Electronic Journal of Probability

Vol. 14 (2009), Paper no. 38, pages 1080–1116.

Journal URL
http://www.math.washington.edu/~ejpecp/

Distance estimates for dependent thinnings of point processes with densities

Dominic Schuhmacher*†
School of Mathematics and Statistics
University of Western Australia, Crawley WA 6009, Australia
dominic@maths.uwa.edu.au

Abstract

In [Schuhmacher, *Electron. J. Probab.* **10** (2005), 165–201] estimates of the Barbour-Brown distance d_2 between the distribution of a thinned point process and the distribution of a Poisson process were derived by combining discretization with a result based on Stein's method. In the present article we concentrate on point processes that have a density with respect to a Poisson process, for which we can apply a corresponding result directly without the detour of discretization. This enables us to obtain better and more natural bounds in the d_2-metric, and for the first time also bounds in the stronger total variation metric.

We give applications for thinning by covering with an independent Boolean model and "Matérn type I"-thinning of fairly general point processes. These applications give new insight into the respective models, and either generalize or improve earlier results.

Key words: Point process, Poisson process approximation, Stein's method, point process density, random field, thinning, total variation distance, Barbour-Brown distance.

AMS 2000 Subject Classification: Primary 60G55; Secondary: 60E99, 60D05.

Submitted to EJP on September 21, 2007, final version accepted April 7, 2009.

*Current address: Institute of Mathematical Statistics and Actuarial Science, Alpeneggstrasse 22, 3012 Bern, Switzerland. E-Mail: schuhmacher@stat.unibe.ch.

†Work supported by the Swiss National Science Foundation under Grant No. PBZH2-111668.

1 Introduction

We consider thinnings of simple point processes on a general compact metric space $\mathscr{X}$, where simple means that the probability of having multiple points at the same location is zero. The thinning of such a process ξ according to a $[0,1]$-valued measurable random field π on $\mathscr{X}$ is the point process ξ_π, unique with regard to its distribution, that can be obtained in the following way: for any realizations $\xi(\omega)$ (a point measure on $\mathscr{X}$) and $\pi(\omega,\cdot)$ (a function $\mathscr{X} \to [0,1]$), look at each point s of $\xi(\omega)$ in turn, and retain it with probability $\pi(\omega,s)$, or delete it with probability $1-\pi(\omega,s)$, independently of any retention/deletion decisions of other points. Regard the points left over by this procedure as a realization of the thinned point process ξ_π. We usually refer to ξ as *the original process*, and to π as *the retention field*.

The following is a well-established fact: if we thin more and more, in the sense that we consider a sequence of retention fields $(\pi_n)_{n\in\mathbb{N}}$ with $\sup_{x\in\mathscr{X}} \pi_n(x) \xrightarrow{\mathscr{D}} 0$ as $n \to \infty$, and compensate for the thinning by choosing point processes ξ_n whose intensities increase as n goes to infinity in a way that is compatible with the retention fields, then we obtain convergence in distribution towards a Cox process. Convergence in distribution for random measures, and in particular for point processes, is defined via the convergence of expectations of bounded continuous functions, where continuity is in terms of the vague topology (for details see [13], Section 4.1).

In order to specify what choice of the sequence (ξ_n) is compatible with (π_n), we introduce the random measure Λ_n that is given by $\Lambda_n(A) := \int_A \pi_n(x)\, \xi_n(dx)$ for every Borel set A in $\mathscr{X}$. The theorem below was shown in [12] for constant deterministic π_n, and in [5] for general π_n (in fact, both times under the more general assumption that $\mathscr{X}$ is a locally compact, second countable Hausdorff space). For a detailed history of this result see [22], p. 167; note also the contribution of Mecke, who established the fundamental relationship between thinnings and Cox processes [16].

Theorem 1.A (Kallenberg [12], Brown [5])**.** *For the sequences $(\pi_n)_{n\in\mathbb{N}}$ and $(\xi_n)_{n\in\mathbb{N}}$ introduced above, we obtain convergence in distribution of the thinned sequence $\big((\xi_n)_{\pi_n}\big)_{n\in\mathbb{N}}$ towards a point process η if and only if there is a random measure Λ on $\mathscr{X}$ such that $\Lambda_n \xrightarrow{\mathscr{D}} \Lambda$ as $n \to \infty$. In this case $\eta \sim \mathrm{Cox}(\Lambda)$, i.e. η is a Cox process with directing measure Λ.*

In [22] the above setting was considered for the situation that the ξ_n are point processes on $[0,1]^D$ that are obtained from a single point process ξ by gradual contraction of $\mathbb{R}_+^D$ using the functions κ_n given by $\kappa_n(x) := (1/n)x$ for every $x \in \mathbb{R}_+^D$ (for notational convenience in the proofs, the order of contracting and thinning was interchanged). Under the additional assumption that ξ and π_n satisfy mixing conditions, which makes it plausible for the limiting process in Theorem 1.A to be Poisson (see the remark after Theorem 1.A of [22]), several upper bounds for the Wasserstein-type distance d_2 between the distribution of the thinned and contracted process and a suitable Poisson process distribution were obtained under various conditions. These results were derived by discretizing the thinned process and the limiting Poisson process, and applying then a discrete version of the "local Barbour-Brown theorem", Theorem 3.6 in [2], which is based on Stein's method (see [24] and [3]).

Although the bounds were of good quality and have proved their usefulness in several applications, they had some shortcomings, which were mainly related to the fact that they could only be expressed in terms of discretization cuboids. This made the results rather unpleasant to apply in many situations where truly non-discrete point processes were considered. However, it had appeared difficult for quite some time to get control over the "long range weak dependences" in a meaningful way without resorting to discretization.

The present article overcomes these difficulties. We derive a slightly adapted version of Theorem 3.6 in [2] that simplifies some of the notation and corrects a minor mistake in the original result, and are able to employ it directly without the detour of discretization in order to obtain meaningful upper bounds in a more elegant way. These upper bounds are much more natural and more easily applied than the previous ones; they hold for the most part under more general conditions and are qualitatively somewhat better. Furthermore, we also get reasonable estimates in the stronger total variation metric d_{TV}, which was not possible with the previous method due to the fact that the distance between a discrete and a continuous point process distribution is always 1 in this metric.

Many of these improvements become evident from looking at the example in Corollary 4.G. In this example, a Poisson process on $\mathbb{R}^D$ that is partly covered by a Boolean model of Euclidean balls is considered when contracting space by a factor n and suitably adapting the L_D-norm of the radii of the balls to stabilize the intensity. We show that the total variation distance for Poisson process approximation is $O\big((\log n)^{-(D-1)}\big)$ as $n \to \infty$. In contrast, all that is obtainable for this setting from the corresponding considerations in [22] is that, for arbitrarily small $\zeta > 0$, there is a constant $c(\zeta)$ such that the (weaker) d_2-distance is bounded by $c(\zeta)n^{\zeta D}$ for every n. The latter is not useful at all and even requires a more involved argument to compute the first and second moment terms of the retention field because suprema over discretization cuboids have to be considered.

In order to apply our version of Theorem 3.6 in [2], we have to restrict ourselves to point processes ξ that have a density with respect to the distribution of a simple Poisson process, which is a natural and common choice for a reference measure and leaves us with a very rich class of processes.

We start out in Section 2 by giving some notation and technical background including our adaptation of Theorem 3.6 of [2] in Section 2.4. The main results are then presented in Section 3, which can be read without detailed knowledge of Section 2. We provide upper bounds for the d_{TV}- and the d_2-distances between $\mathscr{L}(\xi_\pi)$ and a suitable Poisson process distribution, first in a general setting, and then for a number of important special cases. The last of these special cases (see Corollary 3.E) is suitable for comparison with the upper bounds in [22]. Finally, in Section 4, two applications of the main results are studied. The first one is a more general version of the thinning by covering with a Boolean model that was mentioned above, and improves on results in [22]. In the second application, any point of ξ is deleted if there is another point present within a fixed distance r (following the construction of the Matérn type I hard core process, but using a more general ξ), and retained with probability q otherwise. The bounds obtained in this setting generalize a result in [26], where $q = 1$ and ξ had to be a Poisson process.

2 Preliminaries

We first introduce some basic notation and conventions, before giving an overview of some of the theoretical background and presenting the more technical definitions in the various subsections. The reader may want to skip these subsections on first reading, as the gist of non-technical information is repeated where it first appears in Section 3 or else easily accessible by the cross-references given.

Let $(\mathscr{X}, d_0)$ always be a compact metric space with $d_0 \leq 1$ that admits a finite diffuse measure $\alpha \neq 0$, where diffuse means that $\alpha(\{x\}) = 0$ for every $x \in \mathscr{X}$. Denote by $\mathscr{B}$ the Borel σ-algebra on $\mathscr{X}$, and by $\mathscr{B}_A$ the trace σ-algebra $\mathscr{B}|_A = \{B \cap A; B \in \mathscr{B}\}$ for any set $A \subset \mathscr{X}$. Furthermore, write $\mathfrak{M}$ for the space of finite measures on $\mathscr{X}$, and equip it with the vague topology (see [13], Section 15.7) and the corresponding Borel σ-algebra $\mathscr{M}$ (see [13], Lemma 4.1 and Section 1.1). Do the same for the subspace $\mathfrak{N} \subset \mathfrak{M}$ of finite point measures, and denote the corresponding σ-algebra by $\mathscr{N}$. Write furthermore $\mathfrak{N}^* := \{\varrho \in \mathfrak{N};\ \varrho(\{x\}) \leq 1 \text{ for all } x \in \mathscr{X}\}$ for the $\mathscr{N}$-measurable set of simple point measures. A *random measure* on $\mathscr{X}$ is a random element of $\mathfrak{M}$, and a *point process* on $\mathscr{X}$ a random element of $\mathfrak{N}$. A point process ξ is called *simple* if $\mathbb{P}[\xi \in \mathfrak{N}^*] = 1$. By $\mathrm{Po}(\lambda)$ we denote the distribution of the Poisson process on $\mathscr{X}$ with intensity measure λ if $\lambda \in \mathfrak{M}$, and the Poisson distribution with parameter λ if λ is a positive real number.

We think of measures (random or otherwise) always as being defined on all of $\mathscr{X}$. Thus, for any measure μ on $\mathscr{X}$ and any $A \in \mathscr{B}$, we denote by $\mu|_A$ the measure on $\mathscr{X}$ that is given by $\mu|_A(B) := \mu(A \cap B)$ for all $B \in \mathscr{B}$. Let $\mathfrak{M}(A) := \{\mu|_A;\ \mu \in \mathfrak{M}\}$, $\mathfrak{N}(A) := \{\varrho|_A;\ \varrho \in \mathfrak{N}\}$, $\mathscr{M}(A) := \mathscr{M}|_{\mathfrak{M}(A)} = \{C \cap \mathfrak{M}(A);\ C \in \mathscr{M}\}$, and $\mathscr{N}(A) := \mathscr{N}|_{\mathfrak{N}(A)} = \{C \cap \mathfrak{N}(A);\ C \in \mathscr{N}\}$. Furthermore, set $\mathfrak{N}^*(A) := \mathfrak{N}(A) \cap \mathfrak{N}^*$. Sometimes absolute value bars are used to denote the total mass of a measure, i.e. $|\mu| := \mu(\mathscr{X})$ for any $\mu \in \mathfrak{M}$.

For $\sigma \in \mathfrak{N}^*$, we do not notationally distinguish between the point measure and its support. Like that we can avoid having to enumerate the points of the measure, which sometimes saves us from tedious notation. Thus, the notation $\int f\, d\sigma = \sum_{s \in \sigma} f(s)$ may be used instead of writing $\int f\, d\sigma = \sum_{i=1}^{v} f(s_i)$, where $\sigma = \sum_{i=1}^{v} \delta_{s_i}$. The same concept is extended to the realizations of a simple point process, as long as it is not important what happens on the null set of point patterns with multiple points. Hence we may also write $\mathbb{E}\big(\int f\, d\xi\big) = \mathbb{E}\big(\sum_{s \in \xi} f(s)\big)$ if ξ is simple. In order to facilitate the reading of certain formulae, we furthermore make the convention of using the letters $x, \tilde{x}, y$ for general elements of the state space $\mathscr{X}$, while $s, \tilde{s}, t$ are reserved for points of a point pattern in $\mathscr{X}$. Finally, we sometimes omit the addition "almost surely" or "(for) almost every ..." for equations and inequalities between functions on measure spaces if it is evident and of no importance that the relation does not hold pointwise.

2.1 Densities with respect to the standard Poisson process distribution P_1

Let $\alpha \neq 0$ be a fixed diffuse measure in $\mathfrak{M}$, which we will regard as our reference measure on $\mathscr{X}$. Typically, if (a superset of) $\mathscr{X}$ has a suitable group structure, α is chosen to be (the restriction of) the Haar measure. If $\mathscr{X} \subset \mathbb{R}^D$, we usually choose $\alpha = \mathit{Leb}^D|_{\mathscr{X}}$. We write $P_1 := \mathrm{Po}(\alpha)$ for the distribution of what we call the *standard Poisson process* on $\mathscr{X}$, and $P_{1,A} := \mathrm{Po}(\alpha|_A)$ for the distribution of the Poisson process on $\mathscr{X}$ with expectation measure $\alpha|_A$. It is convenient to admit also $\alpha(A) = 0$, in which case $P_{1,A} = \delta_0$, where 0 denotes the zero measure on $\mathscr{X}$.

A popular way of specifying a point process distribution is by giving its Radon-Nikodym density with respect to the distribution of the standard Poisson process (if the density exists; see [17], Sections 6.1 and 6.2 for a number of examples). The following lemma, which is a simple consequence of Theorem 3.1.1 in [21], will be useful.

Lemma 2.A. *For any $a > 0$, a density f_A of $P_{a,A} := \mathrm{Po}(a \cdot \alpha|_A)$ with respect to $P_{1,A}$ is given by*

$$f_A(\varrho) = e^{(1-a)\alpha(A)} a^{|\varrho|}$$

for every $\varrho \in \mathfrak{N}(A)$. □

To avoid certain technical problems, we require our densities to be hereditary whenever we are dealing with conditional densities.

Definition. A function $f : \mathfrak{N} \to \mathbb{R}_+$ is called *hereditary* if $f(\varrho) = 0$ implies $f(\sigma) = 0$ whenever $\varrho, \sigma \in \mathfrak{N}$ with $\varrho \subset \sigma$.

The point processes on $\mathscr{X}$ that have a hereditary density with respect to P_1 form the class of *Gibbs point processes*. These include pairwise and higher order interaction processes (see [17], Section 6.2). The general form of a Gibbs process density is given in Definition 4.2 of [1].

2.2 Thinnings

In what follows, let ξ always be a point process on $\mathscr{X}$ that has density f with respect to P_1, and let $\pi := (\pi(\cdot, x); x \in \mathscr{X})$ be a $[0,1]$-valued random field on $\mathscr{X}$ that is measurable in the sense that the mapping $\Omega \times \mathscr{X} \to [0,1]$, $(\omega, x) \mapsto \pi(\omega, x)$ is $\sigma(\pi) \otimes \mathscr{B}$-measurable. In the main part we strengthen the technical conditions on ξ and π somewhat in order to avoid some tedious detours in the proofs.

We use the definition from [22] for the π-thinning of ξ.

Definition (Thinning)**.** First, assume that $\xi = \sigma = \sum_{i=1}^{v} \delta_{s_i}$ and $\pi = p$ are non-random, where $v \in \mathbb{Z}_+$, $s_i \in \mathscr{X}$, and p is a measurable function $\mathscr{X} \to [0,1]$. Then, a π-thinning of ξ is defined as $\xi_\pi = \sum_{i=1}^{v} X_i \delta_{s_i}$, where the X_i are independent Bernoulli random variables with expectations $p(s_i)$, respectively. Under these circumstances, ξ_π has a distribution $P_{(\sigma,p)}$ that does not depend on the chosen enumeration of σ. We obtain the general π-thinning from this by randomization, that is by the condition $\mathbb{P}[\xi_\pi \in \cdot \,|\, \xi, \pi] = P_{(\xi,\pi)}$ (it is straightforward to see that $P_{(\xi,\pi)}$ is a $\sigma(\xi,\pi)$-measurable family in the sense that $P_{(\xi,\pi)}(D)$ is $\sigma(\xi,\pi)$-measurable for every $D \in \mathfrak{N}$). Note that the distribution of ξ_π is uniquely determined by this procedure.

The following lemma gives the first two factorial moment measures of ξ_π. For $\varrho = \sum_{i=1}^{v} \delta_{s_i} \in \mathfrak{N}$, write $\varrho^{[2]} := \sum_{i,j=1, i\neq j}^{v} \delta_{(s_i,s_j)}$ for the factorial product measure on $\mathscr{X} \times \mathscr{X}$. Remember that the expectation measure μ_1 of ξ is defined by $\mu_1(A) := \mathbb{E}(\xi(A))$ for every $A \in \mathscr{B}$, and that the second factorial moment measure $\mu_{[2]}$ of ξ is defined to be the expectation measure of $\xi^{[2]}$. Let Λ be the random measure on $\mathscr{X}$ that is given by $\Lambda(A) := \int_A \pi(x)\, \xi(dx)$ for $A \in \mathscr{B}$ (cf. Λ_n in Section 1), and write furthermore $\Lambda^{[2]}(B) := \int_B \pi(x)\pi(\tilde{x})\, \xi^{[2]}\big(d(x,\tilde{x})\big)$ for every $B \in \mathscr{B}^2$.

Lemma 2.B. *We obtain for the expectation measure $\mu_1^{(\pi)}$ and the second factorial moment measure $\mu_{[2]}^{(\pi)}$ of ξ_π*

(i) $\mu_1^{(\pi)}(A) = \mathbb{E}\big(\Lambda(A)\big)$ *for every $A \in \mathscr{B}$;*

(ii) $\mu_{[2]}^{(\pi)}(B) = \mathbb{E}\big(\Lambda^{[2]}(B)\big)$ *for every $B \in \mathscr{B}^2$.*

Proof. Write $\xi = \sum_{i=1}^{V} \delta_{S_i}$, where V and S_i are $\sigma(\xi)$-measurable random elements with values in $\mathbb{Z}_+$ and $\mathscr{X}$, respectively. Such a representation exists by Lemma 2.3 in [13].

(i) For every $A \in \mathscr{B}$ we have

$$\begin{aligned}\mu_1^{(\pi)}(A) = \mathbb{E}\big(\xi_\pi(A)\big) &= \mathbb{E}\big(\mathbb{E}\big(\xi_\pi(A) \,\big|\, \xi, \pi\big)\big) \\ &= \mathbb{E}\Bigg(\sum_{i=1}^{V} \pi(S_i)\mathrm{I}[S_i \in A]\Bigg) \\ &= \mathbb{E}\Bigg(\int_A \pi(x)\, \xi(dx)\Bigg).\end{aligned}$$

(ii) For every $B \in \mathscr{B}^2$ we have

$$\begin{aligned}\mu_{[2]}^{(\pi)}(B) = \mathbb{E}\big(\xi_\pi^{[2]}(B)\big) &= \mathbb{E}\big(\mathbb{E}\big(\xi_\pi^{[2]}(B) \,\big|\, \xi, \pi\big)\big) \\ &= \mathbb{E}\Bigg(\sum_{\substack{i,j=1 \\ i\neq j}}^{V} \pi(S_i)\pi(S_j)\mathrm{I}[(S_i, S_j) \in B]\Bigg) \\ &= \mathbb{E}\Bigg(\int_B \pi(x)\pi(\tilde{x})\, \xi^{[2]}\big(d(x,\tilde{x})\big)\Bigg).\end{aligned}$$

□

2.3 Metrics used on the space of point process distributions

We use two metrics on the space $\mathfrak{P}(\mathfrak{N})$ of probability measures on $\mathfrak{N}$. The one that is more widely known is the total variation metric, which can be defined on any space of probability measures. For $P, Q \in \mathfrak{P}(\mathfrak{N})$ it is given by

$$d_{TV}(P,Q) = \sup_{C \in \mathscr{N}} \Big|P(C) - Q(C)\Big|$$

or, equivalently, by

$$d_{TV}(P,Q) = \min_{\substack{\xi_1 \sim P \\ \xi_2 \sim Q}} \mathbb{P}[\xi_1 \neq \xi_2]. \tag{2.1}$$

See [4], Appendix A.1, for this and other general results about the total variation metric.

The second metric we use is a Wasserstein type of metric introduced by Barbour and Brown in [2], and denoted by d_2. It is often a more natural metric to use on $\mathfrak{P}(\mathfrak{N})$ than d_{TV}, because it takes the metric d_0 on $\mathscr{X}$ into account and metrizes convergence in distribution of point processes. The total variation metric on the other hand is strictly stronger, and at times too strong to be useful.

