{aligned}
 (6.14) \quad E_2 &\leq 2n^2 \int_{\mathbb{X}} \int_{\mathbb{X}} \mathbb{P}(R_k(x, \eta_n + \delta_x) \in (r_n(x, b_0), r_n(x, b)]) \\
 &\quad \times \mathbb{P}(R_k(z, \eta_n + \delta_z) \in (r_n(z, b_0), r_n(z, b)]) \\
 &\quad \times \mathbb{1}\{|x - z| \leq r_n(x, b) + r_n(z, b)\} \lambda(x) \lambda(z) dx dz \\
 &\leq 2n^2 \int_{\mathbb{X}} \int_{\mathbb{X}} \mathbb{P}(\eta_n(B_{r_n(x,b_0)}(x)) \leq k-1) \mathbb{P}(\eta_n(B_{r_n(z,b_0)}(z)) \leq k-1) \\
 &\quad \times \mathbb{1}\{|x - z| \leq r_n(x, b) + r_n(z, b)\} \lambda(x) \lambda(z) dx dz, \\
 &\leq 2 \left(e^{-b_0} \sum_{i=0}^{k-1} \frac{(k-1)! (a_n + b_0)^i}{i! (\log n)^{k-1}} \right)^2 \int_{\mathbb{X}} \int_{\mathbb{X}} \mathbb{1}\{|x - z| \leq r_n(x, b) + r_n(z, b)\} \\
 &\quad \times \lambda(x) \lambda(z) dx dz \\
 &\leq C_2 \frac{a_n + b}{n},
 \end{aligned}$$

where in the last bound we have used that

$$\sup_{x \in \mathbb{X}} \omega_d n r_n(x, b)^d \leq \lambda_-^{-1} \sup_{x \in \mathbb{X}} n \mathbf{K}(B_{r_n(x,b)}) = \frac{a_n + b}{\lambda_-}.$$

Note that the constant $C_2 \in (0, \infty)$ depends on b_0, k, λ_+ and λ_- .

Similarly to the above bound, the fourth term in Theorem 4.1 can be bounded by

$$\begin{aligned}
 E_3 &\leq 2n^2 \int_{\mathbb{X}} \int_{\mathbb{X}} \mathbb{P} \left([\eta_n + \delta_z] \left(B_{r_n(x,b_0)}(x) \right) \leq k - 1, [\eta_n + \delta_x] \left(B_{r_n(z,b_0)}(z) \right) \leq k - 1 \right) \\
 &\quad \times \mathbb{1} \{ |x - z| \leq r_n(x, b) + r_n(z, b) \} \lambda(x) \lambda(z) \, dx \, dz \\
 &\leq 2n^2 \lambda_+^2 \int_{\mathbb{X}} \int_{\mathbb{X}} \mathbb{P} \left([\eta_n + \delta_z] \left(B_{r_n(x,b_0)}(x) \right) \leq k - 1, [\eta_n + \delta_x] \left(B_{r_n(z,b_0)}(z) \right) \leq k - 1 \right) \\
 &\quad \times \mathbb{1} \{ |x - z| \leq \min \{ r_n(x, b_0), r_n(z, b_0) \} \} \\
 &\quad \times \mathbb{1} \{ |x - z| \leq r_n(x, b) + r_n(z, b) \} \, dx \, dz \\
 &\quad + 2n^2 \lambda_+^2 \int_{\mathbb{X}} \int_{\mathbb{X}} \mathbb{P} \left([\eta_n + \delta_z] \left(B_{r_n(x,b_0)}(x) \right) \right. \\
 &\quad \quad \quad \left. \leq k - 1, [\eta_n + \delta_x] \left(B_{r_n(z,b_0)}(z) \right) \leq k - 1 \right) \\
 &\quad \times \mathbb{1} \{ |x - z| > \min \{ r_n(x, b_0), r_n(z, b_0) \} \} \\
 &\quad \times \mathbb{1} \{ |x - z| \leq r_n(x, b) + r_n(z, b) \} \, dx \, dz \\
 &=: E_{3,1} + E_{3,2}.
 \end{aligned}$$

We first consider $E_{3,1}$. Since $E_{3,1} = 0$ for $k = 1$, we can assume $k \geq 2$ in the sequel. We obtain that

$$\begin{aligned}
 E_{3,1} &\leq 2n^2 \lambda_+^2 \int_{\mathbb{X}} \int_{\mathbb{X}} \mathbb{P} \left(\eta_n \left(B_{r_n(x,b_0)}(x) \right) \leq k - 2, \eta_n \left(B_{r_n(z,b_0)}(z) \right) \leq k - 2 \right) \\
 &\quad \times \mathbb{1} \{ |x - z| \leq \min \{ r_n(x, b_0), r_n(z, b_0) \} \} \\
 &\quad \times \mathbb{1} \{ |x - z| \leq r_n(x, b) + r_n(z, b) \} \, dx \, dz \\
 &\leq 4n^2 \lambda_+^2 \int_{\mathbb{X}} \int_{\mathbb{X}} \mathbb{P} \left(\eta_n \left(B_{r_n(x,b_0)}(x) \right) \leq k - 2, \eta_n \left(B_{r_n(z,b_0)}(z) \right) \leq k - 2 \right) \\
 &\quad \times \mathbb{1} \{ r_n(x, b_0) \leq r_n(z, b_0) \} \mathbb{1} \{ |x - z| \leq r_n(x, b) + r_n(z, b) \} \, dx \, dz \\
 &\leq 4n^2 \lambda_+^2 \int_{\mathbb{X}} \int_{\mathbb{X}} \mathbb{P} \left(\eta_n \left(B_{r_n(x,b_0)}(x) \right) \leq k - 2 \right) \\
 &\quad \times \mathbb{P} \left(\eta_n \left(B_{r_n(z,b_0)}(z) \setminus B_{r_n(x,b_0)}(x) \right) \leq k - 2 \right) \\
 &\quad \times \mathbb{1} \{ r_n(x, b_0) \leq r_n(z, b_0) \} \mathbb{1} \{ |x - z| \leq r_n(x, b) + r_n(z, b) \} \, dx \, dz.
 \end{aligned}$$

