Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry

Laurent Decreusefond, Matthias Schulte, Christoph Thäle

Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry

Laurent Decreusefond[*], Matthias Schulte[†] and Christoph Thäle[‡]

Abstract

A Poisson or a binomial process on an abstract state space and a symmetric function f acting on k-tuples of its points are considered. They induce a point process on the target space of f. The main result is a functional limit theorem which provides an upper bound for an optimal transportation distance between the image process and a Poisson process on the target space. The technical background are a version of Stein's method for Poisson process approximation, a Glauber dynamics representation for the Poisson process and the Malliavin formalism. As applications of the main result, error bounds for approximations of U-statistics by Poisson, compound Poisson and stable random variables are derived and examples from stochastic geometry are investigated.

Keywords. Binomial process, configuration space, functional limit theorem, Glauber dynamics, Kantorovich-Rubinstein distance, Malliavin formalism, Poisson process, Stein's method, stochastic geometry, U-statistics.
MSC2010. Primary 60F17, 60G55; Secondary 60D05, 60E07, 60H07.

1 Introduction

The arguably most prominent functional limit theorem is Donsker's invariance principle. It asserts that the distribution of a linear interpolation between the points of a suitably re-scaled random walk converges to the Wiener measure on the space of continuous functions on $\mathbb{R}_+$, the non-negative real half-line (see e.g. [24, Corollary 16.7]). Besides the Wiener process, there is another fundamental stochastic process, which plays an important rôle in many branches of probability theory and its applications, namely the Poisson process. However, functional limit theorems involving the Poisson process have found much less attention in the literature. The aim of this paper is to provide a quantitative version of a functional limit theorem for Poisson processes and to derive from it error bounds for the probabilistic approximation of U-statistics by a Poisson, a compound Poisson or a stable random variable. We demonstrate the versatility of our results by applying these bounds to functionals of random geometric graphs, distance-power statistics, non-intersecting flat processes and random polytopes.

Let us informally describe the set-up of this paper, precise definitions and statements follow in Section 3. Let $(\mathbb{X}, \mathcal{X})$ and $(\mathbb{Y}, \mathcal{Y})$ be two measurable spaces (satisfying some mild regularity assumptions, see below), let $\mathbf{K}_1$ be a probability measure on $\mathbb{X}$ and fix an integer $k \geq 1$. Moreover, for each $n \in \mathbb{N}$ let $f_n : \operatorname{dom} f_n \to \mathbb{Y}$ be a symmetric mapping whose domain $\operatorname{dom} f_n$ is a symmetric subset of $\mathbb{X}^k$. Next, consider a collection $\beta_n = \{X_1, \ldots, X_n\}$ of $n \geq k$ i.i.d. random elements $X_1, \ldots, X_n$ of $\mathbb{X}$ with distribution $\mathbf{K}_1$. We apply for each $n \geq k$, f_n to every k-tuple of distinct elements of β_n. This

[*]Telecom ParisTech, Rue Barrault 46, F-75634 Paris cedex 13, France. E-mail: laurent.decreusefond@telecom-paristech.fr

[†]Karlsruhe Institute of Technology, Department of Mathematics, Institute of Stochastics, D-76128 Karlsruhe, Germany. E-mail: matthias.schulte@kit.edu

[‡]Ruhr University Bochum, Faculty of Mathematics, NA 3/68, D-44780 Bochum, Germany. E-mail: christoph.thaele@rub.de

induces a point process ξ_n on $\mathbb{Y}$ of the form

$$\xi_n = \frac{1}{k!} \sum_{(x_1,\dots,x_k)\in\beta_{n,\neq}^k \cap \operatorname{dom} f_n} \delta_{f_n(x_1,\dots,x_k)},$$

where $\beta_{n,\neq}^k = \{(x_1,\dots,x_k) \in \beta_n^k : x_j \neq x_j,\ i \neq j, i, j = 1,\dots,k\}$ and δ_y stands for the unit mass Dirac measure concentrated at $y \in \mathbb{Y}$.

The motivation for studying the point processes ξ_n as defined above comes from the theory of U-statistics and from a class of extreme value problems arising in stochastic geometry. At first, if $\operatorname{dom} f_n = \mathbb{X}^k$ and $\mathbb{Y} = \mathbb{R}$, the points of ξ_n can be regarded as the summands of the U-statistic

$$S_n = \frac{1}{k!} \sum_{(x_1,\dots,x_k)\in\beta_{n,\neq}^k} f_n(x_1,\dots,x_k)\,.$$

These objects play a prominent rôle in large parts of probability theory and mathematical statistics, and an analysis of the point process of summands is helpful for the understanding of their (asymptotic) properties. On the other hand, in several problems arising in stochastic geometry one is interested in extreme values of the type

$$\min_{(x_1,\dots,x_k)\in\beta_{n,\neq}^k} f_n(x_1,\dots,x_k)$$

in case that $\operatorname{dom} f_n = \mathbb{X}^k$ and $\mathbb{Y} = [0,\infty)$. Clearly, this minimum is the distance from the origin to the first point of the point process ξ_n. For these reasons, a study of the point proceses ξ_n unifies both mentioned problems.

The intensity measure $\mathbf{L}_n$ of ξ_n is given by

$$\mathbf{L}_n(A) = \mathbf{E}\xi_n(A) = \frac{(n)_k}{k!} \int_{\operatorname{dom} f_n} \mathbf{1}(f_n(x_1,\dots,x_k) \in A)\, \mathbf{K}_1^k(\mathrm{d}(x_1,\dots,x_k))\,, \quad A \in \mathcal{Y}\,,$$

where $(n)_k$ is the descending factorial. Our main result, Theorem 3.1 below, provides an upper bound for the Kantorovich-Rubinstein distance $\mathbf{d_{KR}}(\xi_n,\zeta)$ between ξ_n and a Poisson process ζ on $\mathbb{Y}$ with finite intensity measure $\mathbf{M}$. Here, the Kantorovich-Rubinstein distance is a variant of an optimal transportation distance, which measures the closeness between two point processes or, more precisely, their distributions. In particular, we show that ξ_n converges in Kantorovich-Rubinstein distance and, thus, in distribution to ζ if

$$\mathrm{d_{TV}}(\mathbf{L}_n,\mathbf{M}) \to 0 \quad \text{and} \quad \mathbf{E}\xi_n(\mathbb{Y})^2 \to \mathbf{M}(\mathbb{Y})^2 + \mathbf{M}(\mathbb{Y})\,, \quad \text{as} \quad n\to\infty\,,$$

where $\mathrm{d_{TV}}(\,\cdot\,,\,\cdot\,)$ denotes the total variation distance of measures on $\mathbb{Y}$. More precisely, the upper bound for the Kantorovich-Rubinstein distance only depends on $\mathrm{d_{TV}}(\mathbf{L}_n,\mathbf{M})$ and the first two moments of $\xi_n(\mathbb{Y})$. This is a functional version of the famous results by Arratia, Chen, Goldstein and Gordon [1, 14] that "two moments suffice for Poisson approximation".

Besides the binomial process β_n of n independent and identically distributed points, we also allow the input process to be a Poisson process on $\mathbb{X}$ with a σ-finite intensity measure. In some instances, an underlying Poisson process is more natural and sometimes even unavoidable, especially if the underlying point process on $\mathbb{X}$ is supposed to have infinitely many points. To exploit this flexibility, we consider both set-ups in parallel.

Poisson process approximation has been studied by several authors by means of Stein's method, but to the best of our knowledge this is the first paper, where the Kantorovich-Rubinstein distance is investigated. The works of Barbour [2], Barbour and Brown [4] and the last chapter of the monograph [8] of Barbour, Holst and Janson concern Poisson process approximation in the total variation distance. But since the total variation distance is not suitable for all problems and since the so-called „Stein magic factors" do not get small if $\mathbf{L}_n(\mathbb{Y})$ is large (in contrast to classical Poisson approximation), one often uses weaker notions of distance. Starting with the work of Barbour and Brown [4] and

Barbour, Holst and Janson [8], this has been done by Brown, Chen, Schuhmacher, Weinberg and Xia [11, 12, 13, 15, 39, 40, 42]. Our work goes in the opposite direction since the Kantorovich-Rubinstein distance between point processes is stronger than the total variation distance in the sense that convergence in Kantorovich-Rubinstein distance implies convergence in total variation distance, but not vice versa. Roughly speaking and in a transferred sense, the Kantorovich-Rubinstein distance is related to the total variation distance between point processes as the Wasserstein distance is related to the total variation distance for integer-valued random variables. Since its test functions are allowed to take values different from zero and one, the Kantorovich-Rubinstein distance is more sensitive to the behaviour and the geometry of the compared point processes than the total variation distance. Let us further remark that in the recent paper [41], Schuhmacher and Stucki consider the total variation distance between two Gibbs processes. This includes Poisson process approximation as a special case. However, the approximated point processes of the present paper do in general no satisfy the technical conditions assumed in [41] since they are not necessarily hereditary.

Besides the notion of distance and its connection to the theory of optimal transportation, the other main ingredient of our approach is a functional version of Stein's method for Poisson process approximation. It relies on a Glauber dynamics representation for Poisson processes and the Malliavin formalism. More precisely, we use an integration-by-parts argument on the target space and then a commutation relation between the discrete gradient on that space and the semi-group associated with the Glauber dynamics. This way we avoid the explicit computation and investigation of a solution of the Stein equation. We would like to highlight that our approach is generic and depends only on the underlying random structure (here, a binomial or a Poisson process) and not on a very specific model so that extensions to other probabilistic frameworks (such as Gaussian random measures or Rademacher sequences) should also be possible. However, they are beyond the scope of this paper and will be treated elsewhere.

To demonstrate the versatility of our new functional limit theorem, we consider probabilistic approximations of U-statistics over binomial or Poisson input processes. In a first regime, we consider the Poisson approximation of U-statistics and provide an error bound for the Wasserstein distance. Our result improves and extends earlier works of Barbour and Eagleson [7], and Peccati [31]. The second regime concerns compound Poisson approximation of U-statistics in total variation distance. Here, we do not impose any conditions on the nature of the compound Poisson distribution, which is allowed to be discrete or continuous. In contrast, previous results for the compound Poisson approximation via Stein's method only deal with the discrete case, see, for example, the work of Barbour, Chen and Loh [5], the survey [6] of Barbour and Chryssaphinou and especially the paper [21] of Eichelsbacher and Roos, who consider U-statistics over a binomial input process. In this light, we generalize the results of [21] to a larger class of limiting distributions and also to the case of an underlying Poisson process. In a third regime, we use our functional limit theorem to investigate probabilistic approximations of U-statistics by α-stable random variables with $0 < \alpha < 1$ and to derive explicit error bounds for the Kolmogorov distance. In the previous work [17], Dabrowski, Dehling, Mikosch and Sharipov also obtained α-stable limits for U-statistics from point process convergence results. However, their technique does not allow any conclusion about a rate of convergence.

Finally, we apply our general result to problems arising in stochastic geometry. Random geometric graphs are one of the fundamental models of spatial stochastics, see [34], for example. We derive limit theorems for several U-statistics of random geometric graphs, where the limiting distributions are Poisson or compound Poisson, and show a new point process limit theorem for the midpoints of short edges. As further examples, we consider distance-power statistics with α-stable limit distributions, midpoints between non-intersecting Poisson m-flats which are close together and the diameter of random polytopes with vertices on the sphere.

In a natural way our paper continues the line of research on point process convergence and extreme values initiated by the second and the third author in [43, 44], where the proofs are based on the main result of [31] and the underlying point process has to be Poisson. In contrast to these previous works our technique also allows to deal with an underlying binomial process and delivers in both cases bounds for the Kantorovich-Rubinstein distance. Furthermore, the bounds derived here improve the rates of

convergence of some of the scalar limit theorems from [43, 44]. Our findings also complement the works [19] and [20] of the first author with Joulin and Savy concerning the Kantorovich-Rubinstein distance on configuration spaces and related notions.

This paper is organised as follows. Before we present our main result for Poisson process convergence in Section 3, we recall in Section 2 some necessary notation and results about point processes and also summarize some facts from convex geometry which are important for our examples from stochastic geometry. The proof of our main result in Section 6 is prepared by a brief discussion of the underlying Stein principle in Section 4 and the Glauber dynamics, a key step in our argument, in Section 5. Section 7 is devoted to applications of our functional limit theorem to probabilistic approximations of U-statistics and to problems from stochastic geometry.

2 Preliminaries

In the present section we introduce some basic notions and notation, which are used in the text. Throughout $(\Omega, \mathcal{F}, \mathbf{P})$ will be an abstract probability space, which is rich enough to carry all the random objects we deal with. Expectation with respect to $\mathbf{P}$ is denoted by $\mathbf{E}$.

2.1 Configuration spaces. Let $(\mathbb{Y}, \mathcal{Y})$ be a lcscH space, that is, $\mathbb{Y}$ is a topological space with countable base such that every point in $\mathbb{Y}$ has a compact neighbourhood and such that any two points of $\mathbb{Y}$ can be separated by disjoint neighbourhoods. Such a space is separable and completely metrizable. Here, $\mathcal{Y}$ denotes the Borel σ-field generated by the topology of $\mathbb{Y}$. By $\mathrm{N}_{\mathbb{Y}}$ we denote the space of σ-finite counting measures (i.e., point configurations) on $\mathbb{Y}$, whereas $\widetilde{\mathrm{N}}_{\mathbb{Y}}$ and $\widehat{\mathrm{N}}_{\mathbb{Y}}$ stand for the sets of all locally finite (i.e., bounded on all relatively compact sets) and finite counting measures on $\mathbb{Y}$, respectively. By a slight abuse of notation we will write $y \in \omega$ if $y \in \mathbb{Y}$ is charged by the measure ω and also use the set-notation $\omega_1 \subset \omega_2$ to indicate that ω_1 is a sub-configuration of ω_2 (with a similar meaning we also understand $\omega_2 \setminus \omega_1$). Let $\mathcal{N}_{\mathbb{Y}}$ be the σ-field on $\mathrm{N}_{\mathbb{Y}}$ generated by the mappings

$$\psi_A : \mathrm{N}_{\mathbb{Y}} \to \mathbb{N}_0 \cup \{\infty\}, \omega \mapsto \omega(A), \qquad A \in \mathcal{Y},$$

where $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ is the set of natural numbers including zero. We equip $\widetilde{\mathrm{N}}_{\mathbb{Y}}$ and $\widehat{\mathrm{N}}_{\mathbb{Y}}$ with the corresponding trace σ-fields of $\mathcal{N}_{\mathbb{Y}}$. The σ-field of $\widetilde{\mathrm{N}}_{\mathbb{Y}}$ is then the Borel σ-field for the vague topology on $\widetilde{\mathrm{N}}_{\mathbb{Y}}$, which is generated by the mappings

$$e_g : \widetilde{\mathrm{N}}_{\mathbb{Y}} \to [0, \infty), \omega \mapsto \int_{\mathbb{Y}} g \, \mathrm{d}\omega,$$

where $g \geq 0$ is a continuous function on $\mathbb{Y}$ with compact support. The space $\widetilde{\mathrm{N}}_{\mathbb{Y}}$ then becomes a Polish space, see Theorem A2.3 in [24]. A point process (or random counting measure) μ is a random element in $\mathrm{N}_{\mathbb{Y}}$. By a locally finite point process and a finite point process we mean random elements in $\widetilde{\mathrm{N}}_{\mathbb{Y}}$ and $\widehat{\mathrm{N}}_{\mathbb{Y}}$, respectively. It follows from [38, Lemma 3.1.3] that a point process μ can almost surely be represented as

$$\mu = \sum_{i=1}^{\mu(\mathbb{Y})} \delta_{x_i} \quad \text{with} \quad x_i \in \mathbb{Y}, \quad i \in \mathbb{N}, \quad \text{and} \quad \mu(\mathbb{Y}) \in \mathbb{N}_0 \cup \{+\infty\},$$

where δ_y stands for the unit mass Dirac measure concentrated at $y \in \mathbb{Y}$. Thus, we may interpret μ also as a random collection of points, taking into account potential multiplicities.

2.2 Poisson processes. Let $\mathbf{M}$ be a σ-finite measure on $\mathbb{Y}$ and let $\mathbf{M}^k$ stand for its k-fold product measure. By a Poisson process on $\mathbb{Y}$ with intensity measure $\mathbf{M}$, we understand a point process ζ with the properties that i) for any $B \in \mathcal{Y}$ the random variable $\zeta(B)$ is Poisson distributed with mean $\mathbf{M}(B)$ and that ii) ζ is independently scattered, i.e., for any $n \in \mathbb{N}$ and disjoint $B_1, \ldots, B_n \in \mathcal{Y}$ the random variables $\zeta(B_1), \ldots, \zeta(B_n)$ are independent. We notice that if $\mathbf{M}$ is a finite measure, ζ charges almost surely only a finite number of points in $\mathbb{Y}$, whose total number follows a Poisson

distribution with mean $\mathbf{M}(\mathbb{Y})$. We will write $\mathbf{P}_\zeta$ for the distribution of ζ on $\mathrm{N}_\mathbb{Y}$. In this paper we will speak about a homogeneous Poisson process on a set $A \in \mathcal{B}(\mathbb{R}^d)$, where $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ-field on $\mathbb{R}^d$, if the intensity measure is a multiple of the restriction of the Lebesgue measure to A. Also, if $d = 1$, a homogeneous Poisson process ζ on $[0, \infty)$ can be thought of as a piecewise deterministic (pure jump) stochastic process in continuous time, starting at zero and having jumps of size one and i.i.d. exponentially distributed waiting times between the jumps. The points of discontinuity of this random process are the jump times of ζ.

One of our main tools to deal with Poisson functionals (by these we mean real-valued random variables depending only on a Poisson process) is the multivariate Mecke formula [38, Corollary 3.2.3]. It says that for any integer $k \geq 1$ and any measurable and non-negative $f : \mathbb{Y}^k \times \mathrm{N}_\mathbb{Y} \to \mathbb{R}$,

$$\mathbf{E} \sum_{(y_1,\ldots,y_k)\in\zeta^k_{\neq}} f(y_1, \ldots, y_k, \zeta) = \int_{\mathbb{Y}^k} \mathbf{E} f(y_1, \ldots, y_k, \zeta + \delta_{y_1} + \ldots + \delta_{y_k})\, \mathbf{M}^k\big(\mathrm{d}(y_1, \ldots, y_k)\big)\,, \tag{2.1}$$

where $\zeta^k_{\neq}$ is the collection of all k-tuples of distinct points charged by ζ. If the point process ζ is simple (i.e., if $\zeta(\{y\}) \in \{0, 1\}$ almost surely for any $y \in \mathbb{Y}$), $\zeta^k_{\neq}$ can be written as

$$\zeta^k_{\neq} = \{(y_1, \ldots, y_k) \in \mathbb{Y}^k : y_i \neq y_j \in \zeta \text{ for } i \neq j,\ i, j = 1, \ldots, k\}\,,$$

while in the non-simple case distinct points can have the same location. We remark that (2.1) with $k = 1$ is even a characterizing property of the Poisson process ζ, cf. Theorem 3.2.5 of [38].

2.3 Binomial processes. Let $\mathbf{M}_1$ be a probability measure on $\mathbb{Y}$. A binomial process with intensity measure $\mathbf{M} := n\mathbf{M}_1$, $n \in \mathrm{N}$, is a collection of n random points, distributed independently according to the measure $\mathbf{M}_1$. This process also arises by conditioning a Poisson process with intensity measure $\mathbf{M}$ on having exactly n points. In this paper we shall denote the random counting measure induced by such a binomial process by β_n. We also write $\beta^k_{n,\neq}$ to indicate the collection of all k-tuples of distinct points charged by β_n. Then, the counterpart to the multivariate Mecke formula (2.1) for a binomial process reads as follows:

$$\begin{aligned}\mathbf{E} \sum_{(y_1,\ldots,y_k)\in\beta^k_{n,\neq}} & f(y_1, \ldots, y_k, \beta_n) \\ &= (n)_k \int_{\mathbb{Y}^k} \mathbf{E} f(y_1, \ldots, y_k, \beta_{n-k} + \delta_{y_1} + \ldots + \delta_{y_k})\, \mathbf{M}_1^k\big(\mathrm{d}(y_1, \ldots, y_k)\big)\,,\end{aligned} \tag{2.2}$$

where $(n)_k := n(n-1)\cdots(n-k+1)$ is the descending factorial and f is a real-valued non-negative measurable function on $\mathbb{Y}^k \times \mathrm{N}_\mathbb{Y}$. This can easily be seen directly and is also a special case of the Georgii-Nguyen-Zessin formula, for which we refer to [18, Proposition 15.5.II].

2.4 Probability distances. In order to compare two real-valued random variables Y_1 and Y_2 (or more precisely their distributions) and to measure their closeness, we use several probability distances in this paper. The Kolmogorov distance of Y_1 and Y_2 is given by

$$\mathbf{d_K}(Y_1, Y_2) := \sup_{z\in\mathbb{R}} |\mathbf{P}(Y_1 \leq z) - \mathbf{P}(Y_2 \leq z)|\,,$$

while the total variation distance is

$$\mathbf{d_{TV}}(Y_1, Y_2) := \sup_{A\in\mathcal{B}(\mathbb{R})} |\mathbf{P}(Y_1 \in A) - \mathbf{P}(Y_2 \in A)|\,,$$

where, recall, $\mathcal{B}(\mathbb{R})$ stands for the Borel σ-field on $\mathbb{R}$. If Y_1 and Y_2 are integer-valued random variables, we can re-write their total variation distance as

$$\mathbf{d_{TV}}(Y_1, Y_2) = \frac{1}{2} \sum_{k\in\mathbb{Z}} |\mathbf{P}(Y_1 = k) - \mathbf{P}(Y_2 = k)|\,.$$

Let us denote by Lip(1) the set of all functions $h : \mathbb{R} \to \mathbb{R}$ whose Lipschitz constant is at most one and define the Wasserstein distance of two real-valued random variables Y_1 and Y_2 by

$$\mathbf{d_W}(Y_1, Y_2) := \sup_{h \in \mathrm{Lip}(1)} |\mathbf{E}h(Y_1) - \mathbf{E}h(Y_2)| \,.$$

All these probability distances have the property that they imply convergence in distribution, meaning that for a sequence $(Y_n)_{n\in\mathbb{N}}$ of random variables convergence in distribution to another random variable Y is implied by

$$\mathbf{d_I}(Y_n, Y) \to 0 \quad \text{as} \quad n \to \infty \,, \qquad \text{for some } \mathbf{I} \in \{\mathbf{K}, \mathbf{TV}, \mathbf{W}\} \,. \tag{2.3}$$

Moreover, for integer-valued random variables Y_1 and Y_2 let us mention the general inequality

$$\mathbf{d_K}(Y_1, Y_2) \le \mathbf{d_{TV}}(Y_1, Y_2) \le \mathbf{d_W}(Y_1, Y_2) \,, \tag{2.4}$$

which directly follows from the definitions of the involved probability distances and the fact that Y_1 and Y_2 are concentrated on the integers. Note that (2.4) does not remain valid for general real-valued random variables.

