



Centrum voor Wiskunde en Informatica

REPORTRAPPORT

PNA

Probability, Networks and Algorithms



Probability, Networks and Algorithms

A J-function for marked point patterns

M.N.M. van Lieshout

REPORT PNA-R0404 JUNE 2004

CWI is the National Research Institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organization for Scientific Research (NWO).

CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

Information Systems (INS)

Copyright © 2004, Stichting Centrum voor Wiskunde en Informatica

P.O. Box 94079, 1090 GB Amsterdam (NL)

Kruislaan 413, 1098 SJ Amsterdam (NL)

Telephone +31 20 592 9333

Telefax +31 20 592 4199

ISSN 1386-3711

A J-function for marked point patterns

ABSTRACT

We propose a new summary statistic for marked point patterns. The underlying principle is to compare the distance from a marked point to the nearest other marked point in the pattern to the same distance seen from an arbitrary point in space. Information about the range of interaction can be inferred, and the statistic is well-behaved under random mark allocation. We develop a range of Hanisch style kernel estimators to tackle the problems of exploding tail variance earlier associated with J -function plug-in estimators, and carry out an exploratory analysis of a forestry data set.

2000 Mathematics Subject Classification: 60D05, 60G55, 62G99, 62M30.

Keywords and Phrases: empty space function, J -function, marked point pattern, mark correlation function, nearest-neighbour distance distribution empty space function, J -function, marked point pattern, mark correlation function, nearest-neighbour distance distribution function, product density, random labelling, reduced second moment measure, spatial interaction, spatial statistics.

A J -function for marked point patterns

M.N.M. van Lieshout

CWI

P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

Abstract

We propose a new summary statistic for marked point patterns. The underlying principle is to compare the distance from a marked point to the nearest other marked point in the pattern to the same distance seen from an arbitrary point in space. Information about the range of interaction can be inferred, and the statistic is well-behaved under random mark allocation. We develop a range of Hanisch style kernel estimators to tackle the problems of exploding tail variance earlier associated with J -function plug-in estimators, and carry out an exploratory analysis of a forestry data set.

Keywords & Phrases: empty space function, J -function, marked point pattern, mark correlation function, nearest-neighbour distance distribution function, product density, random labelling, reduced second moment measure, spatial interaction, spatial statistics.

2000 Mathematics Subject Classification: 60D05, 60G55, 62G99, 62M30.

1. INTRODUCTION

Marked point patterns are spatial point configurations with a mark attached to each point [34]. The points could represent the locations in (Euclidean) space of objects, while the marks capture additional information. The latter could be a type label, in which case we also speak of a multivariate point pattern [11], a continuous measurement or shape descriptor, or a combination of these.

The statistical analysis of such a pattern in general begins by plotting a few summary statistics. Which statistic is used depends on taste and the type of mark. For discrete marks, cross versions of the classic nearest neighbour distance distribution function G [14] or the second order K -function [27] are popular. Alternatively, multivariate J -functions [22] may be used, or the mark correlation functions advocated in [26, 34]. For real-valued marks, mark correlation functions are typically used, at least if the marks are not binned so as to reduce the situation to the case of discrete marks. It is often a good idea to plot a range of summary statistics, as they tend to capture different aspects of the pattern and thus provide complementary information.

In this paper, we focus on generalisations of the J -function [21] to point process with real-valued marks. The underlying idea of this summary statistic is to compare distances seen from an arbitrary fixed point 0 to the nearest point of the pattern – measured by the empty space function F – to those seen from a typical point of the pattern, as captured by the nearest neighbour distance function G . Thus, for (unmarked) point processes,

$$J(t) = \frac{1 - G(t)}{1 - F(t)}$$

defined for all $t \geq 0$ for which $F(t) \neq 1$. For a Poisson process, $J \equiv 1$; values $J(t) > 1$ indicate repulsion at range t , for clustered patterns the J -values tend to be less than 1.

The power of the J -function in hypothesis testing was assessed in [5, 9, 36]. Extensions to multivariate point processes were proposed in [22], and window based J -functions suggested in [5, 9]. For applications in agriculture, astronomy, forestry and geology, see [15, 18, 19, 20, 31], or the recent theses [8, 25].

The plan of this paper is as follows. In Section 2 we recall fundamental concepts from Palm theory, including reduced Campbell measures, product densities, Papangelou conditional intensities and the Nguyen–Zessin formula. In Section 3 we define a J -function for stationary marked point processes and discuss its behaviour under a Poisson or random labelling assumption. The next section concerns representation theorems for the J -function in terms of the conditional intensity and product densities, and explains the relationship with the reduced second moment measure. We turn to estimation in Section 5, and discuss Hanisch style ratio-estimators for the marked J -function, as well as kernel estimators for its derivative statistics. An application of our statistic to a forestry data set of pine saplings is the topic of Section 6. The paper is concluded with a summary and discussion.

2. PRELIMINARIES

In this section, we recall some fundamental concepts from Palm theory. Further details can be found for example in the textbooks by Daley and Vere–Jones [13] or Stoyan et al. [33].

Throughout this paper, let Y be a stationary marked point process on \mathbb{R}^d with marks in a complete, separable metric space M (typically a subset of an Euclidean space) and write P for its distribution.

2.1 Palm theory

The *first order moment measure* of Y is defined as the expected number of points

$$\mu_1(A \times B) := \mathbb{E} \left[\int_A \int_B dY(x, m) \right]$$

in a bounded Borel set A in \mathbb{R}^d with marks in the Borel set $B \subseteq M$. In general, μ_1 need not be finite. If it is, we say the moment measure of Y exists. By the usual measure theoretic arguments, μ_1 can be extended uniquely to a measure on the Borel product σ -algebra on $\mathbb{R}^d \times M$. By the stationarity assumption, the first order moment measure is invariant under translations of the points (leaving the marks as they are), so it can be factorised as $\mu_1 = \lambda \ell \times \nu_M$ where λ is the intensity, ℓ denotes Lebesgue measure, and ν_M is our notation for the *mark distribution* ν_M on M . Throughout this paper, we shall assume that λ is strictly positive to exclude trivial cases.

The *reduced Campbell measure* $C^!$ of Y is defined by

$$C^!((A \times B) \times F) := \mathbb{E} \left[\int_A \int_B \mathbf{1}_F(Y \setminus \{(x, m)\}) dY(x, m) \right]$$

for all bounded Borel sets $A \subset \mathbb{R}^d$, all Borel sets $B \subseteq M$, and all F in the σ -algebra \mathcal{N} generated by the requirement that for all bounded Borel sets, the number of points with marks in Borel subsets of M is a finite random variable. Note that $C^!$ is σ -finite, and can be extended uniquely to a σ -finite measure on the product σ -algebra.