For $x, z \in \mathbb{X}$ we have

$$\mathbb{P} \left(\eta_n \left(B_{r_n(x,b_0)}(x) \right) \leq k - 2 \right) = \sum_{i=0}^{k-2} \frac{(a_n + b_0)^i}{i!} e^{-(a_n+b_0)} \leq \frac{C_{b_0,k}}{n(a_n + b_0)}$$

with some constant $C_{b_0,k} \in (0, \infty)$ only depending on b_0 and k and

$$\begin{aligned}
 &\mathbb{P} \left(\eta_n \left(B_{r_n(z,b_0)}(z) \setminus B_{r_n(x,b_0)}(x) \right) \leq k - 2 \right) \\
 &= \sum_{i=0}^{k-2} \frac{\left(n \mathbf{K} \left(B_{r_n(z,b_0)}(z) \setminus B_{r_n(x,b_0)}(x) \right) \right)^i}{i!} e^{-n \mathbf{K} \left(B_{r_n(z,b_0)}(z) \setminus B_{r_n(x,b_0)}(x) \right)} \\
 &\leq C_k e^{-n \mathbf{K} \left(B_{r_n(z,b_0)}(z) \setminus B_{r_n(x,b_0)}(x) \right) / 2}
 \end{aligned}$$

with some constant $C_k \in (0, \infty)$ only depending on k . By a short computation one can establish that there exists a dimension dependent constant $c_d \in (0, \infty)$ such that

$$|B_1(0) \setminus B_1(v)| \geq c_d |v|$$

for all $v \in B(0, 2)$, see e.g. [PG10, Equation (7.5)]. From this and $r_n(x, b_0) \leq r_n(z, b_0)$ we deduce

$$\begin{aligned} \mathbf{K} \left(B_{r_n(z, b_0)}(z) \setminus B_{r_n(x, b_0)}(x) \right) & \geq \lambda_- \left| B_{r_n(z, b_0)}(z) \setminus B_{r_n(x, b_0)}(x) \right| \\ & \geq \lambda_- \left| B_{r_n(z, b_0)}(z) \setminus B_{r_n(z, b_0)}(x) \right| \geq \lambda_- c_d r_n(z, b_0)^{d-1} |x - z|. \end{aligned}$$

Combining the previous estimates and using spherical coordinates we obtain

$$\begin{aligned} E_{3,1} & \leq 4n^2 \lambda_+^2 \int_{\mathbb{X}} \int_{\mathbb{X}} \frac{C_{b_0, k} C_k}{n(a_n + b_0)} e^{-n\lambda_- c_d r_n(z, b_0)^{d-1} |x-z|/2} \\ & \quad \times \mathbf{1} \{ |x - z| \leq r_n(x, b) + r_n(z, b) \} \, dx \, dz \\ & \leq 4\lambda_+^2 C_{b_0, k} C_k \frac{n}{a_n + b_0} d\omega_d \int_{\mathbb{X}} \int_0^\infty e^{-n\lambda_- c_d r_n(z, b_0)^{d-1} s/2} s^{d-1} \, ds \, dz \\ & \leq 4\lambda_+^2 C_{b_0, k} C_k d\omega_d \frac{n}{a_n + b_0} \int_{\mathbb{X}} \frac{2^d}{(n\lambda_- c_d r_n(z, b_0)^{d-1})^d} \int_0^\infty e^{-u} u^{d-1} \, du \, dz. \end{aligned}$$

Since $n\lambda_+ \omega_d r_n(z, b_0)^d \geq n\mathbf{K}(B_{r_n(z, b_0)}(z)) = a_n + b_0$, we conclude that there exists a constant $C_{3,1} \in (0, \infty)$ only depending on b_0, k, d, λ_+ and λ_- such that

$$(6.15) \quad E_{3,1} \leq \frac{C_{3,1}}{(a_n + b_0)^d}.$$

For $E_{3,2}$ we obtain

$$\begin{aligned} E_{3,2} & \leq 4n^2 \lambda_+^2 \int_{\mathbb{X}} \int_{\mathbb{X}} \mathbb{P} \left(\eta_n \left(B_{r_n(x, b_0)}(x) \right) \leq k - 1, \eta_n \left(B_{r_n(z, b_0)}(z) \right) \leq k - 1 \right) \\ & \quad \times \mathbf{1} \{ |x - z| > r_n(x, b_0), r_n(z, b_0) \geq r_n(x, b_0) \} \\ & \quad \times \mathbf{1} \{ |x - z| \leq r_n(x, b) + r_n(z, b) \} \, dx \, dz \\ & \leq 4n^2 \lambda_+^2 \int_{\mathbb{X}} \int_{\mathbb{X}} \mathbb{P} \left(\eta_n \left(B_{r_n(x, b_0)}(x) \right) \leq k - 1 \right) \\ & \quad \times \mathbb{P} \left(\eta_n \left(B_{r_n(z, b_0)}(z) \setminus B_{r_n(x, b_0)}(x) \right) \leq k - 1 \right) \\ & \quad \times \mathbf{1} \{ |x - z| > r_n(x, b_0), r_n(z, b_0) \geq r_n(x, b_0) \} \\ & \quad \times \mathbf{1} \{ |x - z| \leq r_n(x, b) + r_n(z, b) \} \, dx \, dz. \end{aligned}$$

Let $x, z \in \mathbb{X}$. There exists a constant $C'_{b_0, k} \in (0, \infty)$ only depending on b_0 and k such that

$$\mathbb{P} \left(\eta_n \left(B_{r_n(x, b_0)}(x) \right) \leq k - 1 \right) \leq \sum_{i=0}^{k-1} \frac{(a_n + b_0)^i}{i!} e^{-(a_n + b_0)} \leq \frac{C'_{b_0, k}}{n}.$$

From $|x - z| > r_n(x, b_0)$ and the assumptions on \mathbf{K} it follows that

$$\begin{aligned} \mathbf{K} \left(B_{r_n(z, b_0)}(z) \setminus B_{r_n(x, b_0)}(x) \right) &\geq \lambda_- \left| B_{r_n(z, b_0)}(z) \setminus B_{r_n(x, b_0)}(x) \right| \\ &\geq \frac{\lambda_-}{2} \left| B_{r_n(z, b_0)}(z) \right| \geq \frac{\lambda_-}{2\lambda_+} \mathbf{K} \left(B_{r_n(z, b_0)}(z) \right) \end{aligned}$$

so that

$$n\mathbf{K} \left(B_{r_n(z, b_0)}(z) \setminus B_{r_n(x, b_0)}(x) \right) \geq \gamma (a_n + b_0)$$

with $\gamma := \lambda_- / (2\lambda_+)$. Consequently, we can choose a constant $C''_{b_0, k} \in (0, \infty)$ only depending on b_0 and k such that

$$\mathbb{P} \left(\eta_n \left(B_{r_n(z, b_0)}(z) \setminus B_{r_n(x, b_0)}(x) \right) \leq k - 1 \right) \leq C''_{b_0, k} \frac{(a_n + b_0)^{k-1}}{n^\gamma}.$$

Moreover, we have $n\lambda_- \omega_d r_n(x, b)^d \leq n\mathbf{K}(B_{r_n(x, b)}(x)) = a_n + b$ and $n\lambda_- \omega_d r_n(z, b)^d \leq n\mathbf{K}(B_{r_n(z, b)}(z)) = a_n + b$. Combining all these estimates, we see that there exists a constant $C_{3,2} \in (0, \infty)$ only depending on b_0, k, λ_+ and λ_- such that

$$(6.16) \quad E_{3,2} \leq C_{3,2} n^2 \frac{1}{n} \frac{(a_n + b_0)^{k-1}}{n^\gamma} \frac{a_n + b}{n} = C_{3,2} \frac{(a_n + b_0)^{k-1} (a_n + b)}{n^\gamma}.$$

Combining (6.12), (6.13), (6.14), (6.15) and (6.16) with Theorem 4.1 yields that

$$\begin{aligned} &\mathbf{d}_{\mathbf{KR}} \left(\xi[\eta_n] \cap (\mathbb{X} \times (b_0, \infty)), \zeta \cap (\mathbb{X} \times (b_0, \infty)) \right) \\ &\leq C_0 \frac{\log \log n}{\log n} + C_1 e^{-b} \left(1 + \frac{(k-1) \log \log n + b}{\log n} \right)^{k-1} + C_2 \frac{a_n + b}{n} + C_{3,1} \frac{1}{(a_n + b_0)^d} \\ &\quad + C_{3,2} \frac{(a_n + b_0)^{k-1} (a_n + b)}{n^\gamma}. \end{aligned}$$

Now the choice $b = \log n$ proves (6.11).