2.5 Kantorovich-Rubinstein distance. We define the total variation distance between two measures ν_1 and ν_2 on $\mathbb{Y}$ by

$$\mathrm{d_{TV}}(\nu_1, \nu_2) := \sup_{\substack{A \in \mathcal{Y} \\ \nu_1(A), \nu_2(A) < \infty}} |\nu_1(A) - \nu_2(A)| \,,$$

a notion which should not be confused with the total variation distance between random variables introduced above. Note that $\mathrm{d_{TV}}(\nu_1, \nu_2)$ can in principle take any value in $[0, \infty]$.

We say that a map $h : \widetilde{\mathrm{N}}_{\mathbb{Y}} \to \mathbb{R}$ is 1-Lipschitz if

$$|h(\omega_1) - h(\omega_2)| \le \mathrm{d_{TV}}(\omega_1, \omega_2) \qquad \text{for all } \omega_1, \omega_2 \in \widetilde{\mathrm{N}}_{\mathbb{Y}} \,,$$

and denote by $\mathscr{L}_1$ the set of all these maps which are measurable.

The Kantorovich-Rubinstein distance between two probability measures $\mathbf{Q}_1$ and $\mathbf{Q}_2$ on $\mathrm{N}_{\mathbb{Y}}$ is defined as the optimal transportation cost

$$\mathbf{d_{KR}}(\mathbf{Q}_1, \mathbf{Q}_2) := \inf_{\mathbf{C} \in \Sigma(\mathbf{Q}_1, \mathbf{Q}_2)} \int_{\mathrm{N}_{\mathbb{Y}} \times \mathrm{N}_{\mathbb{Y}}} \mathrm{d_{TV}}(\omega_1, \omega_2)\, \mathbf{C}(\mathrm{d}(\omega_1, \omega_2)) \tag{2.5}$$

for the cost function $\mathrm{d_{TV}}(\,\cdot\,,\,\cdot\,)$, where $\Sigma(\mathbf{Q}_1, \mathbf{Q}_2)$ denotes the set of probability measures on $\mathrm{N}_{\mathbb{Y}} \times \mathrm{N}_{\mathbb{Y}}$ with first marginal $\mathbf{Q}_1$ and second marginal $\mathbf{Q}_2$ (i.e., couplings of $\mathbf{Q}_1$ and $\mathbf{Q}_2$). If $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are concentrated on $\widetilde{\mathrm{N}}_{\mathbb{Y}}$ there is at least one coupling $\mathbf{C} \in \Sigma(\mathbf{Q}_1, \mathbf{Q}_2)$ for which the infimum in (2.5) is attained according to [46, Theorem 4.1], and the Kantorovich duality theorem [46, Theorem 5.10] says that this minimum equals

$$\mathbf{d_{KR}}(\mathbf{Q}_1, \mathbf{Q}_2) = \sup \Big| \int_{\widetilde{\mathrm{N}}_{\mathbb{Y}}} h(\omega)\, \mathbf{Q}_1(\mathrm{d}\omega) - \int_{\widetilde{\mathrm{N}}_{\mathbb{Y}}} h(\omega)\, \mathbf{Q}_2(\mathrm{d}\omega) \Big| \,, \tag{2.6}$$

where the supremum is over all $h \in \mathscr{L}_1$ that are integrable with respect to $\mathbf{Q}_1$ and $\mathbf{Q}_2$.

By abuse of notation we will also write $\mathbf{d_{KR}}(\zeta_n, \zeta)$ instead of $\mathbf{d_{KR}}(\mathbf{Q}_n, \mathbf{Q})$ if the point process ζ_n on $\mathbb{Y}$ has distribution $\mathbf{Q}_n$ for any $n \ge 1$ and the point process ζ on $\mathbb{Y}$ has distribution $\mathbf{Q}$. Note that the integrability condition in (2.6) is automatically fulfilled for all $h \in \mathscr{L}_1$ if $\mathbf{E}\zeta_n(\mathbb{Y}) < \infty$ and $\mathbf{E}\zeta(\mathbb{Y}) < \infty$. The Kantorovich-Rubinstein distance is also called Wasserstein distance, Monge-Kantorovich distance or Rubinstein distance. For a detailed discussion of the terminology we refer to the bibliographic notes of Chapter 6 in [46].

The following result ensures that convergence of locally finite point processes in Kantorovich-Rubinstein distance implies convergence in distribution.

Proposition 2.1. *Assume that $(\zeta_n)_{n\in\mathbb{N}}$ is a sequence of locally finite point processes on $\mathbb{Y}$ and that ζ is another locally finite point process on $\mathbb{Y}$ such that $\mathbf{d_{KR}}(\zeta_n,\zeta)\to 0$, as $n\to\infty$. Then ζ_n converges in distribution to ζ, as $n\to\infty$.*

Proof. The structure of the vague topology on $\widetilde{\mathrm{N}}_{\mathbb{Y}}$ implies that it is necessary and sufficient to prove that for any continuous $g:\mathbb{Y}\to\mathbb{R}$ with compact support, the random variables $\int g\,\mathrm{d}\zeta_n$ converge in distribution to $\int g\,\mathrm{d}\zeta$, see [24, Theorem 16.16]. By (2.3), it is sufficient to show that for all Borel sets $B\subset\mathbb{R}$, we have that

$$\mathbf{E}e_{g,B}(\zeta_n)\to\mathbf{E}e_{g,B}(\zeta)\,,\quad \text{as}\quad n\to\infty\,,$$

where $e_{g,B}:\widetilde{\mathrm{N}}_{\mathbb{Y}}\to\mathbb{R}$, $\omega\mapsto\mathbf{1}(\int g\,\mathrm{d}\omega\in B)$. To show this, we notice that for each g and B as above the mapping $e_{g,B}$ belong to $\mathscr{L}_1$, whence

$$\left|\mathbf{E}e_{g,B}(\zeta_n)-\mathbf{E}e_{g,B}(\zeta)\right|\le\mathbf{d_{KR}}(\zeta_n,\zeta)$$

and the result follows. □

An alternative distance to measure the closeness of two point processes ζ_1 and ζ_2 on $\mathbb{Y}$ is the total variation distance

$$\mathbf{d_{TV}}(\zeta_1,\zeta_2):=\sup_{A\in\mathcal{N}_{\mathbb{Y}}}\left|\mathbf{P}(\zeta_1\in A)-\mathbf{P}(\zeta_2\in A)\right|.$$

It is always dominated by the Kantorovich-Rubinstein distance since

$$\begin{aligned}\mathbf{d_{TV}}(\zeta_1,\zeta_2)&=\sup_{A\in\mathcal{N}_{\mathbb{Y}}}\left|\inf_{\mathbf{C}\in\Sigma(\zeta_1,\zeta_2)}\int_{\mathrm{N}_{\mathbb{Y}}\times\mathrm{N}_{\mathbb{Y}}}\mathbf{1}(\omega_1\in A)-\mathbf{1}(\omega_2\in A)\,\mathbf{C}(\mathrm{d}(\omega_1,\omega_2))\right|\\&\le\inf_{\mathbf{C}\in\Sigma(\zeta_1,\zeta_2)}\int_{\mathrm{N}_{\mathbb{Y}}\times\mathrm{N}_{\mathbb{Y}}}\mathrm{d_{TV}}(\omega_1,\omega_2)\,\mathbf{C}(\mathrm{d}(\omega_1,\omega_2))=\mathbf{d_{KR}}(\zeta_1,\zeta_2)\,.\end{aligned}$$

The following example shows that convergence in Kantorovich-Rubinstein distance is strictly finer than convergence in total variation distance.

Example 2.2. Let ζ be a Poisson process on $\mathbb{Y}$ with finite intensity measure $\mathbf{M}$. Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of independent random elements in $\mathbb{Y}$ with distribution $\mathbf{M}(\mathbb{Y})^{-1}\mathbf{M}(\cdot)$ and let Z be a Bernoulli random variable such that $\mathbf{P}(Z=1)=p$ for some $p\in(0,1)$. Moreover, assume that ζ, $(X_i)_{i\in\mathbb{N}}$ and Z are independent. Now, we consider the point process

$$\zeta_{n,p}:=\zeta+\mathbf{1}(Z=1)\sum_{i=1}^{n}\delta_{X_i}\,.$$

Since ζ and $\zeta_{n,p}$ coincide on an event with probability $1-p$, we have that $\mathbf{d_{TV}}(\zeta,\zeta_{n,p})\le p$. By taking $h(\mu)=\mu(\mathbb{Y})$ as a test function in (2.6), we deduce that $\mathbf{d_{KR}}(\zeta,\zeta_{n,p})\ge np$. Taking $p_n=1/\sqrt{n}$ for p shows that

$$\mathbf{d_{TV}}(\zeta,\zeta_{n,p_n})\to 0\quad\text{and}\quad\mathbf{d_{KR}}(\zeta,\zeta_{n,p_n})\to\infty\,,\quad\text{as}\quad n\to\infty\,,$$

so that $(\zeta_{n,p_n})_{n\in\mathbb{N}}$ converges to ζ in total variation distance but not in Kantorovich-Rubinstein distance.

In the previous example the Kantorovich-Rubinstein distance is more strongly affected by the rare event that $\zeta\neq\zeta_{n,p_n}$ than the total variation distance since the class of test functions is larger and contains functions taking also values different from zero and one. As already mentioned in the introduction, one can say that the difference between the Kantorovich-Rubinstein distance and the total variation distance for point processes is similar to the difference between the Wasserstein and the total variation distance for integer-valued random variables. As particular example we cite the work [10], where Poisson approximation of random variables with respect to the Wasserstein distance has been considered, extending previous results for the total variation distance, see also Section 7.1 below.

2.6 A discrete gradient. For a counting measure $\omega \in \widetilde{\mathrm{N}}_{\mathbb{Y}}$ and a measurable function $h : \widetilde{\mathrm{N}}_{\mathbb{Y}} \to \mathbb{R}$ let us introduce the discrete gradient in direction $y \in \mathbb{Y}$ by

$$D_y h(\omega) := h(\omega + \delta_y) - h(\omega)\,,$$

where, recall, δ_y is the unit-mass Dirac measure charging $y \in \mathbb{Y}$. In our notation we often suppress the dependence of $D_y h(\omega)$ on the underlying counting measure ω and write $D_y h$. Clearly, if $h \in \mathscr{L}_1$, it holds that $|D_y h| \le 1$ for all $y \in \mathbb{Y}$.

2.7 Geometric preparations. For our applications in Section 7, we need some facts from convex geometry. The Euclidean norm in $\mathbb{R}^d$ is denoted by $\|\cdot\|$. The Euclidean distance between two sets $A_1, A_2 \subset \mathbb{R}^d$ is given by

$$\mathrm{dist}(A_1, A_2) = \inf\{\|x_1 - x_2\| : x_1 \in A_1, x_2 \in A_2\}\,.$$

If $A_1 = \{x\}$ with $x \in \mathbb{R}^d$, we write $\mathrm{dist}(x, A_2)$ instead of $\mathrm{dist}(\{x\}, A_2)$. For a measurable set $K \subset \mathbb{R}^d$, we write $\mathrm{vol}(K)$ for the volume (i.e., d-dimensional Lebesgue measure) of K. For the volume of the unit ball $B^d = \{x \in \mathbb{R}^d : \|x\| \le 1\}$ in $\mathbb{R}^d$, we introduce the abbreviation $\kappa_d := \mathrm{vol}(B^d)$. More generally, $B^d(x, r)$ will denote the closed d-dimensional ball of radius $r > 0$ centred at $x \in \mathbb{R}^d$ and we write $B^d(r)$ instead of $B^d(0, r)$ for short. For $r \ge 0$, the Minkowski sum $K_r = K + rB^d$ of K and rB^d is the so-called r-parallel set of K. In particular, if K is a convex set with non-empty interior, Steiner's formula (see e.g. [38, Equation (14.5)]) says that the volume $\mathrm{vol}(K_r)$ is a polynomial of degree d in r. Formally,

$$\mathrm{vol}(K_r) = \sum_{i=0}^{d} \kappa_{d-i}\, V_i(K)\, r^{d-i}\,. \tag{2.7}$$

The coefficients $V_0(K), \ldots, V_d(K)$ are the so-called intrinsic volumes of K, especially $V_0(K) = 1$ whenever $K \neq \emptyset$, $V_1(K)$ is a constant multiple of the mean width of K, $V_{d-1}(K)$ is half of the surface area of K (if K is the closure of its interior) and $V_d(K) = \mathrm{vol}(K)$, cf. [38, Chapter 14.2].

For $1 \le m \le d-1$, we denote in this paper by $\mathbb{G}_m^d$ the space of m-dimensional linear subspaces and by $\mathbb{A}_m^d$ the space of m-dimensional affine subspaces of $\mathbb{R}^d$. For $L, M \in \mathbb{G}_m^d$ let $[L, M]$ be the subspace determinant of L and M, that is the $2m$-volume of a parallelepiped spanned by two orthonormal bases in L and in M. In one of our examples, we will also deal with the integrated subspace determinant and for this reason we recall that

$$\int_{\mathbb{G}_m^d} \int_{\mathbb{G}_m^d} [L, M]\, \mathrm{d}L\, \mathrm{d}M = \frac{\binom{d-m}{m}}{\binom{d}{m}} \frac{\kappa_{d-m}^2}{\kappa_d \kappa_{d-2m}} \tag{2.8}$$

from [23, Lemma 4.4]. Here, $\mathrm{d}L$ and $\mathrm{d}M$ indicate integration with respect to the unique Haar probability measure on $\mathbb{G}_m^d$.

3 Main results

3.1 General estimate

Let $(\mathbb{Y}, \mathcal{Y})$ be a lcscH space and let us fix another lcscH space $(\mathbb{X}, \mathcal{X})$. We adopt the notation introduced in Section 2 and denote by $\mathrm{N}_{\mathbb{X}}$ the space of σ-finite counting measures on $\mathbb{X}$.

Let μ be a point process on $\mathbb{X}$ with a σ-finite intensity measure $\mathbf{K}(\cdot) := \mathbf{E}\mu(\cdot)$. Fix an integer $k \ge 1$ and let $f : \mathrm{dom}\, f \to \mathbb{Y}$ be a symmetric and measurable function, where $\mathrm{dom}\, f$ is a symmetric subset of $\mathbb{X}^k$, i.e., if $(x_1, \ldots, x_k) \in \mathrm{dom}\, f$, then $(x_{\sigma(1)}, \ldots, x_{\sigma(k)}) \in \mathrm{dom}\, f$ for all permutations σ of $\{1, \ldots, k\}$. We now apply f to all k-tuples of distinct points of μ contained in $\mathrm{dom}\, f$ to form a point process ξ, i.e.,

$$\xi := \frac{1}{k!} \sum_{(x_1,\ldots,x_k) \in \mu^k_{\neq} \cap \mathrm{dom}\, f} \delta_{f(x_1,\ldots,x_k)}\,.$$

Since f is symmetric, every $f(x_1,\dots,x_k)$ also appears for the $k!$ permutations of the argument $(x_1,\dots,x_k)$. However, for each subset $\{x_1,\dots,x_k\}\subset\mu$ of distinct points of μ we assign to $f(x_1,\dots,x_k)$ only multiplicity one as can be seen from the above definition of ξ. But ξ might still have points of multiplicity greater than one if there are different combinations of k points in $\mathbb{X}$ that are mapped under f to the same point in $\mathbb{Y}$. The intensity measure of ξ is denoted by $\mathbf{L}$ and is given by

$$\mathbf{L}(A)=\mathbf{E}\xi(A)=\mathbf{E}\sum_{y\in\xi}\mathbf{1}(y\in A)=\frac{1}{k!}\,\mathbf{E}\sum_{(x_1,\dots,x_k)\in\mu^k_{\neq}\cap\operatorname{dom}f}\mathbf{1}(f(x_1,\dots,x_k)\in A)\,,\quad A\in\mathcal{Y}\,.$$

In what follows, we consider for μ two different types of point processes, namely Poisson processes and binomial processes. By η we denote a Poisson process on $\mathbb{X}$ with a σ-finite intensity measure $\mathbf{K}$. By β_n we denote a binomial process of $n\in\mathbb{N}$ points in $\mathbb{X}$, which are independent and identically distributed in $\mathbb{X}$ according to a probability measure $\mathbf{K}_1$ on $\mathbb{X}$. Such a binomial process β_n has intensity measure $\mathbf{K}:=n\mathbf{K}_1$. Now, the multivariate Mecke formula (2.1) and its binomial analogue (2.2) imply that the intensity measure $\mathbf{L}$ of ξ is given by

$$\mathbf{L}(A)=\frac{1}{k!}\int_{\operatorname{dom}f}\mathbf{1}(f(x_1,\dots,x_k)\in A)\,\mathbf{K}^k(\mathrm{d}(x_1,\dots,x_k))\,,\quad A\in\mathcal{Y}\,,\tag{3.1}$$

in the Poisson case and by

$$\mathbf{L}(A)=\frac{(n)_k}{k!}\int_{\operatorname{dom}f}\mathbf{1}(f(x_1,\dots,x_k)\in A)\,\mathbf{K}_1^k(\mathrm{d}(x_1,\dots,x_k))\,,\quad A\in\mathcal{Y}\,,\tag{3.2}$$

if we start with a binomial process (to deal with both cases simultanously we use the same notation for both set-ups). Let us finally introduce $r(\operatorname{dom}f)$ for $k\geq 2$ by

$$r(\operatorname{dom}f):=\max_{1\leq\ell\leq k-1}\int_{\mathbb{X}^\ell}\Big(\int_{\mathbb{X}^{k-\ell}}\mathbf{1}((x_1,\dots,x_k)\in\operatorname{dom}f)\,\mathbf{K}^{k-\ell}(\mathrm{d}(x_{\ell+1},\dots,x_k))\Big)^2\,\mathbf{K}^\ell(\mathrm{d}(x_1,\dots,x_\ell))\,,$$

and, for $k=1$, put $r(\operatorname{dom}f):=0$. Moreover, we use the convention that $(n-k)_k/(n)_k:=0$ if $n<k$.

We can now state our main result, a functional limit theorem, which provides a bound on the Kantorovich-Rubinstein distance between ξ and a suitable Poisson process on $\mathbb{Y}$.

Theorem 3.1. *Let ζ be a Poisson process on $\mathbb{Y}$ with finite intensity measure $\mathbf{M}$. If ξ is induced by the Poisson process η, then*

$$\begin{aligned}\mathbf{d_{KR}}(\xi,\zeta)&\leq\mathrm{d_{TV}}(\mathbf{L},\mathbf{M})+2\big(\mathbf{E}\xi(\mathbb{Y})^2-\mathbf{E}\xi(\mathbb{Y})-(\mathbf{E}\xi(\mathbb{Y}))^2\big)\\&\leq\mathrm{d_{TV}}(\mathbf{L},\mathbf{M})+\frac{2^{k+1}}{k!}r(\operatorname{dom}f)\,.\end{aligned}\tag{3.3}$$

If otherwise ξ is derived from the binomial process β_n, then

$$\begin{aligned}\mathbf{d_{KR}}(\xi,\zeta)&\leq\mathrm{d_{TV}}(\mathbf{L},\mathbf{M})+2\Big(\mathbf{E}\xi(\mathbb{Y})^2-\mathbf{E}\xi(\mathbb{Y})-\frac{(n-k)_k}{(n)_k}(\mathbf{E}\xi(\mathbb{Y}))^2\Big)+\frac{6^k k!}{n}\,(\mathbf{E}\xi(\mathbb{Y}))^2\\&\leq\mathrm{d_{TV}}(\mathbf{L},\mathbf{M})+\frac{2^{k+1}}{k!}r(\operatorname{dom}f)+\frac{6^k k!}{n}\,\mathbf{L}(\mathbb{Y})^2\,.\end{aligned}$$

Remark 3.2. (i) If the underlying point process is a binomial process β_n with n points and if $n<k$, the point process ξ is empty with probability one and $\mathbf{L}\equiv 0$. In this case, $\mathbf{d_{KR}}\big(\xi,\zeta\big)\leq\mathbf{E}\zeta(\mathbb{Y})=\mathrm{d_{TV}}(\mathbf{L},\mathbf{M})$ and the bound on $\mathbf{d_{KR}}\big(\xi,\zeta\big)$ is trivially valid. For this reason, no further restriction on n is necessary.

(ii) In the Poisson case, it can happen that $\mathbf{L}(\mathbb{Y})=\infty$. In this case, we have $\mathrm{d_{TV}}(\mathbf{L},\mathbf{M})=\infty$ and the bound (3.3) is trivial. Hence, Theorem 3.1 is only of interest if $\mathbf{L}(\mathbb{Y})<\infty$, which is equivalent to $\mathbf{K}^k(\operatorname{dom}f)<\infty$, a condition which ensures that ξ is almost surely finite.

(iii) Taking $\mathbf{M} = \mathbf{L}$ in the Poisson case in Theorem 3.1 shows that

$$\mathbf{d_{KR}}(\xi, \zeta) \leq 2\big(\mathbf{E}\xi(\mathbb{Y})^2 - \mathbf{E}\xi(\mathbb{Y}) - (\mathbf{E}\xi(\mathbb{Y}))^2\big) \leq \frac{2^{k+1}}{k!} r(\operatorname{dom} f)\,.$$

In particular, if $k = 1$, this gives $\mathbf{d_{KR}}(\xi, \zeta) = 0$, which in view of Proposition 2.1 implies that ξ is a Poisson process. This is consistent with the well known mapping theorem for Poisson processes, for which we refer to [25, Chapter 2.3].

(iv) If $\mathbb{X} = \mathbb{Y}$ and $f : \mathbb{X} \to \mathbb{X}$ is the identity, Theorem 3.1 yields that, for Poisson processes ξ and ζ with finite intensity measures $\mathbf{L}$ and $\mathbf{M}$, respectively,

$$\mathbf{d_{KR}}(\xi, \zeta) \leq \mathrm{d_{TV}}(\mathbf{L}, \mathbf{M})\,.$$

In other words, the Kantorovich-Rubinstein distance between two Poisson processes is bounded by the total variation distance of their intensity measures. For a similar estimate in a more restricted set-up we refer to [20, Proposition 4.1].