Define the d_1-*distance* between two point measures $\varrho_1 = \sum_{i=1}^{|\varrho_1|} \delta_{s_{1,i}}$ and $\varrho_2 = \sum_{i=1}^{|\varrho_2|} \delta_{s_{2,i}}$ in $\mathfrak{N}$ as

$$d_1(\varrho_1, \varrho_2) := \begin{cases} 1 & \text{if } |\varrho_1| \neq |\varrho_2|, \\ \min_{\tau \in \Sigma_v} \big[\frac{1}{v}\sum_{i=1}^{v} d_0(s_{1,i}, s_{2,\tau(i)})\big] & \text{if } |\varrho_1| = |\varrho_2| = v > 0, \\ 0 & \text{if } |\varrho_1| = |\varrho_2| = 0, \end{cases} \tag{2.2}$$

where Σ_v denotes the set of permutations of $\{1, 2, \ldots, v\}$. It can be seen that $(\mathfrak{N}, d_1)$ is a complete, separable metric space and that d_1 is bounded by 1. Now let $\mathscr{F}_2 := \{f : \mathfrak{N} \to \mathbb{R};\ |f(\varrho_1) - f(\varrho_2)| \leq$

$d_1(\varrho_1, \varrho_2)$ for all $\varrho_1, \varrho_2 \in \mathfrak{N}\}$, and define the d_2-*distance* between two measures $P, Q \in \mathfrak{P}(\mathfrak{N})$ as

$$d_2(P,Q) := \sup_{f \in \mathscr{F}_2} \left| \int f \, dP - \int f \, dQ \right|$$

or, equivalently, as

$$d_2(P,Q) = \min_{\substack{\xi_1 \sim P \\ \xi_2 \sim Q}} \mathbb{E} d_1(\xi_1, \xi_2). \tag{2.3}$$

See [22], [23], and [26] for this and many other results about the Barbour-Brown metric d_2. By (2.1) and (2.3) we obtain that $d_2 \leq d_{TV}$.

2.4 Distance estimates for Poisson process approximation of point processes with a spatial dependence structure

In this subsection, a theorem is presented that provides upper bounds for the total variation and d_2 distances between the distribution of a point process with a spatial dependence structure and a suitable Poisson process distribution. This result is very similar to Theorems 2.4 and 3.6 in [2], but deviates in several minor aspects, two of which are more pronounced: first, we use densities with respect to a Poisson process distribution instead of Janossy densities, which simplifies part of the notation considerably; secondly, we take a somewhat different approach for controlling the long range dependence within ξ (see the term for $\breve{\phi}(x)$ in Equation (2.11)), which avoids the imposition of an unwelcome technical condition (cf. Remark A.C).

Let ξ be a point process on $\mathscr{X}$ whose distribution has a density f with respect to P_1 and whose expectation measure $\mu = \mu_1$ is finite. Then μ has a density u with respect to α that is given by

$$u(x) = \int_{\mathfrak{N}} f(\varrho + \delta_x) \, P_1(d\varrho) \tag{2.4}$$

for α-almost every $x \in \mathscr{X}$, which is obtained as a special case of Equation (2.8) below.

For $A \in \mathscr{B}$, we set

$$f_A(\varrho) := \int_{\mathfrak{N}(A^c)} f(\varrho + \tilde{\varrho}) \, P_{1,A^c}(d\tilde{\varrho}) \tag{2.5}$$

for every $\varrho \in \mathfrak{N}(A)$, which gives a density $f_A : \mathfrak{N} \to \mathbb{R}_+$ of the distribution of $\xi|_A$ with respect to $P_{1,A}$ (we extend f_A on $\mathfrak{N} \setminus \mathfrak{N}(A)$ by setting it to zero). This can be seen by the fact that integrating f_A over an arbitrary set $C \in \mathfrak{N}(A)$ yields

$$\begin{aligned}
\int_C f_A(\varrho) \, P_{1,A}(d\varrho) &= \int_{\mathfrak{N}(A)} \int_{\mathfrak{N}(A^c)} \mathrm{I}[\varrho \in C] f(\varrho + \tilde{\varrho}) \, P_{1,A^c}(d\tilde{\varrho}) \, P_{1,A}(d\varrho) \\
&= \int_{\mathfrak{N}} \mathrm{I}[\sigma|_A \in C] f(\sigma) \, P_1(d\sigma) \\
&= \mathbb{P}[\xi|_A \in C],
\end{aligned}$$

where we used that $(\eta|_A, \eta|_{A^c}) \sim P_{1,A} \otimes P_{1,A^c}$ for $\eta \sim P_1$. Note that the argument remains correct if either $\alpha(A)$ or $\alpha(A^c)$ is zero.

We introduce a *neighborhood structure* $(N_x)_{x\in\mathscr{X}}$ on $\mathscr{X}$, by which we mean any collection of sets that satisfy $x \in N_x$ for every $x \in \mathscr{X}$ (note that we do not assume N_x to be a d_0-neighborhood of x in the topological sense). We require this neighborhood structure to be measurable in the sense that

$$N(\mathscr{X}) := \{(x,y) \in \mathscr{X}^2; y \in N_x\} \in \mathscr{B}^2. \tag{2.6}$$

This measurability condition is slightly stronger than the ones required in [2] (see the discussion before Remark 2.1 in [7] for details), but quite a bit more convenient. The N_x play the role of regions of strong dependence: it is advantageous in view of Theorem 2.C below to choose N_x not too large, but in such a way that the point process ξ around the location x depends only weakly on $\xi|_{N_x^c}$. Write $\dot{N}_x$ for $N_x \setminus \{x\}$.

We use the following crucial formula about point process densities, which is proved as Proposition A.A in the appendix. For any non-negative or bounded measurable function $h : \mathscr{X} \times \mathfrak{N} \to \mathbb{R}$, we have

$$\mathbb{E}\biggl(\int_{\mathscr{X}} h(x, \xi|_{N_x^c})\,\xi(dx)\biggr) = \int_{\mathscr{X}} \int_{\mathfrak{N}(N_x^c)} h(x,\varrho) f_{N_x^c \cup \{x\}}(\varrho + \delta_x)\, P_{1,N_x^c}(d\varrho)\, \alpha(dx). \tag{2.7}$$

As an important consequence we obtain by choosing $h(x,\varrho) := 1_A(x)$ that

$$u(x) = \int_{\mathfrak{N}(N_x^c)} f_{N_x^c \cup \{x\}}(\varrho + \delta_x)\, P_{1,N_x^c}(d\varrho) \tag{2.8}$$

for α-almost every $x \in \mathscr{X}$, which implies Equation (2.4) if we set $N_x = \{x\}$.

In many of the more concrete models, the density f of ξ is hereditary and therefore ξ is a Gibbs process, in which case the above expressions can be simplified by introducing conditional densities. Let $\mathfrak{K} := \bigcup_{x\in\mathscr{X}} (\{x\} \times \mathfrak{N}(N_x^c)) \subset \mathscr{X} \times \mathfrak{N}$, which can be seen to be in $\mathscr{B} \otimes \mathscr{N}$ by Condition (2.6) using a monotone class argument. Define a mapping $g : \mathscr{X} \times \mathfrak{N} \to \mathbb{R}_+$ by

$$g(x;\varrho) := \frac{f_{N_x^c \cup \{x\}}(\varrho + \delta_x)}{f_{N_x^c}(\varrho)} \tag{2.9}$$

for $(x,\varrho) \in \mathfrak{K}$ and $g(x;\varrho) := 0$ otherwise, where the fraction in (2.9) is taken to be zero if the denominator (and hence by hereditarity also the numerator) is zero. For $(x,\varrho) \in \mathfrak{K}$ the term $g(x;\varrho)$ can be interpreted as the conditional density of a point at x given the configuration of ξ outside of N_x is ϱ. Equation (2.7) can then be replaced by the following result, which is a generalization of the Nguyen-Zessin Formula (see [19], Equation (3.3)): for any non-negative or bounded measurable function $h : \mathscr{X} \times \mathfrak{N} \to \mathbb{R}$, we have

$$\mathbb{E}\biggl(\int_{\mathscr{X}} h(x, \xi|_{N_x^c})\,\xi(dx)\biggr) = \int_{\mathscr{X}} \mathbb{E}\bigl(h(x,\xi|_{N_x^c}) g(x;\xi|_{N_x^c})\bigr)\,\alpha(dx). \tag{2.10}$$

This formula was already stated as Equation (2.7) in [2] for functions h that are constant in x and as Equation (2.10) in [7] for general functions, both times however under too wide conditions. See Corollary A.B for the proof and Remark A.C for an example that shows that an additional assumption, such as hereditarity, is required. As an analog to (2.8), we obtain for the density u of the expectation measure of ξ that

$$u(x) = \mathbb{E}\bigl(g(x;\xi|_{N_x^c})\bigr)$$

for α-almost every $x \in \mathscr{X}$.

We are now in a position to derive the required distance bounds.

Theorem 2.C (based on Barbour and Brown [2], Theorems 2.4 and 3.6)**.** *Suppose that ξ is a point process which has density f with respect to P_1 and finite expectation measure μ. Let furthermore $(N_x)_{x\in\mathscr{X}}$ be a neighborhood structure that is measurable in the sense of Condition* (2.6)*. Then*

$$\textit{(i)}\ \ d_{TV}\big(\mathscr{L}(\xi),\mathrm{Po}(\mu)\big) \le \int_{\mathscr{X}} \mu(N_x)\,\mu(dx) + \mathbb{E}\Big(\int_{\mathscr{X}} \xi(\dot{N}_x)\,\xi(dx)\Big) + \int_{\mathscr{X}} \breve{\phi}(x)\,\alpha(dx);$$

$$\begin{aligned}\textit{(ii)}\ \ d_2\big(\mathscr{L}(\xi),\mathrm{Po}(\mu)\big) &\le M_2(\mu)\Big[\int_{\mathscr{X}} \mu(N_x)\,\mu(dx) + \mathbb{E}\Big(\int_{\mathscr{X}} \xi(\dot{N}_x)\,\xi(dx)\Big)\Big] \\ &\quad + M_1(\mu)\int_{\mathscr{X}} \breve{\phi}(x)\,\alpha(dx);\end{aligned}$$

where

$$M_1(\mu) = \min\Big(1, \frac{1.647}{\sqrt{|\mu|}}\Big),\ M_2(\mu) = \min\Big(1, \frac{11}{6|\mu|}\Big(1 + 2\log^+\Big(\frac{6|\mu|}{11}\Big)\Big)\Big),$$

and

$$\begin{aligned}\breve{\phi}(x) &= \int_{\mathfrak{N}(N_x^c)} \Big|f_{N_x^c\cup\{x\}}(\varrho+\delta_x) - f_{N_x^c}(\varrho)u(x)\Big|\,P_{1,N_x^c}(d\varrho) \\ &= 2\sup_{C\in\mathscr{N}(N_x^c)}\Big|\int_C \big[f_{N_x^c\cup\{x\}}(\varrho+\delta_x) - f_{N_x^c}(\varrho)u(x)\big]\,P_{1,N_x^c}(d\varrho)\Big|.\end{aligned} \tag{2.11}$$

If f is hereditary, $\breve{\phi}$ can be rewritten as

$$\breve{\phi}(x) = \mathbb{E}\Big|g(x;\xi|_{N_x^c}) - u(x)\Big|. \tag{2.12}$$

Remark 2.D. We refer to the three summands in the upper bound of Theorem 2.C.(i) as basic term, strong dependence term, and weak dependence term, respectively. The basic term depends only on the first order properties of ξ and on $\alpha(N_x)$. The strong dependence term controls what happens within the neighborhoods of strong dependence and is small if $\alpha(N_x)$ is not too big and if there is not too much positive short range dependence within ξ. Finally, the weak dependence term is small if there is only little long range dependence.

Remark 2.E. Theorem 5.27 in [26] (which is based on several earlier results by various authors) gives an alternative upper bound for the d_2-distance above, which when carefully further estimated is in many situations superior to the one in [2], insofar as the logarithmic term in $M_2(\mu)$ can often be disposed of. After applying the same modifications as in the proof of Theorem 2.C below, it can be seen that

$$\begin{aligned}d_2\big(\mathscr{L}(\xi),\mathrm{Po}(\mu)\big) &\le \int_{\mathscr{X}} \mathbb{E}\Big(\Big(\frac{3.5}{|\mu|} + \frac{2.5}{\xi(N_x^c)+1}\Big)\xi(N_x)\Big)\,\mu(dx) \\ &\quad + \mathbb{E}\Big(\int_{\mathscr{X}} \Big(\frac{3.5}{|\mu|} + \frac{2.5}{\xi(N_x^c)+1}\Big)\xi(\dot{N}_x)\,\xi(dx)\Big) + M_1(\mu)\int_{\mathscr{X}} \breve{\phi}(x)\,\alpha(dx).\end{aligned}$$

Since working with this inequality requires a more specialized treatment of the thinnings in our main result, and since the benefit of removing the logarithmic term is negligible for most practical purposes, we do not use this bound in the present article.

Proof of Theorem 2.C. Following the proof of Theorems 2.4 and 3.6 in [2] (applying Equations (2.9) and (2.10) of [2], but not Equation (2.11)), we obtain by using Stein's method that

$$\begin{aligned}\left|\mathbb{E}\tilde{f}(\xi)-\mathbb{E}\tilde{f}(\eta)\right| \leq \Delta_2\tilde{h}\left[\int_{\mathscr{X}}\mathbb{E}\xi(N_x)\,\mu(dx)+\mathbb{E}\left(\int_{\mathscr{X}}\xi(\dot{N}_x)\,\xi(dx)\right)\right]\\ +\left|\mathbb{E}\int_{\mathscr{X}}\left[\tilde{h}(\xi|_{N_x^c}+\delta_x)-\tilde{h}(\xi|_{N_x^c})\right]\left(\xi(dx)-\mu(dx)\right)\right|\end{aligned} \tag{2.13}$$

for every $\tilde{f}\in\mathscr{F}_{TV}=\{1_C;\,C\in\mathscr{N}\}$ [or $\tilde{f}\in\mathscr{F}_2$ in the case of statement (ii)], where $\eta\sim\mathrm{Po}(\mu)$ and $\tilde{h}:=\tilde{h}_{\tilde{f}}$ are the solutions of the so-called Stein equation (see [2], Equation (2.2)), which have maximal first and second differences

$$\Delta_1\tilde{h}:=\sup_{\varrho\in\mathfrak{N},\,x\in\mathscr{X}}\left|\tilde{h}(\varrho+\delta_x)-\tilde{h}(\varrho)\right|$$

and

$$\Delta_2\tilde{h}:=\sup_{\varrho\in\mathfrak{N},\,x,y\in\mathscr{X}}\left|\tilde{h}(\varrho+\delta_x+\delta_y)-\tilde{h}(\varrho+\delta_x)-\tilde{h}(\varrho+\delta_y)+\tilde{h}(\varrho)\right|$$

that are bounded by 1 [or $\Delta_1\tilde{h}\leq M_1(\mu)$ and $\Delta_2\tilde{h}\leq M_2(\mu)$ in the case of statement (ii); see [26], Propositions 5.16 and 5.17].

All that is left to do is bounding the term in the second line of (2.13), which is done as follows. Setting $C_x^+:=\{\varrho\in\mathfrak{N}(N_x^c);\,f_{N_x^c\cup\{x\}}(\varrho+\delta_x)>f_{N_x^c}(\varrho)u(x)\}\in\mathscr{N}(N_x^c)$, we obtain

$$\begin{aligned}&\left|\mathbb{E}\int_{\mathscr{X}}\left[\tilde{h}(\xi|_{N_x^c}+\delta_x)-\tilde{h}(\xi|_{N_x^c})\right]\left(\xi(dx)-\mu(dx)\right)\right|\\ &\quad=\left|\int_{\mathscr{X}}\int_{\mathfrak{N}(N_x^c)}\left[\tilde{h}(\varrho+\delta_x)-\tilde{h}(\varrho)\right]\left(f_{N_x^c\cup\{x\}}(\varrho+\delta_x)-f_{N_x^c}(\varrho)u(x)\right)P_{1,N_x^c}(d\varrho)\,\alpha(dx)\right|\\ &\quad\leq\Delta_1\tilde{h}\int_{\mathscr{X}}\int_{\mathfrak{N}(N_x^c)}\left|f_{N_x^c\cup\{x\}}(\varrho+\delta_x)-f_{N_x^c}(\varrho)u(x)\right|P_{1,N_x^c}(d\varrho)\,\alpha(dx)\\ &\quad=2\,\Delta_1\tilde{h}\int_{\mathscr{X}}\int_{C_x^+}\left(f_{N_x^c\cup\{x\}}(\varrho+\delta_x)-f_{N_x^c}(\varrho)u(x)\right)P_{1,N_x^c}(d\varrho)\,\alpha(dx)\\ &\quad=2\,\Delta_1\tilde{h}\int_{\mathscr{X}}\sup_{C\in\mathscr{N}(N_x^c)}\left|\int_C\left(f_{N_x^c\cup\{x\}}(\varrho+\delta_x)-f_{N_x^c}(\varrho)u(x)\right)P_{1,N_x^c}(d\varrho)\right|\alpha(dx),\end{aligned} \tag{2.14}$$

where we use Equation (2.7) for the second line and

$$\int_{\mathfrak{N}(N_x^c)}\left(f_{N_x^c\cup\{x\}}(\varrho+\delta_x)-f_{N_x^c}(\varrho)u(x)\right)P_{1,N_x^c}(d\varrho)=0$$

for the fourth line, which follows from Equation (2.8). The integrands with respect to $\alpha(dx)$ in the last three lines of (2.14) are all equal to

$$\int_{\mathfrak{N}}\left|f_{N_x^c\cup\{x\}}(\sigma|_{N_x^c}+\delta_x)-f_{N_x^c}(\sigma|_{N_x^c})u(x)\right|P_1(d\sigma)$$

and hence $\mathscr{B}$-measurable by the definition of f_A and the fact that Condition (2.6) implies the measurability of the mapping $\big[\mathscr{X} \times \mathfrak{N} \to \mathscr{X} \times \mathfrak{N}, (x, \sigma) \mapsto (x, \sigma|_{N_x^c})\big]$ (see [7], after Equation (2.4)).

Plugging (2.14) into (2.13) and taking the supremum over $\tilde{f}$ completes the proof for general f. If f is hereditary, we have $f_{N_x^c \cup \{x\}}(\varrho + \delta_x) = g(x; \varrho) f_{N_x^c}(\varrho)$, so that the additional representation of $\breve{\phi}$ claimed in (2.12) follows from the third line of Inequality (2.14). $\square$

3 The distance bounds

We begin this section by presenting the general setting for our main result, Theorem 3.A, partly compiling assumptions that were already mentioned, partly adding more specialized notation and conditions. Thereafter the main result and a number of corollaries are stated, and in the last subsection the corresponding proofs are given.

3.1 Setting for Theorem 3.A

Let ξ be a point process on the compact metric space $(\mathscr{X}, d_0)$ which has density f with respect to the standard Poisson process distribution $P_1 = \mathrm{Po}(\alpha)$ (see Subsection 2.1) and finite expectation measure $\mu = \mu_1$. For any $A \in \mathscr{B}$, write f_A for the density of $\xi|_A$ that is given by (2.5). Furthermore, let $\pi = \big(\pi(\cdot, x);\, x \in \mathscr{X}\big)$ be a $[0,1]$-valued random field. We assume that π when interpreted as a random function on $\mathscr{X}$ takes values in a space $E \subset [0,1]^{\mathscr{X}}$ which is what we call an evaluable path space (i.e. $\Phi : E \times \mathscr{X} \to [0,1], (p, x) \mapsto p(x)$ is measurable) and that there exists a regular conditional distribution of π given the value of ξ. Neither of these assumptions presents a serious obstacle; we refer to Appendix A.3 for details. Let then Λ be the random measure given by $\Lambda(A) := \int_A \pi(x)\, \xi(dx)$ for $A \in \mathscr{B}$, and set $\Lambda^{[2]}(B) := \int_B \pi(x)\pi(\tilde{x})\, \xi^{[2]}\big(d(x, \tilde{x})\big)$ for any $B \in \mathscr{B}^2$.