Since convergence in KR distance implies convergence in distribution, we have

$$\xi[\eta_n] \cap (\mathbb{X} \times (b_0, \infty)) \xrightarrow{d} \zeta \cap (\mathbb{X} \times (b_0, \infty)) \text{ as } n \rightarrow \infty.$$

By the characterization of convergence in distribution of point processes in [Kal02, Theorem 16.16], this implies $\xi[\eta_n] \xrightarrow{d} \zeta$ as $n \rightarrow \infty$. \square

In the next theorem, we consider the large k -nearest neighbor balls for an underlying binomial point process under some assumptions on the density of the distribution of the points. For an underlying homogeneous binomial point process Poisson process convergence was established in [Pen97]. As mentioned before, we are not aware of any quantitative result for binomial point processes even with a weaker rate of convergence or under weaker approximation distance in the literature. For recent Poisson approximation results concerning k -nearest neighbor balls of a binomial point process in \mathbb{R}^d see [CHO21] and (for $k = 1$) [GHW19]. The results of [CHO21] are in total variation distance for marginal distributions of the point processes that we consider and the convergence rates are the same.

THEOREM 6.5. — *Let $d \geq 2$ and $k \geq 1$. Let β_n be a binomial point process of n points distributed according to a probability measure \mathbf{K} on \mathbb{X} and assume that \mathbf{K} has a density $\lambda : \mathbb{X} \rightarrow (0, \infty)$ such that $0 < \lambda_- \leq \lambda(x) \leq \lambda_+ < \infty$ for all $x \in \mathbb{X}$. Let*

ζ be a Poisson process with intensity measure $\mathbf{M}(dx, dt) = \lambda(x) dx e^{-t} dt$ on $\mathbb{X} \times \mathbb{R}$. Then, for any $b_0 \in \mathbb{R}$, there exists a constant $C \in (0, \infty)$ only depending on b_0, k, d, λ_+ and λ_- such that

$$(6.17) \quad \mathbf{d}_{\mathbf{KR}}(\xi[\beta_n] \cap (\mathbb{X} \times (b_0, \infty)), \zeta \cap (\mathbb{X} \times (b_0, \infty))) \leq C \frac{\log \log n}{\log n}$$

for all $n \geq 3$. Moreover, $\xi[\beta_n] \xrightarrow{d} \zeta$ as $n \rightarrow \infty$.

Proof. — It suffices to prove (6.17) and the remaining statement follows as in Theorem 6.4. The rest of the proof will be concerned about proving (6.17).

We define $f, g, \mathcal{S}, \xi_0, r_n$ and ϱ_n exactly as in the proof of Theorem 6.4. We also define $S_x = B_{r_n(x,b)}(x)$ for a fixed $b \in (\max\{0, b_0\}, a_n]$, which will be chosen later. Our goal is to apply Theorem 5.1.

Throughout this proof we use several times the observation that

$$(6.18) \quad 0 \leq e^{-y} - (1 - y/n)^n \leq e^{-y} \frac{y^2}{n}$$

for $y \in \mathbb{R}$ and $n \in \mathbb{N}$ such that $y/n \leq 1/2$. From now on we assume that n is sufficiently large so that

$$(6.19) \quad 0 \leq \frac{a_n + b_0}{n} \leq \frac{1}{4} \quad \text{and} \quad 0 \leq \frac{2a_n}{n} \leq \frac{1}{4}.$$

We denote by \mathbf{L} the intensity measure of $\xi[\beta_n]$. For $B \in \mathcal{X}$ and $u \in (b_0, \min\{\varrho_n, a_n\})$ we have

$$\begin{aligned} \mathbf{L}(B \times (u, \infty)) &= n \int_B \mathbb{P}(\beta_{n-1}(B_{r_n(x,u)}(x)) \leq k-1) \lambda(x) dx \\ &= n \mathbf{K}(B) \sum_{i=0}^{k-1} \binom{n-1}{i} \frac{(a_n + u)^i}{n^i} \left(1 - \frac{a_n + u}{n}\right)^{n-1-i} \\ &= n \mathbf{K}(B) \sum_{i=0}^{k-1} \frac{(n-1)_i (a_n + u)^i}{n^i i!} \left(1 - \frac{a_n + u}{n}\right)^{n-1-i}. \end{aligned}$$

Thus \mathbf{L} has the density

$$\begin{aligned} \ell(x, u) &= -n\lambda(x) \sum_{i=1}^{k-1} \frac{(n-1)_i (a_n + u)^{i-1}}{n^i (i-1)!} \left(1 - \frac{a_n + u}{n}\right)^{n-1-i} \\ &\quad + n\lambda(x) \sum_{i=0}^{k-1} \frac{(n-1)_i n-1-i (a_n + u)^i}{n^i n i!} \left(1 - \frac{a_n + u}{n}\right)^{n-2-i} \\ &=: \lambda(x)q(u) \end{aligned}$$

for $x \in \mathbb{X}$ and $u \in (b_0, \min\{\varrho_n, a_n\})$, whence

$$\begin{aligned} & \left|q(u) - e^{-u}\right| \\ & \leq n \sum_{i=1}^{k-1} \frac{(a_n + u)^{i-1}}{(i-1)!} \left(1 - \frac{a_n + u}{n}\right)^{n-1-i} + n \sum_{i=0}^{k-2} \frac{(a_n + u)^i}{i!} \left(1 - \frac{a_n + u}{n}\right)^{n-2-i} \\ & \quad + \left|n \frac{(n-1)_{k-1}}{n^{k-1}} \frac{n-k}{n} \frac{(a_n + u)^{k-1}}{(k-1)!} \left(1 - \frac{a_n + u}{n}\right)^{n-k-1} - e^{-u}\right| \\ & =: R_1 + R_2 + R_3. \end{aligned}$$

From (6.18) and (6.19) we deduce that

$$n \left(1 - \frac{a_n + u}{n}\right)^{n-k-1} \leq \frac{4^{k+1}}{3^{k+1}} n e^{-(a_n+u)} = \frac{4^{k+1}(k-1)!}{3^{k+1}} \frac{e^{-u}}{(\log n)^{k-1}}.$$