3.2 The Euclidean case

In this subsection we shall apply our general estimate of Theorem 3.1 to the important situation that the target space $\mathbb{Y}$ is $\mathbb{R}^d$ endowed with the standard Borel σ-field $\mathcal{B}(\mathbb{R}^d)$. This is tailored towards some of our applications in Section 7 and is similar to the set-up in [43, 44]. We let $(\mathbb{X}, \mathcal{X})$ be a lcscH space and let $(\eta_t)_{t\geq 1}$ be a family of Poisson processes in $\mathbb{X}$ with intensity measures $\mathbf{K}_t = t\mathbf{K}$, $t \geq 1$, where $\mathbf{K}$ is a fixed σ-finite measure. By $(\beta_t)_{t\geq 1}$ we denote a family of binomial processes such that $\beta_t = \beta_{\lceil t\rceil}$ and $\beta_{\lceil t\rceil}$ is a process of $\lceil t\rceil$ points chosen independently according to a fixed probability measure $\mathbf{K}_1$. In this situation we use the notation $\mathbf{K}_t := \lceil t\rceil\mathbf{K}$. We write $(\mu_t)_{t\geq 1}$ in the sequel to indicate either $(\eta_t)_{t\geq 1}$ or $(\beta_t)_{t\geq 1}$.

For a fixed integer $k \geq 1$ we consider symmetric and measurable functions $f_t : \mathbb{X}^k \to \mathbb{R}^d$, $t \geq 1$. We are interested in the behaviour of the derived point processes

$$\xi_t := \frac{1}{k!} \sum_{(x_1,\ldots,x_k)\in\mu_{t,\neq}^k} \delta_{f_t(x_1,\ldots,x_k)}\,, \quad t \geq 1\,.$$

For this reason, we consider the re-scaled point processes

$$t^\gamma \bullet \xi_t := \frac{1}{k!} \sum_{(x_1,\ldots,x_k)\in\mu_{t,\neq}^k} \delta_{t^\gamma f_t(x_1,\ldots,x_k)}\,, \quad t \geq 1\,,$$

where $\gamma \in \mathbb{R}$ is a suitable constant. In order to compare $t^\gamma \bullet \xi_t$ with a Poisson process, we need to introduce the following notation. The intensity measure $\mathbf{L}_t$ of the re-scaled point process $t^\gamma \bullet \xi_t$ is given by

$$\mathbf{L}_t(B) := \frac{1}{k!}\,\mathbf{E} \sum_{(x_1,\ldots,x_k)\in\mu_{t,\neq}^k} \mathbf{1}\big(f_t(x_1,\ldots,x_k) \in t^{-\gamma}B\big)\,, \quad B \in \mathcal{B}(\mathbb{R}^d)\,.$$

For $B \in \mathcal{B}(\mathbb{R}^d)$ let $r_t(B)$ be given by $r_t(B) := 0$ for $k = 1$ and

$$r_t(B) := \max_{1\leq\ell\leq k-1} \int_{\mathbb{X}^\ell} \Big(\int_{\mathbb{X}^{k-\ell}} \mathbf{1}(f_t(x_1,\ldots,x_k) \in t^{-\gamma}B)\, \mathbf{K}_t^{k-\ell}(\mathrm{d}(x_{\ell+1},\ldots,x_k))\Big)^2 \mathbf{K}_t^\ell(\mathrm{d}(x_1,\ldots,x_\ell))$$

for $k \geq 2$. Furthermore, for a measure ν on $\mathbb{R}^d$ and $B \in \mathcal{B}(\mathbb{R}^d)$ let $\nu|_B$ be the restriction of ν to B.

Corollary 3.3. *Let ζ be a Poisson process on $\mathbb{R}^d$ with intensity measure $\mathbf{M}$ and let $B \in \mathcal{B}(\mathbb{R}^d)$ be such that $\mathbf{M}(B) < \infty$. If ξ_t is induced by a Poisson process η_t with $t \geq 1$, then*

$$\begin{aligned}\mathbf{d_{KR}}\big((t^\gamma \bullet \xi_t)|_B, \zeta|_B\big) &\leq \mathrm{d_{TV}}(\mathbf{L}_t|_B, \mathbf{M}|_B) + 2\big(\mathbf{E}\xi_t(t^{-\gamma}B)^2 - \mathbf{E}\xi_t(t^{-\gamma}B) - (\mathbf{E}\xi_t(t^{-\gamma}B))^2\big)\\ &\leq \mathrm{d_{TV}}(\mathbf{L}_t|_B, \mathbf{M}|_B) + \frac{2^{k+1}}{k!} r_t(B)\,.\end{aligned}$$

If ξ_t is induced by a binomial process β_t with $t \geq 1$, then

$$\begin{aligned}\mathbf{d_{KR}}\big((t^\gamma \bullet \xi_t)|_B, \zeta|_B\big) &\leq \mathrm{d_{TV}}(\mathbf{L}_t|_B, \mathbf{M}|_B) + 2\Big(\mathbf{E}\xi_t(t^{-\gamma}B)^2 - \mathbf{E}\xi_t(t^{-\gamma}B) - \frac{(\lceil t\rceil - k)_k}{(\lceil t\rceil)_k}(\mathbf{E}\xi_t(t^{-\gamma}B))^2\Big) \\ &\qquad + \frac{6^k k!}{t}(\mathbf{E}\xi_t(t^{-\gamma}B))^2 \\ &\leq \mathrm{d_{TV}}(\mathbf{L}_t|_B, \mathbf{M}|_B) + \frac{2^{k+1}}{k!}r_t(B) + \frac{6^k k!}{t}\mathbf{L}_t(B)^2\,.\end{aligned}$$

Proof. This is a direct consequence of Theorem 3.1 with $t^\gamma \bullet \xi_t|_B$ instead of ξ and $\zeta|_B$ instead of ζ there. □

In view of limit theorems, the most natural choice for $\mathbf{M}$ is to take $\mathbf{M}$ as the strong limit of the measures $\mathbf{L}_t$, as $t \to \infty$. That is,

$$\mathbf{M}(B) = \lim_{t\to\infty} \mathbf{L}_t(B) \qquad \text{for all } B \in \mathcal{B}(\mathbb{R}^d)\,.$$

However, we emphasize that this does not necessarily imply that $\mathrm{d_{TV}}(\mathbf{L}_t, \mathbf{M}) \to 0$, as $t \to \infty$, even though this is true for our applications presented below.

Remark 3.4. (i) The upper bounds in Corollary 3.3 are not uniform in the sense that they depend on the set B. This was to be expected since the re-scaled point processes $t^\gamma \bullet \xi_t$ can be finite for any $t \geq 1$, while a realization of ζ can charge an infinite number of points (compare with our applications in Section 7). This is the reason for introducing the restriction to the set B, which allows us to compare $t^\gamma \bullet \xi_t|_B$ with $\zeta|_B$ using the Kantorovich-Rubinstein distance.

(ii) To allow for an easier comparison with the previous paper [43], we remark that ibidem the Poisson case for $d = 1$ is considered. Moreover, the intensity measure $\mathbf{M}$ there is concentrated on the positive real half-axis and has the form

$$\mathbf{M}(B) = a\,b\int_B \mathbf{1}(u \geq 0)\,u^{b-1}\,\mathrm{d}u\,, \qquad B \in \mathcal{B}(\mathbb{R})\,,$$

for some constants $a, b > 0$. In this case, the Poisson process ζ is a so-called Weibull process since the distance from the origin to the closest point of ζ is Weibull distributed with distribution function $u \mapsto (1 - \exp(-a\,u^b))\,\mathbf{1}(u > 0)$. We remark that this form of $\mathbf{M}$ was tailored to the applications in [43], a more general version is stated without proof in [44].

(iii) Note that $r_t(B)$ is dominated by $\mathbf{L}_t(B)\hat{r}_t(B)$, where $\hat{r}_t(B)$ is defined as

$$\hat{r}_t(B) := \max_{\substack{1\leq\ell\leq k-1,\\ (x_1,\ldots,x_\ell)\in\mathbb{X}^\ell}} \mathbf{K}_t^{k-\ell}\big(\{(y_1,\ldots,y_{k-\ell}) \in \mathbb{X}^{k-\ell} : f_t(x_1,\ldots,x_\ell,y_1,\ldots,y_{k-\ell}) \in t^{-\gamma}B\}\big)$$

for $B \in \mathcal{B}(\mathbb{R}^d)$. A quantity similar to $\hat{r}_t(B)$ has also played a prominent rôle in the previous study [43]. In many applications a bound for $\hat{r}_t(B)$ is already sufficient in order to apply Corollary 3.3. However, there are situations for which $\hat{r}_t(B)$ is an increasing function in t, while $r_t(B)$ tends to zero, as $t \to \infty$. This way, [43, Theorem 1.1], in which $\hat{r}_t$ instead of r_t appears, is not applicable in such cases as erroneously done in Sections 2.5 and 2.6 ibidem. However, in these specific cases it is readily checked that r_t behaves nicely, implying that the results there are correct.

4 A general Stein principle

This section is devoted to a more informal discussion about the method of bounding the Kantorovich-Rubinstein distance between point processes using a Stein principle. This approach is the key argument of our proof of Theorem 3.1 in Section 6. Recall that the aim is to provide an upper bound for the Kantorovich-Rubinstein distance between a Poisson process ζ on a space $\mathbb{Y}$ with finite intensity measure

$\mathbf{M}$ and a second point process ξ on $\mathbb{Y}$, which in turn is derived from another point process μ on a space $\mathbb{X}$ by a transformation.

The first part of Stein's method consists in characterizing the target object, here the Poisson process ζ. The way is to consider a functional operator L which, at a formal level, satisfies for a finite point process ν the identity

$$\mathbf{E}[LF(\nu)] = 0 \qquad \text{for a large class of functions } F : \widehat{\mathrm{N}}_{\mathbb{Y}} \to \mathbb{R} \tag{4.1}$$

if and only if ν is a Poisson process with intensity measure $\mathbf{M}$. It is usually not difficult to construct such an operator for a given target object. What may become challenging, especially in infinite dimensions (compare with [3, 16, 45]), is to prove that the target object is the unique solution of (4.1). In our case, uniqueness follows from the theory of spatial birth-death processes, see [35].

The second step of Stein's method is to solve the so-called Stein equation

$$LF(\omega) = \mathbf{E}h(\zeta) - h(\omega)\,, \qquad \omega \in \widehat{\mathrm{N}}_{\mathbb{Y}}\,, \tag{4.2}$$

for a certain class of test functions $h : \widehat{\mathrm{N}}_{\mathbb{Y}} \to \mathbb{R}$. This means that we have to compute a solution F_h for a given test function h and to evaluate $LF_h(\omega)$.

A prominent method to do this is the so-called generator approach (see the survey article [36] and the references cited therein). The underlying idea is to interpret L as infinitesimal generator of a Markov process with the distribution of ζ as its invariant distribution, whence L satisfies (4.1). If $(P_s)_{s\geq 0}$ is the semi-group associated with this Markov process, one can show that

$$LF_h(\omega) = \int_0^\infty LP_s h(\omega)\,\mathrm{d}s\,, \qquad \omega \in \widehat{\mathrm{N}}_{\mathbb{Y}}\,. \tag{4.3}$$

In order to compare the point process ξ with ζ, we put $\omega = \xi$ and take expectations in (4.2) and (4.3). This leads to

$$\mathbf{E}h(\zeta) - \mathbf{E}h(\xi) = \mathbf{E}LF_h = \mathbf{E}\int_0^\infty LP_s h(\xi)\,\mathrm{d}s\,.$$

To derive this identity rigorously is the content of the subsequent section. In the context of our main result, the point process ξ is induced by an underlying point process μ on another space $\mathbb{X}$. More, formally we have that $\xi = T(\mu)$, where T is a suitable transformation, i.e., a mapping from $\mathrm{N}_{\mathbb{X}}$ to $\widehat{\mathrm{N}}_{\mathbb{Y}}$. Hence, we will have to compute

$$\mathbf{E}\int_0^\infty LP_s h(T(\mu))\,\mathrm{d}s\,.$$

This expression is bounded in Section 6 by exploiting the special structure of the transformation T and the fact that μ is a Poisson or binomial process.

5 Glauber dynamics for the Poisson process

We now specialize the general scheme outlined in Section 4 to our particular situation. Although the approach is similar to [4, Section 2], for example, we prefer to carry out the details here since we consider a different class of test functions, namely Lipschitz functions instead of bounded functions. We assume the same set-up as for Theorem 3.1, that is, ζ is a Poisson process on a lcscH space $\mathbb{Y}$ with a finite intensity measure $\mathbf{M}$ and distribution $\mathbf{P}_\zeta$. We now construct a Glauber dynamics for $\mathbf{P}_\zeta$, that is a continuous-time Markov process $(G(s))_{s\geq 0}$ with state space $\widehat{\mathrm{N}}_{\mathbb{Y}}$ and $\mathbf{P}_\zeta$ as its stationary (i.e., invariant) distribution, see [35]. Its generator L is given by

$$Lh(\omega) := \int_{\mathbb{Y}} h(\omega+\delta_y) - h(\omega)\,\mathbf{M}(\mathrm{d}y) + \int_{\mathbb{Y}} h(\omega-\delta_y) - h(\omega)\,\omega(\mathrm{d}y)\,, \quad \omega \in \widehat{\mathrm{N}}_{\mathbb{Y}}\,, \tag{5.1}$$

where $h : \widehat{\mathrm{N}}_{\mathbb{Y}} \to \mathbb{R}$ is a measurable and bounded function. According to our notational convention, L may be re-written as

$$Lh(\omega) = \int_{\mathbb{Y}} h(\omega+\delta_y) - h(\omega)\,\mathbf{M}(\mathrm{d}y) + \sum_{y\in\omega} \big(h(\omega-\delta_y) - h(\omega)\big)\,.$$

Note that $Lh(\omega)$ is well-defined for all $h \in \mathscr{L}_1$ and $\omega \in \widehat{\mathrm{N}}_{\mathbb{Y}}$ since the Lipschitz property implies that the integrands in (5.1) are bounded by one. Moreover, we notice that the operator L uniquely determines the process $(G(s))_{s\geq 0}$, which has $\mathbf{P}_\zeta$ as its unique invariant distribution, see [18, Proposition 10.4.VII] or [35].

The Markov process $(G(s))_{s\geq 0}$ is a spatial birth-death process in continuous time whose dynamics can be described as follows. If at time s, the system is in state ω_s, each particle charged by ω_s dies at rate 1 and a new particle is born at y with rate $\mathbf{M}(\mathrm{d}y)$. Alternatively, imagine a homogeneous Poisson process ζ_b on $\mathbb{R}_+$ with intensity $\mathbf{M}(\mathbb{Y})$. The jump times of ζ_b determine the birth times of the particles in ζ. At each jump of ζ_b a new particle is born and is placed in $\mathbb{Y}$ according to the distribution $\mathbf{M}(\,\cdot\,)/\mathbf{M}(\mathbb{Y})$, independently of the current configuration. Moreover, each particle has a lifetime which is exponentially distributed with parameter 1, independent of the past and of the rest of the configuration, see again [35].

The semi-group $(P_s)_{s\geq 0}$ associated with the Markov process $(G(s))_{s\geq 0}$ is defined as

$$P_s h(\omega) = \mathbf{E}[\, h(G(s)) \,|\, G(0) = \omega \,], \qquad \omega \in \widehat{\mathrm{N}}_{\mathbb{Y}}\,, \quad h : \widetilde{\mathrm{N}}_{\mathbb{Y}} \to \mathbb{R}\,. \tag{5.2}$$

For $h \in \mathscr{L}_1$ and $\omega \in \widehat{\mathrm{N}}_{\mathbb{Y}}$ the conditional expectation is always well defined since

$$\begin{aligned} \big|P_s h(\omega)\big| &= \big|\mathbf{E}[\, h(G(s)) \,|\, G(0) = \omega \,]\big| \\ &\leq \mathbf{E}[\, \big|\, h(G(s)) - h(\omega)\,\big| \,|\, G(0) = \omega \,] + \big|h(\omega)\big| \\ &\leq \mathbf{E}[\, \mathrm{d_{TV}}(G(s), \omega) \,|\, G(0) = \omega \,] + \big|h(\omega)\big| \\ &\leq \mathbf{E}\zeta_b([0,s]) + \omega(\mathbb{Y}) + \big|h(\omega)\big| < \infty\,, \end{aligned}$$

where ζ_b is the homogeneous Poisson process from the description of the birth-death dynamics above. Below we will need the following lemmas about the process $(G(s))_{s\geq 0}$ and its semi-group $(P_s)_{s\geq 0}$. The first one provides a commutation relation between the discrete gradient and the semi-group.

Lemma 5.1. *For any $s \geq 0$, $\omega \in \widehat{\mathrm{N}}_{\mathbb{Y}}$, $y \in \mathbb{Y}$ and $h \in \mathscr{L}_1$,*

$$D_y P_s h(\omega) = e^{-s}\, P_s(D_y h)(\omega)\,.$$

Proof. To construct a sample path of $(G(s))_{s\geq 0}$, given the initial configuration $G(0) = \omega + \delta_y$, we have to add the independent particle y to a realization of $(G(s))_{s\geq 0}$ starting from the initial configuration ω. These two realizations will be identical after the particle y has died. Thus, denoting by $\ell(y)$ the lifetime of y and using (5.2), we can write

$$\begin{aligned} D_y P_s h(\omega) &= \mathbf{E}[h(G(s)) \,|\, G(0) = \omega + \delta_y] - \mathbf{E}[h(G(s)) \,|\, G(0) = \omega] \\ &= \mathbf{E}[(h(G(s) + \delta_y) - h(G(s)))\, \mathbf{1}(\ell(y) \geq s) \,|\, G(0) = \omega]\,. \end{aligned}$$

Since $\ell(y)$ is independent of everything else and is exponentially distributed with mean one, we can continue with

$$D_y P_s h(\omega) = \mathbf{E}[\mathbf{1}(\ell(y) \geq s)]\, \mathbf{E}[(h(G(s) + \delta_y) - h(G(s))) \,|\, G(0) = \omega] = e^{-s}\, P_s(D_y h)(\omega)\,,$$

where we have used (5.2) again. This completes the proof. □

Lemma 5.2. *Let $\omega_1, \omega_2 \in \widehat{\mathrm{N}}_{\mathbb{Y}}$ with $\omega_2 \subset \omega_1$. If $h \in \mathscr{L}_1$ and $s \geq 0$, then*

$$\big|\mathbf{E}[h(G(s)) \,|\, G(0) = \omega_1] - \mathbf{E}[h(G(s)) \,|\, G(0) = \omega_2]\big| \leq (\omega_1 \setminus \omega_2)(\mathbb{Y})\, e^{-s}\,.$$

Proof. Recall that each particle y of the initial configuration $G(0)$ has an exponentially distributed

lifetime $\ell(y)$ with mean one. Thus, since $h \in \mathscr{L}_1$, it holds that

$$\begin{aligned}
&\big|\mathbf{E}[h(G(s)) \,|\, G(0) = \omega_1] - \mathbf{E}[h(G(s)) \,|\, G(0) = \omega_2]\big| \\
&\quad \le \mathbf{E}\big[\big|h(G(s) + \sum_{y \in \omega_1 \setminus \omega_2} \mathbf{1}(\ell(y) \ge s)\delta_y) - h(G(s))\big| \,|\, G(0) = \omega_2\big] \\
&\quad \le \mathbf{E}\big[\mathrm{d_{TV}}(G(s) + \sum_{y \in \omega_1 \setminus \omega_2} \mathbf{1}(\ell(y) \ge s)\delta_y, G(s)) \,|\, G(0) = \omega_2\big] \\
&\quad \le \mathbf{E} \sum_{y \in \omega_1 \setminus \omega_2} \mathbf{1}(\ell(y) \ge s) \\
&\quad = (\omega_1 \setminus \omega_2)(\mathbb{Y})\, e^{-s}\,,
\end{aligned}$$

which proves the claim. □

Lemma 5.3. *For any $\omega \in \widehat{\mathrm{N}}_{\mathbb{Y}}$ and $h \in \mathscr{L}_1$,*

$$\lim_{s \to \infty} P_s h(\omega) = \mathbf{E}h(\zeta) = \int h \,\mathrm{d}\mathbf{P}_\zeta\,.$$

Proof. We notice first that the expectation on the right-hand side is well defined since $h \in \mathscr{L}_1$ implies that

$$\mathbf{E}|h(\zeta)| \le \mathbf{E}|h(\zeta) - h(\emptyset)| + |h(\emptyset)| \le \mathbf{E}\mathrm{d_{TV}}(\zeta, \emptyset) + |h(\emptyset)| \le \mathbf{E}\zeta(\mathbb{Y}) + |h(\emptyset)| = \mathbf{M}(\mathbb{Y}) + |h(\emptyset)|\,,$$

where $\emptyset$ stands for the counting measure that corresponds to the empty point configuration.

From Lemma 5.2 with $\omega_1 = \omega$ and $\omega_2 = \emptyset$ we have that

$$\big|\mathbf{E}\big[h(G(s)) \,|\, G(0) = \omega\big] - \mathbf{E}\big[h(G(s)) \,|\, G(0) = \emptyset\big]\big| \le \omega(\mathbb{Y})\, e^{-s}\,. \tag{5.3}$$

The number of particles of $G(s)$ starting from the empty configuration follows the evolution of an M/M/∞ queue with arrival (birth) rate $\mathbf{M}(\mathbb{Y})$ and service (death) rate 1, and thus is Poisson distributed with parameter $(1 - e^{-s})\mathbf{M}(\mathbb{Y})$. Since the position of each of the particles is independent of everything else, $G(s)$ has the same distribution as a Poisson process on $\mathbb{Y}$ with intensity measure $(1 - e^{-s})\mathbf{M}$. Since ζ has the same distribution as the superposition of two independent Poisson processes with intensity measures $(1 - e^{-s})\mathbf{M}$ and $e^{-s}\mathbf{M}$, respectively, we obtain that

$$|\mathbf{E}[h(G(s)) \,|\, G(0) = \emptyset] - \mathbf{E}h(\zeta)| \le e^{-s}\mathbf{M}(\mathbb{Y})\,. \tag{5.4}$$

Combining (5.3) and (5.4) and letting $s \to \infty$ completes the proof. □

The next lemma, which can be seen as an integration by parts formula, is the key for the proof of Theorem 3.1 given in Section 6 below.