Suppose the first order moment measure exists. Then, for each fixed $F \in \mathcal{N}$, $C^1(\cdot \times F)$ is absolutely continuous with respect to μ_1 (or equivalently with respect to $\ell \times \nu_M$), so that

$$C^1((A \times B) \times F) = \int_A \int_B P^{l(x,m)}(F) d\mu_1(x, m) \quad (2.1)$$

where the Radon–Nikodym derivative $P^{l(x,m)}(F)$ is measurable and μ_1 -integrable for bounded Euclidean Borel sets A and Borel subsets B of M . As a function of (x, m) , $P^{l(x,m)}(F)$ is defined uniquely up to a μ_1 -null set. A regular version is a choice for which $P^{ly}(\cdot)$ is a probability distribution for each fixed $y = (x, m) \in \mathbb{R}^d \times M$. Such a version exists, but is not unique. It is called a *reduced Palm distribution* of Y at the point x with mark m . For any Borel set $B \subseteq M$ for which $\nu_M(B) > 0$, define a reduced Palm distribution P_B^{lx} of Y with respect to B at the point $x \in \mathbb{R}^d$ by

$$\nu_M(B) P_B^{lx}(F) := \int_B P^{l(x,m)}(F) d\nu_M(m). \quad (2.2)$$

By virtue of Fubini's theorem, $P_B^{lx}(F)$ is a measurable function of x on \mathbb{R}^d .

Write $T_{-u}F = \{\mathbf{y} - u : \mathbf{y} \in F\}$ for $u \in \mathbb{R}^d$. Then the stationarity of Y implies that the reduced Campbell measure is invariant under translations in the sense that

$$C^1(((A - u) \times B) \times T_{-u}F) = C^1((A \times B) \times F)$$

for all u , all bounded Borel sets $A \subset \mathbb{R}^d$, all Borel sets $B \subseteq M$, and all F in the σ -algebra \mathcal{N} . Provided $\nu_M(B) > 0$, it follows from (2.1) that for given $u \in \mathbb{R}^d$ and $F \in \mathcal{N}$,

$$\int_A \int_B P^{l(x,m)}(F) d\ell(x) d\nu_M(m) = \int_A \int_B P^{l(x-u,m)}(T_{-u}F) d\ell(x) d\nu_M(m)$$

for all Borel sets $A \subset \mathbb{R}^d$. Hence $P_B^{lx}(F) = P_B^{lx-u}(T_{-u}F)$ for ℓ -almost all x . Since \mathcal{N} is countably generated, and reduced Palm distributions are probability measures, the exceptional null set may be chosen in such a way that it does not depend on F . In particular for $F = T_xG$, $G \in \mathcal{N}$, $P_B^{lx}(T_xG) = P_B^{lx-u}(T_{x-u}G)$ except on a null set. It follows that the local Palm distributions are translates

$$P_B^{lx}(T_xG) \equiv \tilde{P}_B(G)$$

of a single probability distribution \tilde{P}_B for ℓ -almost all x . The fact that \tilde{P}_B is a probability distribution is inherited from any reduced Palm distribution version. Heuristically, $\tilde{P}_B(G)$ may be interpreted as the conditional probability that the marked pattern Y from which the marked point at the origin is deleted falls in the event G given Y has a point at the origin with mark in B . We shall henceforth write $P_B^{l0}(G)$ with slight abuse of notation.

2.2 The Nguyen–Zessin identity

As before, let Y be a stationary marked point process for which the first order moment measure exists. Additionally, suppose that $C^1((A \times B), \cdot)$ is absolutely continuous with respect to P for all bounded Borel sets $A \subset \mathbb{R}^d$ and all Borel sets $B \subseteq M$. Then

$$C^1((A \times B) \times F) = \int_F \Lambda((A \times B); \mathbf{y}) dP(\mathbf{y})$$

for all $F \in \mathcal{N}$, where $\Lambda((A \times B); \cdot)$ is a Radon–Nikodym derivative which is uniquely defined up to a P -null set. There exist regular versions for which $\Lambda((A \times B); \cdot)$ is a measurable function for any fixed A and B , and a locally finite Borel measure for fixed \mathbf{y} . Such a version is called a Papangelou kernel. If furthermore $\Lambda(\cdot; \mathbf{y})$ has a density $\lambda((x, m); \mathbf{y})$ with respect to the product of Lebesgue measure and ν_M , the latter is called a *Papangelou conditional intensity*. It is uniquely defined up to a μ_1 -null set, and can be interpreted as the conditional probability of finding a point at an infinitesimal product region centred at x with mark in m given that the marked point pattern elsewhere is \mathbf{y} . We shall also consider integrated conditional intensities $\lambda_B(x; \mathbf{y}) = \int_B \lambda((x, m); \mathbf{y}) d\nu_M(m)$ over Borel sets $B \subseteq M$.

As for regular versions of reduced Palm distributions, the translation invariance of the reduced Campbell measure may be exploited to deduce that for fixed $u \in \mathbb{R}^d$, a Borel subset B of M with positive probability $\nu_M(B) > 0$, and $F \in \mathcal{N}$

$$\mathbb{E}[\mathbf{1}_F(Y) \lambda_B(x; Y)] = \mathbb{E}[\mathbf{1}_F(Y + u) \lambda_B(x - u; Y)]$$

for ℓ almost all x , and the exceptional null set may be chosen independently of F . Therefore,

$$\mathbb{E}[\mathbf{1}_G(Y - x) \lambda_B(x; Y)] = \int_G \lambda_B(x; \mathbf{y} + x) dP(\mathbf{y}) \equiv C_B(G)$$

is constant for ℓ -almost all $x \in \mathbb{R}^d$. By the fact that \mathcal{N} is countably generated, the null set may be chosen independently of G , so that

$$\lambda_B(x_1; Y + x_1) = \lambda_B(x_2; Y + x_2) \quad \text{a.s.}$$

for x_1, x_2 not in the exceptional ℓ -null set. With slight abuse of notation, we shall write $\mathbb{E}[\mathbf{1}_G(Y) \lambda_B(0; Y)]$ and $\lambda_B(0; Y)$ as in the previous subsection.