Together with $|a_n + u|^j \leq 2^{j-1}(|a_n|^j + |u|^j)$ and $\sup_{v \in (b_0, \infty)} |v|^j e^{-v/2} < \infty$ for $j \in \mathbb{N}$ this yields

$$R_1 \leq \frac{c_1 e^{-u/2}}{\log n} \quad \text{and} \quad R_2 \leq \frac{c_2 e^{-u/2}}{\log n},$$

where $c_1, c_2 \in (0, \infty)$ depend only on b_0 and k . For R_3 we have

$$\begin{aligned} R_3 & \leq \left| \frac{(n-1)_{k-1}}{n^{k-1}} \frac{n-k}{n} - 1 \right| \frac{n(a_n + u)^{k-1}}{(k-1)!} \left(1 - \frac{a_n + u}{n}\right)^{n-k-1} \\ & \quad + \left| \frac{n(a_n + u)^{k-1}}{(k-1)!} - \frac{na_n^{k-1}}{(k-1)!} \right| \left(1 - \frac{a_n + u}{n}\right)^{n-k-1} \\ & \quad + \left| 1 - \left(1 - \frac{a_n + u}{n}\right)^{k+1} \right| \frac{na_n^{k-1}}{(k-1)!} \left(1 - \frac{a_n + u}{n}\right)^{n-k-1} \\ & \quad + \left| \frac{na_n^{k-1}}{(k-1)!} \left(1 - \frac{a_n + u}{n}\right)^n - e^{-u} \right|. \end{aligned}$$

By the same arguments as used for R_1 and R_2 and

$$\begin{aligned} & \left| \frac{na_n^{k-1}}{(k-1)!} \left(1 - \frac{a_n + u}{n}\right)^n - e^{-u} \right| \\ & \leq \frac{na_n^{k-1}}{(k-1)!} e^{-a_n-u} \frac{(a_n + u)^2}{n} + \left| \frac{na_n^{k-1}}{(k-1)!} e^{-a_n-u} - e^{-u} \right|, \end{aligned}$$

which follows from (6.18), we obtain

$$R_3 \leq \frac{c_3 e^{-u/2}}{n} + \frac{c_4 e^{-u/2}}{\log n} + \frac{c_5 e^{-u/2} a_n}{n} + \frac{c_6 e^{-u} a_n^2}{n} + \frac{c_7 e^{-u} \log \log n}{\log n}$$

with constants $c_3, c_4, c_5, c_6, c_7 \in (0, \infty)$ depending only on b_0 and k . Let $R(u)$ denote the sum of the upper bounds for R_1, R_2 and R_3 and let \mathbf{M}_0 be the intensity measure of $\zeta \cap (\mathbb{X} \times (b_0, \infty))$. The estimates above imply that

$$\begin{aligned}
 (6.20) \quad d_{TV}(\mathbf{L}, \mathbf{M}_0) &\leq \int_{b_0}^{\infty} R(u) \, du + \mathbf{L}(\mathbb{X} \times (\min\{\varrho_n, a_n\}, \infty)) + \mathbf{M}_0(\mathbb{X} \times (\min\{\varrho_n, a_n\}, \infty)) \\
 &\leq C'_0 \frac{\log \log n}{\log n}
 \end{aligned}$$

with a constant $C'_0 \in (0, \infty)$ depending on b_0, k, d and λ_- .

In the following we compute the terms E_1, \dots, E_6 from Theorem 5.1. Using the above expression for \mathbf{L} , we obtain that

$$\begin{aligned}
 (6.21) \quad E_1 &\leq 2n \int_B \mathbb{P}(\beta_{n-1}(B_{r_n(x,b)}(x)) \leq k-1) \lambda(x) \, dx \\
 &\leq C'_1 e^{-b} \left(1 + \frac{(k-1) \log \log n + b}{\log n}\right)^{k-1}
 \end{aligned}$$

with $C'_1 := 2k! \left(\frac{4}{3}\right)^{k+1}$. For $n \in \mathbb{N}$ and $x \in \mathbb{X}$ define $U_x = B_{r_n(x,b_0)}(x)$. It follows from (6.18) and (6.19) that

$$\begin{aligned}
 (6.22) \quad \mathbb{E}\{g(x, \beta_{n-1} + \delta_x)\} &\leq \mathbb{P}(\beta_{n-1}(U_x) \leq k-1) \\
 &= \sum_{i=0}^{k-1} \binom{n-1}{i} \left(\frac{a_n + b_0}{n}\right)^i \left(1 - \frac{a_n + b_0}{n}\right)^{n-1-i} \leq \frac{4^k}{3^k} \sum_{i=0}^{k-1} \frac{(a_n + b_0)^i}{i!} e^{-(a_n + b_0)} \leq \frac{\tilde{c}_1}{n}
 \end{aligned}$$

with a constant $\tilde{c}_1 \in (0, \infty)$ depending only on b_0 and k . For $x, y, y_1, y_2 \in \mathbb{X}$ with $S_x \cap U_{y_1} = \emptyset$ we obtain analogously that there exists a constant $\tilde{c}_2 \in (0, \infty)$ depending only on b_0 and k such that

$$\begin{aligned}
 (6.23) \quad \mathbb{E}\{g(x, \beta_{n-2} + \delta_x + \delta_y)\} &\leq \mathbb{E}\{g(x, \beta_{n-2} + \delta_x)\} \\
 &\leq \mathbb{P}(\beta_{n-2}(U_x) \leq k-1) \leq \frac{\tilde{c}_2}{n},
 \end{aligned}$$

$$\begin{aligned}
 (6.24) \quad \mathbb{E}\{g(y_1, \tilde{\beta}_{x,n-2} + \delta_{y_1} + \delta_{y_2})\} &\leq \mathbb{P}(\tilde{\beta}_{x,n-2}(U_{y_1}) \leq k-1) \\
 &\leq \mathbb{P}(\beta_{n-2}(U_{y_1}) \leq k-1) \leq \frac{\tilde{c}_2}{n}
 \end{aligned}$$

and

$$\begin{aligned}
 (6.25) \quad \mathbb{E}\{g(y_1, \tilde{\beta}_{x,n-1} + \delta_{y_1})\} &\leq \mathbb{P}(\tilde{\beta}_{x,n-1}(U_{y_1}) \leq k-1) \\
 &\leq \mathbb{P}(\tilde{\beta}_{x,n-2}(U_{y_1}) \leq k-1) \leq \frac{\tilde{c}_2}{n}.
 \end{aligned}$$

For $x, y \in \mathbb{X}$ and $i, j \in \{0, 1, \dots, k - 1\}$ we have

$$\begin{aligned}
 & \mathbb{P}(\beta_{n-2}(U_x) = i, \beta_{n-2}(U_y) = j) \\
 &= \sum_{s=0}^{\min\{i,j\}} \mathbb{P}(\beta_{n-2}(U_x \cap U_y^c) = i - s, \beta_{n-2}(U_x^c \cap U_y) = j - s, \beta_{n-2}(U_x \cap U_y) = s) \\
 &= \sum_{s=0}^{\min\{i,j\}} \frac{(n-2)_{i+j-s}}{(i-s)!(j-s)!s!} \mathbf{K}(U_x \cap U_y^c)^{i-s} \mathbf{K}(U_x^c \cap U_y)^{j-s} \\
 &\quad \times \mathbf{K}(U_x \cap U_y)^s (1 - \mathbf{K}(U_x \cup U_y))^{n-2-i-j+s} \\
 &\leq 4^k \sum_{s=0}^{\min\{i,j\}} \frac{(n\mathbf{K}(U_x \cap U_y^c))^{i-s}}{(i-s)!} \frac{(n\mathbf{K}(U_x^c \cap U_y))^{j-s}}{(j-s)!} \\
 &\quad \times \frac{(n\mathbf{K}(U_x \cap U_y))^s}{s!} \exp(-n\mathbf{K}(U_x \cup U_y)) \\
 &= 4^k \sum_{s=0}^{\min\{i,j\}} \mathbb{P}(\eta_n(U_x \cap U_y^c) = i - s, \eta_n(U_x^c \cap U_y) = j - s, \eta_n(U_x \cap U_y) = s) \\
 &= 4^k \mathbb{P}(\eta_n(U_x) = i, \eta_n(U_y) = j),
 \end{aligned}$$