Choose a neighborhood structure $(N_x)_{x \in \mathscr{X}}$ that is measurable in the sense of Condition (2.6). We assume for every $x \in \mathscr{X}$ that $\pi(x)$ and $\pi|_{N_x^c}$ are both strictly locally dependent on ξ in such a way that the corresponding "regions of dependence" do not interfere with one another. More exactly, this means the following: introduce an arbitrary metric $\tilde{d}_0$ on $\mathscr{X}$ that generates the same topology as d_0, and write $\mathbb{B}(x, r)$ for the closed $\tilde{d}_0$-ball at $x \in \mathscr{X}$ with radius $r \geq 0$ and $\mathbb{B}(A, r) := \{y \in \mathscr{X};\, \tilde{d}_0(y, x) \leq r$ for some $x \in A\}$ for the $\tilde{d}_0$-halo set of distance $r \geq 0$ around $A \subset \mathscr{X}$. Suppose then that we can fix a real number $R \geq 0$ such that

$$\mathbb{B}(x, R) \cap \mathbb{B}(N_x^c, R) = \emptyset \tag{3.1}$$

and

$$\pi(x) \perp\!\!\!\perp_{\xi|_{\mathbb{B}(x,R)}} \xi|_{\mathbb{B}(x,R)^c} \;\text{ and }\; \pi|_{N_x^c} \perp\!\!\!\perp_{\xi|_{\mathbb{B}(N_x^c,R)}} \xi|_{\mathbb{B}(N_x^c,R)^c} \tag{3.2}$$

for every $x \in \mathscr{X}$, where $X \perp\!\!\!\perp_Z Y$ denotes conditional independence of X and Y given Z. If Z is almost surely constant, this is just the (unconditional) independence of X and Y; in particular we require $\pi(x) \perp\!\!\!\perp \xi$ if $\alpha(\mathbb{B}(x, R)) = 0$. Set $A_{\mathrm{int}} := A_{\mathrm{int}}(x) := \mathbb{B}(x, R)$ and $A_{\mathrm{ext}} := A_{\mathrm{ext}}(x) := \mathbb{B}(N_x^c, R)$, where we usually suppress the location x when it is clear from the context.

We introduce two functions to control the dependences in (ξ, π). The function $\breve{\beta} : \mathscr{X} \to \mathbb{R}_+$ is given by

$$\breve{\beta}(x) := \int_{\mathfrak{N}(A_{\mathrm{int}})} \mathbb{E}\big(\pi(x) \,\big|\, \xi|_{A_{\mathrm{int}}} = \varrho_{\mathrm{int}} + \delta_x\big) \cdot \int_{\mathfrak{N}(A_{\mathrm{ext}})} \big|\bar{f}(\varrho_{\mathrm{ext}}, \varrho_{\mathrm{int}} + \delta_x)\big| \, P_{1,A_{\mathrm{ext}}}(d\varrho_{\mathrm{ext}}) \, P_{1,A_{\mathrm{int}}}(d\varrho_{\mathrm{int}}), \tag{3.3}$$

where $\bar{f}(\varrho_{\mathrm{ext}}, \varrho_{\mathrm{int}}+\delta_x) := f_{A_{\mathrm{ext}} \cup A_{\mathrm{int}}}(\varrho_{\mathrm{ext}}+\varrho_{\mathrm{int}}+\delta_x) - f_{A_{\mathrm{ext}}}(\varrho_{\mathrm{ext}}) f_{A_{\mathrm{int}}}(\varrho_{\mathrm{int}}+\delta_x)$, and hence controls the long range dependence within ξ, as well as the short range dependence of π on ξ. If $\alpha(A_{\mathrm{int}}) = 0$, the conditional expectation above is to be interpreted as $\mathbb{E}(\pi(x))$ for every $\varrho_{\mathrm{int}} \in \mathfrak{N}(A_{\mathrm{int}})$. The function $\breve{\gamma} : \mathscr{X} \to \mathbb{R}_+$ is taken to be a measurable function that satisfies

$$\int_{\mathfrak{N}} \operatorname*{ess\,sup}_{C \in \sigma(\pi|_{N_x^c})} \Big|\mathrm{cov}\big(\pi(x), 1_C \,\big|\, \xi = \varrho + \delta_x\big)\Big| f(\varrho + \delta_x) \, P_1(d\varrho) \leq \breve{\gamma}(x), \tag{3.4}$$

and hence controls the average long range dependence within π given ξ. For the definition of the essential supremum of an arbitrary set of measurable functions (above, the functions $\big[\varrho \mapsto \big|\mathrm{cov}(\pi(x), 1_C \,|\, \xi = \varrho + \delta_x)\big|\big]$ for $C \in \sigma(\pi|_{N_x^c})$), see [18], Proposition II.4.1.

Special cases:

In order to better understand the somewhat involved requirements on the neighborhood structure $(N_x)_{x \in \mathscr{X}}$, the "catchment radius" R, and the "dependence controlling functions" $\breve{\beta}$ and $\breve{\gamma}$, consider the following special cases, which will be examined in more detail in subsequent parts of the paper:

1) Independence of ξ and π: In this case it is possible to satisfy Conditions (3.1) and (3.2) with any measurable neighborhood structure $(N_x)_{x \in \mathscr{X}}$ by setting $R = 0$. Doing so, we obtain

$$\breve{\beta}(x) = \mathbb{E}\big(\pi(x)\big) \breve{\phi}(x) \quad \text{and} \quad \sup_{C \in \sigma(\pi|_{N_x^c})} \big|\mathrm{cov}(\pi(x), 1_C)\big| \cdot u(x) \leq \breve{\gamma}(x)$$

as conditions for $\breve{\beta}$ and $\breve{\gamma}$, where $\breve{\phi}$ is given in Equation (2.11). We will encounter this situation in Corollary 3.B and in the application in Subsection 4.1, where π is a constant times the indicator of the complement of a Boolean model.

2) Local functional dependence of π on ξ: Assume that, for some $R \geq 0$, we can write $\pi(x) = p(x, \xi|_{\mathbb{B}(x,R)})$ for a (deterministic) measurable function $p : \mathscr{X} \times \mathfrak{N} \to [0,1]$. Then the neighborhoods N_x have to be chosen large enough to satisfy Condition (3.1) (e.g. choose $N_x \supset \mathbb{B}(x, 2R)$ for every $x \in \mathscr{X}$), whereas Condition (3.2) always holds, by the definition of π. We then can set

$$\breve{\beta}(x) = \int_{\mathfrak{N}} p\big(x, \varrho|_{A_{\mathrm{int}}} + \delta_x\big) \big|\bar{f}\big(\varrho|_{A_{\mathrm{ext}}}, \varrho|_{A_{\mathrm{int}}} + \delta_x\big)\big| \, P_1(d\varrho) \quad \text{and} \quad \breve{\gamma}(x) = 0.$$

We will encounter this situation in the application in Subsection 4.2, where $p(x, \varrho)$ is essentially a constant times the indicator of $\big\{\varrho\big(\mathbb{B}(x,R) \setminus \{x\}\big) = 0\big\}$.

3) Constant deterministic retention field: As a specialization of either of situations 1) and 2), we have the case where $\pi \equiv p \in [0,1]$ is deterministic, so that the thinning is obtained by keeping each point with probability p, independently of its position, and of the positions and the retention/deletion decisions of other points. As seen in the first special case above we can have any measurable neighborhood structure and set $R = 0$. We then have $\breve{\beta}(x) = p\breve{\phi}(x)$ and may set $\breve{\gamma}(x) = 0$. This situation is the subject of Corollary 3.C.

3.2 Results

We are now in the position to formulate our main theorem. The metrics d_{TV} and d_2 have been defined in Subsection 2.3. Note that $(\mathbb{E}\Lambda)^2$ denotes the product measure of the expectation measure $\mathbb{E}\Lambda$ with itself.

Theorem 3.A. *Suppose that the assumptions of Subsection 3.1 hold and write $\mu^{(\pi)} = \mu_1^{(\pi)} = \mathbb{E}\Lambda$ for the expectation measure of ξ_π. We then have*

$$\text{(i)}\ \ d_{TV}\big(\mathscr{L}(\xi_\pi), \mathrm{Po}(\mu^{(\pi)})\big) \le \big(\mathbb{E}\Lambda\big)^2\big(N(\mathscr{X})\big) + \mathbb{E}\Lambda^{[2]}\big(N(\mathscr{X})\big) + \int_{\mathscr{X}} \breve{\beta}(x)\,\alpha(dx) + 2\int_{\mathscr{X}} \breve{\gamma}(x)\,\alpha(dx);$$

$$\text{(ii)}\ \ d_2\big(\mathscr{L}(\xi_\pi), \mathrm{Po}(\mu^{(\pi)})\big) \le M_2\big(\mu^{(\pi)}\big)\Big(\big(\mathbb{E}\Lambda\big)^2\big(N(\mathscr{X})\big) + \mathbb{E}\Lambda^{[2]}\big(N(\mathscr{X})\big)\Big) + M_1\big(\mu^{(\pi)}\big)\bigg(\int_{\mathscr{X}} \breve{\beta}(x)\,\alpha(dx) + 2\int_{\mathscr{X}} \breve{\gamma}(x)\,\alpha(dx)\bigg);$$

where

$$M_1(\mu^{(\pi)}) = \min\bigg(1, \frac{1.647}{\sqrt{\mathbb{E}\Lambda(\mathscr{X})}}\bigg), \text{ and}$$
$$M_2(\mu^{(\pi)}) = \min\bigg(1, \frac{11}{6\,\mathbb{E}\Lambda(\mathscr{X})}\Big(1 + 2\log^+\big(\tfrac{6}{11}\mathbb{E}\Lambda(\mathscr{X})\big)\Big)\bigg).$$

If ξ and π are independent, we obtain an interesting special case, where the upper bound can be expressed in terms of essentially the quantities appearing in Theorem 2.C, which are based solely on ξ, and some rather straightforward quantities based on π.

Corollary 3.B (Independent case)**.** *Suppose that ξ is a point process on $(\mathscr{X}, d_0)$ which has density f with respect to P_1 and finite expectation measure $\mu = \mu_1$. Denote the density of μ with respect to the reference measure α by u (cf. Eq.* (2.4)*). Let $\pi = \big(\pi(\cdot, x);\, x \in \mathscr{X}\big)$ be a $[0,1]$-valued random field that has an evaluable path space $E \subset [0,1]^{\mathscr{X}}$ and is independent of ξ. Choose an arbitrary neighborhood structure $(N_x)_{x\in\mathscr{X}}$ that is measurable in the sense of Condition* (2.6)*, and take $\breve{\gamma} : \mathscr{X} \to \mathbb{R}_+$ to be a measurable function that satisfies*

$$\sup_{C \in \sigma(\pi|_{N_x^c})} \Big|\mathrm{cov}(\pi(x), 1_C)\Big| \cdot u(x) \le \breve{\gamma}(x). \tag{3.5}$$

Note that the expectation measure of ξ_π is $\mu^{(\pi)}(\cdot) = \mathbb{E}\Lambda(\cdot) = \int_\cdot \mathbb{E}\pi(x)\,\mu_1(dx)$, and let $M_1(\mu^{(\pi)})$ and

$M_2(\mu^{(\pi)})$ be defined as in Theorem 3.A. We then obtain

(i) $d_{TV}\big(\mathscr{L}(\xi_\pi), \mathrm{Po}(\mu^{(\pi)})\big)$

$$\leq \int_{\mathscr{X}} \int_{N_x} \mathbb{E}\pi(x)\mathbb{E}\pi(\tilde{x})\, \mu_1(d\tilde{x})\, \mu_1(dx) + \int_{N(\mathscr{X})} \mathbb{E}\big(\pi(x)\pi(\tilde{x})\big)\, \mu_{[2]}\big(d(x,\tilde{x})\big)$$
$$+ \int_{\mathscr{X}} \mathbb{E}\pi(x)\, \breve{\phi}(x)\, \alpha(dx) + 2\int_{\mathscr{X}} \breve{\gamma}(x)\, \alpha(dx);$$

(ii) $d_2\big(\mathscr{L}(\xi_\pi), \mathrm{Po}(\mu^{(\pi)})\big)$

$$\leq M_2\big(\mu^{(\pi)}\big)\bigg(\int_{\mathscr{X}} \int_{N_x} \mathbb{E}\pi(x)\mathbb{E}\pi(\tilde{x})\, \mu_1(d\tilde{x})\, \mu_1(dx) + \int_{N(\mathscr{X})} \mathbb{E}\big(\pi(x)\pi(\tilde{x})\big)\, \mu_{[2]}\big(d(x,\tilde{x})\big)\bigg)$$
$$+ M_1\big(\mu^{(\pi)}\big)\bigg(\int_{\mathscr{X}} \mathbb{E}\pi(x)\, \breve{\phi}(x)\, \alpha(dx) + 2\int_{\mathscr{X}} \breve{\gamma}(x)\, \alpha(dx)\bigg);$$

where $\breve{\phi}$ is given by Equation (2.11) *and, if f is hereditary, by Equation* (2.12).

A further corollary is given for the case of a deterministic retention field, which means that the retention decisions are independent of each other given the point process ξ. We formulate only a very special case, where the point process lives on $\mathbb{R}^D$ and various spatial homogeneities are assumed, which leads to a particularly simple upper bound. The corollary also illustrates how we can deal with the issue of boundary effects in the neighborhood structure by extending the N_x beyond $\mathscr{X}$ in such a way that they are translated versions of each other and that the same holds true for their "complements" $M_x \setminus N_x$.

We always tacitly assume that $\mathbb{R}^D$ is equipped with the Euclidean topology. Write $|A| := Leb^D(A)$ for any Borel set $A \subset \mathbb{R}^D$, and $\mathscr{X} + \mathscr{X}' := \{x + x';\, x \in \mathscr{X}, x' \in \mathscr{X}'\}$ and $\mathscr{X} - x := \{\tilde{x} - x;\, \tilde{x} \in \mathscr{X}\}$ for $\mathscr{X}, \mathscr{X}' \subset \mathbb{R}^D$ and $x \in \mathbb{R}^D$. For the definition of point processes on non-compact spaces and elementary concepts such as stationarity, we refer to the standard point process literature (e.g. [13] or [8]).

Corollary 3.C (Constant case). *Let $\mathscr{X}, \mathscr{Y} \subset \mathbb{R}^D$ be two compact sets, where $\mathscr{X}$ has positive volume and $\mathscr{X} + \bigcup_{x \in \mathscr{X}}(\mathscr{X} - x) \subset \mathscr{Y}$, and consider a metric $d_0 \leq 1$ on $\mathscr{X}$ that generates the Euclidean topology (typically, $d_0(x,y) = \min(|x-y|, 1)$ for all $x, y \in \mathscr{X}$).*

Suppose that ζ is a stationary point process on $\mathbb{R}^D$ whose restriction $\xi := \zeta|_{\mathscr{Y}}$ is a Gibbs process with density f with respect to $P_1 = \mathrm{Po}(Leb^D|_{\mathscr{Y}})$ and finite expectation measure $\mu = \mu_1 = m_1 Leb^D|_{\mathscr{Y}}$. Denote by $\mathscr{K}$ the second reduced moment measure of ζ (see around Equation (4.7) *for details). Let $(N_x)_{x \in \mathscr{X}}$ and $(M_x)_{x \in \mathscr{X}}$ be two neighborhood structures whose sets $N_x := x + N$ and $M_x := x + M$ (not necessarily $\subset \mathscr{X}$ now!) are translated copies of single bounded measurable sets $N, M \subset \mathbb{R}^D$ which are chosen in such a way that $N \subset M$ and $\mathscr{X} \subset M_x \subset \mathscr{Y}$ for every $x \in \mathscr{X}$. Choosing our retention field $\pi \equiv p \in [0,1]$ to be deterministic and constant and noting that $\mu^{(p)}|_{\mathscr{X}} = p\, m_1 Leb^D|_{\mathscr{X}}$, we then have*

(i) $d_{TV}\big(\mathscr{L}(\xi_p|_{\mathscr{X}}), \mathrm{Po}(p\, m_1 Leb^D|_{\mathscr{X}})\big) \leq p^2 m_1^2 |\mathscr{X}|(|N| + \mathscr{K}(N)) + p|\mathscr{X}|\,\mathbb{E}|\Gamma - \mathbb{E}\Gamma|;$

(ii) $d_2\big(\mathscr{L}(\xi_p|_{\mathscr{X}}), \mathrm{Po}(p\, m_1 Leb^D|_{\mathscr{X}})\big)$

$$\leq \min\Big(p\, m_1|\mathscr{X}|, \tfrac{11}{6}\big(1 + 2\log^+\big(\tfrac{6p\, m_1|\mathscr{X}|}{11}\big)\big)\Big)\, p\, m_1(|N| + \mathscr{K}(N))$$
$$+ \min\big(\sqrt{p\, m_1|\mathscr{X}|}, 1.647\big)\sqrt{p\, m_1|\mathscr{X}|}\,\tfrac{1}{m_1}\mathbb{E}|\Gamma - \mathbb{E}\Gamma|,$$

where Γ is an arbitrary random variable that has the same distribution as $g'(x;\xi|_{M_x\setminus N_x}) = f_{(M_x\setminus N_x)\cup\{x\}}(\xi|_{M_x\setminus N_x}+\delta_x)/f_{M_x\setminus N_x}(\xi|_{M_x\setminus N_x})$ (for one and therefore every $x\in\mathscr{X}$).

Remark 3.D. While it seems very appealing to admit $M_x=\mathscr{Y}=\mathbb{R}^D$, this case actually requires a different and somewhat technically involved construction for the conditional density $g'(x;\xi|_{\mathbb{R}^D\setminus N_x})$, because it cannot reasonably be assumed that a density of a point process distribution with respect to the standard Poisson process distribution exists if the state space is $\mathbb{R}^D$ (consider for example a hard core process: the hard core event that no two points are closer than some fixed distance $r>0$ is a $\mathrm{Po}(Leb^D)$-null set). As a matter of fact, the natural setting is that of a Gibbs process on the whole of $\mathbb{R}^D$ defined via a stable potential on the set of finite point measures on $\mathbb{R}^D$, which essentially provides us with "compatible" conditional densities for the point process distribution on bounded Borel sets given the point process outside (see [20], from page 6.9 onwards, for a detailed construction). For a fixed bounded Borel set $\mathscr{Y}\subset\mathbb{R}^D$ we write $f_B(\cdot\,|\,\tau):\mathfrak{N}(B)\to\mathbb{R}_+$ for the conditional density of $\xi|_B$ given $\xi|_{\mathscr{Y}^c}=\tau$. It can then be seen that the crucial inequality

$$\mathbb{E}\Big|\tilde{g}(x;\xi|_{\mathscr{X}\setminus N_x})-m_1\Big|\le\mathbb{E}\Big|g'(x;\xi|_{M_x\setminus N_x})-m_1\Big|$$

(see Inequality (3.14)) and hence the proof of Corollary 3.C can be reproduced under very general conditions if $M_x=\mathbb{R}^D$, where

$$\tilde{g}(x;\tilde{\varrho})=\frac{\int_{\mathfrak{N}(\mathscr{Y}^c)}f_{(\mathscr{X}\setminus N_x)\cup\{x\}}(\tilde{\varrho}+\delta_x\,|\,\tau)\,\mathbb{P}(\xi|_{\mathscr{Y}^c})^{-1}(d\tau)}{\int_{\mathfrak{N}(\mathscr{Y}^c)}f_{\mathscr{X}\setminus N_x}(\tilde{\varrho}\,|\,\tau)\,\mathbb{P}(\xi|_{\mathscr{Y}^c})^{-1}(d\tau)}$$

for every $\tilde{\varrho}\in\mathfrak{N}(\mathscr{X}\setminus N_x)$, and

$$g'(x;\varrho)=\frac{f_{(\mathscr{Y}\setminus N_x)\cup\{x\}}(\varrho|_{\mathscr{Y}\setminus N_x}+\delta_x\,|\,\varrho|_{\mathscr{Y}^c})}{f_{\mathscr{Y}\setminus N_x}(\varrho|_{\mathscr{Y}\setminus N_x}\,|\,\varrho|_{\mathscr{Y}^c})}$$

for every $\varrho\in\mathfrak{N}(\mathbb{R}^D\setminus N_x)$ (as earlier we set such fractions to zero if the denominator is zero). By the construction of the Gibbs process on $\mathbb{R}^D$ (using Equation (6.10) in [20]), the term $g'(x;\varrho)$ does not depend on the choice of $\mathscr{Y}\supset N_x$, except for ϱ in a $\mathbb{P}(\xi|_{\mathbb{R}^D\setminus N_x})^{-1}$-null set. It can be interpreted as the conditional density of a point at x given $\xi|_{\mathbb{R}^D\setminus N_x}=\varrho$.

In the next result, the situation of Theorem 3.A and its corollaries is examined for the case where we compensate for the thinning by contracting the state space as it was done in [22].