Under the above assumptions, the following basic formulae

$$\mathbb{E} \left[\sum_{y \in Y} g(y, Y \setminus y) \right] = \int \int \int g((x, m), \mathbf{y}) dC^l((x, m), \mathbf{y}) \tag{2.3}$$

$$= \lambda \int \int \mathbb{E}^{l(x, m)} [g((x, m), Y)] d\ell(x) d\nu_M(m)$$

$$= \lambda \int \int \mathbb{E}^{l(0, m)} [g((x, m), Y + x)] d\ell(x) d\nu_M(m) \tag{2.4}$$

$$= \int \int \mathbb{E} [g((x, m), Y) \lambda((x, m); Y)] d\ell(x) d\nu_M(m) \tag{2.5}$$

hold for all non-negative, measurable functions g (in the sense that the left hand side is finite if and only if the right hand side is). The first equation expresses the Campbell–Mecke formula (see e.g. [13, Prop. 12.1.IV] or [33, p. 107] for marked point processes, which under stationarity reduces to (2.4), see [33, equation (4.4.10)]. The last equation due to Nguyen and Zessin [23, Theorem 2 and 2'] requires the existence of a conditional intensity, in other words that C^l is absolutely continuous with respect to the product measure of $\ell \times \nu_M$ and P .

If $g(y, Y) = f(y)$ is a function of its first argument only, the fundamental formulae reduce to

$$\mathbb{E} \left[\sum_{y \in Y} f(y) \right] = \int \int f(x, m) d\mu_1(x, m) = \int \int f(x, m) \mathbb{E} [\lambda((x, m); Y)] d\ell(x) d\nu_M(m)$$

for any non-negative, measurable function f on $\mathbb{R}^d \times M$. Consequently, $\mathbb{E}[\lambda((x, m); Y)] = \lambda$ for $\ell \times \nu_M$ -almost all (x, m) . Similarly, for any Borel mark set B for which $\nu_M(B) > 0$,

$$\mathbb{E}[g(Y) \lambda_B(0; Y)] = \lambda \nu_M(B) \mathbb{E}_B^{!0} [g(Y)]. \quad (2.6)$$

Example 1. For a stationary unmarked point process X , the Nguyen–Zessin formula (2.6) reduces to

$$\mathbb{E}[g(X) \lambda(0; X)] = \lambda \mathbb{E}^{!0} [g(X)]$$

(cf. [33, equation (5.5.16)] or [23, Theorem 4]). For a discrete mark space $M = \{1, \dots, n_M\}$,

$$\mathbb{E}[g(Y) \lambda((0, i); Y)] = \lambda_i \mathbb{E}^{!(0, i)} [g(Y)]$$

with the notation $\lambda_i = \lambda \nu_M(\{i\}) > 0$.

2.3 Product densities and correlation functions

The n -th order factorial moment measure $\mu^{(n)}$ of a marked point process Y is defined by

$$\mathbb{E} \left[\sum_{y_1, \dots, y_n \in Y}^{\neq} g(y_1, \dots, y_n) \right] = \int \cdots \int g(y_1, \dots, y_n) d\mu^{(n)}(y_1, \dots, y_n)$$

for any non-negative measurable function on $(\mathbb{R}^d \times M)^n$. The sum is over all n -tuples $y_1 = (x_1, m_1), \dots, y_n = (x_n, m_n)$ for which $x_i \neq x_j$ for $i \neq j$. For $n = 1$, $\mu^{(1)} = \mu_1$, the first order moment measure. We shall say that the n -th order factorial moment measure exists, if it is locally finite. If furthermore $\mu^{(n)}$ is absolutely continuous with respect to the n -fold product measure of $\ell \times \nu_M$ with itself, a Radon–Nikodym derivative exists, the *product density* denoted by $\rho^{(n)}$. Note that product densities are permutation invariant functions, and

$$\begin{aligned} \mathbb{E} \left[\sum_{y_1, \dots, y_n \in Y}^{\neq} g(y_1, \dots, y_n) \right] = \\ \int \cdots \int g(y_1, \dots, y_n) \rho^{(n)}(y_1, \dots, y_n) d\ell \times \nu_M(y_1) \cdots d\ell \times \nu_M(y_n) \end{aligned} \quad (2.7)$$

for all non-negative, measurable functions $g \geq 0$. An appealing heuristic interpretation is that $\rho^{(n)}(y_1, \dots, y_n) d\ell(x_1) \cdots d\ell(x_n) d\nu_M(m_1) \cdots d\nu_M(m_n)$ represents the probability of a point falling in each of the infinitesimal regions centred at x_1, \dots, x_n with respective marks around m_1, \dots, m_n .

In the physics literature, for stationary marked point processes Y with intensity $\lambda > 0$, the related n -point *correlation functions* are commonly used. They are defined recursively in terms of product densities as follows:

$$\begin{aligned} \xi_1 &\equiv 1; \\ \rho^{(n)}(y_1, \dots, y_n) &= \lambda^n \sum_{k=1}^n \sum_{D_1, \dots, D_k} \xi_{n(D_1)}(\mathbf{Y}_{D_1}) \cdots \xi_{n(D_k)}(\mathbf{Y}_{D_k}) \end{aligned} \quad (2.8)$$

where the last sum ranges over all $\{D_1, \dots, D_k \neq \emptyset\}$ partitions of $\{1, \dots, n\}$ in k non-empty, disjoint sets, and $\mathbf{y}_{D_j} = \{y_i : i \in D_j\}$ is the corresponding partition of marked points. For a stationary Poisson process, $\rho^{(n)} \equiv \lambda^n$, so that for $n > 1$, the ξ_n account for the excess due to n -tuples in comparison with the reference Poisson process. Thus, $\xi_n > 0$ suggests clustering, while $\xi_n < 0$ tends to correspond to n -th order inhibition.

Suppose that $\mu^{(n)}$ exists, and write $\alpha^{(n)}$ for the n -th order factorial moment measure of the unmarked point process associated with Y . Then, for fixed Borel subsets B_1, \dots, B_n of M , the measure $\mu^{(n)}(\cdot \times B_1 \times \dots \times B_n)$ is absolutely continuous with respect to $\alpha^{(n)}$. Hence, a Radon–Nikodym derivative $M^{x_1, \dots, x_n}(B_1 \times \dots \times B_n)$, can be found such that

$$\mu^{(n)}(A_1 \times \dots \times A_n \times B_1 \times \dots \times B_n) = \int_{A_1} \dots \int_{A_n} M^{x_1, \dots, x_n}(B_1 \times \dots \times B_n) d\alpha^{(n)}(x_1, \dots, x_n).$$

As usual, the Radon–Nikodym derivative is a measurable function of x_1, \dots, x_n defined up to an $\alpha^{(n)}$ null set. A regular version is a choice for which for fixed x_1, \dots, x_n , $M^{x_1, \dots, x_n}(\cdot)$ is a probability measure, known as the n -point *mark distribution*. It may be interpreted as the probability that a mark in B_i is assigned to the point at x_i , given there are points at x_1, \dots, x_n .