where the inequality follows from (6.18) and (6.19). This implies that

$$\begin{aligned}
 (6.26) \quad & \mathbb{E}\{g(x, \beta_{n-2} + \delta_x + \delta_y) g(y, \beta_{n-2} + \delta_x + \delta_y)\} \\
 & \leq \mathbb{P}([\beta_{n-2} + \delta_y](U_x) \leq k - 1, [\beta_{n-2} + \delta_x](U_y) \leq k - 1) \\
 & \leq 4^k \mathbb{P}([\eta_n + \delta_y](U_x) \leq k - 1, [\eta_n + \delta_x](U_y) \leq k - 1)
 \end{aligned}$$

and if $S_x \cap S_y = \emptyset$,

$$\begin{aligned}
 (6.27) \quad & \mathbb{E}\{g(x, \beta_{n-2} + \delta_x) g(y, \beta_{n-2} + \delta_y)\} \\
 & = \mathbb{E}\{g(x, \beta_{n-2} + \delta_x + \delta_y) g(y, \beta_{n-2} + \delta_x + \delta_y)\} \\
 & \leq 4^k \mathbb{P}((\eta_n + \delta_y)(U_x) \leq k - 1, (\eta_n + \delta_x)(U_y) \leq k - 1) \\
 & \leq 4^k \mathbb{P}(\eta_n(U_x) \leq k - 1, \eta_n(U_y) \leq k - 1) \\
 & = 4^k \mathbb{P}(\eta_n(U_x) \leq k - 1) \mathbb{P}(\eta_n(U_y) \leq k - 1) \leq \frac{\tilde{c}_3}{n^2}
 \end{aligned}$$

with a constant $\tilde{c}_3 \in (0, \infty)$ depending only on b_0 and k . A similar computation shows that there exists a constant $\tilde{c}_4 \in (0, \infty)$ depending only on b_0 and k such that

$$\begin{aligned}
 (6.28) \quad & \mathbb{E}\{g(x, \beta_{n-3} + \delta_x + \delta_z) g(y, \beta_{n-3} + \delta_y)\} \\
 & \leq \mathbb{E}\{g(x, \beta_{n-3} + \delta_x) g(y, \beta_{n-3} + \delta_y)\} \\
 & \leq \frac{\tilde{c}_4}{n^2}
 \end{aligned}$$

for $x, y, z \in \mathbb{X}$ with $S_x \cap S_y = \emptyset$. In the sequel, we will use that $\tilde{g} \leq g$. It follows from (6.22) and

$$\begin{aligned} \mathbb{1} \{S_x \cap S_y \neq \emptyset\} &\leq \mathbb{1} \{|x - y| \leq r_n(x, b) + r_n(y, b)\} \\ &\leq \mathbb{1} \{|x - y| \leq 2r_n(x, b)\} + \mathbb{1} \{|x - y| \leq 2r_n(y, b)\} \end{aligned}$$

that

$$\begin{aligned} E_2 &\leq 2\tilde{c}_1^2 \int_{\mathbb{X}^2} \mathbb{1} \{S_x \cap S_y \neq \emptyset\} \mathbf{K}^2(d(x, y)) \\ (6.29) \quad &\leq 4\tilde{c}_1^2 \int_{\mathbb{X}} \mathbf{K}(B_{2r_n(x, b)}(x)) \mathbf{K}(dx) \\ &\leq 4\tilde{c}_1^2 2^d \frac{\lambda_+}{\lambda_-} \frac{a_n + b}{n}. \end{aligned}$$

From (6.26) and the same arguments as in the Poisson case (see (6.15) and (6.16)) we obtain that

$$\begin{aligned} E_3 &\leq 2 \cdot 4^k n^2 \int_{\mathbb{X}^2} \mathbb{1} \{S_x \cap S_y \neq \emptyset\} \mathbb{P}([\eta_n + \delta_y](U_x)) \\ (6.30) \quad &\leq k - 1, [\eta_n + \delta_x](U_y) \leq k - 1 \mathbf{K}^2(d(x, y)) \\ &\leq C'_{3,1} \frac{1}{(a_n + b_0)^d} + C'_{3,2} \frac{(a_n + b_0)^{k-1} (a_n + b)}{n^\gamma} \end{aligned}$$

with constants $C'_{3,1}, C'_{3,2}, \gamma \in (0, \infty)$ only depending on b_0, k, d, λ_+ and λ_- . Because of (6.22) and (6.23) we have, for $x \in \mathbb{X}$,

$$\begin{aligned} (1 + n\mathbf{K}(S_x)) \mathbb{E} \{g(x, \beta_{n-1} + \delta_x)\} + n \int_{S_x} \mathbb{E} \{g(x, \beta_{n-2} + \delta_x + \delta_z)\} \mathbf{K}(dz) \\ \leq \frac{\tilde{c}_1(1 + a_n + b)}{n} + \frac{\tilde{c}_2(a_n + b)}{n}. \end{aligned}$$

From $b \leq a_n$ and (6.19) we derive for all $x \in \mathbb{X}$ that $\mathbf{K}(S_x) \leq \frac{2a_n}{n} \leq \frac{1}{4}$, whence for $\mathbf{Q} = \mathbf{K}$ the measure \mathbf{Q}_x in Theorem 5.1 is dominated by $\frac{4}{3}\mathbf{K}$. Together with (6.24) and (6.25) this implies that

$$\begin{aligned} E_4 &\leq 2n \left(\frac{\tilde{c}_1(1 + a_n + b)}{n} + \frac{\tilde{c}_2(a_n + b)}{n} \right) \left(\frac{4}{3} \frac{\tilde{c}_2}{n} + \frac{16}{9} n \frac{\tilde{c}_2}{n} \int_{\mathbb{X}} \mathbf{K}(S_{y_1}) \mathbf{K}(dy_1) \right) \\ (6.31) \quad &\leq 2(\tilde{c}_1(1 + a_n + b) + \tilde{c}_2(a_n + b)) \left(\frac{4}{3} \frac{\tilde{c}_2}{n} + \frac{16}{9} \tilde{c}_2 \frac{a_n + b}{n} \right). \end{aligned}$$

By (6.22) and (6.23) we have

$$(6.32) \quad E_5 \leq \frac{8}{3}(a_n + b)^2 n \frac{\tilde{c}_1 \tilde{c}_2}{n^2} = \frac{8}{3} \tilde{c}_1 \tilde{c}_2 \frac{(a_n + b)^2}{n}.$$