Corollary 3.E (Thinning and contraction in $\mathbb{R}^D$). *Suppose that $\mathscr{X}$ is a compact subset of $\mathbb{R}^D$ and that $\alpha=Leb^D|_{\mathscr{X}}$. Let $T\ge1$ and $\kappa_T:\mathbb{R}^D\to\mathbb{R}^D, x\mapsto\frac{1}{T}x$. Assume furthermore that the metric d_0 on $\mathscr{X}$ generates the Euclidean topology and satisfies $d_0\big(\kappa_T(x),\kappa_T(y)\big)\le d_0(x,y)$ for every $x,y\in\mathscr{X}$.*
Then Theorem 3.A and Corollaries 3.B and 3.C remain true if the point processes on the left hand sides of the estimates are replaced by their respective image processes under the contraction κ_T. We thus obtain under the general prerequisites of Theorem 3.A

(i) $$d_{TV}\big(\mathscr{L}(\xi_\pi\kappa_T^{-1}),\mathrm{Po}(\mu^{(\pi)}\kappa_T^{-1})\big)\le\big(\mathbb{E}\Lambda\big)^2\big(N(\mathscr{X})\big)+\mathbb{E}\Lambda^{[2]}\big(N(\mathscr{X})\big)+\int_{\mathscr{X}}\breve{\beta}(x)\,dx+2\int_{\mathscr{X}}\breve{\gamma}(x)\,dx;$$

(ii) $$d_2\big(\mathscr{L}(\xi_\pi\kappa_T^{-1}),\mathrm{Po}(\mu^{(\pi)}\kappa_T^{-1})\big)\le M_2\big(\mu^{(\pi)}\big)\Big(\big(\mathbb{E}\Lambda\big)^2\big(N(\mathscr{X})\big)+\mathbb{E}\Lambda^{[2]}\big(N(\mathscr{X})\big)\Big)+M_1\big(\mu^{(\pi)}\big)\bigg(\int_{\mathscr{X}}\breve{\beta}(x)\,dx+2\int_{\mathscr{X}}\breve{\gamma}(x)\,dx\bigg).$$

Remark 3.F (Comparison with [22])**.** The setting of Corollary 3.E corresponds in large parts to the situation in [22], especially if we set $\mathscr{X} := \kappa_T^{-1}(J)$ for a fixed compact set $J \subset \mathbb{R}^D$ and compare our statement (ii) to Theorem 3.B.

It is more strict in essentially two respects. First, of course, we admit only point processes whose distributions are absolutely continuous with respect to a homogeneous Poisson process. Secondly, we require strict local dependence of π on ξ (see Condition (3.2)), which in [22] was only done for Section 4 (in slightly different form), but which also gives the direct benefit of a conceptually simpler and more intuitive control of the long range dependences.

On the other hand, the setting of Corollary 3.E gives us more freedom than we had in [22] in the sense that the objects live on a general compact subset of $\mathbb{R}^D$, that there are only minimal conditions on the moment measures (as opposed to Assumption 1 in [22]), and that the choice of d_0 and of the neighborhoods of strong dependence N_x is much wider.

Regarding the upper bounds obtained we have clearly improved. The terms in our statement (ii) above have their counterparts in the various terms in Theorem 3.B of [22] (with the integrals over $\breve{\beta}(x)$ and $\breve{\gamma}(x)$ being summerized as a single long range dependence term), but have become simpler and more natural, without any suprema or infima over discretization cuboids and with explicit and manageable constants. The bound as a whole is somewhat better (if we use a heuristic approach for comparing the long range dependence terms) and quite a bit more easily applied, which can be seen from comparing the application in Subsection 4.1 with the same application in Subsection 3.3 of [22].

Remark 3.G. As pointed out in [22], it may be desirable to approximate the distribution of the thinned point process by a Poisson process distribution that has a somewhat different expectation measure. Corresponding distance estimates can easily be obtained from upper bounds for distances between Poisson process distributions. We have $d_{TV}\big(\mathrm{Po}(\lambda), \mathrm{Po}(\mu)\big) \leq \|\lambda - \mu\|$ for $\lambda, \mu \in \mathfrak{M}$ by Remark 2.9 in [2], where $\|\cdot\|$ denotes the total variation norm for signed measures. For d_2 an upper bound is given as Inequality (A.3) in [22] (which is the same as Inequality (2.8) in [6]).

3.3 Proofs

Proof of Theorem 3.A. By Lemma A.D in the appendix a density $f^{(\pi)}$ of ξ_π with respect to P_1 exists, and the finiteness of μ implies the finiteness of $\mu^{(\pi)}$. Hence we can apply Theorem 2.C.

The integrals in the upper bound can be further evaluated as follows. For the first two integrals (basic term and strong dependence term), we have by Lemma 2.B that

$$\int_{\mathscr{X}} \mu^{(\pi)}(N_x)\, \mu^{(\pi)}(dx) = \int_{\mathscr{X}} \big(\mathbb{E}\Lambda\big)(N_x)\, \big(\mathbb{E}\Lambda\big)(dx) = \big(\mathbb{E}\Lambda\big)^2\big(N(\mathscr{X})\big) \tag{3.6}$$

and

$$\mathbb{E}\bigg(\int_{\mathscr{X}} \xi_\pi(\dot{N}_x)\, \xi_\pi(dx)\bigg) = \mathbb{E}\big(\xi_\pi^{[2]}\big(N(\mathscr{X})\big)\big) = \mathbb{E}\big(\Lambda^{[2]}\big(N(\mathscr{X})\big)\big). \tag{3.7}$$

For the third integral (weak dependence term) some more work is necessary. The term that we would like to estimate is

$$2\int_{\mathscr{X}} \sup_{C \in \mathscr{N}(N_x^c)} \bigg| \int_C \big[f^{(\pi)}_{N_x^c \cup \{x\}}(\varrho + \delta_x) - f^{(\pi)}_{N_x^c}(\varrho) u^{(\pi)}(x)\big]\, P_{1,N_x^c}(d\varrho) \bigg|\, \alpha(dx), \tag{3.8}$$

where $u^{(\pi)}$ is the density of $\mu^{(\pi)}$. Equations (A.3) and (A.4) from the appendix imply that, for almost every $x \in \mathscr{X}$ and for $C \in \mathscr{N}(N_x^c)$,

$$\int_C f_{N_x^c}^{(\pi)}(\varrho)\, P_{1,N_x^c}(d\varrho) = \int_{\mathfrak{N}} \mathbb{E}\big(Q_C^{(\pi)}(\sigma|_{N_x^c}) \,\big|\, \xi = \sigma\big) f(\sigma)\, P_1(d\sigma) \tag{3.9}$$

and

$$\int_C f_{N_x^c \cup \{x\}}^{(\pi)}(\varrho + \delta_x)\, P_{1,N_x^c}(d\varrho) = \int_{\mathfrak{N}} \mathbb{E}\big(\pi(x) Q_C^{(\pi)}(\sigma|_{N_x^c}) \,\big|\, \xi = \sigma + \delta_x\big) f(\sigma + \delta_x)\, P_1(d\sigma), \tag{3.10}$$

where $Q_C^{(p)}(\sigma) = \sum_{\varrho \subset \sigma, \varrho \in C} \big(\prod_{s \in \varrho} p(s)\big)\big(\prod_{\tilde{s} \in \sigma \setminus \varrho} (1 - p(\tilde{s}))\big)$ for every $\sigma \in \mathfrak{N}^*$ and every $p \in E$, so that $\big[(p, \sigma) \mapsto Q_C^{(p)}(\sigma)\big]$ is $\mathscr{E} \otimes \mathscr{N}$-measurable. By Equation (2.8) we have furthermore that

$$u^{(\pi)}(x) = \int_{\mathfrak{N}(N_x^c)} f_{N_x^c \cup \{x\}}^{(\pi)}(\varrho + \delta_x)\, P_{1,N_x^c}(d\varrho) = \int_{\mathfrak{N}} \mathbb{E}\big(\pi(x) \,\big|\, \xi = \sigma + \delta_x\big) f(\sigma + \delta_x)\, P_1(d\sigma), \tag{3.11}$$

using that $Q_{\mathfrak{N}(N_x^c)}^{(p)}(\sigma) = 1$ for every $\sigma \in \mathfrak{N}^*(N_x^c)$.

The absolute value term in (3.8) can then be estimated as

$$\begin{aligned}
&\bigg| \int_C \big[f_{N_x^c \cup \{x\}}^{(\pi)}(\varrho + \delta_x) - f_{N_x^c}^{(\pi)}(\varrho) u^{(\pi)}(x)\big]\, P_{1,N_x^c}(d\varrho) \bigg| \\
&= \bigg| \int_{\mathfrak{N}} \mathbb{E}\big(\pi(x) Q_C^{(\pi)}(\sigma|_{N_x^c}) \,\big|\, \xi = \sigma + \delta_x\big) f(\sigma + \delta_x)\, P_1(d\sigma) \\
&\quad - \int_{\mathfrak{N}} \mathbb{E}\big(Q_C^{(\pi)}(\sigma|_{N_x^c}) \,\big|\, \xi = \sigma\big) f(\sigma)\, P_1(d\sigma) \cdot \int_{\mathfrak{N}} \mathbb{E}\big(\pi(x) \,\big|\, \xi = \sigma + \delta_x\big) f(\sigma + \delta_x)\, P_1(d\sigma) \bigg| \\
&\le \bigg| \int_{\mathfrak{N}} \mathbb{E}\big(\pi(x) Q_C^{(\pi)}(\sigma|_{N_x^c}) \,\big|\, \xi = \sigma + \delta_x\big) f(\sigma + \delta_x)\, P_1(d\sigma) \\
&\qquad - \int_{\mathfrak{N}} \mathbb{E}\big(Q_C^{(\pi)}(\sigma|_{N_x^c}) \,\big|\, \xi = \sigma + \delta_x\big)\, \mathbb{E}\big(\pi(x) \,\big|\, \xi = \sigma + \delta_x\big) f(\sigma + \delta_x)\, P_1(d\sigma) \bigg| \\
&\quad + \bigg| \int_{\mathfrak{N}} \mathbb{E}\big(Q_C^{(\pi)}(\sigma|_{N_x^c}) \,\big|\, \xi = \sigma + \delta_x\big)\, \mathbb{E}\big(\pi(x) \,\big|\, \xi = \sigma + \delta_x\big) f(\sigma + \delta_x)\, P_1(d\sigma) \\
&\qquad - \int_{\mathfrak{N}} \mathbb{E}\big(Q_C^{(\pi)}(\sigma|_{N_x^c}) \,\big|\, \xi = \sigma\big) f(\sigma)\, P_1(d\sigma) \cdot \int_{\mathfrak{N}} \mathbb{E}\big(\pi(x) \,\big|\, \xi = \sigma + \delta_x\big) f(\sigma + \delta_x)\, P_1(d\sigma) \bigg| \\
&= \bigg| \int_{\mathfrak{N}} \operatorname{cov}\big(\pi(x), Q_C^{(\pi)}(\sigma|_{N_x^c}) \,\big|\, \xi = \sigma + \delta_x\big) f(\sigma + \delta_x)\, P_1(d\sigma) \bigg| \\
&\quad + \bigg| \int_{\mathfrak{N}(A_{\mathrm{ext}})} \int_{\mathfrak{N}(A_{\mathrm{int}})} \mathbb{E}\big(Q_C^{(\pi)}(\varrho_{\mathrm{ext}}|_{N_x^c}) \,\big|\, \xi|_{A_{\mathrm{ext}}} = \varrho_{\mathrm{ext}}\big)\, \mathbb{E}\big(\pi(x) \,\big|\, \xi|_{A_{\mathrm{int}}} = \varrho_{\mathrm{int}} + \delta_x\big) \\
&\qquad\qquad f_{A_{\mathrm{ext}} \cup A_{\mathrm{int}}}(\varrho_{\mathrm{ext}} + \varrho_{\mathrm{int}} + \delta_x)\, P_{1,A_{\mathrm{int}}}(d\varrho_{\mathrm{int}})\, P_{1,A_{\mathrm{ext}}}(d\varrho_{\mathrm{ext}}) \\
&\qquad - \int_{\mathfrak{N}(A_{\mathrm{ext}})} \mathbb{E}\big(Q_C^{(\pi)}(\varrho_{\mathrm{ext}}|_{N_x^c}) \,\big|\, \xi|_{A_{\mathrm{ext}}} = \varrho_{\mathrm{ext}}\big) f_{A_{\mathrm{ext}}}(\varrho_{\mathrm{ext}})\, P_{1,A_{\mathrm{ext}}}(d\varrho_{\mathrm{ext}}) \\
&\qquad\quad \cdot \int_{\mathfrak{N}(A_{\mathrm{int}})} \mathbb{E}\big(\pi(x) \,\big|\, \xi|_{A_{\mathrm{int}}} = \varrho_{\mathrm{int}} + \delta_x\big) f_{A_{\mathrm{int}}}(\varrho_{\mathrm{int}} + \delta_x)\, P_{1,A_{\mathrm{int}}}(d\varrho_{\mathrm{int}}) \bigg|,
\end{aligned} \tag{3.12}$$

where Condition (3.2) was used for the last equality. Note that $Q_C^{(\pi)}(\sigma|_{N_x^c})$ depends on π only via $\pi|_{N_x^c}$, and does so in a $\mathscr{E}(N_x^c)$-measurable way, where $\mathscr{E}(N_x^c) = \sigma\big([\tilde{p} \mapsto \tilde{p}(x)];\, x \in N_x^c\big)$ is the canonical σ-algebra on $E(N_x^c) := \{p|_{N_x^c};\, p \in E\}$.

The first summand on the right hand side of Inequality (3.12) can then be bounded further as

$$\begin{aligned}
&\left|\int_{\mathfrak{N}} \operatorname{cov}\big(\pi(x), Q_C^{(\pi)}(\sigma|_{N_x^c}) \;\big|\; \xi = \sigma + \delta_x\big) f(\sigma + \delta_x)\, P_1(d\sigma)\right| \\
&\qquad\qquad \le \int_{\mathfrak{N}} \operatorname*{ess\,sup}_{h: E(N_x^c) \to [0,1]} \left|\operatorname{cov}\big(\pi(x), h(\pi|_{N_x^c}) \;\big|\; \xi = \sigma + \delta_x\big)\right| f(\sigma + \delta_x)\, P_1(d\sigma) \\
&\qquad\qquad = \int_{\mathfrak{N}} \operatorname*{ess\,sup}_{h: E(N_x^c) \to \{0,1\}} \left|\operatorname{cov}\big(\pi(x), h(\pi|_{N_x^c}) \;\big|\; \xi = \sigma + \delta_x\big)\right| f(\sigma + \delta_x)\, P_1(d\sigma) \\
&\qquad\qquad \le \breve{\gamma}(x),
\end{aligned}$$

where the essential suprema are taken over all $\mathscr{E}(N_x^c)$-measurable functions with values in $[0,1]$ and $\{0,1\}$, respectively. The third line is obtained by

$$\begin{aligned}
\left|\operatorname{cov}(X,Y)\right| &\le \|Y\|_{L_\infty} \left|\operatorname{cov}\big(X, \operatorname{sign}(\tilde{Y})\big)\right| \\
&\le 2\|Y\|_{L_\infty} \max\Big(\left|\operatorname{cov}\big(X, \mathrm{I}[\tilde{Y} > 0]\big)\right|, \left|\operatorname{cov}\big(X, \mathrm{I}[\tilde{Y} < 0]\big)\right|\Big)
\end{aligned}$$

for all random variables $X \in L_1$ and $Y \in L_\infty$, and for $\tilde{Y} := \mathbb{E}(X \,|\, Y) - \mathbb{E}X$ (see [10], Section 1.2, proof of Lemma 3), where we set $X := \pi(x)$ and $Y := h(\pi|_{N_x^c}) - 1/2$.

For the second summand on the right hand side of Inequality (3.12), we use the notation $F_C(\varrho_{\text{ext}}) := \mathbb{E}\big(Q_C^{(\pi)}(\varrho_{\text{ext}}|_{N_x^c}) \;\big|\; \xi|_{A_{\text{ext}}} = \varrho_{\text{ext}}\big)$ and $G(\varrho_{\text{int}} + \delta_x) := \mathbb{E}\big(\pi(x) \;\big|\; \xi|_{A_{\text{int}}} = \varrho_{\text{int}} + \delta_x\big)$, and bound it as

$$\begin{aligned}
&\left|\int_{\mathfrak{N}(A_{\text{ext}})} \int_{\mathfrak{N}(A_{\text{int}})} F_C(\varrho_{\text{ext}}) G(\varrho_{\text{int}} + \delta_x) \bar{f}(\varrho_{\text{ext}}, \varrho_{\text{int}} + \delta_x)\, P_{1,A_{\text{int}}}(d\varrho_{\text{int}})\, P_{1,A_{\text{ext}}}(d\varrho_{\text{ext}})\right| \\
&\le \int_{\mathfrak{N}(A_{\text{int}})} G(\varrho_{\text{int}} + \delta_x) \\
&\qquad\qquad \cdot \sup_{F: \mathfrak{N}(A_{\text{ext}}) \to [0,1]} \left|\int_{\mathfrak{N}(A_{\text{ext}})} F(\varrho_{\text{ext}}) \bar{f}(\varrho_{\text{ext}}, \varrho_{\text{int}} + \delta_x)\, P_{1,A_{\text{ext}}}(d\varrho_{\text{ext}})\right| P_{1,A_{\text{int}}}(d\varrho_{\text{int}}) \\
&= \frac{1}{2} \int_{\mathfrak{N}(A_{\text{int}})} G(\varrho_{\text{int}} + \delta_x) \int_{\mathfrak{N}(A_{\text{ext}})} \left|\bar{f}(\varrho_{\text{ext}}, \varrho_{\text{int}} + \delta_x)\right| P_{1,A_{\text{ext}}}(d\varrho_{\text{ext}})\, P_{1,A_{\text{int}}}(d\varrho_{\text{int}}),
\end{aligned}$$

where the supremum is taken over $\mathscr{N}(A_{\text{ext}})$-measurable functions. The equality is obtained by setting $F_0(\varrho_{\text{ext}}) := \mathrm{I}\big[\bar{f}(\varrho_{\text{ext}}, \varrho_{\text{int}} + \delta_x) > 0\big]$ and noting that $F_0 : \mathfrak{N}(A_{\text{ext}}) \to [0,1]$ is measurable and maximizes the absolute value term after the supremum.