3. A J -FUNCTION FOR MARKED POINT PATTERNS

We begin this section with recalling two well-known summary statistics from spatial statistics.

The *empty space function* F of Y is the cumulative distribution function of the distance from an arbitrarily chosen origin to the nearest point of the process, that is

$$F(t) := P\{Y \cap (B(0, t) \times M) \neq \emptyset\}$$

for $t \geq 0$. Here we write $B(0, t)$ for the closed ball of radius t centred at 0. The *nearest neighbour distance distribution function* from a point with mark in B is defined by

$$G_B(t) := P_B^0\{Y \cap (B(0, t) \times M) \neq \emptyset\}$$

for $t \geq 0$ and Borel sets $B \subseteq M$ of positive ν_M -mass, the cumulative distribution function of the distance from a typical point of the process with mark in B to the nearest other point of Y regardless of its mark.

The J -function compares F to G_B , as made precise in the following definition.

Definition 1. *Let B be a Borel subset of M with $\nu_M(B) > 0$. Then the J -function with respect to mark set B is given by*

$$J_B(t) = \frac{1 - G_B(t)}{1 - F(t)}$$

and defined for all $t \geq 0$ for which $F(t) < 1$.

To interpret $J_B(t)$, note that for an independently marked Poisson point process, Slivnyak’s theorem implies that $G_B(t) = F(t)$ for all B and t , hence $J_B \equiv 1$. Values less than 1 occur when $G_B(t) > F(t)$, that is when nearest neighbour distances are smaller than distances from the origin. Intuitively, such cases suggest clustering. On the other hand, values larger than 1 occur when the empty spaces are small in comparison to the distance from a point with mark in B to its nearest neighbour, an indication of inhibition. Note though [7] that a J -function that is 1 on its domain of definition does not imply that Y is a Poisson process.

Example 2. Suppose the marks are discrete labels, say $1, \dots, n_M$, and let $B = \{i\}$. Then the i -to-any nearest neighbour distance distribution function is given in our notation by

$$G_B(t) = P_B^{l0}\{Y \cap (B(0, t) \times M) \neq \emptyset\};$$

the more usual notation is $G_{i\bullet}(t)$. Upon division by $1 - F(t)$ when allowed, we obtain the multivariate i -to-any J -function $J_{i\bullet}$ [22]. Write (X_1, \dots, X_{n_M}) for the vector of point patterns of types $1, \dots, n_M$. If X_i is independent of $\cup_{j \neq i} X_j$, then $J_{i\bullet}(t) = J_{ii}(t)$, the marginal J -function of X_i . If furthermore X_i is a Poisson process, $J_{i\bullet}(t) \equiv 1$.

If one recalls that $\nu_M(B) G_B^{l0}(t) = \nu_M(B) P_B^{l0}\{Y \cap (B(0, t) \times M) \neq \emptyset\}$ is the almost everywhere constant value of $\int_B P^{l(x, m)}\{Y \cap (B(x, t) \times M) \neq \emptyset\} d\nu_M(m)$, the definition of $J_B(t)$ may be rewritten as a mixture

$$J_B(t) = \frac{1}{\nu_M(B)} \int_B \left[\frac{P^{l(x, m)}\{Y \cap (B(x, t) \times M) \neq \emptyset\}}{1 - F(t)} \right] d\nu_M(m) \quad (3.1)$$

for ℓ -almost all $x \in \mathbb{R}^d$. Thus, $J_B(t)$ may be interpreted as an average over B of J -functions with respect to a point marked $m \in B$ at an arbitrarily chosen origin.

Example 3. For the multivariate point processes of Examples 1–2, the mixture formula reads

$$J_M(t) = \sum_{i=1}^{n_M} \frac{\lambda_i}{\lambda} J_{i\bullet}(t)$$

for all t for which $F(t) < 1$. If all components of (X_1, \dots, X_{n_M}) are mutually independent, $J_M(t)$ is a convex combination of the J_{ii} with weights λ_i/λ , cf. [22].

Definition 2. The marked point process Y has the random labelling property if the marks of the points are conditionally i.i.d. given the point locations.

For marked point processes with the random labelling property, the J -function is of a convenient form, as stated more precisely in the following result.

Proposition 1. Let X be a stationary point process on \mathbb{R}^d with finite positive intensity λ , randomly labelled with mark distribution ν_M , and write Y for the marked point process thus obtained. Then, for all $t \geq 0$ with $F(t) < 1$, the J -function of Y with respect to a Borel mark set $B \subseteq M$ with $\nu_M(B) > 0$ is given by

$$J_B(t) = J_X(t),$$

the J -function of X .

Proof: One needs to prove that the nearest neighbour distance distribution function of X coincides with that of Y with respect to any mark set, i.e. that

$$1 - G_B(t) = 1 - G_X(t) \quad (3.2)$$

for all Borel sets $B \subseteq M$ with $\nu_M(B) > 0$. Here we use the notation G_X for the nearest neighbour distance distribution function of X , and shall use similar notations P_X and \mathbb{E}_X for

the distribution of X and its expectation below. To prove (3.2), fix such a B and let A be any bounded Borel set of positive d -dimensional volume. Consider the measurable function

$$g((x, m), Y) = \mathbf{1}_A(x) \mathbf{1}_B(m) \mathbf{1}\{Y \cap (B(x, t) \times M) \neq \emptyset\}.$$

Now, the number of marked points falling in A is finite almost surely, with a finite expectation. Hence the expected sum of g over the points in Y is finite and

$$\begin{aligned} \mathbb{E} \left[\sum_{(x, m(x))=y \in Y} g(y, Y \setminus \{y\}) \right] &= \mathbb{E} \left[\sum_{x \in X \cap A} \mathbf{1}\{(X \setminus \{x\}) \cap B(x, t) \neq \emptyset\} \mathbf{1}_B(m(x)) \right] \\ &= \nu_M(B) \mathbb{E}_X \left[\sum_{x \in X \cap A} \mathbf{1}\{(X \setminus \{x\}) \cap B(x, t) \neq \emptyset\} \right] \\ &= \lambda \nu_M(B) \int_A P_X^{!0} \{(X + x) \cap B(x, t) \neq \emptyset\} d\ell(x) \\ &= \lambda \nu_M(B) \ell(A) G_X(t) \end{aligned}$$

because of the conditionally independent mark assignments and the Campbell–Mecke formula for P_X .