It follows from (6.27) and (6.28) that

$$(6.33) \quad E_6 \leq 2n^2 \frac{\tilde{c}_3}{n^2} \frac{4(a_n + b)}{3n} + 2n^3 \frac{\tilde{c}_4}{n^2} \frac{4(a_n + b)^2}{3n^2} = \frac{8}{3} (\tilde{c}_3 + \tilde{c}_4(a_n + b)) \frac{a_n + b}{n}.$$

Choosing $b = a_n$ and combining the bounds (6.20), (6.21), (6.29), (6.30), (6.31), (6.32) and (6.33) with Theorem 5.1 proves (6.17). \square

Appendix A. Additional details for the proof of Theorem 6.2

Recall the definition of $g(\mathbf{x}, \eta)$ in (6.3),

$$g(\mathbf{x}, \eta) := \mathbb{1} \{ \mathbf{x} \subset \eta \text{ generates a critical point of index } k, \text{ with } \rho \in (r_n, R_n] \}.$$

Next, recall that $c(\mathbf{x})$ and $\rho(\mathbf{x})$ are defined as the center and radius of the unique $(k - 1)$ -dimensional sphere containing \mathbf{x} . In addition, define $\sigma(\mathbf{x})$ to be the k -dimensional simplex spanned by \mathbf{x} and let $B(\mathbf{x})$ be the open ball with center $c(\mathbf{x})$ and radius $r(\mathbf{x})$. Setting

$$\begin{aligned} h_{\text{crit}}(\mathbf{x}) &:= \mathbb{1} \{ c(\mathbf{x}) \in \sigma(\mathbf{x}) \}, \\ g_{\text{crit}}(\mathbf{x}, \eta) &:= \mathbb{1} \{ B(\mathbf{x}) \cap \eta = \emptyset \}, \end{aligned}$$

then according to [BW17, Lemma 2.4], \mathbf{x} generates a critical point of index k (for the distance function from η) if and only if $h_{\text{crit}}(\mathbf{x}) = g_{\text{crit}}(\mathbf{x}, \eta) = 1$. Further defining

$$h_r(\mathbf{x}) := h_{\text{crit}}(\mathbf{x}) \mathbb{1} \{ \rho(\mathbf{x}) \in (r, R_n] \},$$

we can write

$$(A.1) \quad g(\mathbf{x}, \eta) = g_{\text{crit}}(\mathbf{x}, \eta) h_{r_n}(\mathbf{x}).$$

Finally, recall the definition of ξ_k ,

$$\xi_k = \xi_k[\eta_n] := \frac{1}{(k + 1)!} \sum_{\mathbf{x} \in (\eta_n)_{\neq}^{k+1}} g(\mathbf{x}, \eta_n) \delta_{f(\mathbf{x})},$$

where $f(\mathbf{x}) = (c(\mathbf{x}), \alpha(\mathbf{x}))$, and

$$\alpha(\mathbf{x}) := \omega_d n \rho(\mathbf{x})^d - \log n - (k - 1) \log \log n.$$

LEMMA A.1. — *Let $A \subset \mathbb{Y} = \mathbb{T}^d \times \mathbb{R}_0$. Then,*

$$\mathbf{L}(A) = \mathbb{E} \{ \xi_k(A) \} = D_k n \int_{\hat{A}} s^{k-1} e^{-s} \mathbb{1} \{ s \leq \omega_d n R_n^d \} ds dc,$$

where

$$\hat{A} := \{ (c, s) : (c, \alpha) \in A, s = \log n + (k - 1) \log \log n + \alpha \}.$$

Proof. — Recall that,

$$\xi_k(A) = \frac{1}{(k + 1)!} \sum_{\mathbf{x} \in (\eta_n)_{\neq}^{k+1}} g(\mathbf{x}, \eta_n) \mathbb{1} \{ (c(\mathbf{x}), \alpha(\mathbf{x})) \in A \},$$

and using Mecke’s formula and (A.1),

$$(A.2) \quad \begin{aligned} \mathbf{L}(A) &= \frac{n^{k+1}}{(k + 1)!} \int_{(\mathbb{T}^d)^{k+1}} h_{r_n}(\mathbf{x}) \mathbb{1} \{ (c(\mathbf{x}), \alpha(\mathbf{x})) \in A \} \mathbb{E} \{ g_{\text{crit}}(\mathbf{x}, \eta_n + \delta_{\mathbf{x}}) \} d\mathbf{x} \\ &= \frac{n^{k+1}}{(k + 1)!} \int_{(\mathbb{T}^d)^{k+1}} h_{r_n}(\mathbf{x}) \mathbb{1} \{ (c(\mathbf{x}), \alpha(\mathbf{x})) \in A \} e^{-\omega_d n \rho^d(\mathbf{x})} d\mathbf{x}. \end{aligned}$$

In order to evaluate the last integral we will use the Blaschke–Petkantschin (BP) formula, discussed in [Bob22, Appendix C]. The main idea is to use a generalized polar-coordinate system, using the map $\mathbf{x} \mapsto (c(\mathbf{x}), \rho(\mathbf{x}), \Pi(\mathbf{x}), \theta(\mathbf{x}))$, where $c(\mathbf{x})$ and

$\rho(\mathbf{x})$ are defined above, $\Pi(\mathbf{x})$ is the affine k -dimensional plane containing \mathbf{x} , and $\theta(\mathbf{x})$ are the coordinates of \mathbf{x} on the sphere lying on $\Pi(\mathbf{x})$ with center $c(\mathbf{x})$ and radius $\rho(\mathbf{x})$. Notice that using these new variables, h_{crit} is independent of c, ρ, Π , and therefore we will denote $h_{\text{crit}}(\mathbf{x}) = h_{\text{crit}}(\theta)$. Thus, applying the BP-formula to the integral in (A.2), yields

$$D_{\text{bp}} \int_{\mathbb{T}^d} \int_{r_n}^{R_n} \mathbb{1} \{(c, \alpha(\rho)) \in A\} \rho^{dk-1} e^{-n\omega_d \rho^d} d\rho dc \int_{(\mathbb{S}^{k-1})^{k+1}} h_{\text{crit}}(\theta) (V_{\text{simp}}(\theta))^{d-k+1} d\theta,$$

where V_{simp} is the volume of the k -simplex spanned by θ , and abusing notation, we set

$$\alpha(\rho) := \omega_d n \rho^d - \log n - (k - 1) \log \log n.$$

Taking the change of variables $\omega_d n \rho^d \rightarrow s$ we have

$$\begin{aligned} \int_{\mathbb{T}^d} \int_{r_n}^{R_n} \mathbb{1} \{(c, \alpha(\rho)) \in A\} \rho^{dk-1} e^{-\omega_d n \rho^d} d\rho dc \\ = \frac{1}{d(n\omega_d)^k} \int_{\mathbb{T}^d} \int_{\omega_d n r_n^d}^{\omega_d n R_n^d} \mathbb{1} \{(c, s) \in \widehat{A}\} s^{k-1} e^{-s} ds dc. \end{aligned}$$