Thus the total estimate for the weak dependence term is

$$\begin{aligned}
&2 \int_{\mathscr{X}} \sup_{C \in \mathscr{N}(N_x^c)} \left|\int_C \big[f^{(\pi)}_{N_x^c \cup \{x\}}(\varrho + \delta_x) - f^{(\pi)}_{N_x^c}(\varrho) u^{(\pi)}(x)\big]\, P_{1,N_x^c}(d\varrho)\right| \alpha(dx) \\
&\qquad\qquad \le 2 \int_{\mathscr{X}} \breve{\gamma}(x)\, \alpha(dx) + \int_{\mathscr{X}} \breve{\beta}(x)\, \alpha(dx).
\end{aligned} \tag{3.13}$$

Plugging (3.6), (3.7), and (3.13) into Theorem 2.C yields statement (i), and, since $|\mu^{(\pi)}| = \mathbb{E}\Lambda(\mathscr{X})$, also statement (ii). $\square$

Proof of Corollary 3.B. We aim at applying Theorem 3.A for $R = 0$. Clearly Condition (3.1) holds for any neighborhood structure. By the independence of ξ and π we have $\mathscr{L}(\pi)$ as a regular conditional distribution of π given the value of ξ and we see that Condition (3.2) is satisfied, that $\breve{\beta}(x) = \mathbb{E}(\pi(x))\breve{\phi}(x)$ for almost every x, and that Inequality (3.4) simplifies to (3.5) by Equation (2.4). Using the representation $\xi = \sum_{i=1}^{V} \delta_{S_i}$ from the proof of Lemma 2.B, we have furthermore that

$$\begin{aligned}
\mathbb{E}\Lambda(A) &= \mathbb{E}\biggl(\mathbb{E}\biggl(\sum_{i=1}^{V} \pi(S_i)\mathrm{I}[S_i \in A] \biggm| \xi\biggr)\biggr) \\
&= \mathbb{E}\biggl(\sum_{i=1}^{V} \mathbb{E}\bigl(\pi(S_i) \bigm| \xi\bigr)\, \mathrm{I}[S_i \in A]\biggr) \\
&= \mathbb{E}\biggl(\sum_{i=1}^{V} \bigl(\mathbb{E}\pi(x)\bigr)\bigr|_{x=S_i} \mathrm{I}[S_i \in A]\biggr) \\
&= \int_A \mathbb{E}\pi(x)\, \mu_1(dx)
\end{aligned}$$

for every $A \in \mathscr{B}$, and by the analogous computations that

$$\mathbb{E}\Lambda^{[2]}(B) = \int_B \mathbb{E}\bigl(\pi(x)\pi(\tilde{x})\bigr)\, \mu_{[2]}\bigl(d(x, \tilde{x})\bigr)$$

for every $B \in \mathscr{B}^2$. Based on these results we can apply Theorem 3.A and obtain the upper bounds stated. $\square$

Proof of Corollary 3.C. We apply Corollary 3.B for the point process $\tilde{\xi} := \xi|_{\mathscr{X}}$, which has hereditary density $\tilde{f} := f_{\mathscr{X}}$ with respect to $\mathrm{Po}(Leb^D|_{\mathscr{X}})$ and expectation measure $\tilde{\mu} = m_1 Leb^D|_{\mathscr{X}}$, where all of these objects are interpreted as living on $\mathscr{X}$ (as opposed to living on $\mathscr{Y}$ and being trivial on $\mathscr{Y} \setminus \mathscr{X}$). Consider as neighborhood structure $(\tilde{N}_x)_{x\in\mathscr{X}}$ given by $\tilde{N}_x := N_x \cap \mathscr{X}$, write $\tilde{N}_x^c$ for the complement of $\tilde{N}_x$ in $\mathscr{X}$, and set $N(\mathscr{X}) := \{(x, y) \in \mathscr{X} \times \mathscr{Y};\ y \in N_x\}$ and $\tilde{N}(\mathscr{X}) := \{(x, y) \in \mathscr{X}^2;\ y \in \tilde{N}_x\}$, which are measurable by the fact that the N_x are translated copies of a single measurable set. Denoting the conditional density based on $\tilde{f}$ by $\tilde{g}$, we obtain for the $\breve{\phi}(x)$-term

$$\begin{aligned}
&\mathbb{E}\bigl|\tilde{g}(x; \tilde{\xi}|_{\tilde{N}_x^c}) - m_1\bigr| \\
&\quad= \int_{\mathfrak{N}(\tilde{N}_x^c)} \bigl|\tilde{f}_{\tilde{N}_x^c \cup\{x\}}(\varrho + \delta_x) - \tilde{f}_{\tilde{N}_x^c}(\varrho)\, m_1\bigr|\, P_{1,\tilde{N}_x^c}(d\varrho) \\
&\quad= \int_{\mathfrak{N}(\mathscr{X}\setminus N_x)} \bigl|f_{(\mathscr{X}\setminus N_x)\cup\{x\}}(\varrho + \delta_x) - f_{\mathscr{X}\setminus N_x}(\varrho)\, m_1\bigr|\, P_{1,\mathscr{X}\setminus N_x}(d\varrho) \\
&\quad= \int_{\mathfrak{N}(\mathscr{X}\setminus N_x)} \biggl| \int_{\mathfrak{N}(M_x\setminus(\mathscr{X}\cup N_x))} \bigl(f_{(M_x\setminus N_x)\cup\{x\}}(\varrho + \tilde{\varrho} + \delta_x) \\
&\qquad\qquad - f_{(M_x\setminus N_x)}(\varrho + \tilde{\varrho})\, m_1\bigr)\, P_{1,M_x\setminus(\mathscr{X}\cup N_x)}(d\tilde{\varrho}) \biggr|\, P_{1,\mathscr{X}\setminus N_x}(d\varrho) \qquad (3.14) \\
&\quad\le \int_{\mathfrak{N}(M_x\setminus N_x)} \bigl|f_{(M_x\setminus N_x)\cup\{x\}}(\sigma + \delta_x) - f_{M_x\setminus N_x}(\sigma)\, m_1\bigr|\, P_{1,M_x\setminus N_x}(d\sigma) \\
&\quad= \mathbb{E}\bigl|g'(x; \xi|_{M_x\setminus N_x}) - m_1\bigr|,
\end{aligned}$$

and thus by Corollary 3.B that

$$\begin{aligned}
&d_{TV}\big(\mathscr{L}(\xi_p|_{\mathscr{X}}), \mathrm{Po}(p\, m_1 Leb^D|_{\mathscr{X}})\big) \\
&\qquad = d_{TV}\big(\mathscr{L}(\tilde{\xi}_p), \mathrm{Po}(\tilde{\mu}^{(p)})\big) \\
&\qquad \leq p^2 \int_{\mathscr{X}} \mu_1(\tilde{N}_x)\, \mu_1(dx) + p^2 \mu_{[2]}(\tilde{N}(\mathscr{X})) + p \int_{\mathscr{X}} \mathbb{E}\big|\tilde{g}(x;\tilde{\xi}|_{\tilde{N}_x^c}) - m_1\big|\, dx + 0 \\
&\qquad \leq p^2 \int_{\mathscr{X}} \mu_1(N_x)\, \mu_1(dx) + p^2 \mu_{[2]}(N(\mathscr{X})) + p \int_{\mathscr{X}} \mathbb{E}\big|g'(x;\xi|_{M_x \setminus N_x}) - m_1\big|\, dx.
\end{aligned}$$

Statement (i) follows from this by noting that $\mu_{[2]}(N(\mathscr{X})) = m_1^2 |\mathscr{X}| \mathscr{K}(N)$ (see Equation (4.5)) and using the various spatial homogeneities that were required. Statement (ii) is obtained likewise, using additionally that $\mathbb{E}\Lambda(\mathscr{X}) = p\, m_1 |\mathscr{X}|$. □

Proof of Corollary 3.E. From the definition it is clear that the total variation metric is not affected by changes of scale of the state space, so that

$$d_{TV}\big(\mathscr{L}(\xi_\pi \kappa_T^{-1}), Po(\mu^{(\pi)} \kappa_T^{-1})\big) = d_{TV}\big(\mathscr{L}(\xi_\pi), Po(\mu^{(\pi)})\big). \tag{3.15}$$

The definition of d_1 and the inequality required for d_0 imply that $d_1(\varrho_1 \kappa_T^{-1}, \varrho_2 \kappa_T^{-1}) \leq d_1(\varrho_1, \varrho_2)$ for all $\varrho_1, \varrho_2 \in \mathfrak{N}$, whence, by Equation (2.3),

$$d_2\big(\mathscr{L}(\xi_\pi \kappa_T^{-1}), Po(\mu^{(\pi)} \kappa_T^{-1})\big) \leq d_2\big(\mathscr{L}(\xi_\pi), Po(\mu^{(\pi)})\big). \tag{3.16}$$

With Equations (3.15) and (3.16) it is seen that all the results from Theorem 3.A to Corollary 3.C remain correct if we do the proposed replacements; in particular, the upper bounds stated follow immediately from Theorem 3.A. □

4 Applications

We study two applications for a fairly general point process ξ here. The first one concerns the thinning of ξ by covering it with an independent Boolean model. This is up to a few technical adjustments the setting that was used in Section 3.3 of [22]. We present this application in order to illustrate to what degree the results of the current article improve on the main distance bounds in [22], and give new insight into the high intensity limit behavior. The second application deals with a Matérn type I thinning of ξ. We present it as an example where the rather involved $\breve{\beta}$-term is non-zero and can be reasonably simplified. The bound is compared to a result in [26], where the same thinning was considered for the special case that ξ is a Poisson process.

In this whole section we consider a metric $\tilde{d}_0$ on $\mathbb{R}^D$ that is generated by a norm, and use notation of the form $\mathbb{B}(x,r)$ for closed $\tilde{d}_0$-balls in $\mathbb{R}^D$ and $\mathbb{B}^c(x,r)$ for their complements. Write furthermore $\mathbb{B}_{\mathscr{X}}(x,r) := \mathbb{B}(x,r) \cap \mathscr{X}$ for the corresponding balls in $\mathscr{X}$ and $\mathbb{B}_{\mathscr{X}}^c(x,r) := \mathscr{X} \setminus \mathbb{B}_{\mathscr{X}}(x,r)$. We call the subset $\mathscr{X}$ of $\mathbb{R}^D$ *admissible* if it is compact, of positive volume, and has a boundary $\partial\mathscr{X}$ that is of volume zero.

Remark 4.A ($\tilde{d}_0$-balls)**.** It can be shown quite easily that, for any symmetric convex body, i.e. any compact convex set $C \subset \mathbb{R}^D$ that contains the origin as an interior point and is symmetric with respect to it, there is a norm in $\mathbb{R}^D$ whose unit ball $\mathbb{B}(0,1)$ is C. Thus the sets $\mathbb{B}(Y_i, R_i)$ which the Boolean model is composed of in the first application (see Equation (4.1)) and the "competition sets" $\mathbb{B}(s,r)$ in the second application (see Equation (4.9)) are shifted and scaled versions of an abitrary symmetric convex body.

4.1 Thinning by covering with an independent Boolean model

The details for this application are as follows.

$\Diamond$ *Model Setting 1.* Suppose that $\mathscr{X} \subset \mathbb{R}^D$ is admissible and that ξ is a point process on $\mathscr{X}$ which has a density f with respect to $P_1 := \mathrm{Po}(Leb^D|_{\mathscr{X}})$ and finite expectation measure $\mu = \mu_1$ with density u. Let $q \in [0,1]$, and take Ξ to be a stationary Boolean model (see [25], Section 3.1) on $\mathbb{R}^D$ whose grains are $\tilde{d}_0$-balls of random but essentially bounded radius, denoting by $l_1 > 0$ the intensity of the germ process and by $R_i \in L_\infty$ the radii of the grains (which are i.i.d.). This means that Ξ takes the form

$$\Xi = \bigcup_{i=1}^{\infty} \mathbb{B}(Y_i, R_i), \tag{4.1}$$

where Y_i are the points of a $\mathrm{Po}(l_1 Leb^D)$-process that is independent of $(R_i)_{i\in\mathbb{N}}$. Assume furthermore that ξ and Ξ are independent, and define a retention field by $\pi(\omega, x) := q\,\mathrm{I}[x \notin \Xi(\omega)]$ for $\omega \in \Omega$ and $x \in \mathscr{X}$. Thinning with respect to π corresponds to deleting all the points that are covered by Ξ, while retaining uncovered points independently of one another with probability q. $\Diamond$

We aim at applying Corollary 3.B in this setting. Assume without loss of generality that $\mathbb{P}[R_1 > 0] > 0$ (otherwise Proposition 4.B below is easily checked directly), and remove from $\Xi(\omega) \cap \mathscr{X}$ any lower-dimensional parts, stemming either from balls with radius zero or from balls that only just touch $\mathscr{X}$ from the outside, by taking the closure of its interior in $\mathbb{R}^D$. Note that this does not alter the distribution of the obtained thinning, because only a set of volume zero is removed in this way and because ξ and π are independent. As a consequence of Proposition A.E(iv), where $\mathscr{Y} = \mathscr{X}$ and $\Sigma = \mathbb{Q}^D \cap \mathscr{X}$, we obtain then that π has an evaluable path space. Let $N_x := \mathbb{B}_{\mathscr{X}}(x, \bar{r})$ for some $\bar{r} \geq 2\|R_1\|_{L_\infty}$ and every $x \in \mathscr{X}$, which implies independence of $\pi(x)$ and $\pi|_{N_x^c}$ and hence that we can choose $\breve{\gamma} \equiv 0$ in Inequality (3.5). We set furthermore $r := \|R_1\|_{L_D}$, so that $r^D = \mathbb{E}(R_1^D)$. By the fact that the capacity functional T_Ξ of the Boolean model Ξ is given by

$$T_\Xi(C) := \mathbb{P}[\Xi \cap C \neq \emptyset] = 1 - \exp\bigl(-l_1\, \mathbb{E}\bigl(Leb^D(\mathbb{B}(0,R_1) + C)\bigr)\bigr)$$

for any compact set $C \subset \mathbb{R}^D$ (see [25], Equation (3.1.2)), we obtain for the expectations in the upper bound of Corollary 3.B

$$\mathbb{E}\pi(x) = q\bigl(1 - T_\Xi(\{x\})\bigr) = q e^{-l_1 \mathbb{E}|\mathbb{B}(0,R_1)|} \tag{4.2}$$

and

$$\mathbb{E}\bigl(\pi(x)\pi(\tilde{x})\bigr) = q^2\bigl(1 - T_\Xi(\{x,\tilde{x}\})\bigr) = q^2 e^{-l_1 \mathbb{E}|\mathbb{B}(0,R_1) \cup \mathbb{B}(\tilde{x}-x,R_1)|}. \tag{4.3}$$

As earlier, we use absolute value bars for a measurable subset of $\mathbb{R}^D$ to denote its Lebesgue mass. Defining $\alpha_D := |\mathbb{B}(0,1)|$ and $b : \mathbb{R}^D \to [0,1]$ by $b(y) := \mathbb{E}|\mathbb{B}(0,R_1) \setminus \mathbb{B}(y,R_1)| / \mathbb{E}|\mathbb{B}(0,R_1)|$, we then have the following result.

Proposition 4.B. *Under Model Setting 1 laid down above and letting $\bar{r} \geq 2\|R_1\|_{L_\infty}$, $r := \|R_1\|_{L_D}$, and $N_{\bar{r}}(\mathscr{X}) := \{(x,\tilde{x}) \in \mathscr{X}^2;\, \tilde{d}_0(x,\tilde{x}) \leq \bar{r}\}$, we obtain that*

$$d_{TV}\big(\mathscr{L}(\xi_\pi), \mathrm{Po}(\mu^{(\pi)})\big) \leq q^2 e^{-2l_1\alpha_D r^D} \mu_1^2\big(N_{\bar{r}}(\mathscr{X})\big) + q^2 \int_{N_{\bar{r}}(\mathscr{X})} e^{-l_1(1+b(\tilde{x}-x))\alpha_D r^D}\, \mu_{[2]}\big(d(x,\tilde{x})\big) + 2q e^{-l_1\alpha_D r^D} |\mathscr{X}| \breve{\beta}_{\bar{r}}^{(\mathrm{sup})},$$

where $\mu^{(\pi)} = q e^{-l_1\alpha_D r^D}\mu_1$ and

$$\breve{\beta}_{\bar{r}}^{(\mathrm{sup})} := \sup_{x\in\mathscr{X}} \sup_{C\in\mathscr{N}(\mathbb{B}^c_{\mathscr{X}}(x,\bar{r}))} \left| \int_C \big[f_{\mathbb{B}^c_{\mathscr{X}}(x,\bar{r})\cup\{x\}}(\varrho+\delta_x) - f_{\mathbb{B}^c_{\mathscr{X}}(x,\bar{r})}(\varrho)u(x)\big]\, P_{1,\mathbb{B}^c_{\mathscr{X}}(x,\bar{r})}(d\varrho) \right|. \tag{4.4}$$

□

Remark 4.C. Under Assumption 1 made for Proposition 3.G in [22], the above estimate can be bounded by const$\cdot\big(\bar{r}^D|\mathscr{X}|q^2e^{-l_1\alpha_D r^D} + |\mathscr{X}|qe^{-l_1\alpha_D r^D}\breve{\beta}_{\bar{r}}^{(\mathrm{sup})}\big)$, which makes it somewhat better than the one in Proposition 3.G (if we accept $\breve{\beta}_{\bar{r}}^{(\mathrm{sup})}$ as a natural substitute for $\breve{\beta}^{(\mathrm{ind})}(m)$ in [22] and apply Equation (3.15)), also since the result is formulated in the stronger d_{TV}-metric instead of d_2.

However, the main point worth noting here is that the derivation above is considerably simpler and more elegant than the one for Proposition 3.6, because we do not have to worry about covering discretization cuboids. For the same reason we are easily able to work with balls that are based on other metrics than the Euclidean one and can write down the explicit constants in the upper bound.

Remark 4.D (Generalizations of the Boolean model used)**.** By Remark 4.A above, Proposition 4.B allows statements about Boolean models of randomly scaled symmetric convex bodies of any fixed shape. More general Boolean models can be treated quite easily, in principle up to stationary Boolean models $\Xi = \bigcup_{i=1}^\infty (Y_i + \Xi_i)$ with arbitrary random compact sets Ξ_i (always i.i.d.), as long as it is ensured that there is an $\bar{r} \in \mathbb{R}_+$ such that $\Xi_1 \subset \mathbb{B}(0,\bar{r}/2)$ almost surely (otherwise we cannot set $\breve{\gamma} \equiv 0$). We then choose neighborhoods of the form $N_x = x + N$, where N contains the origin, is compact, and satisfies $\mathbb{P}[\{0\}\cap(y+\Xi_1) \neq \emptyset, N^c \cap (y+\Xi_1) \neq \emptyset] = 0$ for every $y \in \mathbb{R}^D$ (e.g. $N = \mathbb{B}(0,\bar{r})$). An upper bound in the spirit of Proposition 4.B follows by applying Equation (3.1.2) of [25] in a similar way as for Equations (4.2) and (4.3) above.

If we assume that ξ is second order stationary (i.e. the restriction to $\mathscr{X}$ of a second order stationary point process ζ on $\mathbb{R}^D$), the rather complex second factorial moment measure can be replaced by a term involving the corresponding reduced moment measure. Second order stationarity means that the second moment measure μ_2 of ζ is locally finite ($\mu_2(B) < \infty$ for every bounded measurable $B \subset \mathbb{R}^D$) and invariant under translations along the diagonal $\{(x,x);\, x\in\mathbb{R}^D\}$ (see [8], Definition 10.4.I), and implies stationarity of the expectation measure, so that $\mu_1 = m_1 Leb^D$ for some $m_1 \in \mathbb{R}_+$. It follows from Lemma 10.4.III in [8] that there is a measure $\mathscr{K}$ on $\mathbb{R}^D$ (unique if $m_1 > 0$) such that

$$\int_{\mathbb{R}^D\times\mathbb{R}^D} h(x,\tilde{x})\, \mu_{[2]}\big(d(x,\tilde{x})\big) = m_1^2 \int_{\mathbb{R}^D}\int_{\mathbb{R}^D} h(x,x+y)\, \mathscr{K}(dy)\, dx \tag{4.5}$$

for every measurable function $h : \mathscr{X}^2 \to \mathbb{R}_+$. Hence

$$\begin{aligned}\int_{N_{\bar{r}}(\mathscr{X})} e^{-l_1(1+b(\tilde{x}-x))\alpha_D r^D}\, \mu_{[2]}\big(d(x,\tilde{x})\big) &= m_1^2 \int_{\mathscr{X}} \int_{\mathbb{B}_{\mathscr{X}-x}(0,\bar{r})} e^{-l_1[1+b(y)]\alpha_D r^D}\, \mathscr{K}(dy)\, dx \\ &\leq m_1^2 |\mathscr{X}| \int_{\mathbb{B}(0,\bar{r})} e^{-l_1[1+b(y)]\alpha_D r^D}\, \mathscr{K}(dy).\end{aligned} \tag{4.6}$$

If ζ is stationary and $m_1 > 0$, it can be seen by Equation (4.10) (see [25], beginning of Section 4.5) that the measure $\mathscr{K}$ is given as

$$\mathscr{K}(B) = \frac{1}{m_1} \mathbb{E}\zeta_0^!(B) \tag{4.7}$$

for every Borel set $B \subset \mathbb{R}^D$, where $\zeta_0^!$ denotes the reduced Palm process of ζ given a point in 0 (see [13], Lemma 10.2 and Section 12.3), so that $\mathbb{E}\zeta_0^!(B)$ can be interpreted as the expected number of points of ζ in B given there is a point in 0. The measure $\mathscr{K}$ is usually referred to as *second reduced moment measure*, although some authors prefer defining it as m_1 (or even m_1^2) times the above measure. Set furthermore $K(\tilde{r}) := \mathscr{K}\big(\mathbb{B}(0,\tilde{r})\big)$ for every $\tilde{r} \in \mathbb{R}_+$, which, if $\tilde{d}_0$ is the Euclidean metric, defines Ripley's K-function.

We examine the situation of Corollary 3.H in [22], waiving two technical conditions that were needed there, but insisting on second-order stationarity in order to bring the second summand in the upper bound in a nicer form.