On the other hand, since Y is stationary with $\mu_1 = \lambda \ell \times \nu_M$, where λ is the intensity of X , by the Campbell–Mecke formula (2.4)

$$\begin{aligned} \mathbb{E} \left[\sum_{y \in Y} g(y, Y \setminus \{y\}) \right] &= \lambda \int_A \int_B P^{!(0, m)} \{(Y + x) \cap (B(x, t) \times M) \neq \emptyset\} d\ell(x) d\nu_M(m) \\ &= \lambda \nu_M(B) \ell(A) G_B(t) \end{aligned}$$

We conclude that $G_X(t) = G_B(t)$, and the desired result follows upon division by $1 - F(t)$ on both sides of the equation (3.2). \square

4. REPRESENTATION THEOREMS

In this section, relationships between the J -function and fundamental marked point processes descriptors, namely the Papangelou conditional intensity and n -point correlation functions, are explored. We shall obtain a connection with the widely used second order K -function, and prove that the J -function with respect to any mark set becomes flat beyond the effective range of interaction.

4.1 Representation in terms of Papangelou conditional intensity

The Nguyen–Zessin formula motivated the definition of the J -function, as it relates expectations under the reduced Palm distribution to ones under the distribution of Y itself. Hence it should not come as a surprise that the J -function with respect to a mark set can be expressed explicitly in terms of conditional intensities.

Proposition 2. *Let Y be a stationary marked point process with finite positive intensity λ for which a regular version of the conditional intensity exists that satisfies the Nguyen–Zessin*

formula (2.6), and B a Borel set in the mark space M with $\nu_M(B) > 0$. Then $G_B(t) < 1$ for some $t \geq 0$ implies $F(t) < 1$ and

$$J_B(t) = \mathbb{E} \left[\frac{\lambda_B(0; Y)}{\lambda \nu_M(B)} \middle| Y \cap (B(0, t) \times M) = \emptyset \right] \quad (4.1)$$

$$= \left(\mathbb{E}_B^{!0} \left[\frac{\lambda \nu_M(B)}{\lambda_B(0; Y)} \middle| Y \cap (B(0, t) \times M) = \emptyset \right] \right)^{-1}. \quad (4.2)$$

Proof: Let A be the event $\{Y \cap (B(0, t) \times M) = \emptyset\}$, so that $F(t) = 1$ if and only if $P(A) = 0$. Apply the Nguyen–Zessin formula (2.6) for the measurable function $g(Y) = \mathbf{1}_A(Y) \geq 0$ to obtain

$$\mathbb{E} [\mathbf{1}_A(Y) \lambda_B(0; Y)] = \lambda \nu_M(B) \mathbb{E}_B^{!0} [\mathbf{1}_A(Y)] = \lambda \nu_M(B) (1 - G_B(t)).$$

If $P(A) = 0$, so is the expectation in the left hand side of the above equality. Since by assumption $\lambda \nu_M(B) > 0$, necessarily $G_B(t) = 1$. Hence $G_B(t) < 1$ implies $P(A) > 0$, or, equivalently, $F(t) < 1$. Thus, we may divide the left and right hand side by $\lambda \nu_M(B) (1 - F(t))$ to obtain (4.1).

Next, apply the Nguyen–Zessin formula to the function $g(Y) = \mathbf{1}_A(Y)/\lambda_B(0; Y)$ (we shall show below that the function is well-defined). Then, from

$$\lambda \nu_M(B) \mathbb{E}_B^{!0} [\mathbf{1}_A(Y)/\lambda_B(0; Y)] = \mathbb{E} [\mathbf{1}_A(Y)] = 1 - F(t),$$

we obtain (4.2) upon dividing both sides of the equation by $1 - G_B(t)$. Note that $\lambda_B(0; Y) > 0$ almost surely with respect to the reduced Palm distribution $P_B^{!0}$, since

$$\lambda \nu_M(B) P_B^{!0} \{\lambda_B(0; Y) = 0\} = \mathbb{E} [\mathbf{1}\{\lambda_B(0; Y) = 0\} \lambda_B(0; Y)] = 0,$$

hence the conditional expectation in (4.2) is well-defined. It follows that

$$\begin{aligned} \lambda \nu_M(B) &= \lambda \nu_M(B) \mathbb{E}_B^{!0} [\mathbf{1}\{\lambda_B(0; Y) > 0\}] = \mathbb{E} [\mathbf{1}\{\lambda_B(0; Y) > 0\} \lambda_B(0; Y)] \\ &\leq \mathbb{E} [\lambda_B(0; Y)] = \lambda \nu_M(B). \end{aligned}$$

Hence

$$\mathbf{1}\{\lambda_B(0; Y) > 0\} \lambda_B(0; Y) = \lambda_B(0; Y) \quad P - \text{a.s.},$$

or, in other words, $P\{\lambda_B(0; Y) = 0\} = 0$, and the function $g(Y)$ is well-defined. \square

Example 4. For a stationary point process X with intensity $\lambda > 0$, Proposition 2 holds with

$$J(t) = \mathbb{E} \left[\frac{\lambda(0; X)}{\lambda} \middle| X \cap B(0, t) = \emptyset \right] = \left(\mathbb{E}^{!0} \left[\frac{\lambda}{\lambda(0; X)} \middle| X \cap B(0, t) = \emptyset \right] \right)^{-1};$$

in the multivariate case (Examples 1–2)

$$J_{i\bullet}(t) = \mathbb{E} \left[\frac{\lambda((0, i); Y)}{\lambda_i} \middle| X_{\bullet} \cap B(0, t) = \emptyset \right] = \left(\mathbb{E}^{!(0, i)} \left[\frac{\lambda_i}{\lambda((0, i); Y)} \middle| X_{\bullet} \cap B(0, t) = \emptyset \right] \right)^{-1}$$

where $X_{\bullet} = \cup_{i=1}^{n_M} X_i$ (cf. [21, 22]).

The following corollary gives a useful interpretation of the J -statistic.

Corollary 1. *Let Y be a stationary marked point process with finite positive intensity λ for which a regular version of the conditional intensity exists that satisfies the Nguyen–Zessin formula (2.6). Then $J_B(t) \geq 1$ (respectively ≤ 1) if and only if*

$$\text{Cov}(\lambda_B(0; Y), \mathbf{1}\{Y \cap (B(0, t) \times M) = \emptyset\}) \geq 0$$

(respectively is non-positive).