Defining

$$D_k := \frac{D_{\text{bp}}}{(k + 1)! d \omega_d^k} \int_{(\mathbb{S}^{k-1})^{k+1}} h_{\text{crit}}(\theta) (V_{\text{simp}}(\theta))^{d-k+1} d\theta,$$

and putting everything back into (A.2) completes the proof of Lemma A.1. □

LEMMA A.2. —

$$\text{Var}(\xi_k(\mathbb{Y})) - \mathbb{E}\{\xi_k(\mathbb{Y})\} = O\left((\log \log n)^d (\log n)^{-\frac{d-k}{d+1}}\right).$$

Proof. — We start by evaluating the second moment,

$$\begin{aligned} \mathbb{E}\{(\xi_k(\mathbb{Y}))^2\} &= \frac{1}{((k + 1)!)^2} \mathbb{E}\left\{ \sum_{\mathbf{x}_1, \mathbf{x}_2 \in (\eta_n)^{k+1}_{\neq}} g(\mathbf{x}_1, \eta_n) g(\mathbf{x}_2, \eta_n) \right\} \\ &= \frac{1}{((k + 1)!)^2} \sum_{j=0}^{k+1} \mathbb{E}\left\{ \sum_{|\mathbf{x}_1 \cap \mathbf{x}_2|=j} g(\mathbf{x}_1, \eta_n) g(\mathbf{x}_2, \eta_n) \right\} \\ &= \frac{1}{((k + 1)!)^2} \sum_{j=0}^{k+1} I_j, \end{aligned}$$

where by $\mathbf{x}_1 \cap \mathbf{x}_2$ we refer to the coordinates both tuples share in common. Therefore, we have

$$\text{Var}(\xi_k(\mathbb{Y})) - \mathbb{E}\{\xi_k(\mathbb{Y})\} = \sum_{j=1}^k I_j + (I_0 - \mathbb{E}\{\xi_k(\mathbb{Y})\}^2).$$

Recall the definition of a_n (6.1)

$$a_n := \omega_d n r_n^d = \log n + (k - 1) \log \log n + \alpha_0.$$

In [Bob22, Equation (8.15)] it is shown that for $1 \leq j \leq k$, and for small enough δ_j, ϵ_j we have

$$I_j \leq C_1 n a_n^{k-1} e^{-a_n} \left(\epsilon_j^{k+2-j} a_n^{k+1-j} + (\delta_j / \epsilon_j)^{d-k} (\delta_j a_n)^{k+1-j} + a_n^{k+1-j} e^{-C_2 \delta_j a_n} \right),$$

and in [Bob22, Equation (8.17)] it is shown that

$$I_0 - \mathbb{E} \{\xi_k(\mathbb{Y})\}^2 \leq C_3 n a_n^{k-1} e^{-a_n} \left(\epsilon_0^{d+2} a_n^{k+1} + a_n^{k+1} e^{-C_4 \epsilon_0 a_n} \right),$$

where C_1, \dots, C_4 are some positive constants. The choice of $\epsilon_j = a_n^{-(1-\frac{1}{d+2-j})}$ and $\delta_j = \frac{k+2-j}{C_2} \frac{\log a_n}{a_n}$ yields that for $1 \leq j \leq k$

$$\begin{aligned} I_j &\leq C_1 n a_n^{k-1} e^{-a_n} \left(a_n^{-\frac{d-k}{d+2-j}} + \left(\frac{k+2-j}{C_2} \log a_n \right)^{d-j+1} a_n^{-\frac{d-k}{d+2-j}} + a_n^{-1} \right) \\ &= O \left((\log a_n)^{d-j+1} a_n^{-\frac{d-k}{d+2-j}} \right) = O \left((\log \log n)^d (\log n)^{-\frac{d-k}{d+1}} \right). \end{aligned}$$

In addition, taking $\epsilon_0 = \frac{k+2}{C_4} \frac{\log a_n}{a_n}$ yields

$$I_0 - \mathbb{E} \{\xi_k(\mathbb{Y})\}^2 = O \left(a_n^{-1} \right) = o \left((\log \log n)^d (\log n)^{-\frac{d-k}{d+1}} \right).$$

This completes the proof of Lemma A.2. \square

BIBLIOGRAPHY

- [AGG89] Richard A. Arratia, Larry Goldstein, and Louis Gordon, *Two moments suffice for Poisson approximations: the Chen–Stein method*, Ann. Probab. **17** (1989), no. 1, 9–25. \uparrow 1513, 1519
- [BA14] Omer Bobrowski and Robert J. Adler, *Distance functions, critical points, and the topology of random čech complexes*, Homology Homotopy Appl. **16** (2014), no. 2, 311–344. \uparrow 1514
- [Bar88] Andrew D. Barbour, *Stein’s method and Poisson process convergence*, J. Appl. Probab. **25** (1988), no. A, 175–184, “A celebration of applied probability”, Spec. Vol. \uparrow 1490
- [BB92] Andrew D. Barbour and Timothy C. Brown, *Stein’s method and point process approximation*, Stochastic Processes Appl. **43** (1992), no. 1, 9–31. \uparrow 1490, 1491, 1493, 1496, 1497, 1501
- [BBK20] François Baccelli, Bartłomiej Błaszczyszyn, and Mohamed Karray, *Random Measures, Point Processes and Stochastic Geometry*, <https://hal.inria.fr/hal-02460214/>, 2020. \uparrow 1494, 1495, 1500
- [BHJ92] Andrew D. Barbour, Lars Holst, and Svante Janson, *Poisson approximation*, Oxford Studies in Probability, vol. 2, Oxford University Press, 1992. \uparrow 1490, 1491, 1496
- [BK18] Omer Bobrowski and Matthew Kahle, *Topology of random geometric complexes: a survey*, J. Appl. Comput. Topol. **1** (2018), no. 3–4, 331–364. \uparrow 1514
- [BL09] Volker Baumstark and Günter Last, *Gamma distributions for stationary Poisson flat processes*, Adv. Appl. Probab. **41** (2009), no. 4, 911–939. \uparrow 1500
- [BM22] Chinmoy Bhattacharjee and Ilya Molchanov, *Gaussian approximation for sums of region-stabilizing scores*, Electron. J. Probab. **27** (2022), article no. 111. \uparrow 1500
- [Bob22] Omer Bobrowski, *Homological Connectivity in Random čech Complexes*, Probab. Theory Relat. Fields **183** (2022), no. 3–4, 715–788. \uparrow 1493, 1514, 1515, 1516, 1529, 1530, 1531
- [BW17] Omer Bobrowski and Shmuel Weinberger, *On the vanishing of homology in random čech complexes*, Random Struct. Algorithms **51** (2017), no. 1, 14–51. \uparrow 1529
- [CC14] Pierre Calka and Nicolas Chenavier, *Extreme values for characteristic radii of a Poisson–Voronoi tessellation*, Extremes **17** (2014), no. 3, 359–385. \uparrow 1513