◊ *Model Setting 1′.* Suppose that $J \subset \mathbb{R}^D$ is admissible, that $n \in \mathbb{N}$, and that $\mathscr{X} = \mathscr{X}_n = \kappa_n^{-1}(J)$, where $\kappa_n(x) = (1/n)x$ for every $x \in \mathbb{R}^D$. Let ξ be a second order stationary point process on $\mathscr{X}$ which has density f with respect to $P_1 = \mathrm{Po}(Leb^D|_{\mathscr{X}})$ and expectation measure $\mu = \mu_1 = m_1 Leb^D|_{\mathscr{X}}$ for some $m_1 \in \mathbb{R}_+$. We assume that ξ is the restriction to $\mathscr{X}$ of one and the same point process ζ on $\mathbb{R}^D$ for every n, and suppress the index n in any quantities that depend on n only by virtue of this restriction. Choose a sequence $(q_n)_{n\in\mathbb{N}}$ with $1/n^D \leq q_n \leq 1$, and a sequence $(\Xi_n)_{n\in\mathbb{N}}$ of stationary Boolean models on $\mathbb{R}^D$ of $\tilde{d}_0$-balls with radii $R_{n,i} \in L_\infty$ (i.i.d. for every n) and germ process intensity $l_1 > 0$ such that

$$r_n := \big\|R_{n,1}\big\|_{L_D} = \Big(\frac{1}{l_1\alpha_D} \log(q_n n^D)\Big)^{1/D}. \tag{4.8}$$

Assume that ξ and Ξ_n are independent for every $n \in \mathbb{N}$, and define retention fields by $\pi_n(\omega,x) := q_n \mathrm{I}[x \notin \Xi_n(\omega)]$ for $\omega \in \Omega$ and $x \in \mathscr{X}$. Let furthermore $\bar{r}_n \geq 2\|R_{n,1}\|_{L_\infty}$, and note that $\mu^{(\pi)}\kappa_n^{-1} = \frac{1}{n^D}\mu_1\kappa_n^{-1} = m_1 Leb^D|_J$. ◊

By (4.6) and since $d_{TV}\big(\mathscr{L}(\xi_{\pi_n}\kappa_n^{-1}), \mathrm{Po}(m_1 Leb^D|_J)\big) = d_{TV}\big(\mathscr{L}(\xi_{\pi_n}), \mathrm{Po}(\frac{m_1}{n^D} Leb^D|_{\mathscr{X}})\big)$ by Equation (3.15), we have the following consequence of Proposition 4.B.

Corollary 4.E. *Under Model Setting 1′ we obtain that*

$$\begin{aligned}&d_{TV}\big(\mathscr{L}(\xi_{\pi_n}\kappa_n^{-1}), \mathrm{Po}(m_1 Leb^D|_J)\big) \\ &\quad\leq m_1^2|J|\alpha_D\Big(\frac{\bar{r}_n}{n}\Big)^D + m_1^2|J|q_n \int_{\mathbb{B}(0,\bar{r}_n)} (q_n n^D)^{-b(y)}\, \mathscr{K}(dy) + 2|J|\breve{\beta}_{\bar{r}_n}^{(\sup)} \\ &\quad\leq |J|\bigg(m_1^2\alpha_D\Big(\frac{\bar{r}_n}{n}\Big)^D + m_1^2 q_n K(\bar{r}_n) + 2\breve{\beta}_{\bar{r}_n}^{(\sup)}\bigg),\end{aligned}$$

where $\breve{\beta}_{\bar{r}}^{(\sup)}$ was defined in Equation (4.4). □

Remark 4.F. Note that Assumption 1b) in [22] implies that $K(\bar{r}_n) = O(\bar{r}_n^D)$ for $n \to \infty$. Compared with the corresponding result in [22] (Corollary 3.H) we therefore have again somewhat better bounds with explicit constants that were obtained in a more direct way.

A rather nice result can be derived from Corollary 4.E in the Poisson case.

Corollary 4.G. *Under Model Setting 1′ and the additional assumptions that ζ is a Poisson process with expectation measure $m_1 Leb^D$, that $q_n = 1$ for every $n \in \mathbb{N}$, that $\tilde{d}_0$ is the Euclidean metric, and that $\|R_{n,1}\|_{L_\infty} = O(r_n)$ for $n \to \infty$, we have*

$$d_{TV}\big(\mathscr{L}(\xi_{\pi_n}\kappa_n^{-1}), \mathrm{Po}(m_1 Leb^D|_J)\big) = O\big((\log n)^{-(D-1)}\big) \quad \text{for } n \to \infty.$$

Remark 4.H. The following (partly heuristical) arguments suggest that the order claimed in Corollary 4.G is sharp for $m_1 > 0$. Assume for simplicity that $J = [0,1]^D$.

For $D = 1$ it is readily understood that $\xi_{\pi_n}\kappa_n^{-1}$ cannot converge in distribution to a Poisson process. The reason is that the uncovered part of $\mathbb{R}$ in the domain of the contraction, i.e. $\mathbb{R} \setminus \Xi_n$, is made up of intervals whose lengths are exponentially distributed with mean $1/l_1$ (no matter how the $R_{n,i}$ are distributed), so that the probability of having two or more points within the first uncovered interval that lies completely in $\mathbb{R}_+$ does not depend on n. Hence we have a constant positive probability that the first two points of the thinned contracted point process in $\mathbb{R}_+$ are within distance $1/(l_1 n)$, say, which cannot be true for a sequence that converges towards a homogeneous Poisson process.

For $D \geq 2$ the situation is more complicated. By Theorem 1 in [11] (compare also statement (ii) on page 244), it can be seen that the uncovered "chinks" of Ξ_n in the domain of the contraction have a volume that is of order $1/\big(\log(n^D)\big)^{D-1}$ for large n, so that the argument of the constant-sized chinks is not valid for general D. Heuristically, the order of the chink volumes together with the fact that the number of chinks in a bounded measurable set is Poisson distributed (see [11], p. 244, statement (i); note the slightly different scaling) suggest that we can think of the process $\xi_{\pi_n}\kappa_n^{-1}$ for n large as a compound Poisson process $\sum_{i=1}^{V_n} Z_i^{(n)} \delta_{S_i^{(n)}}$ with intensity of the Poisson process $\sum_{i=1}^{V_n} \delta_{S_i^{(n)}}$ of order $\big(\log(n^D)\big)^{D-1}$ and i.i.d. clump sizes $Z_i^{(n)}$, for which $\mathbb{P}[Z_1^{(n)} \geq 1]$ is of order $\big(\log(n^D)\big)^{-(D-1)}$ and $\mathbb{P}[Z_1^{(n)} \geq 2]$ is of order $\big(\log(n^D)\big)^{-2(D-1)}$ (by the fact that ξ is a Poisson process). It is easily seen that such a process converges towards a homogeneous Poisson process η as $n \to \infty$ by noting that $d_{TV}\big(\mathscr{L}(\sum_{i=1}^{V_n} Z_i^{(n)} \delta_{S_i^{(n)}}), \mathscr{L}(\sum_{i=1}^{V_n} \mathrm{I}[Z_i^{(n)} \geq 1] \delta_{S_i^{(n)}})\big) \to 0$, and $\sum_{i=1}^{V_n} \mathrm{I}[Z_i^{(n)} \geq 1] \delta_{S_i^{(n)}} \xrightarrow{\mathscr{D}} \eta$, but that its convergence rate in the total variation metric is bounded from below by $\mathbb{P}[\exists i \in \{1, \ldots, V_n\} : Z_i^{(n)} \geq 2]$, which is of order $\big(\log(n^D)\big)^{-(D-1)}$ or, what is the same, order $(\log n)^{-(D-1)}$.

Proof of Corollary 4.G. Our starting point is the first upper bound in Corollary 4.E, where we set $\bar{r}_n := 2\|R_{n,1}\|_{L_\infty}$. Since ζ is Poisson, the third summand is zero, whereas the first summand is clearly $O\big((\log n)^{-(D-1)}\big)$. We investigate the integral in the second summand. We have $\mathscr{K} = Leb^D$ by Equation (4.7) in combination with $\mathscr{L}(\zeta_0^!) = \mathscr{L}(\zeta)$ (see [17], Proposition C.2). Define $\tilde{b} : [0,2] \to [0,1]$ by $\tilde{b}(u) = |\mathbb{B}(0,1) \setminus \mathbb{B}(y,1)|/|\mathbb{B}(0,1)|$, where y is an arbitrary element of $\mathbb{R}^D$ with $|y| = u$, and note that $b(y) \geq \tilde{b}(2|y|/\bar{r}_n)$ for $y \in \mathbb{B}(0, \bar{r}_n)$. Since $\tilde{d}_0$ is the Euclidean metric, it can be shown that there is a constant $\kappa > 0$ such that $\tilde{b}(u) \geq \frac{\kappa}{2} u$ for every $u \in [0,2]$. Writing ω_D for the surface area of the unit sphere in $\mathbb{R}^D$, we then can bound the required integral as

$$\begin{aligned}\int_{\mathbb{B}(0,\bar{r}_n)} (q_n n^D)^{-b(y)}\, \mathscr{K}(dy) &\leq \int_{\mathbb{B}(0,\bar{r}_n)} (n^D)^{-\kappa|y|/\bar{r}_n}\, Leb^D(dy)\\ &= \frac{\omega_D}{\kappa^D}\bar{r}_n^D \int_0^{\kappa} (n^D)^{-r} r^{D-1}\, dr\\ &\leq \frac{(D-1)!\,\omega_D}{\kappa^D}\bar{r}_n^D \bigl(\log(n^D)\bigr)^{-D}(1-n^{-\kappa D}),\end{aligned}$$

where the last inequality follows from integration by parts. Since $\bar{r}_n = O\bigl((\log(n^D))^{1/D}\bigr)$, we thus obtain that also the second summand in the first upper bound of Corollary 4.E is $O\bigl((\log n)^{-(D-1)}\bigr)$. □

4.2 Thinning by Matérn type I competition

Again we base our retention field on a random closed set Ξ, but this time we choose a situation where Ξ is completely determined by the point process ξ. The resulting procedure is the one used for the construction of the Matérn type I hard core process, in which a point is deleted whenever there is any other point within a fixed distance r. The details are as follows.

◊ *Model Setting 2.* Suppose that $r > 0$ and that $\mathscr{X}, \mathscr{X}' \subset \mathbb{R}^D$ are two compact sets, where $\mathscr{X}$ is admissible and $\mathbb{B}(\mathscr{X}, r) \subset \mathscr{X}'$. Let furthermore ξ be a stationary point process on $\mathscr{X}'$ (i.e. the restriction of a stationary point process ζ on $\mathbb{R}^D$) which has density f with respect to $P_1 := \mathrm{Po}(Leb^D|_{\mathscr{X}'})$ and a finite expectation measure $\mu = \mu_1 = m_1 Leb^D|_{\mathscr{X}'}$ for some $m_1 \in \mathbb{R}_+$. By defining ξ on $\mathscr{X}'$, but considering the thinned point process only on $\mathscr{X}$, we avoid boundary effects, which would lead to more complicated notation because of spatial inhomogeneities in the thinned process.

In order to have a Ξ whose realizations are closed sets that are jointly separable, we proceed as follows. Write ξ as $\sum_{i=1}^{V} \delta_{S_i}$, where V and S_i are $\sigma(\xi)$-measurable random elements with values in $\mathbb{Z}_+$ and $\mathscr{X}'$, respectively, and denote by T_i the $\tilde{d}_0$-distance between S_i and its nearest neighbor. Let then

$$B_i(\omega) := \bigl\{y \in \mathscr{X}';\ \tfrac{1}{3}\min(T_i(\omega), r) \leq \tilde{d}_0(y, S_i(\omega)) \leq r\bigr\}$$

for every $\omega \in \Omega$, and set

$$\Xi := \bigcup_{i=1}^{V} B_i.$$

Choose $q \in [0,1]$ and define a retention field on $\mathscr{X}'$ by setting $\pi(\omega, x) := q\mathrm{I}[x \notin \Xi(\omega)]$ if $\omega \in \Omega$ and $x \in \mathscr{X}$ and $\pi(\omega, x) := 0$ if $\omega \in \Omega$ and $x \in \mathscr{X}' \setminus \mathscr{X}$. Note that

$$\pi(\omega, s) = q\mathrm{I}\bigl[s \in \mathscr{X}, \xi(\omega)(\dot{\mathbb{B}}(s, r)) = 0\bigr] \tag{4.9}$$

for $\omega \in \Omega$ and $s \in \xi(\omega)$, where $\dot{\mathbb{B}}(x, r) := \mathbb{B}(x, r) \setminus \{x\}$ for every $x \in \mathscr{X}$. Hence, on $\mathscr{X}$, thinning with respect to π corresponds to deleting all those points that see any other point of the process within distance r (regardless whether this point is itself deleted or not), while retaining points that do not have this property independently of one another with probability q. ◊

This time, we aim at applying Theorem 3.A for the state space $\mathscr{X}'$. By Proposition A.E(iv), π has an evaluable path space (after removing from $\Xi(\omega)\cap\mathscr{X}$ possible lower-dimensional parts by taking the closure of its interior in $\mathbb{R}^D$, which has no influence on the distribution of the resulting thinning). Since π is completely determined by ξ we have the corresponding Dirac measure as a regular conditional distribution of π given the value of ξ. Condition (3.2) is satisfied for a catchment radius of $R = r$, so that $N_x := \mathbb{B}_{\mathscr{X}'}(x, \bar{r})$ for some $\bar{r} \geq 2r$ is a legitimate choice for the neighborhoods of strong dependence. We can furthermore choose $\breve{\gamma} \equiv 0$ in Inequality (3.4).

Write $\xi_x^!$ for the reduced Palm process of ξ given a point in x, and $\xi_{x,\tilde{x}}^!$ for the second-order reduced Palm process of ξ given points in x and $\tilde{x}$ (see [13], Section 12.3, pp. 109 & 110; note that $\nu_n' = \mu_{[n]}$ for obtaining the distributions of the n-th order reduced Palm processes). The first and second order Campbell-Mecke equations state that

$$\mathbb{E}\biggl(\int_{\mathscr{X}'} h(x, \xi - \delta_x)\, \xi(dx)\biggr) = \int_{\mathscr{X}'} \mathbb{E}h(x, \xi_x^!)\, \mu_1(dx) \tag{4.10}$$

and

$$\mathbb{E}\biggl(\int_{\mathscr{X}'^2} h(x, \tilde{x}, \xi - \delta_x - \delta_{\tilde{x}})\, \xi^{[2]}\bigl(d(x,\tilde{x})\bigr)\biggr) = \int_{\mathscr{X}'^2} \mathbb{E}h(x, \tilde{x}, \xi_{x,\tilde{x}}^!)\, \mu_{[2]}\bigl(d(x,\tilde{x})\bigr) \tag{4.11}$$

for non-negative measurable functions h. These equations follow immediately by standard extension arguments from the defining equations of Palm processes (see e.g. [17], Equation (C.4), for the first one). We then obtain by Equation (4.9) that

$$\begin{aligned}\mathbb{E}\Lambda(A) &= \mathbb{E}\biggl(\int_{A\cap\mathscr{X}} q\mathrm{I}\bigl[\xi(\dot{\mathbb{B}}(x,r)) = 0\bigr]\, \xi(dx)\biggr) \\ &= q\int_{A\cap\mathscr{X}} \mathbb{P}\bigl[\xi_x^!\bigl(\mathbb{B}(x,r)\bigr) = 0\bigr]\, \mu_1(dx) = m_1 q\bigl(1 - G(r)\bigr)|A\cap\mathscr{X}|\end{aligned} \tag{4.12}$$

for any $A \in \mathscr{B} = \mathscr{B}(\mathscr{X}')$, where $G : \mathbb{R}_+ \to [0,1]$, $G(\tilde{r}) := \mathbb{P}\bigl[\zeta_0^!\bigl(\mathbb{B}(0,\tilde{r})\bigr) \geq 1\bigr] = \mathbb{P}\bigl[\zeta_x^!\bigl(\mathbb{B}(x,\tilde{r})\bigr) \geq 1\bigr]$ for arbitrary $x \in \mathbb{R}^D$, denotes the nearest neighbor function of ζ, i.e. the distribution function of the distance from a "typical point" to its nearest neighbor, which is a frequently used tool in spatial statistics; see e.g. [1] (Section 3.4), [9], or [17]. In a very similar way, using in addition Equation (4.5) to obtain the last equality, we have with $N_{\bar{r}}(\mathscr{X}') := \{(x,\tilde{x}) \in \mathscr{X}'^2;\ \tilde{d}_0(x,\tilde{x}) \leq \bar{r}\}$ and $N_{\bar{r}}(\mathscr{X}) := N_{\bar{r}}(\mathscr{X}') \cap \mathscr{X}^2$ that

$$\begin{aligned}\mathbb{E}\Lambda^{[2]}\bigl(N_{\bar{r}}(\mathscr{X}')\bigr) &= \mathbb{E}\biggl(\int_{N_{\bar{r}}(\mathscr{X})} q^2\mathrm{I}\bigl[\xi(\dot{\mathbb{B}}(x,r)) = 0\bigr]\mathrm{I}\bigl[\xi(\dot{\mathbb{B}}(\tilde{x},r)) = 0\bigr]\, \xi^{[2]}\bigl(d(x,\tilde{x})\bigr)\biggr) \\ &= q^2\int_{N_{\bar{r}}(\mathscr{X})} \mathrm{I}\bigl[\tilde{d}_0(x,\tilde{x}) > r\bigr]\mathbb{P}\bigl[\xi_{x,\tilde{x}}^!\bigl(\mathbb{B}(x,r)\cup\mathbb{B}(\tilde{x},r)\bigr) = 0\bigr]\, \mu_{[2]}\bigl(d(x,\tilde{x})\bigr) \\ &= m_1^2 q^2 \int_{\mathscr{X}} \int_{(\mathbb{B}(0,\bar{r})\setminus\mathbb{B}(0,r))\cap(\mathscr{X}-x)} \bigl(1 - G_{2,y}(r)\bigr)\, \mathscr{K}(dy)\, dx,\end{aligned} \tag{4.13}$$

where $G_{2,y} : \mathbb{R}_+ \to [0,1]$, $G_{2,y}(\tilde{r}) := \mathbb{P}\bigl[\zeta_{0,y}^!\bigl(\mathbb{B}(0,\tilde{r})\cup\mathbb{B}(y,\tilde{r})\bigr) \geq 1\bigr] = \mathbb{P}\bigl[\zeta_{x,x+y}^!\bigl(\mathbb{B}(x,\tilde{r})\cup\mathbb{B}(x+y,\tilde{r})\bigr) \geq 1\bigr]$ for arbitrary $x \in \mathbb{R}^D$, are the natural two-point analogs of the G-function, with $y \in \mathbb{R}^D$.

Finally, the term $\breve{\beta}(x) = \breve{\beta}_{\bar{r}}(x)$ is zero for $x \in \mathscr{X}' \setminus \mathscr{X}$, and can be evaluated for $x \in \mathscr{X}$ as

$$\begin{aligned}
\breve{\beta}_{\bar{r}}(x) &= \int_{\mathfrak{N}(\mathbb{B}(x,r))} q\mathrm{I}[\varrho_{\mathrm{int}} = 0] \int_{\mathfrak{N}(\mathbb{B}^c_{\mathscr{X}'}(x,\bar{r}-r))} \bigl|\bar{f}(\varrho_{\mathrm{ext}}, \varrho_{\mathrm{int}} + \delta_x)\bigr| \\
&\qquad\qquad P_{1,\mathbb{B}^c_{\mathscr{X}'}(x,\bar{r}-r)}(d\varrho_{\mathrm{ext}})\, P_{1,\mathbb{B}(x,r)}(d\varrho_{\mathrm{int}}) \\
&= qe^{-\alpha_D r^D} \int_{\mathfrak{N}(\mathbb{B}^c_{\mathscr{X}'}(x,\bar{r}-r))} \bigl|\bar{f}(\varrho_{\mathrm{ext}}, \delta_x)\bigr|\, P_{1,\mathbb{B}^c_{\mathscr{X}'}(x,\bar{r}-r)}(d\varrho_{\mathrm{ext}}) \\
&= qe^{-\alpha_D r^D} \mathbb{E}\Bigl|f_{\mathbb{B}(x,r)\,|\,\mathbb{B}^c_{\mathscr{X}'}(x,\bar{r}-r)}(\delta_x; \xi|_{\mathbb{B}^c_{\mathscr{X}'}(x,\bar{r}-r)}) - f_{\mathbb{B}(x,r)}(\delta_x)\Bigr|,
\end{aligned} \tag{4.14}$$

where we set $f_{A_{\mathrm{int}}\,|\,A_{\mathrm{ext}}}(\varrho_{\mathrm{int}}; \varrho_{\mathrm{ext}}) := f_{A_{\mathrm{ext}} \cup A_{\mathrm{int}}}(\varrho_{\mathrm{ext}} + \varrho_{\mathrm{int}}) / f_{A_{\mathrm{ext}}}(\varrho_{\mathrm{ext}})$ if $f_{A_{\mathrm{ext}}}(\varrho_{\mathrm{ext}}) > 0$ and $f_{A_{\mathrm{int}}\,|\,A_{\mathrm{ext}}}(\varrho_{\mathrm{int}}; \varrho_{\mathrm{ext}}) := 0$ otherwise.