Another corollary states that $J_B(t)$ is constant beyond the effective range of interaction.

Definition 3. *A marked point process Y has effective interaction range s if for all Borel sets $B \subseteq M$ with $\nu_M(B) > 0$ its conditional intensity $\lambda_B(0; Y)$ is constant for all realisations which contain no points in $B(0, s)$.*

Thus, for t greater than the effective interaction range, given that $Y \cap (B(0, t) \times M) = \emptyset$, the conditional intensity $\lambda_B(0; Y) = \lambda_B(0; \emptyset)$.

Corollary 2. *If Y has joint interaction range s , $0 < s < \infty$, then $J_B(t)$ is constant for all $t \geq s$ for which it is defined, and*

$$J_B(t) \equiv \frac{\lambda_B(0; \emptyset)}{\lambda \nu_M(B)}.$$

A widely used family of a marked point process models with finite effective interaction range is that of the pairwise interaction models. Such models have a conditional intensity of the form

$$\lambda((u, m); \mathbf{y}) = \beta(m) \prod_{y_j \in \mathbf{y}} \gamma((u, m), y_j)$$

for $(u, m) \notin \mathbf{y}$, where β and γ are non-negative measurable functions (cf. [3, 29, 28, 24]). Typically β is bounded, and $\gamma \leq 1$. It is easily verified that

$$J_B(t) = \frac{1}{\lambda \nu_M(B)} \int_B \beta(m) \mathbb{E} \left[\prod_{y \in Y} \gamma((0, m), y) \middle| Y \cap (B(0, t) \times M) = \emptyset \right] d\nu_M(m)$$

wherever defined. If $\gamma((u, m), (v, n)) \equiv 1$ for $\|u - v\| > s$, the J -function with respect to B reduces to

$$J_B(t) = \frac{1}{\lambda \nu_M(B)} \int_B \beta(m) d\nu_M(m)$$

for $t \geq s$ such that $F(t) < 1$.

4.2 Representation in terms of product densities

In astronomical folklore, the connection between the empty space function and the product densities, or equivalently the n -point correlation functions, is well-known. Indeed, White [37] argues that the fact that F uses product densities of all orders makes it particularly appropriate to detect clustering in galaxy catalogues. Here, we shall give an expression of the J -function for marked point processes in terms of n -point correlation functions, and consider in detail how to obtain a classic second order analysis [34] by truncation.

We shall need the following Lemma.

Lemma 1. (*White (1979)*)

Let Y be a stationary marked point process. Suppose that all order factorial moment measures exist as locally finite measures, and have a Radon–Nikodym derivative $\rho^{(n)}$ with respect to the n -fold product of $\ell \times \nu_M$, $n \in \mathbb{N}$. Then the empty space function can be written as

$$F(t) = - \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \int_{B(0,t) \times M} \cdots \int_{B(0,t) \times M} \rho^{(n)}(y_1, \dots, y_n) d\ell \times \nu_M(y_1) \cdots d\ell \times \nu_M(y_n) =$$

$$1 - \exp \left[\sum_{n=1}^{\infty} \frac{(-\lambda)^n}{n!} \int_{B(0,t) \times M} \cdots \int_{B(0,t) \times M} \xi_n(y_1, \dots, y_n) d\ell \times \nu_M(y_1) \cdots d\ell \times \nu_M(y_n) \right]$$

where ξ_n are given by (2.8).

The main result of this subsection is the following.

Proposition 3. Let Y be a stationary marked point process. Suppose that all order factorial moment measures exist as locally finite measures, and have a Radon–Nikodym derivative $\rho^{(n)}$ with respect to the n -fold product of $\ell \times \nu_M$, $n \in \mathbb{N}$. Then the J -function with respect to any Borel mark set $B \subseteq M$ with $\nu_M(B) > 0$ can be written as

$$J_B(t) = \frac{1}{\nu_M(B)} \left[\nu_M(B) + \sum_{n=1}^{\infty} \frac{(-\lambda)^n}{n!} J_n^B(t) \right]$$

for all $t \geq 0$ for which $F(t) < 1$, where $J_n^B(t)$ is the common value of

$$\int_B \int_{B(0,t) \times M} \cdots \int_{B(0,t) \times M} \xi_{n+1}((a, m), y_1 + a, \dots, y_n + a) d\nu_M(m) d\ell \times \nu_M(y_1) \cdots d\ell \times \nu_M(y_n)$$

for ℓ -almost all $a \in \mathbb{R}^d$.

Proof: By the Campbell–Mecke formula (2.3), for any bounded Borel set $A \subset \mathbb{R}^d$ of strictly positive Lebesgue measure $\ell(A) > 0$,

$$\lambda \nu_M(B) \ell(A) (1 - G_B(t)) = \mathbb{E} \left[\sum_{(a,m) \in Y \cap (A \times B)} \mathbf{1}\{Y \setminus \{(a, m)\} \cap (B(a, t) \times M) = \emptyset\} \right].$$

The expectation on the right hand side is well-defined and finite, since the first order (factorial) moment measure exists as a locally finite measure. Now, by the inclusion-exclusion formula, the expectation may be rewritten as

$$\mathbb{E} \left[\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \sum_{(a,m), y_1, \dots, y_n \in Y}^{\neq} \mathbf{1}_{A \times B}(a, m) \mathbf{1}_{B(a,t) \times M}(y_1) \cdots \mathbf{1}_{B(a,t) \times M}(y_n) \right]$$

which by (2.7) is equal to $\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_A f_n^B(a, t) d\ell(a)$ with

$$f_n^B(a, t) = \int_B \int_{B(0,t) \times M} \cdots \int_{B(0,t) \times M} \rho^{(n+1)}((a, m), y_1 + a, \dots, y_n + a) d\nu_M(m) d\ell \times \nu_M(y_1) \cdots d\ell \times \nu_M(y_n).$$

By stationarity, the factorial moment measures are translation invariant, hence for all $n \in \mathbb{N}$, $f_n^B(\cdot, t)$ is almost everywhere constant, say $f_n^B(t)$. Of course, $f_0^B \equiv \lambda \nu_M(B)$. From the recursion relation (2.8), it follows that $J_n^B(t)$ is well-defined too, with $J_0^B \equiv \nu_M(B)$. In summary

$$\lambda \nu_M(B) (1 - G_B(t)) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} f_n^B(t). \quad (4.3)$$