- [Che14] Nicolas Chenavier, *A general study of extremes of stationary tessellations with examples*, Stochastic Processes Appl. **124** (2014), no. 9, 2917–2953. ↑1513
- [CHO21] Nicolas Chenavier, Norbert Henze, and Moritz Otto, *Limit laws for large k -nearest neighbor balls*, <https://arxiv.org/abs/2105.00038>, 2021. ↑1523
- [CX04] Louis H. Chen and Aihua Xia, *Stein’s method, Palm theory and Poisson process approximation*, Ann. Probab. **32** (2004), no. 3B, 2545–2569. ↑1490, 1491, 1496
- [DST16] Laurent Decreusefond, Matthias Schulte, and Christoph Thäle, *Functional Poisson approximation in Kantorovich–Rubinstein distance with applications to U -statistics and stochastic geometry*, Ann. Probab. **44** (2016), no. 3, 2147–2197. ↑1490, 1491, 1492, 1493, 1494, 1495, 1496, 1498, 1499
- [DVJ07] Daryl J. Daley and David Vere-Jones, *An introduction to the theory of point processes: volume II: general theory and structure*, 2nd ed., Probability and Its Applications, Springer, 2007. ↑1496
- [ERS15] Peter Eichelsbacher, Martin Raič, and Tomacz Schreiber, *Moderate deviations for stabilizing functionals in geometric probability*, Ann. Inst. Henri Poincaré, Probab. Stat. **51** (2015), no. 1, 89–128. ↑1492
- [GHW19] Lázló Györfi, Norbert Henze, and Harro Walk, *The limit distribution of the maximum probability nearest-neighbour ball*, J. Appl. Probab. **56** (2019), no. 2, 574–589. ↑1523
- [GR97] V. Ya. Gershkovich and J. Hyam Rubinsten, *Morse theory for min-type functions*, Asian J. Math. **1** (1997), no. 4, 696–715. ↑1514
- [Hat02] Allen E. Hatcher, *Algebraic topology*, Cambridge University Press, 2002. ↑1514
- [IY20] Srikanth K. Iyer and D. Yogeshwaran, *Thresholds for vanishing of ‘Isolated’ faces in random \check{c} ech and Vietoris–Rips complexes*, Ann. Inst. Henri Poincaré, Probab. Stat. **56** (2020), no. 3, 1869–1897. ↑1499, 1513
- [Kah11] Matthew Kahle, *Random geometric complexes*, Discrete Comput. Geom. **45** (2011), no. 3, 553–573. ↑1513
- [Kal02] Olav Kallenberg, *Foundations of modern probability*, 2nd ed., Probability and Its Applications, Springer, 2002. ↑1516, 1523
- [LP17] Günter Last and Mathew Penrose, *Lectures on the Poisson Process*, Institute of Mathematical Statistics Textbooks, vol. 7, Cambridge University Press, 2017. ↑1494, 1495
- [LPY21] Günter Last, Giovanni Peccati, and D. Yogeshwaran, *Phase transitions and noise sensitivity on the Poisson space via stopping sets and decision trees*, <https://arxiv.org/abs/2101.07180>, 2021. ↑1500
- [LRSY19] Raphaël Lachièze-Rey, Matthias Schulte, and Joseph E. Yukich, *Normal approximation for stabilizing functionals*, Ann. Appl. Probab. **29** (2019), no. 2, 931–993. ↑1492
- [Mil63] John W. Milnor, *Morse theory. Based on lecture notes by M. Spivak and R. Wells*, Annals of Mathematics Studies, vol. 51, Princeton University Press, 1963. ↑1514
- [OA17] Takhashi Owada and Robert J. Adler, *Limit theorems for point processes under geometric constraints (and topological crackle)*, Ann. Probab. **45** (2017), no. 3, 2004–2055. ↑1513
- [Ott20] Moritz Otto, *Poisson approximation of Poisson-driven point processes and extreme values in stochastic geometry*, <https://arxiv.org/abs/2005.10116>, 2020. ↑1490, 1491, 1492, 1493, 1497, 1499, 1500, 1501, 1513, 1518, 1519
- [Ott21] ———, *Extremal behavior of large cells in the Poisson hyperplane mosaic*, <https://arxiv.org/abs/2106.14823>, 2021. ↑1502
- [Pen97] Matthew D. Penrose, *The longest edge of the random minimal spanning tree*, Ann. Appl. Probab. **7** (1997), no. 2, 340–361. ↑1518, 1519, 1523
- [Pen07] ———, *Laws of large numbers in stochastic geometry with statistical applications*, Bernoulli **13** (2007), no. 4, 1124–1150. ↑1500

- [Pen18] ———, *Inhomogeneous random graphs, isolated vertices, and Poisson approximation*, J. Appl. Probab. **55** (2018), no. 1, 112–136. ↑1492, 1498, 1499, 1513
- [PG10] Matthew D. Penrose and Larry Goldstein, *Normal approximation for coverage processes over binomial point processes*, Ann. Appl. Probab. **20** (2010), no. 2, 696–721. ↑1522
- [Pre75] Chris Preston, *Spatial birth and death processes*, Adv. Appl. Probab. **7** (1975), no. 3, 465–466. ↑1496
- [PS22] Federico Pianoforte and Matthias Schulte, *Criteria for Poisson process convergence with applications to inhomogeneous Poisson–Voronoi tessellations*, Stochastic Processes Appl. **147** (2022), 388–422. ↑1491, 1513
- [Ros11] Nathan Ross, *Fundamentals of Stein’s method*, Probab. Surv. **8** (2011), 210–293. ↑1496
- [Sch05] Dominic Schuhmacher, *Distance estimates for dependent superpositions of point processes*, Stochastic Processes Appl. **115** (2005), no. 11, 1819–1837. ↑1491, 1496
- [Sch09] ———, *Stein’s method and Poisson process approximation for a class of Wasserstein metrics*, Bernoulli **15** (2009), no. 2, 550–568. ↑1491, 1496
- [Sch10] Tomasz Schreiber, *Limit theorems in stochastic geometry*, New perspectives in stochastic geometry (Wilfrid S. Kendall and I. Molchanov, eds.), Oxford University Press, 2010, pp. 111–144. ↑1492
- [Xia05] Aihua Xia, *Stein’s method and Poisson process approximation*, An introduction to Stein’s method (A. D. Barbour and L. H. Y. Chen, eds.), Lecture Notes Series. Institute for Mathematical Sciences, vol. 4, Singapore University Press, Singapore, 2005, pp. 115–181. ↑1490, 1491, 1496
- [Yuk13] Joseph E. Yukich, *Limit theorems in discrete stochastic geometry*, Stochastic geometry, spatial statistics and random fields. Asymptotic methods (Evgeny Spodarev, ed.), Lecture Notes in Mathematics, vol. 2068, Springer, 2013, pp. 239–275. ↑1492

Manuscript received on 21st May 2021,
revised on 29th March 2022,
accepted on 24th June 2022.

Recommended by Editors S. Gouézel and N. Privault.
Published under license CC BY 4.0.



eISSN: 2644–9463

This journal is a member of Centre Mersenne.



Omer BOBROWSKI
Viterbi Faculty of Electrical and Computer Engineering
Technion – Israel Institute of Technology, Haifa, Israel
and
School of Mathematical Sciences
Queen Mary University of London, London, UK
omer@ee.technion.ac.il

Matthias SCHULTE
Institute of Mathematics
Hamburg University of Technology,

Hamburg (Germany)
matthias.schulte@tuhh.de

D. YOGESHWARAN
Theoretical Statistics
and Mathematics Unit
Indian Statistical Institute,
Bangalore (India)
d.yogesh@isibang.ac.in