Lemma 5.4. *If $h \in \mathscr{L}_1$ and $\omega \in \widehat{\mathrm{N}}_{\mathbb{Y}}$, then*

$$\mathbf{E}h(\zeta) - h(\omega) = \int_0^\infty L P_s h(\omega)\,\mathrm{d}s\,. \tag{5.5}$$

Proof. For an arbitrary $h \in \mathscr{L}_1$ we define $h_n : \widehat{\mathrm{N}}_{\mathbb{Y}} \to \mathbb{R}$, $n \in \mathbb{N}$, by

$$h_n(\omega) = \begin{cases} n & : h(\omega) > n \\ h(\omega) & : -n \le h(\omega) \le n \\ -n & : h(\omega) < -n\,. \end{cases}$$

Clearly, each of the functions h_n is bounded and belongs to $\mathscr{L}_1$. Since h_n is bounded, the forward-backward equation stated as Theorem 12.22 in [24] implies that

$$P_t h_n(\omega) - h_n(\omega) = \int_0^t L P_s h_n(\omega)\,\mathrm{d}s\,, \quad t \ge 0\,. \tag{5.6}$$

By construction, we have $h_n(\omega) \to h(\omega)$, as $n \to \infty$. The dominated convergence theorem implies that $P_s h_n(\omega) \to P_s h(\omega)$ and $LP_s h_n(\omega) \to LP_s h(\omega)$, as $n \to \infty$, for all $s \geq 0$. By (5.1) and Lemma 5.1, we have that, for $g = h$ or $g = h_n$ and $s \geq 0$,

$$\begin{aligned} |LP_s g(\omega)| &\leq \int_{\mathbb{Y}} e^{-s} |P_s(D_y g)(\omega)| \, \mathbf{M}(\mathrm{d}y) + \int_{\mathbb{Y}} e^{-s} |P_s(D_y g)(\omega - \delta_y)| \, \omega(\mathrm{d}y) \\ &\leq e^{-s}(\mathbf{M}(\mathbb{Y}) + \omega(\mathbb{Y}))\,. \end{aligned} \tag{5.7}$$

In the last step we used the fact that $|P_s(D_y g)| \leq 1$. Now, a further application of the dominated convergence theorem shows that

$$\lim_{n\to\infty} \int_0^t LP_s h_n(\omega)\,\mathrm{d}s = \int_0^t LP_s h(\omega)\,\mathrm{d}s\,, \quad t \geq 0\,,$$

so that, letting $n \to \infty$ in (5.6), yields

$$P_t h(\omega) - h(\omega) = \int_0^t LP_s h(\omega)\,\mathrm{d}s\,, \quad t \geq 0\,. \tag{5.8}$$

Because of (5.7) and the dominated convergence theorem, the right-hand side of (5.8) converges to the right-hand side of (5.5), as $t \to \infty$. Together with Lemma 5.3 for the left-hand side, this completes the proof. □

Remark 5.5. The operator L and the associated semi-group $(P_s)_{s\geq 0}$ on the Poisson space can be also defined via the Wiener-Itô chaos expansion, which we recall now for completeness. We still denote by ζ a Poisson process with intensity measure $\mathbf{M}$ on a lcscH space $\mathbb{Y}$. A crucial property of ζ is that any square integrable functional $F \in L^2(\mathbf{P}_\zeta)$ of ζ can be written as

$$F = \mathbf{E}F + \sum_{n=1}^{\infty} I_n(f_n) \tag{5.9}$$

with

$$f_n(y_1, \ldots, y_n) = \frac{1}{n!}\,\mathbf{E}D^n_{y_1,\ldots,y_n} F(\zeta)\,, \quad y_1, \ldots, y_n \in \mathbb{Y}\,, \quad n \geq 1\,,$$

where $D^n := D \circ D^{n-1}$ with $D^1 := D$ is the n-th iteration of the discrete gradient D introduced in Section 2, and where $I_n(f_n)$ stands for the n-fold Wiener-Itô integral of the square integrable and symmetric function f_n with respect to the signed random measure $\zeta - \mathbf{M}$. Moreover, the series in (5.9) converges in $L^2(\mathbf{P}_\zeta)$ and is called the Wiener-Itô chaos expansion of F (we refer to [28] for further details). We can now define the Ornstein-Uhlenbeck generator L on the Poisson space by

$$LF = -\sum_{n=1}^{\infty} n I_n(f_n)\,,$$

whenever F belongs to $\operatorname{dom} L$, i.e., F is such that $\sum_{n=1}^{\infty} n^2\, n! \|f_n\|^2_{L^2(\mathbf{M}^n)} < \infty$, where $\|\cdot\|_{L^2(\mathbf{M}^n)}$ stands for the usual norm in $L^2(\mathbf{M}^n)$. We remark that LF can equivalently be written as in (5.1) as a consequence of identity (3.19) in [28] and of the relation stated in [33, Lemma 2.11] between the discrete gradient, the Ornstein-Uhlenbeck generator and the so-called Skorohod-integral on the Poisson space, another operator, which is not needed in the sequel. In [27] the relation between the inverse of the Ornstein-Uhlenbeck generator and the associated semi-group is investigated. The semi-group $(P_s)_{s\geq 0}$ can be written in terms of the Wiener-Itô chaos expansion as

$$P_s F = \mathbf{E}F + \sum_{n=1}^{\infty} e^{-ns} I_n(f_n)\,, \qquad s \geq 0\,,$$

where $F \in \operatorname{dom} L$ is assumed to have a chaotic expansion as in (5.9) (see, for example, [27, Equation (3.13)]). Lemma 5.1 is a special case of [27, Lemma 3.1] and Lemma 5.2, Lemma 5.3 and Lemma 5.4 can also be derived via the approach sketched in this remark. However, we preferred to give proofs not relying on Wiener-Itô chaos expansions rather than on trajectorial properties.

Remark 5.6. In [41] a spatial birth-death process is constructed whose invariant distribution is a Gibbs process. This includes the birth-death process in the present paper as a special case and the generator in [41] is a generalization of the generator in (5.1). However, the results in [41] do not cover the results of this section since only the test functions for the total variation distance are considered, while we use Lipschitz functions, which are needed for the Kantorovich-Rubinstein distance.

6 Proof of Theorem 3.1

Before going into the details of the proof of Theorem 3.1, we explain the strategy informally in case of an underlying Poisson process η. Applying the multivariate Mecke formula (2.1) in Equation (6.4) below, we are lead to estimate the integral with respect to $\mathbf{K}^k$ of

$$\mathbf{E}\big[F(\xi(\eta+\delta_{x_1}+\ldots+\delta_{x_k})-\delta_{f(x_1,\ldots,x_k)})-F(\xi(\eta+\delta_{x_1}+\ldots+\delta_{x_k}))\big]\,, \qquad x_1,\ldots,x_k\in\mathbb{X}\,, \tag{6.1}$$

with $F:\widehat{\mathrm{N}}_{\mathbb{Y}}\to\mathbb{R}$ being a certain point process functional and where we write $\xi(\mu)$ instead of ξ to underpin the dependence of ξ on the underlying point configuration μ. The difficulty comes from the fact that adding $\delta_{x_1}+\ldots+\delta_{x_k}$ to the Poisson process η, amounts not only in adding $\delta_{f(x_1,\ldots,x_k)}$ to $\xi(\eta)$ but also all atoms of the form $f(x_{i_1},\ldots,x_{i_\ell},\tilde{x}_{\ell+1},\ldots,\tilde{x}_k)$ with $\ell\in\{1,\ldots,k\}$, pairwise different indices $i_1,\ldots,i_\ell\in\{1,\ldots,k\}$ and $(\tilde{x}_{\ell+1},\ldots,\tilde{x}_k)\in\eta^{k-\ell}_{\neq}$. We denote by $\hat{\xi}(x_1,\ldots,x_k,\eta)$ the collection of these extra atoms. The difference in (6.1) is now decomposed as

$$\begin{aligned}\mathbf{E}\big[&\big(F(\xi(\eta)+\hat{\xi}(x_1,\ldots,x_k,\eta))-F(\xi(\eta))\big)+\big(F(\xi(\eta))-F(\xi(\eta)+\delta_{f(x_1,\ldots,x_k)})\big)\\&+\big(F(\xi(\eta)+\delta_{f(x_1,\ldots,x_k)})-F(\xi(\eta)+\hat{\xi}(x_1,\ldots,x_k,\eta)+\delta_{f(x_1,\ldots,x_k)})\big)\big]\,.\end{aligned} \tag{6.2}$$

The middle term in (6.2) contributes to the total variation distance of the intensity measures in (3.3) in Theorem 3.1. Since F is Lipschitz, the expectation and the integral with respect to $x_1,\ldots,x_k$ of the first and the third term in (6.2) are bounded (up to a constant) by

$$\mathbf{E}\int_{\mathbb{X}^k}\hat{\xi}(x_1,\ldots,x_k,\eta)(\mathbb{Y})\,\mathbf{K}^k(\mathrm{d}(x_1,\ldots,x_k))\,,$$

which in turn is bounded by $\mathbf{E}\xi(\mathbb{Y})^2-\mathbf{E}\xi(\mathbb{Y})-(\mathbf{E}\xi(\mathbb{Y}))^2$ and $r(\mathrm{dom}\,f)$. This effect contributes to the second term of the bounds in Theorem 3.1. For $k=1$, only the middle term in (6.2) is present. This explains, why for $k=1$ the Kantorovich-Rubinstein distance between the transformation of a Poisson process (which is again a Poisson process) and a second Poisson process is bounded by the total variation distance of the intensity measures and the second term in (3.3) in Theorem 3.1 vanishes.

Throughout this section we use the same notation as in Subsection 3.1. Moreover, let $[k]$ be shorthand for $\{1,\ldots,k\}$. For $x=(x_1,\ldots,x_k)\in\mathbb{X}^k$, $I=\{i_1,\ldots,i_{|I|}\}\subset[k]$ and $z=(z_1,\ldots,z_{k-|I|})\in\mathbb{X}^{k-|I|}$, let $(x_I,z)=(x_{i_1},\ldots,x_{i_{|I|}},z_1,\ldots,z_{k-|I|})$. We prepare the proof of Theorem 3.1 with the following lemma.

Lemma 6.1. *Let the assumptions of Theorem 3.1 prevail. If ξ is induced by a Poisson process, then*

$$\begin{aligned}\mathbf{E}\xi(\mathbb{Y})^2=\frac{1}{k!}\sum_{I\subset[k]}\frac{1}{(k-|I|)!}\int_{\mathbb{X}^k}\int_{\mathbb{X}^{k-|I|}}&\mathbf{1}((x_1,\ldots,x_k)\in\mathrm{dom}\,f)\\&\times\mathbf{1}((x_I,z)\in\mathrm{dom}\,f)\,\mathbf{K}^{k-|I|}(\mathrm{d}z)\,\mathbf{K}^k(\mathrm{d}(x_1,\ldots,x_k))\,.\end{aligned}$$

If ξ is derived from a binomial process of n points, then

$$\begin{aligned}\mathbf{E}\xi(\mathbb{Y})^2=\frac{1}{k!}\sum_{I\subset[k]}\frac{(n)_{2k-|I|}}{(k-|I|)!}\int_{\mathbb{X}^k}\int_{\mathbb{X}^{k-|I|}}&\mathbf{1}((x_1,\ldots,x_k)\in\mathrm{dom}\,f)\\&\times\mathbf{1}((x_I,z)\in\mathrm{dom}\,f)\,\mathbf{K}_1^{k-|I|}(\mathrm{d}z)\,\mathbf{K}_1^k(\mathrm{d}(x_1,\ldots,x_k))\,.\end{aligned}$$

Proof. We have that

$$\begin{aligned}\xi(\mathbb{Y})^2 &= \frac{1}{(k!)^2}\Bigg(\sum_{(x_1,\dots,x_k)\in\mu^k_{\neq}} \mathbf{1}((x_1,\dots,x_k)\in\operatorname{dom} f)\Bigg)^2\\ &= \frac{1}{(k!)^2}\sum_{I\subset[k]}\sum_{(x_1,\dots,x_k,z)\in\mu^{2k-|I|}_{\neq}} \frac{k!}{(k-|I|)!}\mathbf{1}((x_1,\dots,x_k)\in\operatorname{dom} f)\mathbf{1}((x_I,z)\in\operatorname{dom} f)\,,\end{aligned}$$

where we have used that two points occurring in different sums can be either equal or distinct and that $\operatorname{dom} f$ is symmetric. Now the multivariate Mecke (2.1) and its binomial analogue (2.2) conclude the proof. □

Proof of Theorem 3.1. Throughout this proof we write $\xi(\eta)$ and $\xi(\beta_n)$ to emphasize the dependence of ξ on the underlying point process. Whenever we do not need special properties of η or β_n, we write $\xi(\mu)$ with the dummy variable μ standing for either η or β_n. As discussed in Remark 3.2 (ii), we can assume for the Poisson case that $\mathbf{L}(\mathbb{Y})<\infty$ and, hence, that $\xi(\eta)$ is almost surely finite since (3.3) is obviously true otherwise. For an underlying binomial process it is sufficient to consider only the case $n\geq k$ since, otherwise, the statement is obviously true as explained in Remark 3.2 (i).

Lemma 5.4 says that, for $h\in\mathscr{L}_1$ and $\omega\in\widehat{\mathrm{N}}_{\mathbb{Y}}$,

$$\mathbf{E}h(\zeta)-h(\omega)=\int_0^\infty LP_sh(\omega)\,\mathrm{d}s\,. \tag{6.3}$$

The Stein-type identity (6.3) is the starting point for our proof. Combining (6.3) with the representation of the generator L in (5.1), choosing $\omega=\xi(\mu)$ and taking expectations results into

$$\begin{aligned}\mathbf{E}h(\zeta)-\mathbf{E}h(\xi(\mu)) &= \mathbf{E}\int_0^\infty LP_sh(\xi(\mu))\,\mathrm{d}s\\ &= \mathbf{E}\int_0^\infty\int_{\mathbb{Y}}\big(P_sh(\xi(\mu)+\delta_y)-P_sh(\xi(\mu))\big)\,\mathbf{M}(\mathrm{d}y)\,\mathrm{d}s\\ &\quad+\mathbf{E}\int_0^\infty\sum_{y\in\xi(\mu)}\big(P_sh(\xi(\mu)-\delta_y)-P_sh(\xi(\mu))\big)\,\mathrm{d}s\,.\end{aligned} \tag{6.4}$$

Let us denote the first and the second term on the right-hand side by $T_{1,\mu}$ and $T_{2,\mu}$, respectively. By Fubini's theorem and the definition of $\xi(\mu)$, we obtain that

$$T_{2,\mu}=\frac{1}{k!}\int_0^\infty\mathbf{E}\sum_{(x_1,\dots,x_k)\in\mu^k_{\neq}\cap\operatorname{dom} f}\big(P_sh(\xi(\mu)-\delta_{f(x_1,\dots,x_k)})-P_sh(\xi(\mu))\big)\,\mathrm{d}s\,.$$

By the multivariate Mecke formula (2.1) and its analogue (2.2) for binomial processes, we see that

$$\begin{aligned}T_{2,\eta}=\frac{1}{k!}\int_0^\infty\int_{\operatorname{dom} f}\mathbf{E}\big[&P_sh(\xi(\eta+\delta_{x_1}+\ldots+\delta_{x_k})-\delta_{f(x_1,\dots,x_k)})\\ &-P_sh(\xi(\eta+\delta_{x_1}+\ldots+\delta_{x_k}))\big]\,\mathbf{K}^k(\mathrm{d}(x_1,\dots,x_k))\,\mathrm{d}s\end{aligned}$$

and

$$\begin{aligned}T_{2,\beta_n}=\frac{(n)_k}{k!}\int_0^\infty\int_{\operatorname{dom} f}\mathbf{E}\big[&P_sh(\xi(\beta_{n-k}+\delta_{x_1}+\ldots+\delta_{x_k})-\delta_{f(x_1,\dots,x_k)})\\ &-P_sh(\xi(\beta_{n-k}+\delta_{x_1}+\ldots+\delta_{x_k}))\big]\,\mathbf{K}^k_1(\mathrm{d}(x_1,\dots,x_k))\,\mathrm{d}s\,.\end{aligned}$$

Let us write $\hat\xi(x_1,\dots,x_k,\mu)$ for the point process

$$\hat\xi(x_1,\dots,x_k,\mu):=\sum_{\emptyset\neq I\subsetneq[k],z\in\mu^{k-|I|}_{\neq}}\frac{1}{(k-|I|)!}\,\mathbf{1}((x_I,z)\in\operatorname{dom} f)\,\delta_{f(x_I,z)}$$

on $\mathbb{Y}$, where $\subsetneq$ denotes proper set-inclusion and where the notation (x_I, z) has been introduced before Lemma 6.1 above. Then

$$\begin{aligned}
T_{2,\eta} &= \frac{1}{k!}\int_0^\infty \int_{\mathrm{dom}\, f} \mathbf{E}\big[P_s h(\xi(\eta) + \hat{\xi}(x_1,\dots,x_k,\eta)) \\
&\qquad\qquad - P_s h(\xi(\eta) + \hat{\xi}(x_1,\dots,x_k,\eta) + \delta_{f(x_1,\dots,x_k)})\big]\, \mathbf{K}^k(\mathrm{d}(x_1,\dots,x_k))\,\mathrm{d}s \\
&= -\frac{1}{k!}\int_0^\infty \int_{\mathrm{dom}\, f} \mathbf{E}\big[P_s h(\xi(\eta) + \delta_{f(x_1,\dots,x_k)}) - P_s h(\xi(\eta))\big]\, \mathbf{K}^k(\mathrm{d}(x_1,\dots,x_k))\,\mathrm{d}s \\
&\quad + \frac{1}{k!}\int_0^\infty \int_{\mathrm{dom}\, f} \mathbf{E}\big[P_s h(\xi(\eta) + \hat{\xi}(x_1,\dots,x_k,\eta)) - P_s h(\xi(\eta)) + P_s h(\xi(\eta) + \delta_{f(x_1,\dots,x_k)}) \\
&\qquad\qquad - P_s h(\xi(\eta) + \hat{\xi}(x_1,\dots,x_k,\eta) + \delta_{f(x_1,\dots,x_k)})\big]\, \mathbf{K}^k(\mathrm{d}(x_1,\dots,x_k))\,\mathrm{d}s \\
&=: \hat{T}_{2,\eta} + R_\eta
\end{aligned}$$

and

$$\begin{aligned}
T_{2,\beta_n} &= \frac{(n)_k}{k!}\int_0^\infty \int_{\mathrm{dom}\, f} \mathbf{E}\big[P_s h(\xi(\beta_{n-k}) + \hat{\xi}(x_1,\dots,x_k,\beta_{n-k})) \\
&\qquad\qquad - P_s h(\xi(\beta_{n-k}) + \hat{\xi}(x_1,\dots,x_k,\beta_{n-k}) + \delta_{f(x_1,\dots,x_k)})\big]\, \mathbf{K}_1^k(\mathrm{d}(x_1,\dots,x_k))\,\mathrm{d}s \\
&= -\frac{(n)_k}{k!}\int_0^\infty \int_{\mathrm{dom}\, f} \mathbf{E}\big[P_s h(\xi(\beta_{n-k}) + \delta_{f(x_1,\dots,x_k)}) - P_s h(\xi(\beta_{n-k}))\big]\, \mathbf{K}_1^k(\mathrm{d}(x_1,\dots,x_k))\,\mathrm{d}s \\
&\quad + \frac{(n)_k}{k!}\int_0^\infty \int_{\mathrm{dom}\, f} \mathbf{E}\big[P_s h(\xi(\beta_{n-k}) + \hat{\xi}(x_1,\dots,x_k,\beta_{n-k})) - P_s h(\xi(\beta_{n-k})) \\
&\qquad\qquad + P_s h(\xi(\beta_{n-k}) + \delta_{f(x_1,\dots,x_k)}) \\
&\qquad\qquad - P_s h(\xi(\beta_{n-k}) + \hat{\xi}(x_1,\dots,x_k,\beta_{n-k}) + \delta_{f(x_1,\dots,x_k)})\big]\, \mathbf{K}_1^k(\mathrm{d}(x_1,\dots,x_k))\,\mathrm{d}s \\
&=: \hat{T}_{2,\beta_n} + R_{\beta_n}\,.
\end{aligned}$$

Together with (6.4) and the formulas for $\mathbf{L}$ in (3.1) and (3.2), we see that

$$\mathbf{E}h(\zeta) - \mathbf{E}h(\xi(\eta)) = \int_0^\infty \int_{\mathbb{Y}} \mathbf{E}\big[D_y P_s h(\xi(\eta))\big]\,(\mathbf{M}-\mathbf{L})(\mathrm{d}y)\,\mathrm{d}s + R_\eta$$

and

$$\begin{aligned}
\mathbf{E}h(\zeta) - \mathbf{E}h(\xi(\beta_n)) &= \int_0^\infty \int_{\mathbb{Y}} \mathbf{E}\big[D_y P_s h(\xi(\beta_n))\big]\,(\mathbf{M}-\mathbf{L})(\mathrm{d}y)\,\mathrm{d}s \\
&\qquad + \int_0^\infty \int_{\mathbb{Y}} \mathbf{E}\big[D_y P_s h(\xi(\beta_n))\big] - \mathbf{E}\big[D_y P_s h(\xi(\beta_{n-k}))\big]\,\mathbf{L}(\mathrm{d}y)\,\mathrm{d}s + R_{\beta_n}\,.
\end{aligned}$$

We now determine the remainder terms R_η and R_{β_n}. For $(x_1,\dots,x_k) \in \mathrm{dom}\, f$ let us define $\widetilde{h}_{x_1,\dots,x_k} : \widehat{\mathrm{N}}_{\mathbb{Y}} \to \mathbb{R}$ by

$$\widetilde{h}_{x_1,\dots,x_k}(\mu) = \frac{1}{2}\left(h(\mu) - h(\mu + \delta_{f(x_1,\dots,x_k)})\right)\,.$$

We can then re-write R_η and R_{β_n} as

$$R_\eta = \frac{2}{k!}\int_0^\infty \int_{\mathrm{dom}\, f} \mathbf{E}\big[P_s\widetilde{h}_{x_1,\dots,x_k}(\xi(\eta) + \hat{\xi}(x_1,\dots,x_k,\eta)) - P_s\widetilde{h}_{x_1,\dots,x_k}(\xi(\eta))\big]\,\mathbf{K}^k(\mathrm{d}(x_1,\dots,x_k))\,\mathrm{d}s$$

and

$$\begin{aligned}
R_{\beta_n} = \frac{2(n)_k}{k!}\int_0^\infty \int_{\mathrm{dom}\, f} &\mathbf{E}\big[P_s\widetilde{h}_{x_1,\dots,x_k}(\xi(\beta_{n-k}) + \hat{\xi}(x_1,\dots,x_k,\beta_{n-k})) \\
&- P_s\widetilde{h}_{x_1,\dots,x_k}(\xi(\beta_{n-k}))\big]\,\mathbf{K}_1^k(\mathrm{d}(x_1,\dots,x_k))\,\mathrm{d}s\,.
\end{aligned}$$