In order to complete the proof, we need to show that the right hand side of equation (4.3) is equal to $\lambda (1 - F(t)) \sum_{n=0}^{\infty} \frac{(-\lambda)^n}{n!} J_n^B(t)$. To do so, use the definition of the correlation functions, and split the partition into those terms that contain the first marked point and those that do not. More precisely, (4.3) equals

$$\lambda \nu_M(B) + \lambda \sum_{n=1}^{\infty} \frac{(-\lambda)^n}{n!} \sum_{D \subseteq \{1, \dots, n\}} J_{n(D)}^B(t) \sum_{k=1}^{n-n(D)} \sum_{D_1, \dots, D_k \neq \emptyset; \cup D_j = \{1, \dots, n\} \setminus D} I_{n(D_1)} \cdots I_{n(D_k)} =$$

$$\left[\lambda J_0^B(t) + \lambda \sum_{n=1}^{\infty} \frac{(-\lambda)^n}{n!} J_n^B(t) \right] \times \left[1 + \sum_{m=1}^{\infty} \frac{(-\lambda)^m}{m!} \sum_{k=1}^m \sum_{D_1, \dots, D_k \neq \emptyset; \cup D_j = \{1, \dots, m\}} I_{n(D_1)} \cdots I_{n(D_k)} \right]$$

where I_n is the n -fold integral over the Cartesian product of $B(0, t) \times M$ of the n -point correlation functions with respect to the appropriate $\ell \times \nu_M$ product measure. Note that the second term is $1 - F(t)$, by White's Lemma, and the desired representation holds. \square

Proposition 3 requires the existence of factorial moment measures of all orders. If such an assumption is not valid, one could truncate the alternating series at some fixed n . Indeed, the approximation in terms of product densities up to second order reads

$$J_B(t) - 1 \approx \frac{-\lambda}{\nu_M(B)} \int_B \left[\int_{B(0,t) \times M} \xi_2((0, m), y) d\ell \times \nu_M(y) \right] d\nu_M(m)$$

$$= \frac{-1}{\lambda \nu_M(B)} \int_B \left[\int_{B(0,t) \times M} \left(\rho^{(2)}((0, m), y) - \lambda^2 \right) d\ell \times \nu_M(y) \right] d\nu_M(m).$$

Thus,

$$J_B(t) - 1 \approx -\lambda [\mathcal{K}_B(t) - \ell(B(0, t))] \quad (4.4)$$

where \mathcal{K}_B is the reduced second moment measure (see e.g. the textbooks [27, 34]) with respect to the mark set B . In other words, $\lambda\mathcal{K}_B(t)$ is the expected number of further points within a radius t of a typical point with mark in B (i.e. under $P_B^{!0}$). In terms of the 2-point mark distribution function, (4.4) reads

$$\begin{aligned} J_B(t) &\approx 1 - \lambda \int_{B(0,t)} \left[\frac{M^{0x}(B \times M)}{\nu_M(B)} g(0, x) - 1 \right] d\ell(x) \\ &= -\lambda \int_{B(0,t)} [k_f(0, x) g(0, x) - 1] d\ell(x) \end{aligned}$$

where g is the pair correlation function of the unmarked point process associated with Y , which is proportional by a factor $1/\lambda^2$ to the second order product density, and $k_f(0, x) = \int_M \int_M f(m_1, m_2) dM^{0,x}(m_1, m_2)$ with $f(m_1, m_2) = \mathbf{1}_B(m_1)/\nu_M(B)$. Note that f is normalised, as $\int_M \int_M f(m_1, m_2) d\nu_M(m_1) d\nu_M(m_2) = 1$.

A typical second order analysis of marked point processes [26, 34] plots estimates of the pair correlation function and $k_f(\cdot, \cdot)$ for a suitably chosen non-negative, Borel measurable and integrable function f . For a single positive real-valued mark, the function $f(m_1, m_2) = m_1 m_2 / \mu^2$ may be used, where $\mu = \int_{\mathbb{R}^+} m d\nu_M(m)$ is the mean mark. For discrete or binned labels, $f(m_1, m_2) = \mathbf{1}_A(m_1) \mathbf{1}_B(m_2) / (\nu_M(A) \nu_M(B))$ is a convenient choice. The latter amounts to a J -function analysis with truncation at second order product densities, see also the appendix.

Example 5. *Let X be a stationary, isotropic, planar point process for which a second order product density exists. Then*

$$J(t) \approx 1 - \lambda \int_{B(0,t)} (g(\|x\|) - 1) d\ell(x) = 1 - 2\pi\lambda \int_0^t r (g(r) - 1) dr,$$

which is known as the ‘Gaussian approximation’ [18] in astronomy.

5. ESTIMATION

Throughout this section, assume that Y is a stationary marked point process on \mathbb{R}^d with marks in a complete, separable metric space M with finite, positive intensity λ . Thus, the first moment measure and hence Palm kernels exist. Below, we propose Hanisch style kernel estimators for the J -function and some associated characteristics, when Y is observed within a compact set $W \subseteq \mathbb{R}^d$ of positive volume $\ell(W)$. We shall rely on the principle that Palm characteristics may be estimated by averages over points of the marked point process [33, p. 130], i.e. the estimator

$$\lambda \nu_M(\widehat{B}) \widehat{\mathbb{E}}_B^{!0} f(Y) = \sum_{y=(x,m) \in Y} \frac{\mathbf{1}_W(x) \mathbf{1}_B(m) f((Y-x) \setminus \{(0,m)\})}{\ell(W)}$$

is unbiased for any non-negative measurable function f by virtue of the Campbell–Mecke formula (2.3). For example, for the mark distribution we have

$$\widehat{\lambda \nu_M(B)} = \sum_{(x,m) \in Y} \frac{\mathbf{1}_W(x) \mathbf{1}_B(m)}{\ell(W)}.$$

for any Borel subset B of M and any compact set $W \subseteq \mathbb{R}^d$ of positive volume $\ell(W) > 0$. For many f , care has to be taken with regard to edge effects caused by the fact that not Y itself is observed, but rather $Y \cap (W \times M)$. For such f , the estimator described above cannot be computed based on the available data. This phenomenon known as the ‘edge effect’ is particularly irksome for irregularly shaped windows W , and in higher dimensions.