The expectation in the last line of (4.14) is much simpler than the general expression we have for $\breve{\beta}$, but is typically still hard to estimate. A more directly applicable estimate, which, however, is very rough and works only with point processes that are not too extreme in a certain sense, is given as follows. Choose $\bar{r} := 2r$, and assume that $f > 0$ P_1-almost surely. We then obtain with $\eta_{\mathrm{ext}} \sim P_{1,\mathbb{B}^c_{\mathscr{X}'}(x,r)}$ that

$$\begin{aligned}
\breve{\beta}(x) &= \int_{\mathfrak{N}(\mathbb{B}(x,r))} q\mathrm{I}[\varrho_{\mathrm{int}} = 0] \int_{\mathfrak{N}(\mathbb{B}^c_{\mathscr{X}'}(x,r))} \bigl|\bar{f}(\varrho_{\mathrm{ext}}, \varrho_{\mathrm{int}} + \delta_x)\bigr| \\
&\qquad\qquad P_{1,\mathbb{B}^c_{\mathscr{X}'}(x,r)}(d\varrho_{\mathrm{ext}})\, P_{1,\mathbb{B}(x,r)}(d\varrho_{\mathrm{int}}) \\
&= q \int_{\mathfrak{N}} \mathrm{I}[\sigma|_{\mathbb{B}(x,r)} = 0] \Bigl| f(\sigma|_{\mathbb{B}^c_{\mathscr{X}'}(x,r)} + \delta_x) - f_{\mathbb{B}^c_{\mathscr{X}'}(x,r)}(\sigma|_{\mathbb{B}^c_{\mathscr{X}'}(x,r)}) f_{\mathbb{B}(x,r)}(\delta_x) \Bigr|\, P_1(d\sigma) \\
&= q \int_{\mathfrak{N}} \mathrm{I}[\sigma(\mathbb{B}(x,r)) = 0] \biggl| 1 - \frac{f_{\mathbb{B}^c_{\mathscr{X}'}(x,r)}(\sigma|_{\mathbb{B}^c_{\mathscr{X}'}(x,r)}) f_{\mathbb{B}(x,r)}(\delta_x)}{f(\sigma|_{\mathbb{B}^c_{\mathscr{X}'}(x,r)} + \delta_x)} \biggr| f(\sigma + \delta_x)\, P_1(d\sigma) \\
&\leq m_1 q \bigl(1 - G(r)\bigr) \biggl\| 1 - \frac{f_{\mathbb{B}^c_{\mathscr{X}'}(x,r)}(\eta_{\mathrm{ext}}) f_{\mathbb{B}(x,r)}(\delta_x)}{f(\eta_{\mathrm{ext}} + \delta_x)} \biggr\|_{L_\infty}
\end{aligned} \tag{4.15}$$

for $x \in \mathscr{X}$, where we used in the last line that

$$m_1 \mathbb{E}\bigl(h(x, \xi^!_x)\bigr) = \int_{\mathfrak{N}} h(x,\sigma) f(\sigma + \delta_x)\, P_1(d\sigma)$$

for almost every $x \in \mathscr{X}'$ and every non-negative measurable function h, which is a consequence of the first Campbell-Mecke equation and of Equation (2.7) with $N_x = \{x\}$. We assume that the $\|\cdot\|_{L_\infty}$-term is bounded by a constant $M \in \mathbb{R}_+$ (that depends neither on r nor on x). Two examples where this is satisfied are given at the end of this subsection.

Plugging (4.12) to (4.15) into Theorem 3.A(i) and choosing now $\bar{r} := 2r$ everywhere yields the following result.

Proposition 4.I. *Under Model Setting 2 laid down above, we obtain that*

$$\begin{aligned}
d_{TV}\bigl(\mathscr{L}(\xi_\pi), \mathrm{Po}(\mu^{(\pi)})\bigr) &\leq m_1^2 |\mathscr{X}| 2^D \alpha_D r^D q^2 \bigl(1 - G(r)\bigr)^2 \\
&\quad + m_1^2 |\mathscr{X}| q^2 \int_{\mathbb{B}(0,2r)\setminus\mathbb{B}(0,r)} \bigl(1 - G_{2,y}(r)\bigr)\, \mathscr{K}(dy) \\
&\quad + qe^{-\alpha_D r^D} \int_{\mathscr{X}} \mathbb{E}\Bigl| f_{\mathbb{B}(x,r)\,|\,\mathbb{B}^c_{\mathscr{X}'}(x,r)}(\delta_x; \xi|_{\mathbb{B}^c_{\mathscr{X}'}(x,r)}) - f_{\mathbb{B}(x,r)}(\delta_x) \Bigr|\, dx,
\end{aligned}$$

where $\mu^{(\pi)} = m_1 q(1-G(r))Leb^D|_{\mathscr{X}}$. If in addition $f > 0$ P_1-almost surely and the $\|\cdot\|_{L_\infty}$-term in Inequality (4.15) *is uniformly bounded by $M \in \mathbb{R}_+$, then the last summand can be estimated further by $m_1|\mathscr{X}|q\big(1-G(r)\big)M$.* □

Remark 4.J. In order to obtain an integrand that does not depend on x in the last summand of the above bound, we can either proceed as in Corollary 3.C, applying Inequality (3.14) and as a consequence replace $\mathbb{B}^c_{\mathscr{X}'}(x,r)$ by $M_x \setminus \mathbb{B}(x,r)$ for bounded "outer neighborhoods" M_x that are shifted copies of one another and that all contain the set $\mathscr{X}'$; or we can proceed as in Remark 3.D, using a Gibbs construction on the whole of $\mathbb{R}^D$ and as a consequence replace $\mathbb{B}^c_{\mathscr{X}'}(x,r)$ by $\mathbb{R}^D \setminus \mathbb{B}(x,r)$.

Remark 4.K (Generalizations of the "competition rule")**.** An equivalent way to describe the Matérn type I competition rule that lends itself to natural generalizations is as follows: assume that each point s of the point process ξ has a "required territory" $C(s) = \mathbb{B}(s, r/2)$ and competition takes place (resulting in the deletion of the competing points) whenever two such territories overlap. By Remark 4.A it is clear that Proposition 4.I allows statements about required territories $C(s) = s + C$ where C is an arbitrary symmetric convex body. General compact sets C can be treated in the analog way. Note in this case that

$$\begin{aligned}\pi(\omega,s) &= q\mathrm{I}\big[s \in \mathscr{X}, (s+C) \cap \textstyle\bigcup_{\tilde{s}\in\xi(\omega)}(\tilde{s}+C) = \emptyset\big]\\ &= q\mathrm{I}\big[s \in \mathscr{X}, \xi(\omega)\big((s+C+C^*)\setminus\{s\}\big) = 0\big],\end{aligned}$$

for $\omega \in \Omega$ and $s \in \xi(\omega)$, where C^* denotes the reflection $\{-x;\, x \in C\}$ of C. We choose the metric $\tilde{d}_0$ and $R \geq 0$ such that $\mathbb{B}(0,R) \supset C + C^*$ and furthermore $N_x \supset \mathbb{B}(x, 2R)$, in order to satisfy Conditions (3.1) and (3.2). Then we can proceed in essentially the same way as for Proposition 4.I, but finish with a result that is slightly more complicated (because of the "gap" between $\mathbb{B}(0,R)$ and $C + C^*$). While it would be very interesting to consider i.i.d. random compact sets instead of the single deterministic set C, it appears to be difficult to obtain useful upper bounds in this case.
A more complex competition rule is used for the Matérn type II thinning and its various generalizations (see for example [14]). In this variant, not all competing points are deleted, but survival is determined according to i.i.d. weights assigned to the points. Upper bounds for Poisson process approximation of thinnings under a traditional Matérn type II competition rule were studied in Section 4.2 of [22], using a somewhat different approach. Under the present approach, obtaining useful bounds for such thinnings based on a global assignment of weights appears to be difficult. What can be treated much more easily are thinnings based on a pairwise assignment of weights (Thinning number 1 in [14]).

If ξ is a homogeneous Poisson process and $q = 1$, then ξ_π is the usual Matérn type I hard core process restricted to $\mathscr{X}$, and the above bound takes especially simple form.

Corollary 4.L. *Under Model Setting 2 and the additional assumptions that ζ is a Poisson process and $q = 1$, we have*

$$\begin{aligned}d_{TV}\big(\mathscr{L}(\xi_\pi), \mathrm{Po}(l_1 Leb^D|_{\mathscr{X}})\big) &\leq |\mathscr{X}|2^D\alpha_D r^D l_1^2 + m_1^2|\mathscr{X}|\int_{\mathbb{B}(0,2r)\setminus\mathbb{B}(0,r)} e^{-m_1|\mathbb{B}(0,r)\cup\mathbb{B}(y,r)|}\,dy\\ &\leq |\mathscr{X}|2^D\alpha_D r^D l_1^2(1 + e^{\frac{1}{2}m_1\alpha_D r^D}),\end{aligned}$$

where $l_1 := m_1 e^{-m_1\alpha_D r^D}$.

Proof. The first inequality follows directly from Propositon 4.I by the fact that ζ is a Poisson process, and hence the last summand is zero and the one- and two-point G-functions can be easily computed by $\mathscr{L}(\zeta_0^!) = \mathscr{L}(\zeta)$ and $\mathscr{L}(\zeta_{0,y}^!) = \mathscr{L}(\zeta)$ (see [17], Proposition C.2, for the one-point case; the two-point case is proved in the analogous way).

The second inequality is a consequence of $|\mathbb{B}(0,r) \cup \mathbb{B}(y,r)| \geq \frac{3}{2}|\mathbb{B}(0,r)|$ for $|y| \geq r$. The latter is due to the fact that all the $\tilde{d}_0$-balls of fixed radius are translated copies of one another that are convex and symmetric with respect to their centers, and can be seen as follows. The symmetry implies that any hyperplane through the origin divides the volume of $\mathbb{B}(0,r)$ in half, while the convexity implies the existence of a supporting hyperplane H_x at every point x of the boundary of $\mathbb{B}(0,r)$, which means that H_x contains x and that $\mathbb{B}(0,r)$ lies completely in one of the closed half-spaces defined by H_x. Thus $H_x - x$ and H_x divide $\mathbb{R}^D$ into three parts, each of which contains half of $\mathbb{B}(0,r)$ or $\mathbb{B}(x,r)$, whence we obtain that $|\mathbb{B}(0,r) \cup \mathbb{B}(x,r)| \geq \frac{3}{2}|\mathbb{B}(0,r)|$ for $|x| = r$. Clearly, $|\mathbb{B}(0,r) \cup \mathbb{B}(y,r)| \geq |\mathbb{B}(0,r) \cup \mathbb{B}(x,r)|$ if $|y| \geq |x|$. □

Remark 4.M. In Theorem 6.6 of [26] a situation very similar to the one in Corollary 4.L was considered in the special case that we choose $\tilde{d}_0$ to be the Euclidean metric. The only substantial difference is that in [26] the Poisson process ξ is defined on $\mathscr{X}$ instead of the superset $\mathscr{X}'$, which leads to less competition near the boundary of $\mathscr{X}$ and consequently to a non-stationary thinned process. However, this difference enters the upper bounds in [26] only insofar as balls are always restricted to $\mathscr{X}$ instead of being balls in $\mathbb{R}^D$.

Disregarding these boundary effects, we see that the estimates in Corollary 4.L are slightly better than the one for the total variation in [26], because our second estimate above is bounded by $2|\mathscr{X}|m_1\alpha_D(2r)^D l_1$, which is exactly the estimate in [26] if we adapt it to our notation.

The main reason for formulating Corollary 4.L, however, was not to improve on this earlier bound, but to demonstrate that Proposition 4.I, which holds for a much greater class of point processes, provides a reasonable estimate in the special case of a Poisson process.

We end this subsection by giving two classes of point processes for which the additional boundedness condition in Proposition 4.I is satisfied.

Example 1. Consider a point process density of the form $f(\varrho) = \lambda^{|\varrho|}\tilde{f}(\varrho)$ for a function $\tilde{f}$ that is bounded and bounded away from zero. One particular instance is the density of the area-interaction process, which is given by

$$f(\varrho) := \kappa\lambda^{|\varrho|}\gamma^{-A_{\tilde{r}}(\varrho)}$$

for every $\varrho \in \mathfrak{N} = \mathfrak{N}(\mathscr{X}')$, where $A_{\tilde{r}}(\varrho) := Leb^D\big(\bigcup_{s\in\varrho}\mathbb{B}_{\mathscr{X}'}(s,\tilde{r})\big)$ and $\lambda, \gamma, \tilde{r} > 0$ are parameters while $\kappa > 0$ is a normalizing constant. The parameter γ governs the type (repulsive or attractive) and the strength of interaction, whereas $\tilde{r}$ and λ control the range of the interaction and the intensity of the point process, respectively. For more details on area-interaction processes see Example 6.5 in [17].

Example 2 (Strauss process). Consider the Strauss process with range $\tilde{r} \in [0,r]$ and interaction strength parameter $\gamma \in (0,1]$ (see [17], Section 6.2.2). This process has a density given by

$$f(\varrho) := \kappa\lambda^{|\varrho|}\gamma^{c_{\tilde{r}}(\varrho)}$$

for every $\varrho \in \mathfrak{N}$, where $c_{\tilde{r}}(\varrho) := \sum_{s,\tilde{s}\in\varrho, s\neq\tilde{s}} \mathrm{I}[\tilde{d}_0(s,\tilde{s}) \leq \tilde{r}]$ counts the pairs of points that lie within distance $\tilde{r}$ of one another, $\lambda > 0$ is an intensity parameter, and $\kappa > 0$ is a normalizing constant.

Then, for $x \in \mathscr{X}$ and $\sigma_{\mathrm{ext}} \in \mathfrak{N}(\mathbb{B}^c_{\mathscr{X}'}(x,r))$,

$$\begin{aligned}
&\frac{f_{\mathbb{B}^c_{\mathscr{X}'}(x,r)}(\sigma_{\mathrm{ext}}) f_{\mathbb{B}(x,r)}(\delta_x)}{f(\sigma_{\mathrm{ext}} + \delta_x)} \\
&\le \frac{1}{\kappa \lambda^{|\sigma_{\mathrm{ext}}|+1} \gamma^{c_{\tilde r}(\sigma_{\mathrm{ext}})}} \int_{\mathfrak{N}(\mathbb{B}(x,r))} \kappa \lambda^{|\sigma_{\mathrm{ext}}|+|\varrho_{\mathrm{int}}|} \gamma^{c_{\tilde r}(\sigma_{\mathrm{ext}})} \, P_{1,\mathbb{B}(x,r)}(d\varrho_{\mathrm{int}}) \\
&\qquad\qquad \cdot \int_{\mathfrak{N}(\mathbb{B}^c_{\mathscr{X}'}(x,r))} \kappa \lambda^{|\varrho_{\mathrm{ext}}|+1} \, P_{1,\mathbb{B}^c_{\mathscr{X}'}(x,r)}(d\varrho_{\mathrm{ext}}) \\
&= \kappa \int_{\mathfrak{N}(\mathbb{B}(x,r))} \int_{\mathfrak{N}(\mathbb{B}^c_{\mathscr{X}'}(x,r))} \lambda^{|\varrho_{\mathrm{int}}|} \lambda^{|\varrho_{\mathrm{ext}}|} \, P_{1,\mathbb{B}^c_{\mathscr{X}'}(x,r)}(d\varrho_{\mathrm{ext}}) \, P_{1,\mathbb{B}(x,r)}(d\varrho_{\mathrm{int}}) \\
&= \kappa \int_{\mathfrak{N}} \lambda^{|\sigma|} \, P_1(d\sigma) = \kappa e^{\lambda - 1},
\end{aligned}$$

where we used for the inequality that $c_{\tilde r}(\sigma + \varrho) \ge c_{\tilde r}(\sigma)$ for all $\varrho, \sigma \in \mathfrak{N}$ and $c_{\tilde r}(\sigma + \delta_x) = c_{\tilde r}(\sigma)$ for $\sigma \in \mathfrak{N}$ and $x \in \mathscr{X}$ with $\sigma(\mathbb{B}_{\mathscr{X}'}(x, \tilde r)) = 0$. Thus, M may be chosen to be $\max(1, \kappa e^{\lambda-1} - 1)$. Although in most cases κ cannot be computed explicitly, it is easy to find reasonable upper bounds.

Appendix

In what follows we formulate and prove some of the more technical results needed in the main part of this article.

A.1 Density formulae used for Theorem 2.C

Proposition A.A. *For a point process ξ on $\mathscr{X}$ with density f with respect to P_1 and finite expectation measure, and for a neighborhood structure $(N_x)_{x \in \mathscr{X}}$ that satisfies Condition* (2.6)*, we have*

$$\mathbb{E}\biggl(\int_{\mathscr{X}} h(x, \xi|_{N_x^c}) \, \xi(dx)\biggr) = \int_{\mathscr{X}} \int_{\mathfrak{N}(N_x^c)} h(x, \varrho) f_{N_x^c \cup \{x\}}(\varrho + \delta_x) \, P_{1,N_x^c}(d\varrho) \, \alpha(dx)$$

for every non-negative or bounded measurable function $h : \mathscr{X} \times \mathfrak{N} \to \mathbb{R}$.

Corollary A.B (Generalized Nguyen-Zessin formula on compact spaces)**.** *Suppose that the conditions of Proposition A.A hold and that f is hereditary. We then have*

$$\mathbb{E}\biggl(\int_{\mathscr{X}} h(x, \xi|_{N_x^c}) \, \xi(dx)\biggr) = \int_{\mathscr{X}} \mathbb{E}\bigl(h(x, \xi|_{N_x^c}) g(x, \xi|_{N_x^c})\bigr) \, \alpha(dx)$$

for every non-negative or bounded measurable function $h : \mathscr{X} \times \mathfrak{N} \to \mathbb{R}$, where g is given in Equation (2.9)*.*

Remark A.C. Note that the statement of Corollary A.B is wrong in the case $N_x = \{x\}$ for all $x \in \mathscr{X}$ if f is not hereditary and $\mathbb{P}[\xi \neq 0] > 0$. As a counterexample consider the process that scatters a fixed number $n \ge 1$ of points uniformly over $\mathscr{X}$ (cf. [8], Example 14.2(a)). This process has a

density given by $f(\varrho) = e^{\alpha(\mathscr{X})}\frac{n!}{\alpha(\mathscr{X})^n}\mathrm{I}[|\varrho| = n]$ and hence satisfies $g(x,\varrho) = 0$ for all $x \in \mathscr{X}$ and $\varrho \in \mathfrak{N}$, which makes the right hand side in Corollary A.B zero for every function h, whereas, with $h(x,\varrho) \equiv 1$, the left hand side is equal to $\mathbb{E}\xi(\mathscr{X}) > 0$.
Since the corollary does not hold generally, its use in the proofs of Theorem 2.4 and 3.6 of [2] and in the proof of Theorem 2.3 of [7] is not justified unless an additional condition (such as hereditarity) is imposed.

Proof of Proposition A.A. We proof the statement for non-negative h; the statement for bounded h follows in the usual way by decomposing h into its positive and negative parts. The Slivnyak-Mecke theorem (see [17], Theorem 3.2, for the special case $\mathscr{X} \subset \mathbb{R}^D$, and [15], Section 3, for the result on a general measurable space $\mathscr{X}$) states that, for $\eta \sim P_1$ and measurable $\tilde{h} : \mathscr{X} \times \mathfrak{N} \to \mathbb{R}_+$,

$$\mathbb{E}\biggl(\int_{\mathscr{X}} \tilde{h}(x,\eta - \delta_x)\, \eta(dx)\biggr) = \int_{\mathscr{X}} \mathbb{E}\tilde{h}(x,\eta)\, \alpha(dx). \tag{A.1}$$

Hence, setting $\tilde{h}(x,\sigma) := h(x,\sigma|_{N_x^c})f(\sigma + \delta_x)$, we obtain

$$\begin{aligned}
\int_{\mathscr{X}} \int_{\mathfrak{N}(N_x^c)} & h(x,\varrho)f_{N_x^c\cup\{x\}}(\varrho + \delta_x)\, P_{1,N_x^c}(d\varrho)\, \alpha(dx) \\
&= \int_{\mathscr{X}} \int_{\mathfrak{N}} h(x,\sigma|_{N_x^c})f(\sigma + \delta_x)\, P_1(d\sigma)\, \alpha(dx) \\
&= \int_{\mathfrak{N}} \int_{\mathscr{X}} h(x,\sigma|_{N_x^c})f(\sigma)\, \sigma(dx)\, P_1(d\sigma) \\
&= \mathbb{E}\biggl(\int_{\mathscr{X}} h(x,\xi|_{N_x^c})\, \xi(dx)\biggr).
\end{aligned}$$

□

Proof of Corollary A.B. The statement follows immediately from Proposition A.A, using that $f_{N_x^c\cup\{x\}}(\varrho + \delta_x) = g(x,\varrho)f_{N_x^c}(\varrho)$ for every $x \in \mathscr{X}$ and every $\varrho \in \mathfrak{N}(N_x^c)$. □

A.2 Density of the thinned process

Let the point process ξ and the random field π be as for the definition of the thinning in Subsection 2.2. We assume additionally, as in Section 3, that all the realizations of π lie in an evaluable path space $E \subset [0,1]^D$ and that there is a regular conditional distribution of π given the value of ξ (see Appendix A.3). It is essential for the construction below that we use the same such distribution throughout (i.e. without changing it in between on $\mathbb{P}\xi^{-1}$-null sets), but insignificant, of course, which one we use.
Set then

$$q(\varrho \,|\, \sigma) := \mathbb{E}\biggl(\prod_{s\in\varrho} \pi(s) \prod_{\tilde{s}\in\sigma\setminus\varrho} \bigl(1 - \pi(\tilde{s})\bigr) \,\bigg|\, \xi = \sigma\biggr)$$

for almost every $\sigma \in \mathfrak{N}$ and for $\varrho \subset \sigma$. It can be easily seen that the mapping $\bigl[\mathfrak{N}^2 \times E \to [0,1], (\varrho,\tilde{\varrho},p) \mapsto \prod_{s\in\varrho} p(s) \prod_{\tilde{s}\in\tilde{\varrho}} \bigl(1-p(\tilde{s})\bigr)\bigr]$ is $\mathscr{N}^2\otimes\mathscr{E}$-measurable, whence we obtain that $q(\varrho \,|\, \sigma)$ is well-defined and that $\varphi : \mathfrak{N}^2 \to [0,1], (\varrho,\tilde{\varrho}) \mapsto q(\varrho \,|\, \varrho + \tilde{\varrho})$ is measurable.