Because of $\widetilde{h}_{x_1,\dots,x_k} \in \mathscr{L}_1$, we obtain by the definition of the semi-group $(P_s)_{s\geq 0}$ in (5.2) and Lemma 5.2 that

$$\begin{aligned}|R_\eta| &\leq \frac{2}{k!}\int_0^\infty \int_{\mathrm{dom}\, f} e^{-s}\, \mathbf{E}\hat{\xi}(x_1,\dots,x_k,\eta)(\mathbb{Y})\, \mathbf{K}^k(\mathrm{d}(x_1,\dots,x_k))\,\mathrm{d}s\\ &= \frac{2}{k!}\int_{\mathrm{dom}\, f} \mathbf{E}\hat{\xi}(x_1,\dots,x_k,\eta)(\mathbb{Y})\, \mathbf{K}^k(\mathrm{d}(x_1,\dots,x_k))\end{aligned}$$

and

$$\begin{aligned}|R_{\beta_n}| &\leq \frac{2(n)_k}{k!}\int_0^\infty \int_{\mathrm{dom}\, f} e^{-s}\, \mathbf{E}\hat{\xi}(x_1,\dots,x_k,\beta_{n-k})(\mathbb{Y})\, \mathbf{K}_1^k(\mathrm{d}(x_1,\dots,x_k))\,\mathrm{d}s\\ &\leq \frac{2(n)_k}{k!}\int_{\mathrm{dom}\, f} \mathbf{E}\hat{\xi}(x_1,\dots,x_k,\beta_{n-k})(\mathbb{Y})\, \mathbf{K}_1^k(\mathrm{d}(x_1,\dots,x_k))\,.\end{aligned}$$

Now, from the Mecke formula (2.1) and its analogue (2.2) for binomial processes it follows that

$$\begin{aligned}\mathbf{E}\hat{\xi}(x_1,\dots,x_k,\eta)(\mathbb{Y}) &= \mathbf{E}\sum_{\emptyset\neq I\subsetneq[k],\, z\in\eta_{\neq}^{k-|I|}} \frac{1}{(k-|I|)!}\,\mathbf{1}(f(x_I,z)\in\mathrm{dom}\, f)\\ &= \sum_{\emptyset\neq I\subsetneq[k]} \frac{1}{(k-|I|)!}\int_{\mathbb{X}^{k-|I|}} \mathbf{1}((x_I,z)\in\mathrm{dom}\, f)\,\mathbf{K}^{k-|I|}(\mathrm{d}z)\end{aligned}$$

and

$$\begin{aligned}\mathbf{E}\hat{\xi}(x_1,\dots,x_k,\beta_{n-k})(\mathbb{Y}) &= \mathbf{E}\sum_{\emptyset\neq I\subsetneq[k],\, z\in\beta_{n-k,\neq}^{k-|I|}} \frac{1}{(k-|I|)!}\,\mathbf{1}(f(x_I,z)\in\mathrm{dom}\, f)\\ &= \sum_{\emptyset\neq I\subsetneq[k]} \frac{(n-k)_{k-|I|}}{(k-|I|)!}\int_{\mathbb{X}^{k-|I|}} \mathbf{1}((x_I,z)\in\mathrm{dom}\, f)\,\mathbf{K}_1^{k-|I|}(\mathrm{d}z)\,.\end{aligned} \tag{6.5}$$

Together with Lemma 6.1, we obtain

$$\begin{aligned}|R_\eta| &\leq \frac{2}{k!}\int_{\mathbb{X}^k}\sum_{\emptyset\neq I\subsetneq[k]} \frac{1}{(k-|I|)!}\int_{\mathbb{X}^{k-|I|}} \mathbf{1}((x_1,\dots,x_k)\in\mathrm{dom}\, f)\\ &\qquad\qquad \times \mathbf{1}((x_I,z)\in\mathrm{dom}\, f)\,\mathbf{K}^{k-|I|}(\mathrm{d}z)\,\mathbf{K}^k(\mathrm{d}(x_1,\dots,x_k))\\ &= 2\big(\mathbf{E}\xi(\mathbb{Y})^2 - \mathbf{L}(\mathbb{Y}) - \mathbf{L}(\mathbb{Y})^2\big) = 2\big(\mathbf{E}\xi(\mathbb{Y})^2 - \mathbf{E}\xi(\mathbb{Y}) - (\mathbf{E}\xi(\mathbb{Y}))^2\big)\end{aligned} \tag{6.6}$$

and

$$\begin{aligned}|R_{\beta_n}| &\leq \frac{2}{k!}\int_{\mathbb{X}^k}\sum_{\emptyset\neq I\subsetneq[k]} \frac{(n)_k(n-k)_{k-|I|}}{(k-|I|)!}\int_{\mathbb{X}^{k-|I|}} \mathbf{1}((x_1,\dots,x_k)\in\mathrm{dom}\, f)\\ &\qquad\qquad \times \mathbf{1}((x_I,z)\in\mathrm{dom}\, f)\,\mathbf{K}_1^{k-|I|}(\mathrm{d}z)\,\mathbf{K}_1^k(\mathrm{d}(x_1,\dots,x_k))\\ &= 2\big(\mathbf{E}\xi(\mathbb{Y})^2 - \mathbf{L}(\mathbb{Y}) - \frac{(n-k)_k}{(n)_k}\mathbf{L}(\mathbb{Y})^2\big) = 2\big(\mathbf{E}\xi(\mathbb{Y})^2 - \mathbf{E}\xi(\mathbb{Y}) - \frac{(n-k)_k}{(n)_k}(\mathbf{E}\xi(\mathbb{Y}))^2\big)\,.\end{aligned} \tag{6.7}$$

The inequalities in (6.6) and (6.7) together with the definition of $r(\mathrm{dom}\, f)$ imply that

$$|R_\eta| \leq \frac{2^{k+1}}{k!}r(\mathrm{dom}\, f) \quad \text{and} \quad |R_{\beta_n}| \leq \frac{2^{k+1}}{k!}r(\mathrm{dom}\, f)\,. \tag{6.8}$$

Next, it follows from Lemma 5.2 that, for $s\geq 0$,

$$|\,\mathbf{E}D_yP_sh(\xi(\mu))\,| \leq \mathbf{E}[\,|P_sh(\xi(\mu)+\delta_y) - P_sh(\xi(\mu))|\,] \leq e^{-s}\,. \tag{6.9}$$

For $y_1, y_2 \in \mathbb{Y}$ and $\tilde{\xi} \in \widetilde{\mathrm{N}}_{\mathbb{Y}}$ we have $\mathrm{d_{TV}}(\tilde{\xi} + \delta_{y_1}, \tilde{\xi} + \delta_{y_2}) \le 1$ so that $h \in \mathscr{L}_1$ leads to

$$|D_{y_1} h(\tilde{\xi}) - D_{y_2} h(\tilde{\xi})| = |h(\tilde{\xi} + \delta_{y_1}) - h(\tilde{\xi} + \delta_{y_2})| \le 1\,.$$

Together with Lemma 5.1, we obtain that

$$|\mathbf{E} D_{y_1} P_s h(\xi(\mu)) - \mathbf{E} D_{y_2} P_s h(\xi(\mu))| = e^{-s} |\mathbf{E} P_s (D_{y_1} h - D_{y_2} h)(\xi(\mu))| \le e^{-s} \tag{6.10}$$

for all $y_1, y_2 \in \mathbb{Y}$ and $s \ge 0$. The estimates in (6.9) and (6.10) show that

$$\left| \int_0^\infty \int_{\mathbb{Y}} \mathbf{E}\left[D_y P_s h(\xi(\mu)) \right] (\mathbf{M} - \mathbf{L})(\mathrm{d}y)\, \mathrm{d}s \right| \le \mathrm{d_{TV}}(\mathbf{M}, \mathbf{L}) \int_0^\infty e^{-s}\, \mathrm{d}s \le \mathrm{d_{TV}}(\mathbf{M}, \mathbf{L})\,. \tag{6.11}$$

Combining (6.6) and (6.8) with (6.11) completes the proof of the Poisson case.

When considering a binomial process, we additionally need to take care of the term

$$\int_0^\infty \int_{\mathbb{Y}} \mathbf{E}\big[D_y P_s h(\xi(\beta_n)) \big] - \mathbf{E}\big[D_y P_s h(\xi(\beta_{n-k})) \big]\, \mathbf{L}(\mathrm{d}y)\, \mathrm{d}s\,.$$

For this, we use Lemma 5.1, the fact that $\frac{1}{2} D_y h \in \mathscr{L}_1$ whenever $h \in \mathscr{L}_1$ and Lemma 5.2 to obtain that

$$\begin{aligned}
&\big| \mathbf{E}\big[D_y P_s h(\xi(\beta_n)) \big] - \mathbf{E}\big[D_y P_s h(\xi(\beta_{n-k})) \big] \big| \\
&\le \int_{\mathbb{X}^k} \big| \mathbf{E}\big[D_y P_s h(\xi(\beta_{n-k} + \delta_{x_1} + \ldots + \delta_{x_k})) \big] - \mathbf{E}\big[D_y P_s h(\xi(\beta_{n-k})) \big] \big|\, \mathbf{K}_1^k(\mathrm{d}(x_1, \ldots, x_k)) \\
&= \int_{\mathbb{X}^k} e^{-s} \big| \mathbf{E}\big[P_s(D_y h)(\xi(\beta_{n-k}) + \hat{\xi}(x_1, \ldots, x_k, \beta_{n-k}) + \delta_{f(x_1,\ldots,x_k)}) - P_s(D_y h)(\xi(\beta_{n-k})) \big] \big| \\
&\qquad\qquad \mathbf{K}_1^k(\mathrm{d}(x_1, \ldots, x_k)) \\
&\le \frac{1}{n^k} \int_{\mathbb{X}^k} 2e^{-2s} \left(\mathbf{E}\hat{\xi}(x_1, \ldots, x_k, \beta_{n-k})(\mathbb{Y}) + \mathbf{1}((x_1, \ldots, x_k) \in \operatorname{dom} f) \right) \mathbf{K}^k(\mathrm{d}(x_1, \ldots, x_k))
\end{aligned}$$

for any $s \ge 0$. It follows from (6.5) and $(n-k)_{k-|I|} \le n^{k-|I|}$ that

$$\begin{aligned}
&\frac{1}{n^k} \int_{\mathbb{X}^k} \mathbf{E}\hat{\xi}(x_1, \ldots, x_k, \beta_{n-k})(\mathbb{Y})\, \mathbf{K}^k(\mathrm{d}(x_1, \ldots, x_k)) \\
&\le \frac{1}{n^k} \int_{\mathbb{X}^k} \sum_{\emptyset \ne I \subsetneq [k]} \frac{1}{(k - |I|)!} \int_{\mathbb{X}^{k-|I|}} \mathbf{1}((x_I, z) \in \operatorname{dom} f)\, \mathbf{K}^{k-|I|}(\mathrm{d}z)\, \mathbf{K}^k(\mathrm{d}(x_1, \ldots, x_k)) \\
&= \frac{1}{n^k} \sum_{\emptyset \ne I \subsetneq [k]} \frac{1}{(k - |I|)!} \int_{\mathbb{X}^k} \mathbf{1}((x_1, \ldots, x_k) \in \operatorname{dom} f)\, \mathbf{K}^k(\mathrm{d}(x_1, \ldots, x_k))\, \mathbf{K}(\mathbb{X})^{|k|-|I|} \\
&\le \frac{(2^k - 2)}{n} \int_{\mathbb{X}^k} \mathbf{1}((x_1, \ldots, x_k) \in \operatorname{dom} f)\, \mathbf{K}^k(\mathrm{d}(x_1, \ldots, x_k))\,.
\end{aligned}$$

Now, (3.2) implies that

$$\int_{\mathbb{X}^k} \mathbf{1}((x_1, \ldots, x_k) \in \operatorname{dom} f)\, \mathbf{K}^k(\mathrm{d}(x_1, \ldots, x_k)) = \frac{k!\, n^k}{(n)_k} \mathbf{L}(\mathbb{Y}) \le k! e^k \mathbf{E}\xi(\mathbb{Y})\,,$$

where we have used that $n^k/(n)_k \le k^k/k! \le e^k$ for $n \ge k$. Hence, using that $2^k e^k \le 6^k$, we find

$$\int_0^\infty \int_{\mathbb{Y}} \big| \mathbf{E}\big[D_y P_s h(\xi(\beta_n)) \big] - \mathbf{E}\big[D_y P_s h(\xi(\beta_{n-k})) \big] \big|\, \mathbf{L}(\mathrm{d}y)\, \mathrm{d}s \le 6^k k!\, \frac{\mathbf{L}(\mathbb{Y})^2}{n} = 6^k k!\, \frac{(\mathbf{E}\xi(\mathbb{Y}))^2}{n}\,.$$

Together with (6.7), (6.8) and (6.11) this completes the proof in the binomial case. □

Remark 6.2. Bounds for the total variation distance between ξ and ζ that are similar to the bounds for the Kantorovich-Rubinstein distance in Theorem 3.1 can be deduced from Theorem 2.6 in [4]. This result implies that

$$\begin{aligned}\mathbf{d_{TV}}(\xi(\eta),\zeta) &\le 2\mathrm{d_{TV}}(\mathbf{L},\mathbf{M})\\ &\quad+\frac{2}{k!}\int_{\mathrm{dom}\, f}\mathbf{E}\mathrm{d_{TV}}(\xi(\eta),\xi(\eta+\delta_{x_1}+\ldots+\delta_{x_k})-\delta_{f(x_1,\ldots,x_k)})\,\mathbf{K}^k(\mathrm{d}(x_1,\ldots,x_k))\end{aligned}$$

and

$$\begin{aligned}\mathbf{d_{TV}}(\xi(\beta_n),\zeta) &\le 2\mathrm{d_{TV}}(\mathbf{L},\mathbf{M})\\ &\quad+\frac{2(n)_k}{k!}\int_{\mathrm{dom}\, f}\mathbf{E}\mathrm{d_{TV}}(\xi(\beta_n),\xi(\beta_{n-k}+\delta_{x_1}+\ldots+\delta_{x_k})-\delta_{f(x_1,\ldots,x_k)})\,\mathbf{K}_1^k(\mathrm{d}(x_1,\ldots,x_k))\,.\end{aligned}$$

Since the integrands are bounded by

$$\mathbf{E}\hat{\xi}(x_1,\ldots,x_k,\eta)(\mathbb{Y})\qquad\text{and}\qquad\mathbf{E}\hat{\xi}(x_1,\ldots,x_k,\beta_{n-k})(\mathbb{Y})+\mathbf{E}\mathrm{d_{TV}}(\xi(\beta_n),\xi(\beta_{n-k}))\,,$$

respectively, the integrals on the right-hand sides can be controlled as in the proof of Theorem 3.1 above.

7 Applications

7.1 Poisson approximation of U-statistics

In this subsection we present a first application of Theorem 3.1 to U-statistics of Poisson or binomial processes. Let $(\mathbb{X},\mathcal{X})$ and $(\mathbb{Y},\mathcal{Y})$ be two lcscH spaces and let for some fixed integer $k\ge 1$, $f_t:\mathbb{X}^k\to\mathbb{Y}$, $t\ge 1$, be symmetric measurable functions. Furthermore, for a σ-finite measure $\mathbf{K}$ and a probability measure $\mathbf{K}_1$ on $\mathbb{X}$ we denote by η_t a Poisson process with intensity measure $\mathbf{K}_t:=t\mathbf{K}$, $t\ge 1$, and by β_t, $t\ge 1$, a binomial process of $\lceil t\rceil$ points with intensity measure $\mathbf{K}_t:=\lceil t\rceil\mathbf{K}_1$, respectively. If μ_t is either η_t or β_t and if B is a measurable subset of $\mathbb{Y}$, we define the U-statistics

$$S_t(B):=\frac{1}{k!}\sum_{(x_1,\ldots,x_k)\in\mu_{t,\neq}^k}\mathbf{1}(f_t(x_1,\ldots,x_k)\in B)\,,\quad t\ge 1\,,$$

which count the number of k-tuples $(x_1,\ldots,x_k)\in\mu_{t,\neq}^k$ for which $f_t(x_1,\ldots,x_k)\in B$. To compare $S_t(B)$ with a Poisson random variable we define

$$r_t(B):=\max_{1\le\ell\le k-1}\int_{\mathbb{X}^\ell}\left(\int_{\mathbb{X}^{k-\ell}}\mathbf{1}(f_t(x_1,\ldots,x_k)\in B)\,\mathbf{K}_t^{k-\ell}(\mathrm{d}(x_{\ell+1},\ldots,x_k))\right)^2\mathbf{K}_t^\ell(\mathrm{d}(x_1,\ldots,x_\ell))$$

if $k>1$ and $r_t(B):=0$ if $k=1$.

Theorem 7.1. *Let $B\in\mathcal{Y}$ and let Z be a Poisson distributed random variable with mean $\lambda\in[0,\infty)$. Suppose that $\mathbf{E}S_t(B)^2<\infty$. If $S_t(B)$ is induced by a Poisson process η_t with $t\ge 1$, then*

$$\begin{aligned}\mathbf{d_W}(S_t(B),Z)&\le|\mathbf{E}S_t(B)-\lambda|+2\big(\mathbf{E}S_t(B)^2-\mathbf{E}S_t(B)-(\mathbf{E}S_t(B))^2\big)\\&\le|\mathbf{E}S_t(B)-\lambda|+\frac{2^{k+1}}{k!}\,r_t(B)\,.\end{aligned}$$

If $S_t(B)$ is induced by a binomial process β_t with $t\ge 1$, then

$$\begin{aligned}\mathbf{d_W}(S_t(B),Z)&\le|\mathbf{E}S_t(B)-\lambda|+2\Big(\mathbf{E}S_t(B)^2-\mathbf{E}S_t(B)-\frac{(\lceil t\rceil-k)_k}{(\lceil t\rceil)_k}(\mathbf{E}S_t(B))^2\Big)+\frac{6^k k!}{t}\,(\mathbf{E}S_t(B))^2\\&\le|\mathbf{E}S_t(B)-\lambda|+\frac{2^{k+1}}{k!}\,r_t(B)+\frac{6^k k!}{t}\,(\mathbf{E}S_t(B))^2\,.\end{aligned}$$

Proof. We define the point processes

$$\xi_t := \frac{1}{k!} \sum_{(x_1,\dots,x_k)\in \mu_{t,\neq}^k} \delta_{f_t(x_1,\dots,x_k)}\,, \quad t \geq 1\,,$$

and denote their intensity measures by $\mathbf{L}_t$, $t \geq 1$. By construction, $S_t(B)$ and $\xi_t(B)$ follow the same distribution. We notice that for any fixed $h \in \mathrm{Lip}(1)$ (recall that these are all $h : \mathbb{R} \to \mathbb{R}$ whose Lipschitz constant is at most one) and $B \in \mathcal{Y}$ the mapping $\omega \mapsto h(\omega(B))$ from $\widetilde{\mathrm{N}}_{\mathbb{Y}}$ to $\mathbb{R}$ satisfies

$$|h(\omega_1(B)) - h(\omega_2(B))| \leq |\omega_1(B) - \omega_2(B)| \leq \mathrm{d_{TV}}(\omega_1,\omega_2), \qquad \omega_1,\omega_2 \in \widetilde{\mathrm{N}}_{\mathbb{Y}}\,,$$

and, thus, belongs to $\mathscr{L}_1$. Consequently, if ζ_t is a Poisson process on $\mathbb{Y}$ with intensity measure $\mathbf{L}_t$, the definition of the Wasserstein distance and (2.6) yield

$$\begin{aligned}\mathbf{d_W}(S_t(B),\zeta_t(B)) = \mathbf{d_W}(\xi_t(B),\zeta_t(B)) &= \sup_{h\in\mathrm{Lip}(1)} \big|\mathbf{E}h(\xi_t(B)) - \mathbf{E}h(\zeta_t(B))\big| \\ &\leq \sup_{g\in\mathscr{L}_1} \big|\mathbf{E}g(\xi_t|_B) - \mathbf{E}g(\zeta_t|_B)\big| = \mathbf{d_{KR}}(\xi_t|_B,\zeta_t|_B)\,.\end{aligned}$$

Now Theorem 3.1 and the observation that $\mathbf{L}_t(B) = \mathbf{E}S_t(B)$ imply the result for the choice $\lambda = \mathbf{E}S_t(B)$. The general case follows from the triangle inequality for the Wasserstein distance and the fact that the Wasserstein distance between a Poisson random variable with mean $\mathbf{E}S_t(B)$ and another Poisson random variable with mean λ is bounded by $|\mathbf{E}S_t(B) - \lambda|$. □

We emphasize that Theorem 7.1 deals with Poisson approximation in Wasserstein distance. As already stated in (2.4), this is stronger than approximation in total variation distance, which is usually considered in the literature (see [10] for the only exception we are aware of). This is possible thanks to our functional limit Theorem 3.1, which deals with the Kantorovich-Rubinstein distance rather than the total variation distance for point processes.

The Poisson approximation in total variation distance of U-statistics over binomial input was considered in [7]. If we assume that $\mathbf{E}S_t(B) = \lambda$ for $t \geq 1$ for the binomial case in Theorem 7.1, we obtain up to a constant, which may depend on λ, the same bound as in [7, Theorem 2] for the total variation distance.

In [31], an abstract bound for the Poisson approximation of Poisson functionals (i.e., random variables depending on a Poisson process) is derived, which is also applicable to U-statistics over Poisson input. Our Theorem 7.1 yields better rates of convergence for this special class of Poisson functionals. In fact, the bound in [43, Proposition 4.1], which is derived from [31], involves the square root of $\hat{r}_t(B)$ (see Remark 3.4 (iii)), while in the bound for the Poisson case in Theorem 7.1 only $\hat{r}_t(B)$ enters.

To illustrate the use of Theorem 7.1 let us consider a particular example, which will recur also in the following subsections. Let $K \subset \mathbb{R}^d$ ($d \geq 1$) be a compact convex set with volume one. For $t \geq 1$ let η_t be a homogeneous Poisson process in K of intensity t and denote by β_t a binomial process in K with $\lceil t \rceil$ points distributed according to the uniform distribution on K. For a family $(\theta_t)_{t\geq 1}$ of positive real numbers let us construct the random geometric graph with vertex set μ_t, where μ_t is η_t or β_t, by drawing an edge between two distinct vertices y_1 and y_2 whenever their Euclidean distance $\|y_1 - y_2\|$ is bounded by θ_t. These random graphs are the natural geometric counterparts to the classical Erdös-Rényi models for combinatorial random graphs. For background material we refer the reader to the monograph [34] and also to the recent paper [37] as well as the references cited therein.

For the random geometric graph introduced above let E_t be the number of edges. Note that E_t is a U-statistic of the form

$$E_t = \frac{1}{2} \sum_{(y_1,y_2)\in\mu_{t,\neq}^2} \mathbf{1}(\|y_1 - y_2\| \leq \theta_t)\,.$$

The multivariate Mecke formula (2.1) and a computation using spherical coordinates show that E_t has expectation $t^2\big(\kappa_d\theta_t^d + O(\theta_t^{d+1})\big)/2$ in the Poisson case, as $\theta_t \to 0$. For an underlying binomial process

the expected number of edges is $\lceil t\rceil(\lceil t\rceil-1)\big(\kappa_d\theta_t^d+O(\theta_t^{d+1})\big)/2$, as $\theta_t\to 0$. If the expectation of E_t converges to a constant, as $t\to\infty$, E_t can be well approximated by a Poisson random variable. In contrast to [31, Theorem 5.1], whose proof involves various non-trivial computations, we can deduce a corresponding approximation result from Theorem 7.1, the proof is postponed to Subsection 7.4.

Corollary 7.2. *Assume that* $\lim_{t\to\infty} t^2\theta_t^d=\lambda\in[0,\infty)$ *and let* Z *be a Poisson distributed random variable with mean* $\kappa_d\lambda/2$. *Then there is a constant* $c>0$ *only depending on the space dimension* d, *the set* K *and* $\sup_{t\ge1} t^2\theta_t^d$ *such that*

$$\mathbf{d_W}(E_t,Z)\le c\left(|t^2\theta_t^d-\lambda|+t^{-\min\{2/d,1\}}\right),\quad t\ge1\,.$$

Remark 7.3. Using the classical Chen-Stein method for Poisson approximation, Theorem 3.4 in [34] delivers a version of Corollary 7.2 in which the rate of convergence is measured by the total variation distance in case of an underlying binomial process. Theorem 3.12 (iii) in [32] is a qualitative version of Corollary 7.2, which has been established by the method of moments. Moreover, Theorem 5.1 in [31] adds a total variation bound. Corollary 7.2 extends these results to a stronger probability metric and at the same time improves the rates of convergence in [31]. Namely, for space dimensions $d\in\{1,2\}$, Corollary 7.2 yields an upper bound of order $|t^2\theta_t^d-\lambda|+t^{-1}$ (for the Wasserstein distance), while Theorem 5.1 in [31] delivers an upper bound of order $|t^2\theta_t^d-\lambda|+t^{-1/2}$ (for the total variation distance).

7.2 Compound Poisson approximation of U-statistics

As in the previous subsection, we denote by μ_t, $t\ge1$, a Poisson process η_t or a binomial process β_t on a lcscH space $\mathbb{X}$. For $k\in\mathbb{N}$ and measurable functions $h_t:\mathbb{X}^k\to\mathbb{R}$, $t\ge1$, we consider the family of U-statistics

$$S_t:=\frac{1}{k!}\sum_{(x_1,\dots,x_k)\in\mu_{t,\neq}^k} h_t(x_1,\dots,x_k)\,,\qquad t\ge1\,.$$

Since the sum runs also over all permutations of a fixed $(x_1,\dots,x_k)\in\mu_{t,\neq}^k$, we assume without loss of generality that h_t is symmetric for any $t\ge1$. For a fixed constant $\gamma\in\mathbb{R}$ and $t\ge1$, we define

$$\mathbf{L}_t(A):=\frac{1}{k!}\mathbf{E}\sum_{(x_1,\dots,x_k)\in\mu_{t,\neq}^k}\mathbf{1}(h_t(x_1,\dots,x_k)\in t^{-\gamma}A\setminus\{0\})\,,\qquad A\in\mathcal{B}(\mathbb{R})\,,$$

and

$$r_t:=\max_{1\le\ell\le k-1}\int_{\mathbb{X}^\ell}\left(\int_{\mathbb{X}^{k-\ell}}\mathbf{1}(h_t(x_1,\dots,x_k)\neq0)\,\mathbf{K}_t^{k-\ell}(\mathrm{d}(x_{\ell+1},\dots,x_k))\right)^2\mathbf{K}_t^{\ell}(\mathrm{d}(x_1,\dots,x_\ell))$$

for $k\ge2$, and put $r_t:=0$ if $k=1$. The following result compares the U-statistic S_t with a compound Poisson random variable. Most of the existing literature is based on a direct use of Stein's method, but only for discrete compound Poisson random variables. This approach is technically sophisticated and also needs in general certain monotonicity assumptions. Moreover, there are even situations in which the solution of the so-called Stein equation cannot be controlled appropriately, and hence in which Stein's method is of little use, see [9]. Being a consequence of the functional limit theorem (Theorem 3.1), our approach circumvents such technicalities and also allows us to deal with compound Poisson random variables having a discrete or continuous distribution.

Theorem 7.4. *Let* ζ *be a Poisson process on* $\mathbb{R}$ *with a finite intensity measure* $\mathbf{M}$, *let* $Z:=\sum_{x\in\zeta}x$ *and let* $\gamma\in\mathbb{R}$. *Then*

$$\mathbf{d_{TV}}(t^\gamma S_t,Z)\le\mathrm{d_{TV}}(\mathbf{L}_t,\mathbf{M})+\frac{2^{k+1}}{k!}r_t\,,\quad t\ge1\,,$$

if in the definition of S_t *a Poisson process* η_t *is used, and*

$$\mathbf{d_{TV}}(t^\gamma S_t,Z)\le\mathrm{d_{TV}}(\mathbf{L}_t,\mathbf{M})+\frac{2^{k+1}}{k!}r_t+\frac{6^k k!}{t}\,\mathbf{L}_t(\mathbb{R})^2\,,\quad t\ge1\,,$$

if the underlying point process is a binomial process β_t.

Proof. We consider the point processes

$$t^\gamma \bullet \xi_t := \frac{1}{k!} \sum_{(x_1,\dots,x_k)\in\mu_{t,\neq}^k} \mathbf{1}(h_t(x_1,\dots,x_k)\neq 0)\,\delta_{t^\gamma h_t(x_1,\dots,x_k)}\,, \quad t\geq 1\,.$$

It follows from the definition of the total variation distance and (2.6) that

$$\mathbf{d_{TV}}(t^\gamma S_t, Z) = \sup_{A\in\mathcal{B}(\mathbb{R})} \left|\mathbf{E1}\Big(\sum_{x\in t^\gamma\bullet\xi_t} x\in A\Big) - \mathbf{E1}\Big(\sum_{x\in\zeta} x\in A\Big)\right| \leq \mathbf{d_{KR}}(t^\gamma\bullet\xi_t,\zeta)$$

since the maps $\omega\to\mathbf{1}(\sum_{x\in\omega}x\in A)$ belong to $\mathscr{L}_1$. Now Theorem 3.1 implies that

$$\mathbf{d_{KR}}(t^\gamma\bullet\xi_t,\zeta)\leq \mathrm{d_{TV}}(\mathbf{L}_t,\mathbf{M}) + \frac{2^{k+1}}{k!}r_t\,,\quad t\geq 1\,,$$

and

$$\mathbf{d_{KR}}(t^\gamma\bullet\xi_t,\zeta)\leq \mathrm{d_{TV}}(\mathbf{L}_t,\mathbf{M}) + \frac{2^{k+1}}{k!}r_t + \frac{6^k k!}{t}\,\mathbf{L}_t(\mathbb{R})^2\,,\quad t\geq 1\,,$$

for the Poisson and the binomial case, respectively. This completes the proof. □

Remark 7.5. A compound Poisson random variable Z can alternatively be written as $Z=\sum_{i=1}^N X_i$, where N is a Poisson distributed random variable and $(X_i)_{i\in\mathbb{N}}$ is a sequence of independent and identically distributed random variables such that N and $(X_i)_{i\in\mathbb{N}}$ are independent. However, the representation of Z in terms of the Poisson process ζ fits better into our general framework.

For the compound-Poisson approximation of U-statistics in the binomial case a similar bound as in Theorem 7.4 is derived in [21, Section 3.6]. However, in that paper h_t is required to take values in the non-negative integers, whereas we do not need to impose such a condition. In addition, we are not aware of any analogous result for an underlying Poisson process.

As an application of Theorem 7.4 we consider general edge-length functionals of the random geometric graph introduced in the course of the previous subsection. Fix a parameter $b\in\mathbb{R}$ and define

$$L_t^{(b)} := \frac{1}{2}\sum_{(x_1,x_2)\in\mu_{t,\neq}^2} \mathbf{1}(\mathrm{dist}(x_1,x_2)\leq\theta_t)\,\mathrm{dist}(x_1,x_2)^b\,,\quad t\geq 1\,,$$

where μ_t stands either for a Poisson process η_t or a binomial process β_t. In particular, $L_t^{(0)}$ is the number of edges in the random geometric graph and $L_t^{(1)}$ is its total edge length. As in Subsection 7.1, we consider the situation that the distance parameters $(\theta_t)_{t\geq1}$ are chosen in such a way that the expected number of edges converges to a constant, as $t\to\infty$. Recall that in Corollary 7.2 the number of edges $L_t^{(0)}$ has been approximated by a Poisson random variable. For general exponents b we approximate $L_t^{(b)}$ by a suitable compound Poisson random variable. The proof of the next result is postponed to Subsection 7.4 below.

Corollary 7.6. *Fix $b\in\mathbb{R}$ and assume that $\lim_{t\to\infty} t^2\theta_t^d=\lambda\in[0,\infty)$. Define $Z:=\sum_{i=1}^N\|X_i\|^b$, where N is a Poisson distributed random variable with mean $\kappa_d\lambda/2$ and $(X_i)_{i\in\mathbb{N}}$ are independent and uniformly distributed points in $B^d(\lambda^{1/d})$, which are independent of N. Then, there is a constant $c>0$ only depending on the space dimension d, the set K and $\sup_{t\geq1}t^2\theta_t^d$ such that*

$$\mathbf{d_{TV}}(t^{2b/d}L_t^{(b)},Z)\leq c\left(|t^2\theta_t^d-\lambda| + t^{-\min\{2/d,1\}}\right),\quad t\geq1\,.$$

Remark 7.7. Corollary 7.6 without a rate of convergence has been derived in [37, Theorem 3.5] by combining a point process convergence result with the continuous mapping theorem. Thanks to Theorem 7.4 we were able to add a rate of convergence for the total variation distance.

7.3 Approximation of U-statistics by α-stable random variables

Let us denote by μ_t, $t \geq 1$, a Poisson process η_t or a binomial process β_t as in the previous subsections. For fixed $k \in \mathbb{N}$ and measurable functions $h_t : \mathbb{X}^k \to \mathbb{R}$, $t \geq 1$, let

$$S_t := \frac{1}{k!} \sum_{(x_1,\ldots,x_k) \in \mu_{t,\neq}^k} h_t(x_1, \ldots, x_k)\,, \quad t \geq 1\,.$$

Here, we can and will assume without loss of generality that h_t is symmetric for any $t \geq 1$. We are interested in the limiting behaviour of these U-statistics in situations, where their summands are heavy tailed, and approximate S_t by an α-stable random variable Z. Recall that this means that for any $n \in \mathbb{N}$ there are independent copies $Z_1, \ldots, Z_n$ of Z satisfying the distributional equality $n^{-1/\alpha}(Z_1 + \ldots + Z_n) \stackrel{D}{=} Z$. We fix $\alpha \in (0,1)$ and $\gamma \in \mathbb{R}$ and apply our functional limit theorem to the point processes

$$t^\gamma \bullet \xi_t := \frac{1}{k!} \sum_{(x_1,\ldots,x_k) \in \mu_{t,\neq}^k} \mathbf{1}(h_t(x_1, \ldots, x_k) \neq 0)\, \delta_{\mathrm{sign}(h_t(x_1,\ldots,x_k))\, t^\gamma |h_t(x_1,\ldots,x_k)|^{-\alpha}}\,, \quad t \geq 1\,,$$

on $\mathbb{R}$, where $\mathrm{sign}(a) = \mathbf{1}(a \geq 0) - \mathbf{1}(a < 0)$. If μ_t is a binomial process, the convergence of the U-statistic S_t to an α-stable random variable has been considered in [17] without giving rates of convergence. Thanks to our quantitative bound for the Kantorovich-Rubinstein distance in Theorem 3.1 we are in the position to add a rate of convergence for the Kolmogorov distance. The statement of our result is prepared by introducing some notation. For $A \in \mathcal{B}(\mathbb{R})$ and $t \geq 1$ we define

$$\mathbf{L}_t(A) := \frac{1}{k!}\mathbf{E} \sum_{(x_1,\ldots,x_k) \in \mu_{t,\neq}^k} \mathbf{1}(h_t(x_1, \ldots, x_k) \neq 0)\; \mathbf{1}(\mathrm{sign}(h_t(x_1, \ldots, x_k))\, |h_t(x_1, \ldots, x_k)|^{-\alpha} \in t^{-\gamma} A)\,,$$

which is the intensity measure of $t^\gamma \bullet \xi_t$, and

$$r_t(A) := \max_{1 \leq \ell \leq k-1} \int_{\mathbb{X}^\ell} \left(\int_{\mathbb{X}^{k-\ell}} \mathbf{1}(h_t(x_1, \ldots, x_k) \neq 0)\; \mathbf{1}(\mathrm{sign}(h_t(x_1, \ldots, x_k))\, |h_t(x_1, \ldots, x_k)|^{-\alpha} \in t^{-\gamma} A) \right.$$
$$\left. \mathbf{K}_t^{k-\ell}(\mathrm{d}(x_{\ell+1}, \ldots, x_k)) \right)^2 \mathbf{K}_t^\ell(\mathrm{d}(x_1, \ldots, x_\ell))$$

if $k \geq 2$ and $r_t(A) := 0$ if $k = 1$. The following result contains a quantitative bound for the approximation of U-statistics by an α-stable random variable with $\alpha \in (0,1)$.

Theorem 7.8. *Let $\alpha \in (0,1)$ and let $\mathbf{M}$ be either the Lebesgue measure on $\mathbb{R}$ or its restriction to $\mathbb{R}_+$. Define $Z := \sum_{x \in \zeta} \mathrm{sign}(x)\, |x|^{-1/\alpha}$, where ζ is a Poisson process with intensity measure $\mathbf{M}$. Assume that there are a constant $\gamma \in \mathbb{R}$ and functions $g_1, g_2, g_3 : \mathbb{R}_+^2 \to \mathbb{R}_+$ such that, for any $a > 0$ and $t \geq 1$,*

$$\mathrm{d_{TV}}(\mathbf{L}_t|_{[-a,a]}, \mathbf{M}|_{[-a,a]}) \leq g_1(a,t)\,, \qquad r_t([-a,a]) \leq g_2(a,t) \tag{7.1}$$

and

$$\frac{t^{-\gamma/\alpha}}{k!}\mathbf{E} \sum_{(x_1,\ldots,x_k) \in \mu_{t,\neq}^k} \mathbf{1}(|h_t(x_1, \ldots, x_k)| < t^{\gamma/\alpha} a^{-1/\alpha})\, |h_t(x_1, \ldots, x_k)| \leq g_3(a,t)\,. \tag{7.2}$$

Then there is a constant $C > 0$ only depending on α and k such that

$$\mathbf{d_K}(t^{-\gamma/\alpha} S_t, Z) \leq C g(t)\,, \quad t \geq 1\,,$$

where

$$g(t) := \begin{cases} \inf_{a>0} \max\{a^{1/2-1/(2\alpha)}, g_1(a,t), g_2(a,t), \sqrt{g_3(a,t)}\} & : \mu_t = \eta_t \\ \inf_{a>0} \max\{a^{1/2-1/(2\alpha)}, g_1(a,t), g_2(a,t), \sqrt{g_3(a,t)}, a^2/t\} & : \mu_t = \beta_t\,. \end{cases}$$

Proof. For $a > 0$ we define the random variables

$$S_{t,a} := \frac{1}{k!} \sum_{(x_1,\dots,x_k)\in\mu^k_{t,\neq}} \mathbf{1}(|h_t(x_1,\dots,x_k)| \geq t^{\gamma/\alpha} a^{-1/\alpha})\, h_t(x_1,\dots,x_k)\,, \quad t \geq 1\,,$$

and

$$Z_a := \sum_{x\in\zeta} \mathbf{1}(|x| \leq a)\ \mathrm{sign}(x)\,|x|^{-1/\alpha}\,.$$

Then, for any $a > 0$ and $\varepsilon > 0$, we find that

$$\begin{aligned}\mathbf{d_K}(t^{-\gamma/\alpha}S_t, Z) &\leq \mathbf{P}(t^{-\gamma/\alpha}|S_t - S_{t,a}| \geq \varepsilon) + \mathbf{d_K}(t^{-\gamma/\alpha}S_{t,a}, Z) + \sup_{z\in\mathbb{R}} |\mathbf{P}(Z \leq z) - \mathbf{P}(Z \leq z+\varepsilon)| \\ &\leq \mathbf{P}(t^{-\gamma/\alpha}|S_t - S_{t,a}| \geq \varepsilon) + \mathbf{P}(|Z - Z_a| \geq \varepsilon) + \mathbf{d_K}(t^{-\gamma/\alpha}S_{t,a}, Z_a) \\ &\quad + 2\sup_{z\in\mathbb{R}} |\mathbf{P}(Z \leq z) - \mathbf{P}(Z \leq z+\varepsilon)|\,.\end{aligned}$$

Combining Markov's inequality with the multivariate Mecke formula (2.1) and assumption (7.2), we obtain that, for all $\varepsilon > 0$,

$$\mathbf{P}(|Z - Z_a| \geq \varepsilon) \leq \frac{2}{\varepsilon}\int_a^\infty x^{-1/\alpha}\,\mathrm{d}x = \frac{2a^{1-1/\alpha}}{(1/\alpha - 1)\varepsilon} \quad \text{and} \quad \mathbf{P}(t^{-\gamma/\alpha}|S_t - S_{t,a}| \geq \varepsilon) \leq \frac{g_3(a,t)}{\varepsilon}\,.$$

As α-stable random variable, Z has a bounded density, see [47, page 13]. Hence, there is a constant $C_\alpha > 0$ only depending on α such that

$$\sup_{z\in\mathbb{R}} |\mathbf{P}(Z \leq z) - \mathbf{P}(Z \leq z+\varepsilon)| \leq C_\alpha \varepsilon\,, \quad \varepsilon \geq 0\,.$$

It follows from the definition of the Kolmogorov distance and (2.6) that

$$\begin{aligned}\mathbf{d_K}(t^{-\gamma/\alpha}S_{t,a}, Z_a) &= \sup_{z\in\mathbb{R}} \Big|\mathbf{P}\Big(\sum_{x\in t^\gamma\bullet\xi_t} \mathbf{1}(x\in[-a,a])\ \mathrm{sign}(x)\,|x|^{-1/\alpha} \leq z\Big) \\ &\qquad\qquad - \mathbf{P}\Big(\sum_{x\in\zeta} \mathbf{1}(x\in[-a,a])\ \mathrm{sign}(x)\,|x|^{-1/\alpha} \leq z\Big)\Big| \\ &\leq \mathbf{d_{KR}}(t^\gamma \bullet \xi_t|_{[-a,a]}, \zeta|_{[-a,a]})\,.\end{aligned}$$

Now, we consider the Poisson case and the binomial case separately. For an underlying Poisson process, Theorem 3.1 and the assumptions in (7.1) show that

$$\mathbf{d_{KR}}(t^\gamma \bullet \xi_t|_{[-a,a]}, \zeta|_{[-a,a]}) \leq g_1(a,t) + \frac{2^{k+1}}{k!} g_2(a,t)\,, \quad t \geq 1\,.$$

Combining this with the previous estimates, we see that

$$\mathbf{d_K}(t^{-\gamma/\alpha}S_t, Z) \leq \frac{2a^{1-1/\alpha}}{(1/\alpha-1)\varepsilon} + \frac{g_3(a,t)}{\varepsilon} + 2C_\alpha\varepsilon + g_1(a,t) + \frac{2^{k+1}}{k!}g_2(a,t)\,.$$

Thus, choosing $\varepsilon = \sqrt{\max\{a^{1-1/\alpha}, g_3(a,t)\}}$ yields the assertion. For the binomial case, Theorem 3.1 and the assumptions in (7.1) imply that

$$\mathbf{d_K}(S_{t,a}, Z_a) \leq \frac{2a^{1-1/\alpha}}{(1/\alpha-1)\varepsilon} + \frac{g_3(a,t)}{\varepsilon} + 2C_\alpha\varepsilon + g_1(a,t) + \frac{2^{k+1}}{k!}g_2(a,t) + \frac{6^k k!}{t}(8a^2 + 2g_1(a,t)^2)\,,$$

where we have used that $\mathbf{L}_t([-a,a])^2 \leq (2a + g_1(a,t))^2 \leq 8a^2 + 2g_1(a,t)^2$. Now, the same choice for ε as in the Poisson case and the fact that the Kolmogorov distance is bounded by one conclude the proof. □

Remark 7.9. For all choices of $\alpha \in (0,2]$ there are α-stable random variables and one can think of U-statistics converging to such variables. For $\alpha \in (1,2]$ and the binomial case this problem has been considered in [17, 22, 29]. A similar technique as that used in the proof of Theorem 7.8 should also be applicable if $\alpha \in (1,2]$. In this case the limiting random variable is given by $Z := \lim_{a\to\infty} Z_a - \mathbf{E}Z_a$, whence an additional centring is necessary. In order to derive bounds similar to that of Theorem 7.8, one has to control the distance between Z and Z_a, which might be difficult to tackle. We would like to mention that the bounds derived in [22] also involve a quantity similar to $\mathbf{d_K}(Z, Z_a)$.

To give an application of Theorem 7.8, let us consider the following distance-power statistics, which are closely related to the edge functionals of random geometric graphs considered above. Let for some $d \geq 1$, $K \subset \mathbb{R}^d$ be a compact convex set with volume one and let $\mathbf{K}$ be the restriction of the Lebesgue measure to K. Let η_t be a Poisson process in K with intensity measure $\mathbf{K}_t = t\mathbf{K}$, $t \geq 1$, and let β_t, $t \geq 1$, be a binomial process of $\lceil t \rceil$ points, which are independent and uniformly distributed in K. Our aim is to investigate the limiting behaviour of the U-statistics

$$S_t := \frac{1}{2} \sum_{(x_1,x_2)\in\mu_{t,\neq}^k} \operatorname{dist}(x_1,x_2)^{-\tau}\,, \qquad t \geq 1\,,$$

where $\tau > 0$ and μ_t stands for η_t or β_t. The following result, whose proof will be given in Subsection 7.4 below, deals with the case $\tau > d$.