Lemma A.D. *A density of the thinned process ξ_π with respect to P_1 is given by*

$$f^{(\pi)}(\varrho) := e^{\alpha(\mathscr{X})} \int_{\mathfrak{N}} q(\varrho \mid \varrho + \tilde{\varrho}) f(\varrho + \tilde{\varrho}) \, P_1(d\tilde{\varrho})$$

for almost every $\varrho \in \mathfrak{N}$.

Proof. The well-definedness and the measurability of $f^{(\pi)}$ follow from the measurability of φ defined above.

Consider two independent Poisson processes $\eta, \tilde{\eta} \sim P_1$. Take furthermore $\chi \sim P_2$ and let $\chi_{1/2}$ be a thinning of χ with retention function $p \equiv 1/2$, which corresponds to picking a subset of the points of χ uniformly at random. Note that $(\eta, \tilde{\eta}) \stackrel{\mathscr{D}}{=} (\chi_{1/2}, \chi \setminus \chi_{1/2})$ (see e.g. [17], Proposition 3.7, for $\mathscr{X} \subset \mathbb{R}^D$; the proof can easily be adapted for general compact metric spaces).

Integration of the proposed density over an arbitrary set $C \in \mathscr{N}$, using Lemma 2.A for the fifth line, yields

$$\begin{aligned}
\int_C f^{(\pi)}(\varrho) \, P_1(d\varrho) &= e^{\alpha(\mathscr{X})} \int_C \int_{\mathfrak{N}} q(\varrho \mid \varrho + \tilde{\varrho}) f(\varrho + \tilde{\varrho}) \, P_1(d\tilde{\varrho}) \, P_1(d\varrho) \\
&= e^{\alpha(\mathscr{X})} \mathbb{E}\Big(\mathrm{I}[\eta \in C] q(\eta \mid \eta + \tilde{\eta}) f(\eta + \tilde{\eta}) \Big) \\
&= e^{\alpha(\mathscr{X})} \mathbb{E}\Big(\mathrm{I}[\chi_{1/2} \in C] q(\chi_{1/2} \mid \chi) f(\chi) \Big) \\
&= e^{\alpha(\mathscr{X})} \int_{\mathfrak{N}} \frac{1}{2^{|\sigma|}} \sum_{\varrho \subset \sigma} \mathrm{I}[\varrho \in C] q(\varrho \mid \sigma) f(\sigma) \, P_2(d\sigma) \\
&= \int_{\mathfrak{N}} \sum_{\varrho \subset \sigma} \mathrm{I}[\varrho \in C] q(\varrho \mid \sigma) f(\sigma) \, P_1(d\sigma) \\
&= \int_{\mathfrak{N}} \mathbb{E}\big(Q_C^{(\pi)}(\sigma) \bigm| \xi = \sigma \big) f(\sigma) \, P_1(d\sigma) \\
&= \int_{\mathfrak{N}} \mathbb{P}[\xi_\pi \in C \bigm| \xi = \sigma] f(\sigma) \, P_1(d\sigma) \\
&= \mathbb{P}[\xi_\pi \in C],
\end{aligned} \tag{A.2}$$

where $Q_C^{(p)}(\sigma) := \sum_{\varrho \subset \sigma, \varrho \in C} \big(\prod_{s \in \varrho} p(s) \big) \big(\prod_{\tilde{s} \in \sigma \setminus \varrho} (1 - p(\tilde{s})) \big)$ for every $\sigma \in \mathfrak{N}^*$ and every $p \in E$, so that $\big[(p, \sigma) \mapsto Q_C^{(p)}(\sigma)\big]$ is $\mathscr{E} \otimes \mathscr{N}$-measurable and $Q_C^{(\pi)}(\xi) = \mathbb{P}[\xi_\pi \in C \mid \xi, \pi]$. From Equation (A.2) the claim follows. □

The proof above yields that

$$\int_C f^{(\pi)}(\varrho) \, P_1(d\varrho) = \int_{\mathfrak{N}} \mathbb{E}\big(Q_C^{(\pi)}(\sigma) \mid \xi = \sigma \big) f(\sigma) \, P_1(d\sigma)$$

for every $C \in \mathcal{N}$, and hence, by Equation (2.5), that more generally, with $A \in \mathcal{B}$,

$$\begin{aligned}
\int_C f_A^{(\pi)}(\varrho)\, P_{1,A}(d\varrho) &= \int_C \int_{\mathfrak{N}(A^c)} f^{(\pi)}(\varrho + \tilde{\varrho})\, P_{1,A^c}(d\tilde{\varrho})\, P_{1,A}(d\varrho) \\
&= \int_{\widetilde{C}} f^{(\pi)}(\sigma)\, P_1(d\sigma) \\
&= \int_{\mathfrak{N}} \mathbb{E}\big(Q_{\widetilde{C}}^{(\pi)}(\sigma) \,\big|\, \xi = \sigma\big) f(\sigma)\, P_1(d\sigma) \\
&= \int_{\mathfrak{N}} \mathbb{E}\big(Q_C^{(\pi)}(\sigma|_A) \,\big|\, \xi = \sigma\big) f(\sigma)\, P_1(d\sigma)
\end{aligned} \tag{A.3}$$

for every $C \in \mathcal{N}(A)$, where $\widetilde{C} := \{\tilde{\sigma} \in \mathfrak{N};\, \tilde{\sigma}|_A \in C\}$. For the last equality we used that $\sum_{\varrho \subset \sigma|_{A^c}} \big(\prod_{s \in \varrho} p(s)\big)\big(\prod_{\tilde{s} \in (\sigma|_{A^c}) \setminus \varrho}(1 - p(\tilde{s}))\big) = 1$.

The various computations in (A.2) remain correct (after the obvious minor modifications) if we add an extra point to ϱ, yielding in the unrestricted case

$$\begin{aligned}
\int_C f^{(\pi)}(\varrho + \delta_x)\, P_1(d\varrho) &= \int_{\mathfrak{N}} \sum_{\varrho \subset \sigma} \mathrm{I}[\varrho \in C]\, q(\varrho + \delta_x \,|\, \sigma + \delta_x) f(\sigma + \delta_x)\, P_1(d\sigma) \\
&= \int_{\mathfrak{N}} \mathbb{E}\big(\pi(x) Q_C^{(\pi)}(\sigma) \,\big|\, \xi = \sigma + \delta_x\big) f(\sigma + \delta_x)\, P_1(d\sigma)
\end{aligned}$$

for every $C \in \mathcal{N}$, which holds for α-almost every $x \in \mathcal{X}$. Hence we obtain in a very similar fashion as in Equation (A.3) that

$$\int_C f_{A \cup \{x\}}^{(\pi)}(\varrho + \delta_x)\, P_{1,A}(d\varrho) = \int_{\mathfrak{N}} \mathbb{E}\big(\pi(x) Q_C^{(\pi)}(\sigma|_A) \,\big|\, \xi = \sigma + \delta_x\big) f(\sigma + \delta_x)\, P_1(d\sigma) \tag{A.4}$$

for every $A \in \mathcal{B}$ and every $C \in \mathcal{N}(A)$, which holds for α-almost every $x \in \mathcal{X}$.

A.3 Technical conditions on ξ and π: evaluable path space and regular conditional distribution

Consider a locally compact, second countable Hausdorff space $\mathcal{Y}$ that is equipped with its Borel σ-algebra $\mathcal{B} = \mathcal{B}(\mathcal{Y})$. This is the most common type of space on which general point processes are defined. Any such space is separable, and a complete metric $\tilde{d}$ can be introduced that generates its topology. With regard to the main part of this article, $\mathcal{Y}$ is usually just our compact state space $\mathcal{X}$, but it is sometimes useful to consider a natural superset of $\mathcal{X}$ (e.g. $\mathbb{R}^D$ if $\mathcal{X} \subset \mathbb{R}^D$). For sets of functions $\mathcal{Y} \to [0,1]$, we introduce the concept of (locally) evaluable path spaces.

Definition. Let $E \subset [0,1]^{\mathcal{Y}}$ and let $\mathcal{E}$ be the canonical σ-algebra on E, which is generated by the evaluation mappings $\Psi_x : E \to [0,1]$, $p \mapsto p(x)$, where $x \in \mathcal{Y}$. For $U \in \mathcal{B}$ set furthermore $E(U) := \{p|_U;\, p \in E\}$ and write $\mathcal{E}(U)$ for the corresponding σ-algebra generated by $\Psi_{U,x} : E(U) \to [0,1]$, $\tilde{p} \mapsto \tilde{p}(x)$, where $x \in U$.

(i) Call E an *evaluable path space* if the mapping $\Phi : E \times \mathcal{Y} \to [0,1], (p,x) \mapsto p(x)$ is $\mathcal{E} \otimes \mathcal{B}$-measurable.

(ii) Call E a *locally evaluable path space* if the mapping $\Phi_U : E(U) \times U \to [0,1], (p,x) \mapsto p(x)$ is $\mathscr{E}(U) \otimes \mathscr{B}_U$-measurable for every $U \subset \mathscr{Y}$ that is open and relatively compact.

It can be easily seen that every locally evaluable path space is also an evaluable path space.

For the main results of this article we assume that π takes values in an evaluable path space and that there exists a regular conditional distribution of π given the value of ξ. Neither of these assumptions presents a serious restriction, because they are both naturally satisfied in many practical applications, and if they are not, we can modify π accordingly (provided it is measurable in the sense of Subsection 2.2) without changing the distribution of the resulting thinning. To see this let R be a third of the minimal interpoint distance in ξ, which is positive except on a null set, and let $\tilde{\pi}(\omega, x) := \pi(\omega, S(\omega))$ if there is a point $S(\omega)$ of $\xi(\omega)$ within distance $R(\omega)$ of x and $\tilde{\pi}(\omega, x) := 0$ otherwise. We have as path space E for π the space of all functions $p : \mathscr{X} \to [0,1]$ that are zero except on finitely many non-overlapping closed balls of positive radius, on each of which they are constant. By Proposition A.E(iii) below it can be seen, using the separability of $\mathscr{Y}$, that this is an evaluable path space. A regular conditional distribution of $\tilde{\pi}$ given the value of ξ can then be defined in a very straightforward manner, using the regular conditional distribution of $(\pi(s))_{s\in\sigma}$ given $\xi = \sigma$.

Since the above construction looks rather artificial in many situations, we provide a few manageable conditions under each of which a path space is (locally) evaluable, and hereby substantiate the statement that an evaluable path space is naturally present in many practical applications. The proposition below is essentially the "path space version" of Proposition A.D in [22]. Where it was conveniently possible, we have generalized the conditions from $\mathbb{R}_+^D$ to the space $\mathscr{Y}$.

Definition. We call a set $E \subset [0,1]^{\mathscr{Y}}$ *separable from above [or below]* if there exists a countable set $\Sigma \subset \mathscr{Y}$ such that for every $p \in E$, every open $\tilde{d}$-ball $B \subset \mathscr{Y}$ and every $y \in \mathbb{R}$ we have that $p(x) > y$ for all $x \in B \cap \Sigma$ implies $p(x) > y$ for all $x \in B$ [or $p(x) < y$ for all $x \in B \cap \Sigma$ implies $p(x) < y$ for all $x \in B$, respectively].

Proposition A.E. *A set $E \subset [0,1]^{\mathscr{Y}}$ is a locally evaluable path space if it satisfies any one of the following conditions.*

(i) $\mathscr{Y} = \mathbb{R}^D$ and there is a closed convex cone $A \subset \mathbb{R}^D$ of positive volume such that every $p \in E$ is continuous from A (see the definition in [22], Appendix A.3);

(ii) Every $p \in E$ is lower semicontinuous, and E is separable from above;

(iii) Every $p \in E$ is upper semicontinuous, and E is separable from below;

(iv) Every $p \in E$ is the indicator of a closed subset of $\mathscr{Y}$, and the family $\mathscr{C}$ of these closed subsets is separable in the sense that the subsets are jointly separable, i.e. there is a countable set $\Sigma \subset \mathscr{Y}$ such that $C = \overline{C \cap \Sigma}$ for every $C \in \mathscr{C}$, where the bar denotes topological closure.

Proof. (i) This follows with some minor adaptations from Proposition A.D in [22] by letting $\Omega := E$ and defining π to be the identity mapping on Ω, so that $\pi(\omega, x) = \omega(x)$ for every $\omega \in \Omega$ and every $x \in \mathbb{R}^D$. Proposition A.D(i) in [22] yields then that $\Phi_R : E(R) \times R \to [0,1], (p,x) \mapsto p(x)$ is $\mathscr{E}(R) \otimes \mathscr{B}_R$-measurable for every bounded open rectangle $R \subset \mathbb{R}^D$ (it is easy to see that the result from [22] carries over from $\mathbb{R}_+^D$ to $\mathbb{R}^D$). The same result for a general bounded open set U instead of R follows by writing U as a countable union of open rectangles.

(ii) We can essentially reproduce the prove of Proposition A.D(ii) in [22]. Let $y \in [0,1)$ and $U \subset \mathscr{Y}$ be open and relatively compact. We show that $\Phi_U^{-1}\big((y,\infty)\big) \in \mathscr{E}(U) \otimes \mathscr{B}_U$. Choose a separant Σ as in the definition of the separability from above and set $\mathscr{G} := \{\mathbb{B}^{\circ}(x,1/n) \subset U; x \in \Sigma', n \in \mathbb{N}\}$, where $\mathbb{B}^{\circ}(x,1/n)$ denotes the open $\tilde{d}$-ball with center x and radius $1/n$, and Σ' is an arbitrary countable dense subset of $\mathscr{Y}$. Noting that $p^{-1}\big((y,\infty)\big) \cap U$ is open by the lower semicontinuity of $p \in E$, we have

$$\Phi_U^{-1}\big((y,\infty)\big) = \bigcup_{B \in \mathscr{G}} \left(\bigcap_{x \in B \cap \Sigma} \Psi_{U,x}^{-1}\big((y,\infty)\big) \right) \times B \in \mathscr{E}(U) \otimes \mathscr{B}_U.$$

(iii) Apply (ii) to $E' := \{1 - p\,;\, p \in E\}$.

(iv) We apply (iii). It is evident that every $p \in E$ is upper semicontinuous. Separability of E from below is inferred from the separability of $\mathscr{C}$ as follows. First note that the definition has to be checked only for $y = 1$, because $p \in E$ takes only the values 0 and 1. Let Σ be a separant for $\mathscr{C}$ as in statement (iv), and let B be an open ball in $\mathscr{Y}$. Then, for $C \in \mathscr{C}$ and $p = 1_C$, $p(x) < 1$ for all $x \in B \cap \Sigma$ implies that $B \cap \Sigma \cap C = \emptyset$, hence $B \subset (C \cap \Sigma)^c$. Since B is open, this implies $B \subset \mathrm{int}\big((C \cap \Sigma)^c\big) = \big(\overline{C \cap \Sigma}\big)^c = C^c$, where $\mathrm{int}(A)$ denotes the interior of the set A for any $A \subset \mathscr{Y}$. Thus $p(x) < 1$ for all $x \in B$. $\square$

Acknowledgement

I would like to thank Adrian Baddeley for helpful discussions and an anonymous referee for many pertinent comments, both of which have led to a considerable improvement of this article. My further thanks go to the School of Mathematics and Statistics at the University of Western Australia for the continued hospitality I have received over the two years of my stay.

References

[1] Baddeley, A. (2007). Spatial point processes and their applications, *Stochastic geometry*, Vol. 1892 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, pp. 1–75.

[2] Barbour, A. D. and Brown, T. C. (1992). Stein's method and point process approximation, *Stochastic Process. Appl.* **43**(1): 9–31.

[3] Barbour, A. D. and Chen, L. H. Y. (eds) (2005). *An introduction to Stein's method*, Vol. 4 of *Lecture Notes Series. Institute for Mathematical Sciences. National University of Singapore*, Singapore University Press, Singapore.

[4] Barbour, A. D., Holst, L. and Janson, S. (1992). *Poisson approximation*, Vol. 2 of *Oxford Studies in Probability*, Oxford University Press, Oxford.

[5] Brown, T. (1979). Position dependent and stochastic thinning of point processes, *Stochastic Process. Appl.* **9**(2): 189–193.

[6] Brown, T. C. and Xia, A. (1995). On metrics in point process approximation, *Stochastics Stochastics Rep.* **52**(3-4): 247–263.

[7] Chen, L. H. Y. and Xia, A. (2004). Stein's method, Palm theory and Poisson process approximation, *Ann. Probab.* **32**(3B): 2545–2569.

[8] Daley, D. J. and Vere-Jones, D. (1988). *An introduction to the theory of point processes*, Springer Series in Statistics, Springer-Verlag, New York.

[9] Diggle, P. J. (2003). *Statistical analysis of spatial point patterns*, second edn, Arnold, London.

[10] Doukhan, P. (1994). *Mixing*, Vol. 85 of *Lecture Notes in Statistics*, Springer-Verlag, New York.

[11] Hall, P. (1985). Distribution of size, structure and number of vacant regions in a high-intensity mosaic, *Z. Wahrsch. Verw. Gebiete* **70**(2): 237–261.

[12] Kallenberg, O. (1975). Limits of compound and thinned point processes, *J. Appl. Probability* **12**: 269–278.

[13] Kallenberg, O. (1986). *Random measures*, fourth edn, Akademie-Verlag, Berlin.

[14] Månsson, M. and Rudemo, M. (2002). Random patterns of nonoverlapping convex grains, *Adv. in Appl. Probab.* **34**: 718–738.

[15] Mecke, J. (1967). Stationäre zufällige Maße auf lokalkompakten Abelschen Gruppen, *Z. Wahrsch. Verw. Gebiete* **9**: 36–58.

[16] Mecke, J. (1968). Eine charakteristische Eigenschaft der doppelt stochastischen Poissonschen Prozesse, *Z. Wahrsch. Verw. Gebiete* **11**: 74–81.

[17] Møller, J. and Waagepetersen, R. P. (2004). *Statistical inference and simulation for spatial point processes*, Vol. 100 of *Monographs on Statistics and Applied Probability*, Chapman & Hall/CRC, Boca Raton, FL.

[18] Neveu, J. (1965). *Mathematical foundations of the calculus of probability*, Translated by Amiel Feinstein, Holden-Day Inc., San Francisco, Calif.

[19] Nguyen, X.-X. and Zessin, H. (1979). Integral and differential characterizations of the Gibbs process, *Math. Nachr.* **88**: 105–115.

[20] Preston, C. (1976). *Random fields*, Vol. 534 of *Lecture Notes in Mathematics*, Springer-Verlag, Berlin.

[21] Reiss, R.-D. (1993). *A course on point processes*, Springer Series in Statistics, Springer-Verlag, New York.

[22] Schuhmacher, D. (2005a). Distance estimates for Poisson process approximations of dependent thinnings, *Electron. J. Probab.* **10**: no. 5, 165–201 (electronic).

[23] Schuhmacher, D. (2005b). *Estimation of distances between point process distributions*, PhD thesis, University of Zurich.
http://www.dissertationen.unizh.ch/2006/schuhmacher/diss.pdf.

[24] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, Univ. California Press, Berkeley, Calif., pp. 583–602.

[25] Stoyan, D., Kendall, W. S. and Mecke, J. (1987). *Stochastic geometry and its applications*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons Ltd., Chichester.

[26] Xia, A. (2005). Stein's method and Poisson process approximation, *An introduction to Stein's method*, Vol. 4 of *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.*, Singapore Univ. Press, Singapore, pp. 115–181.