Corollary 7.10. *Let $\tau > d$, let ζ be a homogeneous Poisson process on $\mathbb{R}_+$ with intensity one and let $Z := (\kappa_d/2)^{\tau/d} \sum_{x\in\zeta} x^{-\tau/d}$. Then there is a constant $C > 0$ only depending on K, τ and d such that*

$$\mathbf{d_K}(t^{-2\tau/d} S_t, Z) \leq C\, t^{\varrho}\,, \quad t \geq 1\,,$$

with

$$\varrho := \inf_{u>0} \max\left\{ \frac{1}{2}u - \frac{\tau}{2d}u, 2u-1, u + \frac{1}{d}u - \frac{2}{d} \right\}.$$

Example 7.11. To have a more specific example, take $\tau = 2d$ in Corollary 7.6, in which case ϱ has the form

$$\varrho = \inf_{u>0} \max\Big\{ -\frac{u}{2}, 2u-1, u + \frac{u-2}{d} \Big\}.$$

For $d \in \{1,2\}$ the infimum is attained at $u = \frac{2}{5}$, giving that $\varrho = -\frac{1}{5}$. For $d \geq 3$, the infimum is attained at $u = \frac{4}{3d+2}$ so that $\varrho = -\frac{2}{3d+2}$ in this case. Thus,

$$\mathbf{d_K}(t^{-4} S_t, Z) \leq \begin{cases} C\, t^{-1/5} & : d \in \{1,2\} \\ C\, t^{-2/(3d+2)} & : d \geq 3\,, \end{cases}$$

where the 1/2-stable random variable Z is of the form $Z = c_d \sum_{x\in\zeta} x^{-2}$ for a unit-intensity homogeneous Poisson process ζ on $\mathbb{R}_+$ and with $c_d = \kappa_d^2/4$. The distribution of Z can be characterized more explicitly. Namely, applying [24, Lemma 12.2 (i)] we see that, for all $t \in \mathbb{R}$,

$$\mathbf{E}\exp(\mathrm{i}tZ) = \mathbf{E}\exp\Big(\mathrm{i}t\, c_d \sum_{x\in\zeta} x^{-2}\Big) = \exp\Big(\int_0^\infty (e^{\mathrm{i}tc_d x^{-2}} - 1)\,\mathrm{d}x \Big) = \exp\big(-\sqrt{-\mathrm{i}t\pi c_d} \big)\,,$$

where i is the imaginary unit. This is the characteristic function of a centred Lévy distribution with scale parameter $\pi c_d/2$. Thus, Z has density $x \mapsto \frac{1}{2}\sqrt{c_d/x^3}\,\exp(-\pi c_d/(4x))\,\mathbf{1}(x>0)$.

Remark 7.12. Note that if $\tau < d/2$, then S_t satisfies a central limit theorem as shown in Theorem 3.1 of [37]. Moreover, the choice $d/2 \leq \tau \leq d$ corresponds to the situation $\alpha \in [1,2]$, to which Remark 7.9 applies.

7.4 Random geometric graphs

Let $K \subset \mathbb{R}^d$ ($d \geq 1$) be a compact convex set with volume one. For $t \geq 1$ let μ_t either be a homogeneous Poisson process η_t of intensity $t \geq 1$ in K or a binomial process β_t of $\lceil t \rceil$ independent and uniformly distributed points in K and let $(\theta_t)_{t\geq 1}$ be a family of positive real numbers. Based on this data we construct a random geometric graph as explained in Section 7.1. In contrast to Corollary 7.2 and Corollary 7.6, where $\lim_{t\to\infty} t^2\theta_t^d = \lambda \in [0,\infty)$, we assume at first that $\lim_{t\to\infty} t^2\theta_t^d = \infty$ and are interested in the point process $\xi_{t,a}$ on K defined by

$$\xi_{t,a} := \frac{1}{2} \sum_{(x,y)\in\mu_{t,\neq}^2} \mathbf{1}(\|x-y\| \leq \min\{\theta_t, t^{-2/d}a\})\, \delta_{(x+y)/2}$$

for some $a > 0$. In other words, $\xi_{t,a}$ charges the collection of all midpoints of edges of the random geometric graph whose length does not exceed $t^{-2/d}a$.

Theorem 7.13. *Let $a > 0$, let ζ be a Poisson process on K with intensity measure $\frac{\kappa_d}{2}\, a^d \,\mathrm{vol}|_K$ and let $\xi_{t,a}$ be constructed from a Poisson process η_t or a binomial process β_t with $t \geq 1$. Also suppose that $\lim_{t\to\infty} t^2\theta_t^d = \infty$. Then $t_0 := \sup\{t \geq 1 : t^2\theta_t^d < a^d\} \cup \{1\} < \infty$ and there is a constant $C > 0$ only depending on a, d and K such that*

$$\mathbf{d_{KR}}(\xi_{t,a}, \zeta) \leq C\, t^{-\min\{2/d,1\}}\,, \qquad t > t_0\,.$$

The rest of this subsection is devoted to the proofs of Theorem 7.13 as well as Corollary 7.2, Corollary 7.6 and Corollary 7.10. We prepare this by the following Lemma. In order to deal with the Poisson and the binomial case in parallel, we define $\chi(t) = t^2$ and $\widetilde{\chi}(t) = t^3$ if $\mu_t = \eta_t$ and $\chi(t) = \lceil t \rceil(\lceil t \rceil - 1)$ and $\widetilde{\chi}(t) = (\lceil t \rceil)^3$ if $\mu_t = \beta_t$.

Lemma 7.14. *There is a constant $C_K > 0$ only depending on d and K such that*

$$\begin{aligned} &\left| \frac{1}{2}\mathbf{E} \sum_{(x,y)\in\mu_{t,\neq}^2} \mathbf{1}((x+y)/2 \in B, \|x-y\| \in \tilde{A}) - \frac{\kappa_d}{2}\mathrm{vol}(B) t^2 d \int_0^\infty \mathbf{1}(r \in \tilde{A})\, r^{d-1}\,\mathrm{d}r \right| \\ &\qquad \leq 2C_K\kappa_d t^2(\tilde{a}^{d+1} + \tilde{a}^{2d}) + \frac{\kappa_d}{2} t\tilde{a}^d \end{aligned} \tag{7.3}$$

for all Borel sets $B \subset K$ and $\tilde{A} \subset [0,\tilde{a}]$ with $\tilde{a} > 0$. Moreover,

$$\widetilde{\chi}(t) \int_K \left(\int_K \mathbf{1}((x+y)/2 \in B, \|x-y\| \leq u)\,\mathrm{d}x \right)^2 \mathrm{d}y \leq 8t^3\kappa_d^2 u^{2d} \tag{7.4}$$

for all Borel sets $B \subset K$ and $u \geq 0$.

Proof. By the multivariate Mecke formula (2.1) for the Poisson process and its analogue (2.2) for the binomial case we obtain that

$$\begin{aligned} &\frac{1}{2}\,\mathbf{E} \sum_{(x,y)\in\mu_{t,\neq}^2} \mathbf{1}((x+y)/2 \in B, \|x-y\| \in \tilde{A}) \\ &= \frac{\chi(t)}{2} \int_K \int_K \mathbf{1}((x+y)/2 \in B, \|x-y\| \in \tilde{A})\,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{\chi(t)}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}((x+y)/2 \in B, \|x-y\| \in \tilde{A})\,\mathrm{d}x\,\mathrm{d}y \\ &\qquad - \frac{\chi(t)}{2} \int_{(\mathbb{R}^d)^2\setminus K^2} \mathbf{1}((x+y)/2 \in B, \|x-y\| \in \tilde{A})\,\mathrm{d}(x,y)\,. \end{aligned}$$

To the first term in the last expression we apply the change of variables $u = x - y$, $v = (x + y)/2$, which has Jacobian one, and spherical coordinates to see that

$$\begin{aligned}\frac{\chi(t)}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathbf{1}\big((x+y)/2\in B,\,\|x-y\|\in\tilde{A}\big)\,\mathrm{d}x\,\mathrm{d}y &= \frac{\chi(t)}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathbf{1}\big(v\in B,\,\|u\|\in\tilde{A}\big)\,\mathrm{d}u\,\mathrm{d}v\\ &= \frac{\chi(t)}{2}\mathrm{vol}(B)d\kappa_d\int_0^\infty\mathbf{1}(r\in\tilde{A})\,r^{d-1}\,\mathrm{d}r\,.\end{aligned}$$

A straightforward compuatation shows that

$$\left|(t^2-\chi(t))\mathrm{vol}(B)\frac{\kappa_d}{2}d\int_0^\infty\mathbf{1}(r\in\tilde{A})\,r^{d-1}\,\mathrm{d}r\right|\le\frac{\kappa_d}{2}t\tilde{a}^d\,.$$

For the second term we have, independently of B, the upper bound

$$\begin{aligned}&\frac{\chi(t)}{2}\int_{(\mathbb{R}^d)^2\setminus K^2}\mathbf{1}\big((x+y)/2\in B,\,\|x-y\|\in\tilde{A}\big)\,\mathrm{d}(x,y)\\ &\qquad\le 2t^2\,\mathrm{vol}\big(\{x\in\mathbb{R}^d\setminus K:\mathrm{dist}(x,K)\le\tilde{a}\}\big)\,\kappa_d\,\tilde{a}^d\,.\end{aligned}$$

From Steiner's formula (2.7) it follows that there is a constant $C_K > 0$ only depending on d and K such that

$$\mathrm{vol}\big(\{x\in\mathbb{R}^d\setminus K:\mathrm{dist}(x,K)\le\tilde{a}\}\big)\le C_K\,(\tilde{a}+\tilde{a}^d)\,.$$

Combining these estimates yields the first bound. On the other hand, we have

$$\widetilde{\chi}(t)\int_K\bigg(\int_K\mathbf{1}((x+y)/2\in B,\|x-y\|\le u)\,\mathrm{d}x\bigg)^2\mathrm{d}y\le 8t^3\int_K(\kappa_d u^d)^2\,\mathrm{d}y = 8t^3\kappa_d^2u^{2d}\,,$$

which is the second bound. □

Proof of Theorem 7.13. Due to our assumption that $\lim_{t\to\infty}t^2\theta_t^d=\infty$, we have that $t_0 := \sup\{t\ge 1 : t^2\theta_t^d < a^d\}\cup\{1\} < \infty$. Note that $\min\{\theta_t, t^{-2/d}a\} = t^{-2/d}a$ for $t > t_0$. We denote by $\mathbf{L}_{t,a}$ the intensity measure of $\xi_{t,a}$. For $t > t_0$ the choice $\tilde{A} = [0, \min\{\theta_t, t^{-2/d}a\}] = [0, t^{-2/d}a]$ in (7.3) leads to

$$\big|\mathbf{L}_{t,a}(B)-\frac{\kappa_d}{2}\mathrm{vol}(B)t^2(t^{-2/d}a)^d\big|\le 2C_K\kappa_d t^2(t^{-2-2/d}a^{d+1}+t^{-4}a^{2d})+\frac{\kappa_d}{2}t^{-1}a^d$$

so that $\mathrm{d_{TV}}(\mathbf{L}_{t,a},\frac{\kappa_d}{2}a^d\mathrm{vol}|_K)\le C_1t^{-\min\{2/d,1\}}$ for $t > t_0$ with a constant $C_1 > 0$ only depending on a, d and K. Moreover, there is a constant $C_2 > 0$ only depending on a, d and K such that $\mathbf{L}_{t,a}(K)\le C_2$ for all $t > t_0$. The inequality (7.4) implies that, for $t > t_0$,

$$\widetilde{\chi}(t)\int_K\bigg(\int_K\mathbf{1}(\|x-y\|\le\min\{\theta_t,t^{-2/d}a\})\,\mathrm{d}x\bigg)^2\mathrm{d}y\le 8t^3\kappa_d^2(t^{-2/d}a)^{2d}=8\kappa_d^2a^{2d}t^{-1}.$$

Now, application of Theorem 3.1 completes the proof. □

Proof of Corollary 7.2. The choice $B = K$ and $\tilde{A} = [0, \theta_t]$ in (7.3) leads to

$$\begin{aligned}\big|\mathbf{E}E_t-\frac{\kappa_d}{2}\lambda\big| &\le \big|\frac{\kappa_d}{2}\lambda-\frac{\kappa_d}{2}t^2\theta_t^d\big|+\big|\mathbf{E}E_t-\frac{\kappa_d}{2}t^2\theta_t^d\big|\\ &\le \frac{\kappa_d}{2}\big|\lambda-t^2\theta_t^d\big|+2C_K\kappa_dt^2(\theta_t^{d+1}+\theta_t^{2d})+\frac{\kappa_d}{2}t\theta_t^d\\ &\le \frac{\kappa_d}{2}|\lambda-t^2\theta_t^d|+2C_K\kappa_d\bigg(\frac{(\sup_{t\ge1}t^2\theta_t^d)^{1+1/d}}{t^{2/d}}+\frac{(\sup_{t\ge1}t^2\theta_t^d)^2}{t^2}\bigg)+\frac{\kappa_d}{2}\frac{\sup_{t\ge1}t^2\theta_t^d}{t}\end{aligned}$$

for $t \ge 1$, which also implies that $\mathbf{E}E_t$ is bounded by a constant only depending on d, K and $\sup_{t\ge1}t^2\theta_t^d$ for $t \ge 1$. It follows from (7.4) that

$$\widetilde{\chi}(t)\int_K\bigg(\int_K\mathbf{1}(\|x-y\|\le\theta_t)\,\mathrm{d}x\bigg)^2\mathrm{d}y\le 8t^3\kappa_d^2\theta_t^{2d}\le 8\kappa_d^2\frac{(\sup_{t\ge1}t^2\theta_t^d)^2}{t}\,.$$

Now, the assertion is a consequence of Theorem 7.1. □

Proof of Corollary 7.6. We assume that $b \neq 0$ in the following since for $b = 0$ the assertion follows from Corollary 7.2. For a Borel set $A \subset [0,\infty)$ we define $A^{1/b} := \{a^{1/b} : a \in A \setminus \{0\}\}$. Hence, we have that

$$\begin{aligned}\mathbf{L}_t(A) &:= \frac{1}{k!}\mathbf{E}\sum_{(x,y)\in\mu_{t,\neq}^2} \mathbf{1}(\|x-y\| \le \theta_t, \|x-y\|^b \in t^{-2b/d}A \setminus \{0\}) \\ &= \frac{1}{k!}\mathbf{E}\sum_{(x,y)\in\mu_{t,\neq}^2} \mathbf{1}(\|x-y\| \in t^{-2/d}A^{1/b} \cap [0,\theta_t])\,.\end{aligned}$$

Moreover, we define

$$\mathbf{M}(A) := \frac{\kappa_d}{2} d \int_0^{\lambda^{1/d}} \mathbf{1}(r \in A^{1/b})\, r^{d-1}\,\mathrm{d}r\,, \quad A \in \mathcal{B}(\mathbb{R})\,.$$

For a Borel set $A \subset [0,\infty)$ inequality (7.3) with $B = K$ and $\tilde{A} = t^{-2/d}A^{1/b} \cap [0,\theta_t]$ implies that

$$\begin{aligned}|\mathbf{L}_t(A) - \mathbf{M}(A)| &\le \left|\frac{\kappa_d}{2}t^2 d\int_0^\infty \mathbf{1}(r \in t^{-2/d}A^{1/b} \cap [0,\theta_t])\, r^{d-1}\,\mathrm{d}r - \frac{\kappa_d}{2} d\int_0^{\lambda^{1/d}} \mathbf{1}(r \in A^{1/b})\, r^{d-1}\,\mathrm{d}r\right| \\ &\quad + 2C_K\kappa_d t^2(\theta_t^{d+1} + \theta_t^{2d}) + \frac{\kappa_d}{2}t\theta_t^d \\ &\le \frac{\kappa_d}{2}|\lambda - t^2\theta_t^d| + 2C_K\kappa_d\left(\frac{(\sup_{t\ge1} t^2\theta_t^d)^{1+1/d}}{t^{2/d}} + \frac{(\sup_{t\ge1} t^2\theta_t^d)^2}{t^2}\right) + \frac{\kappa_d}{2}\frac{\sup_{t\ge1} t^2\theta_t^d}{t}\,.\end{aligned}$$

Hence, there are constants $C_1, C_2 > 0$ only depending on d, K and $\sup_{t\ge1} t^2\theta_t^d$ such that $\mathrm{d_{TV}}(\mathbf{L}_t, \mathbf{M}) \le C_1 t^{-\min\{2/d,1\}}$ for $t \ge 1$ and $\mathbf{L}_t(\mathbb{R}) \le C_2$ for $t \ge 1$. It follows from (7.4) that

$$\widetilde{\chi}(t)\int_K\left(\int_K \mathbf{1}(\|x-y\| \le \theta_t)\,\mathrm{d}x\right)^2 \mathrm{d}y \le 8t^3\kappa_d^2\theta_t^{2d} \le 8\kappa_d^2\frac{(\sup_{t\ge1} t^2\theta_t^d)^2}{t}\,.$$

Now, application of Theorem 7.4 completes the proof. □

Proof of Corollary 7.10. In the following, we check that the assumptions of Theorem 7.8 are satisfied for $h_t(x,y) = (2/\kappa_d)^{\tau/d}\|x-y\|^{-\tau}$ with $\alpha = d/\tau$ and $\gamma = 2$. For a Borel set $A \subset [0,\infty)$ and $t \ge 1$ we have that

$$\mathbf{L}_t(A) := \frac{1}{2}\mathbf{E}\sum_{(x,y)\in\mu_{t,\neq}^2} \mathbf{1}(\kappa_d\|x-y\|^d/2 \in t^{-2}A) = \frac{1}{2}\mathbf{E}\sum_{(x,y)\in\mu_{t,\neq}^2} \mathbf{1}(\|x-y\| \in (2/\kappa_d)^{1/d}t^{-2/d}A^{1/d})$$

with $A^{1/d} := \{x^{1/d} : x \in A\}$. In the following let $\mathbf{M}$ be the restriction of the Lebesgue measure to $\mathbb{R}_+$ and let $a > 0$. Since

$$\begin{aligned}&\frac{\kappa_d}{2}t^2 d\int_0^\infty \mathbf{1}(r \in (2/\kappa_d)^{1/d}t^{-2/d}(A\cap[0,a])^{1/d})\, r^{d-1}\,\mathrm{d}r \\ &= \frac{\kappa_d}{2}\int_0^\infty \mathbf{1}(u \in (2/\kappa_d)t^{-2}(A\cap[0,a]))\,\mathrm{d}u = \int_0^\infty \mathbf{1}(u \in A\cap[0,a])\,\mathrm{d}u = \mathbf{M}|_{[0,a]}(A)\,,\end{aligned}$$

application of (7.3) with $B = K$ and $\tilde{A} = (2/\kappa_d)^{1/d}t^{-2/d}(A\cap[0,a])^{1/d}$ yields that

$$\left|\mathbf{L}_t|_{[0,a]}(A) - \mathbf{M}|_{[0,a]}(A)\right| \le 2C_K\kappa_d t^2(c_a^{d+1}t^{-2-2/d} + c_a^{2d}t^{-4}) + \frac{\kappa_d}{2}t^{-1}c_a^d$$

with $c_a = (2a/\kappa_d)^{1/d}$. Consequently, there is a constant $C_1 > 0$ only depending on d and K such that

$$\mathrm{d_{TV}}(\mathbf{L}_t|_{[0,a]}, \mathbf{M}|_{[0,a]}) \le C_1(a^{1+1/d}t^{-2/d} + a^2t^{-2} + at^{-1}) =: g_1(a,t)\,, \quad t \ge 1\,.$$

It follows from (7.4) that

$$\widetilde{\chi}(t)\int_K\Big(\int_K \mathbf{1}(\kappa_d\|x-y\|^d/2\le t^{-2}a)\,\mathrm{d}x\Big)^2\,\mathrm{d}y$$
$$=\widetilde{\chi}(t)\int_K\Big(\int_K \mathbf{1}(\|x-y\|\le(2/\kappa_d)^{1/d}t^{-2/d}a^{1/d})\,\mathrm{d}x\Big)^2\,\mathrm{d}y\le 8t^3\kappa_d^2(2/\kappa_d)^2t^{-4}a^2=32t^{-1}a^2=:g_2(a,t)\,.$$

Moreover, we have that

$$\frac{t^{-2\tau/d}}{2}\mathbf{E}\sum_{(x,y)\in\mu_{t,\neq}^2}\mathbf{1}((2/\kappa_d)^{\tau/d}\|x-y\|^{-\tau}\le t^{2\tau/d}a^{-\tau/d})\,\|x-y\|^{-\tau}$$
$$\le d\kappa_d t^{2-2\tau/d}\int_{(\kappa_d/2)^{1/d}t^{-2/d}a^{1/d}}^{\infty} r^{-\tau}r^{d-1}\,\mathrm{d}r=\frac{d\kappa_d}{\tau-d}(\kappa_d/2)^{1-\tau/d}a^{1-\tau/d}=:g_3(a,t)\,.$$

Now, Theorem 7.8 completes the proof. □

7.5 Proximity of Poisson flats

For a space dimension $d\ge 2$ and a dimension parameter $m\ge 1$ satisfying $m<d/2$, we investigate the mutual arrangement of the flats of a Poisson m-flat process, that is, a Poisson process on the space of m-dimensional affine subspaces of $\mathbb{R}^d$, which are called m-flats. In order to define such a Poisson m-flat process in a rigorous way, recall that $\mathbb{G}_m^d$ and $\mathbb{A}_m^d$ stand for the space of m-dimensional linear and m-dimensional affine subspaces of $\mathbb{R}^d$, respectively. Let $\mathbb{Q}$ be a probability measure on $\mathbb{G}_m^d$ with the property that two independent random subspaces $L,M\in\mathbb{G}_m^d$ with distribution $\mathbb{Q}$ are almost surely in general position, meaning that the dimension of the linear hull of L and M is $2m$ with probability one. Note that this is satisfied, for example, if $\mathbb{Q}$ is absolutely continuous with respect to the unique Haar probability measure on $\mathbb{G}_m^d$, cf. [38, Theorem 4.4.5 (c)]. The measure $\mathbb{Q}$ induces a translation-invariant measure $\mathbf{K}_t$ on $\mathbb{A}_m^d$ via

$$\int_{\mathbb{A}_m^d} g(E)\,\mathbf{K}_t(\mathrm{d}E)=t\int_{\mathbb{G}_m^d}\int_{E_0^\perp} g(E_0+x)\,\mathrm{vol}_{E_0^\perp}(\mathrm{d}x)\,\mathbb{Q}(\mathrm{d}E_0)\,, \tag{7.5}$$

where $t\ge 1$ is an intensity parameter, $g\ge 0$ is a measurable function on $\mathbb{A}_m^d$ and $\mathrm{vol}_{E_0^\perp}$ denotes the Lebesgue measure on $E_0^\perp$, the orthogonal complement of E_0. We use the convention $\mathbf{K}:=\mathbf{K}_1$ and can re-write $\mathbf{K}_t$ as $\mathbf{K}_t=t\mathbf{K}$. We now consider a Poisson process η_t with intensity measure $\mathbf{K}_t$. This is what is usually called a Poisson m-flat process in stochastic geometry [38, Chapter 4.4]. One particular problem for such m-flat processes is to describe the mutual arrangement of the flats in space. Since $m<d/2$, any two different flats E,F of η_t do not intersect each other with probability one. Thus, they have a well defined distance $\mathrm{dist}(E,F)$, and we denote by $m(E,F)$ the midpoint of the almost surely uniquely determined line segment realizing this distance (the perpendicular of E and F). We are interested here in the point process of the midpoints $m(E,F)$ such that the flats E,F are close together and $m(E,F)$ is in a compact convex set $K\subset\mathbb{R}^d$ of volume $0<\mathrm{vol}(K)<\infty$. To the best of our knowledge, Theorem 7.15 is the first result describing its asymptotic behaviour, as $t\to\infty$. To do so, we define for $t\ge 1$ and $a>0$, $\xi_{t,a}$ on K by

$$\xi_{t,a}:=\frac{1}{2}\sum_{(E,F)\in\eta_{t,\neq}^2}\delta_{m(E,F)}\,\mathbf{1}(\mathrm{dist}(E,F)\le at^{-2/(d-2m)},\,m(E,F)\in K)\,.$$

The intensity measure $\mathbf{L}_{t,a}(B)$ of $\xi_{t,a}$ for a Borel set $B\subset K$ is given by

$$\mathbf{L}_{t,a}(B)=\frac{t^2}{2}\int_{\mathbb{A}_m^d}\int_{\mathbb{A}_m^d}\mathbf{1}(m(E,F)\in B,\,\mathrm{dist}(E,F)\le at^{-2/(d-2m)})\,\mathbf{K}(\mathrm{d}E)\,\mathbf{K}(\mathrm{d}F)$$

due to the multivariate Mecke formula (2.1). It follows from [44, Theorem 1] (it is readily checked that the identity there extends from compact convex sets to general Borel sets) that

$$\mathbf{L}_{t,a}(B) = \frac{t^2}{2}\kappa_{d-2m}(at^{-2/(d-2m)})^{d-2m}\mathrm{vol}(B)\int_{\mathbb{G}_k^d}\int_{\mathbb{G}_k^d}[M,L]\,\mathbb{Q}(\mathrm{d}L)\,\mathbb{Q}(\mathrm{d}M)\,,$$

where $[M,L]$ stands for the subspace determinant of M and L introduced in Section 2. This leads to

$$\mathbf{L}_{t,a}(B) = \frac{\kappa_{d-2m}}{2}\,\mathrm{vol}(B)\,a^{d-2m}\int_{\mathbb{G}_m^d}\int_{\mathbb{G}_m^d}[L,M]\,\mathbb{Q}(\mathrm{d}L)\,\mathbb{Q}(\mathrm{d}M)\,.$$

Now, putting

$$\mathscr{C} := \frac{\kappa_{d-2m}}{2}\int_{\mathbb{G}_m^d}\int_{\mathbb{G}_m^d}[L,M]\,\mathbb{Q}(\mathrm{d}L)\,\mathbb{Q}(\mathrm{d}M)\,, \tag{7.6}$$

we see that

$$\mathrm{d_{TV}}\left(\mathbf{L}_{t,a}, \mathscr{C}\,a^{d-2m}\,\mathrm{vol}|_K\right) = 0\,,$$

where $\mathrm{vol}|_K$ stands for the restriction of the Lebesgue measure on $\mathbb{R}^d$ to K. Moreover, the proof of [44, Theorem 3] shows that there is a constant $\hat{C} > 0$ only depending on a, d, m, $\mathbb{Q}$ and K such that

$$\hat{r}_t := \sup_{E\in\mathbb{A}_m^d} t\int_{\mathbb{A}_m^d}\mathbf{1}(m(E,F)\in K, \mathrm{dist}(E,F)\le at^{-2/(d-2m)})\,\mathbf{K}(\mathrm{d}F) \le \hat{C}\,t^{-1}\,.$$

From this we conclude that

$$r_t := t^3\int_{\mathbb{A}_m^d}\left(\int_{\mathbb{A}_m^d}\mathbf{1}(m(E,F)\in K, \mathrm{dist}(E,F)\le at^{-2/(d-2m)})\,\mathbf{K}(\mathrm{d}E)\right)^2\mathbf{K}(\mathrm{d}F) \le \hat{C}\,\mathbf{L}_{t,a}(K)\,t^{-1}$$

and in view of Theorem 3.1 the following result for the midpoint process $\xi_{t,a}$.

Theorem 7.15. *Let $a > 0$ and let ζ be a Poisson process with intensity measure $\mathscr{C}\,a^{d-2m}\,\mathrm{vol}|_K$, where $\mathscr{C}$ is as at (7.6). Then there is a constant $C > 0$ depending on a, d, m, $\mathbb{Q}$ and K such that*

$$\mathbf{d_{KR}}(\xi_{t,a},\zeta) \le C\,t^{-1}, \quad t \ge 1\,.$$

Remark 7.16. (i) Note that because of (2.8), the constant $\mathscr{C}$ takes the particularly appealing form

$$\mathscr{C} = \frac{1}{2}\frac{\binom{d-m}{m}}{\binom{d}{m}}\frac{\kappa_{d-m}^2}{\kappa_d}$$

if $\mathbb{Q}$ is the invariant Haar probability measure on $\mathbb{G}_m^d$ (or, equivalently, if the m-flat process is stationary and isotropic, see [38]).

(ii) As opposed to our previous applications, we do not consider a binomial counterpart to Theorem 7.15. The reason for that is that there is no normalization, which would turn the measure $\mathbf{K}_1$ defined at (7.5) into a probability measure.

(iii) Theorem 7.15 extends Theorem 7.13 from $m = 0$ (which has been excluded here for technical reasons) to arbitrary m satisfying $m < d/2$. However, due to the slightly different set-ups (an underlying point process on the compact set K vs. a point process on the non-compact space $\mathbb{A}_m^d$), there are boundary effects in the context of Theorem 7.13, implying that the total variation distance $\mathrm{d_{TV}}(\mathbf{L}_{t,a},\mathbf{M})$ is not identically zero there. These boundary effects are not present for $m \ge 1$, which eventually leads to the rate $O(t^{-1})$ for the Kantorovich-Rubinstein distance in this case.

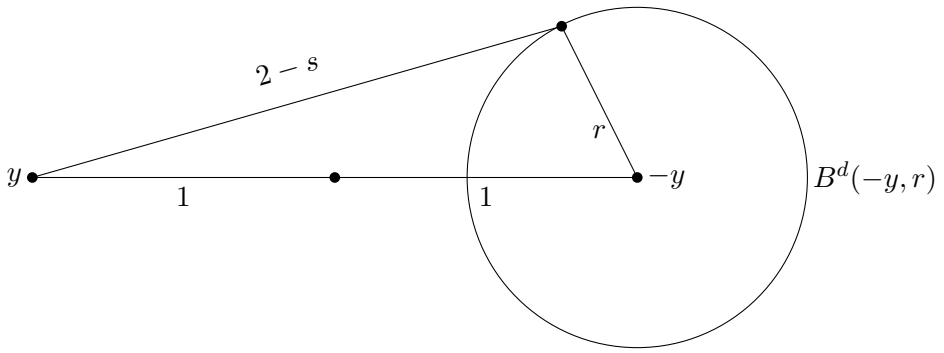

Figure 1: Illustration of the argument used in the derivation of Theorem 7.17

7.6 Random polytopes with vertices on the sphere

Let $\mathbb{S}^{d-1}$ be the unit sphere of dimension $d-1$ ($d \geq 2$). Let μ_t be a Poisson process η_t on $\mathbb{S}^{d-1}$ whose intensity measure is a constant multiple $t \geq 1$ of the normalized spherical Lebesgue measure or a binomial process β_t of $\lceil t \rceil$ independent and uniformly chosen points on $\mathbb{S}^{d-1}$. The convex hull $\mathrm{conv}(\mu_t)$ of μ_t is a random polytope with vertices on $\mathbb{S}^{d-1}$, and we denote by D_t the diameter of $\mathrm{conv}(\mu_t)$, that is,

$$D_t := \max_{(x,y)\in\mu_{t,\neq}^2} \|x-y\|\,.$$

More generally, define the point process of all reversed interpoint distances by

$$\xi_t = \frac{1}{2} \sum_{(x,y)\in\mu_{t,\neq}^2} \delta_{2-\|x-y\|}\,.$$

Clearly, D_t is then two minus the distance from the origin to the closest point of ξ_t. We define

$$\mathbf{L}_t(A) := \frac{1}{2}\,\mathbf{E} \sum_{(x,y)\in\mu_{t,\neq}^2} \mathbf{1}(t^{4/(d-1)}(2-\|x-y\|)\in A)\,, \qquad A\subset\mathbb{R}_+ \quad \text{Borel.}$$

Let $\chi(t) := t^2$ in the Poisson case and $\chi(t) := \lceil t\rceil(\lceil t\rceil - 1)$ in the binomial case. Applying the Mecke formula (2.1) or its analogue (2.2) for binomial processes, respectively, we see that

$$\mathbf{L}_t([0,a]) = \frac{\chi(t)}{2(d\kappa_d)^2} \int_{\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} \mathbf{1}(\|x-y\| \geq 2 - at^{-4/(d-1)})\,\mathcal{H}^{d-1}(\mathrm{d}x)\,\mathcal{H}^{d-1}(\mathrm{d}y)\,,$$

where $d\kappa_d$ is the surface area of $\mathbb{S}^{d-1}$ and $\mathcal{H}^{d-1}$ stands for the $(d-1)$-dimensional Hausdorff measure. For fixed $y\in\mathbb{S}^{d-1}$, the indicator function is one if and only if the point x is contained in a certain spherical cap $\mathbb{S}^{d-1}\cap B^d(-y,r)$ centred at the antipodal point $-y$ of y, whose radius r has to be determined. For this, we refer to Figure 1 and notice that $(2-s)^2 + r^2 = 4$ so that $r = \sqrt{4s-s^2}$. Hence, the $(d-1)$-dimensional volume of $\mathbb{S}^{d-1}\cap B^d(-y,r)$ is given by

$$(d-1)\kappa_{d-1}\int_0^{2s-s^2/2} (2h-h^2)^{(d-3)/2}\,\mathrm{d}h\,,$$

independently of y. Using the substitution $h = 2ut^{-4/(d-1)} - u^2t^{-8/(d-1)}/2$ this means that

$$\begin{aligned}\mathbf{L}_t([0,a]) &= \frac{\chi(t)}{2d\kappa_d}\,(d-1)\kappa_{d-1}\int_0^{2at^{-4/(d-1)}-a^2t^{-8/(d-1)}/2} (2h-h^2)^{(d-3)/2}\,\mathrm{d}h\\ &= \frac{\chi(t)}{2d\kappa_d}\,(d-1)\kappa_{d-1}\int_0^a \big(4ut^{-4/(d-1)} - u^2t^{-8/(d-1)} - (2ut^{-4/(d-1)} - u^2t^{-8/(d-1)}/2)^2\big)^{(d-3)/2}\\ &\qquad\qquad \times (2t^{-4/(d-1)} - ut^{-8/(d-1)})\,\mathrm{d}u\end{aligned}$$

$$= \frac{1}{2d\kappa_d}\,\frac{\chi(t)}{t^2}\,(d-1)\kappa_{d-1}\int_0^a \big(4u - u^2t^{-4/(d-1)} - t^{-4/(d-1)}(2u-u^2t^{-4/(d-1)}/2)^2\big)^{(d-3)/2} \times (2-ut^{-4/(d-1)})\,\mathrm{d}u\,.$$

Hence, we have for any Borel set $A \subset \mathbb{R}_+$ that

$$\mathbf{L}_t(A) = \frac{(d-1)\kappa_{d-1}}{2d\kappa_d}\,\frac{\chi(t)}{t^2}\int_A \big(4u - u^2t^{-4/(d-1)} - t^{-4/(d-1)}(2u-u^2t^{-4/(d-1)}/2)^2\big)^{(d-3)/2} \times (2-ut^{-4/(d-1)})\,\mathrm{d}u\,.$$

The measure $\mathbf{L}_t$ converges, as $t\to\infty$ and in the strong sense, to a measure $\mathbf{M}$ on $\mathbb{R}_+$ given by

$$\mathbf{M}(A) := \frac{d-1}{d\kappa_d}\kappa_{d-1}2^{d-3}\int_A u^{(d-3)/2}\,\mathrm{d}u\,, \qquad A\subset\mathbb{R}_+ \quad \text{Borel}\,. \tag{7.7}$$

Moreover, for any bounded Borel set $B\subset\mathbb{R}_+$ there is a constant $c_{1,B}>0$ only depending on B and the space dimension d such that

$$\mathrm{d_{TV}}(\mathbf{L}_t|_B, \mathbf{M}|_B) \le c_{1,B}t^{-\min\{4/(d-1),1\}}, \quad t\ge 1\,.$$

Here, we have used that $|\chi(t)/t^2-1|\le t^{-1}$ for $t\ge 1$. Let $\widetilde{\chi}(t):=t$ in the Poisson case and $\widetilde{\chi}(t):=\lceil t\rceil$ in the binomial case. The same arguments as above also show that

$$\hat{r}_t(B) := \sup_{x\in\mathbb{S}^{d-1}}\frac{\widetilde{\chi}(t)}{d\kappa_d}\int_{\mathbb{S}^{d-1}}\mathbf{1}(2-\|x-y\|\in t^{-4/(d-1)}B)\,\mathcal{H}^{d-1}(\mathrm{d}y) \le c_{2,B}\,t^{-1}$$

with a constant $c_{2,B}>0$ only depending on B and d so that $\mathbf{L}_t(B)\,\hat{r}_t(B)\le c_{2,B}\,\mathbf{L}_t(B)\,t^{-1}$. Combining Corollary 3.3 and Remark 3.4 (iii) we conclude the following result.

Theorem 7.17. *Let ζ be a Poisson process on $\mathbb{R}_+$ with intensity measure given by (7.7) and let ξ_t be derived from a Poisson process η_t or a binomial process β_t on $\mathbb{S}^{d-1}$. Then, for any bounded Borel set $B\subset\mathbb{R}_+$ there is a constant $C_{B,d}>0$ only depending on B and d such that*

$$\mathbf{d_{KR}}\big((t^{4/(d-1)}\bullet\xi_t)|_B, \zeta|_B\big) \le C_{B,d}\,t^{-\min\{4/(d-1),1\}}, \quad t\ge 1\,.$$

In particular, for the diameter D_t of the random polytope, constructed from a Poisson process η_t or a binomial process β_t, we have

$$\big|\mathbf{P}(t^{4/(d-1)}(2-D_t)>a) - e^{-\frac{1}{d\kappa_d}\kappa_{d-1}2^{d-2}a^{(d-1)/2}}\big| \le C_{a,d}\,t^{-\min\{4/(d-1),1\}}, \quad t\ge 1,$$

with a constant $C_{a,d}>0$ only depending on $a>0$ and d.

Remark 7.18. The limiting distribution for the diameter is also derived in [30, Theorem 5.2] and [26, Theorem 3.1], where the latter one allows the underlying random points to have distributions different from the uniform distribution. While the result in [30] does not give any rates of convergence, in [26, Theorem 3.1] it has erroneously been claimed that the rate of convergence for D_t to its limiting Weibull random variable is of order t^{-1}. However, in our notation the rate of convergence stated in (2.5) in [26] concerns only the difference to a Weibull random variable with parameter $\mathbf{L}_t([0,a])$ and not to a Weibull random variable with parameter $\mathbf{M}([0,a])$ as stated by the authors. For the difference to a Weibull random variable with parameter $\mathbf{L}_t([0,a])$ our result also yields a rate of order t^{-1} since $\mathrm{d_{TV}}(\mathbf{L}_t|_{[0,a]}, \mathbf{L}_t|_{[0,a]}) = 0$ in this case.

Acknowledgements

We would like to thank an anonymous referee for valuable hints and comments, which helped us to improve the text.

This research has been initiated during the Oberwolfach mini-workshop "Stochastic Analysis for Poisson Point Processes". All support is gratefully acknowledged.

LD has partially been supported by ANR Masterie. MS has been funded by the German Research Foundation (DFG) through the research unit "Geometry and Physics of Spatial Random Systems" under the grant HU 1874/3-1. CT has been supported by the German research foundation (DFG) via SFB-TR 12.

References

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

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

[3] Barbour, A.D.: *Stein's method for diffusion approximations*, Probab. Theory Related Fields **84**, 297–322 (1990).

[4] Barbour, A.D.; Brown, T.C.: *Stein's method and point process approximation*, Stochastic Process. Appl. **43**, 9–31 (1992).

[5] Barbour, A.D.; Chen, L.; Loh, W.-L.: *Compound Poisson approximation of nonnegative random variables via Stein's method*, Ann. Probab. **20**, 1843–1866 (1992).

[6] Barbour, A.D.; Chryssaphinou, O.: *Compound Poisson approximation: a user's guide*, Ann. Probab. **11**, 964–1002 (2001).

[7] Barbour, A.D.; Eagleson, G.K.: *Poisson convergence for dissociated statistics*, J. Royal Statist. Soc. B **46**, 397–402 (1984).

[8] Barbour, A.D.; Holst, L.; Janson, S.: *Poisson Approximation*, Oxford University Press, Oxford (1992).

[9] Barbour, A.D.; Utev, S.: *Solving the Stein equation in compound Poisson approximation*, Adv. in Appl. Probab. **30**, 449–475 (1998).

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

[11] Brown, T.C.; Weinberg, G.V.; Xia, A.: *Removing logarithms from Poisson process error bounds*, Stochastic Process. Appl. **87**, 149–165 (2000).

[12] Brown, T.C.; Xia, A.: *On metrics in point process approximation*, Stochastics Stochastics Rep. **52**, 247–263 (1995).

[13] Brown, T.C.; Xia, A.: *Stein's method and birth-death processes*, Ann. Probab. **29**, 1373–1403 (2001).

[14] Chen, L.H.Y: *Poisson approximation for dependent trials*, Ann. Probab. **3**, 534–545 (1975).

[15] Chen, L.H.Y.; Xia, A.: *Stein's method, Palm theory and Poisson process approximation*, Ann. Probab. **32**, 2545–2569 (2004).

[16] Coutin, L.; Decreusefond, L.: *Stein's method for Brownian approximations*, Commun. Stoch. Anal. **7**, 349–372 (2013).

[17] Dabrowski, A.R.; Dehling, H.G.; Mikosch, T.; Sharipov, O.: *Poisson limits for U-statistics*, Stochastic Process. Appl. **99**, 137–157 (2002).

[18] Daley, D.J.; Vere-Jones, D.: *An Introduction to the Theory of Point Processes: Vol. II. General Theory and Structure.* Second edition, Springer, New York (2008).

[19] Decreusefond, L.: *Wasserstein distance on configuration space*, Potential Anal. **28**, 283–300 (2008).

[20] Decreusefond, L.; Joulin, A.; Savy, N.: *Upper bounds on Rubinstein distances on configuration spaces and applications*, Commun. Stoch. Anal. **4**, 377–399 (2010).

[21] Eichelsbacher, P.; Roos, M.: *Compound Poisson approximation for dissociated random variables via Stein's method*, Combin. Probab. Comput. **8**, 335–346 (1999).

[22] Heinrich, L.; Wolf, W.: *On the convergence of U-statistics with stable limit distributions*, J. Multivariate Anal. **44**, 266–278 (1993).

[23] Hug, D.; Schneider, R.; Schuster, R.: *Integral geometry of tensor valuations*, Adv. in Appl. Math. **41**, 482–509 (2008).

[24] Kallenberg, O.: *Foundations of Modern Probability*, Second edition, Springer, New York (2002).

[25] Kingman, J.F.C.: *Poisson Processes*, Oxford University Press, Oxford (1993).

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

[27] Last, G.; Peccati, G.; Schulte, M.: *Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization*, arXiv: 1401.7568 (2014).

[28] Last, G.; Penrose, M.D.: *Poisson process Fock space representation, chaos expansion and covariance inequalities*, Probab. Theory Related Fields **150**, 663–690 (2011).

[29] Malevich, T.L.; Abdalimov, B.: *Stable limit distributions for U-statistics*, Theory Probab. Appl. **22**, 370–377 (1977).

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

[31] Peccati, G: *The Chen-Stein method for Poisson functionals*, arXiv: 1112.5051 (2011).

[32] Lachièze-Rey, R.; Peccati, G.: *Fine Gaussian fluctuations on the Poisson space, I: contractions, cumulants and random geometric graphs*, Electron. J. Probab. 18, article 32, 1–35 (2013).

[33] Peccati, G.; Solé, J.L.; Taqqu, M.S.; Utzet, F.: *Stein's method and normal approximation of Poisson functionals*, Ann. Probab. **38**, 443–478 (2010).

[34] Penrose, M.D.: *Random Geometric Graphs*, Oxford University Press, Oxford (2003).

[35] Preston, C.: *Spatial birth-and-death processes*, in Proceedings of the 40th Session of the International Statistical Institute (Warsaw, 1975), Vol. 2, Bull. Inst. Internat. Statist., Vol. **46**, 371–391 (1975).

[36] Reinert, G.: *Three general approaches to Stein's method*, in: An introduction to Stein's method, Singapore University Press, Singapore (2005).

[37] Reitzner, M.; Schulte, M.; Thäle, C.: *Limit theory for the Gilbert graph*, arXiv: 1312.4861 (2013).

[38] Schneider, R.; Weil, W.: *Stochastic and Integral Geometry*, Springer, Berlin (2008).

[39] Schuhmacher, D.: *Upper bounds for spatial point process approximations*, Ann. Appl. Probab. **15**, 615–651 (2005).

[40] Schuhmacher, D.: *Stein's method and Poisson process approximation for a class of Wasserstein metrics*, Bernoulli **15**, 550–568 (2009).

[41] Schuhmacher, D.; Stucki, K.: *Gibbs point process approximation: total variation bounds using Stein's method*, Ann. Prob. **42**, 1911–1951 (2014).

[42] Schuhmacher, D.; Xia, A.: *A new metric between distributions of point processes*, Adv. in Appl. Probab. **40**, 651–672 (2008).

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

[44] Schulte, M.; Thäle, C.: *Distances between Poisson k-flats*, Methodol. Comput. Appl. Probab. **16**, 311–329 (2014).

[45] Shih, H.-H.: *On Stein's method for infinite-dimensional Gaussian approximation in abstract Wiener spaces*, J. Funct. Anal. **261**, 1236–1283 (2011).

[46] Villani, C.: *Optimal Transport: Old and New*, Springer, Berlin (2009).

[47] Zolotarev, V.M.: *One-dimensional Stable Distributions*, American Mathematical Society, Providence (1986).