In the approach of Hanisch [16] for functions based on inter-point distances such as the J -function, a solution for the edge effect problem lies in the observation that on the event $\{d(x, \partial W) \geq d(x, Y)\}$ the observed distance $d(x, Y \cap (W \times M))$ is equal to the true one $d(x, Y)$. Other types of edge-corrected estimators for F and G are reviewed in many textbooks, including [12, chap. 8], [27, chap. 3], or [33, pp. 122–131], as well as in [2]. We have chosen the Hanisch approach, as it leads to estimators that – in contrast to other estimators – respect the monotonicity and continuity properties of F and G_B , and do not discard too much data. The plug-in principle then yields ratio-unbiased estimators for J_B . Unfortunately though, empirical evidence (see eg. [21, 22]) suggests that the variance of $\widehat{J_B(t)}$ will increase with t , causing a rather fluctuating tail behaviour of $\widehat{J_B(t)}$. To solve this problem, we propose to combine edge correction with smoothing [30], to obtain more robust estimators.

5.1 Cumulative distribution functions

The first result of this section concerns an edge-corrected estimator for the nearest neighbour distance distribution function with respect to a mark set.

Proposition 4. *For any Borel set $B \subseteq M$ with $\nu_M(B) > 0$, define*

$$\sum_{y_k \in Y \cap (W \times B)} \left[\frac{\mathbf{1}\{b_k \geq s_k\} \mathbf{1}\{s_k \leq t\}}{\ell(W^{\ominus s_k})} \right] \tag{5.1}$$

for $t \geq 0$ such that $\ell(W^{\ominus t}) > 0$. Here, $y_k = (x_k, m_k)$ is a marked point, and $s_k = d(x_k, Y \setminus \{(x_k, m_k)\})$ respectively $b_k = d(x_k, \partial W)$ are the Euclidean distances from x_k to the nearest other point in Y and to the boundary of W . The notation $W^{\ominus s_k}$ is used for the morphological erosion of W with a ball of radius s_k centred at the origin, that is, for the set of points in W for which the distance to the boundary is at least s_k . Then (5.1) is an unbiased estimator of $\lambda \nu_M(B) G_B(t)$ for all $t \geq 0$ for which it is defined; it is a non-negative function that is increasing in t .

Proof: The mapping $(x, Y) \mapsto d(x, Y)$ on $\mathbb{R}^d \times N$, the product space of \mathbb{R}^d and the configuration space N of locally finite marked point patterns \mathbf{y} in $\mathbb{R}^d \times M$ is jointly measurable when restricted to W , and non-negative. Therefore

$$\mathbb{E} \left[\sum_{(x,m) \in Y \cap (W \times B)} \mathbf{1}\{x \in W^{\ominus d(x, Y \setminus \{(x, m)\})}\} \mathbf{1}\{d(x, Y \setminus \{(x, m)\}) \leq t\} / \ell(W^{\ominus d(x, Y \setminus \{(x, m)\})}) \right] =$$

$$\lambda \nu_M(B) \int_W \mathbb{E}_B^{10} \left[\frac{\mathbf{1}\{x \in W^{\ominus d(0,Y)}\} \mathbf{1}\{d(0,Y) \leq t\}}{\ell(W^{\ominus d(0,Y)})} \right] d\ell(x)$$

by stationarity and the Campbell–Mecke theorem (2.3). An application of Fubini’s theorem yields that the expectation of (5.1) reduces to

$$\lambda \nu_M(B) \mathbb{E}_B^{10} [\mathbf{1}\{d(0,Y) \leq t\}] = \lambda \nu_M(B) G_B(t).$$

Clearly, (5.1) is non-negative and increasing in t in the range for which it is well-defined. \square

For the rectangular or circular windows that are typically encountered in spatial statistics, the term $\ell(W^{\ominus s})$ can be evaluated explicitly.

Two variations deserve to be mentioned. Firstly,

$$\sum_{y_k \in Y \cap (W \times B)} \left[\frac{\mathbf{1}\{b_k \geq s_k\}}{\ell(W^{\ominus s_k})} \right] \tag{5.2}$$

is an unbiased estimator of $\lambda \nu_M(B)$, so division of (5.1) by (5.2) yields a ratio-unbiased estimator of $G_B(t)$, in the spirit of [35] who advocated to use intensity estimators similar to estimators of the numerator.

The Hanisch estimators for the empty space function [10] is well known:

$$\widehat{F}(t) = \sum_{x_k \in L} \left[\frac{\mathbf{1}\{b_k \geq s_k\} \mathbf{1}\{s_k \leq t\}}{\#\{i : b_i \geq s_k\}} \right] \tag{5.3}$$

with the convention that $0/0 = 0$. The sum is over a finite lattice L in W . It should be noted that there is no need to delete x_k from the point pattern in the computation of s_k , as it will almost surely be no part of a realisation of Y . The range of t is restricted so as to make sure never to divide by zero. Clearly, $\widehat{F}(t)$ is increasing; it is also unbiased.

In a range of papers [5, 8, 9], uncorrected estimators of J -functions were considered. These may be seen as unbiased for ‘window averaged J -functions’, and can be surprisingly powerful as test statistic. As is our J -function, for Poisson processes the windowed J -function is identically equal to 1, but in general explicit evaluation seems to be more cumbersome, no representation theorems have been found, and the behaviour under random labelling is unknown. The theoretical windowed values are typically closer to 1 than the classic ones, which may be understood as ‘Poissonisation due to window averaging’.

5.2 Densities and hazard rates

For exploratory purposes, density and hazard rates often convey more information than cumulative statistics [1, 34]. Indeed, suppose a Papangelou conditional intensity exists so that the Nguyen–Zessin formula (2.6) holds. Then, provided $G_B(t) < 1$,

$$J'_B(t) = [h_F(t) - h_{G_B}(t)] J_B(t)$$

where h_I denotes the hazard rate of statistic I . Thus, the derivative statistic $J'(t, B)/J(t, B)$ is a signed measure of spatial association.

The hazard rates exist, whence $J_B(t)$ is differentiable, under the Nguyen–Zessin condition. Indeed, by theorems from [2, 17], the empty space function of a stationary point process is absolutely continuous with density and hazard rate given by

$$f_F(t) = \frac{\mathbb{E}h_{d-1}(\partial(X^{\oplus t}) \cap Z)}{\ell(Z)}; \quad h_F(t) = \frac{\mathbb{E}[h_{d-1}(\partial(X^{\oplus t}) \cap Z)]}{\mathbb{E}[\ell(Z \setminus X^{\oplus t})]} \quad (5.4)$$

for any compact regular set Z (i.e. $\text{cl}(Z^{\text{int}}) = Z$). Here h_{d-1} is the $d - 1$ dimensional Hausdorff measure, and $Y^{\oplus t} = \cup_{(x,m) \in Y} B(x, t)$. Chiu and Stoyan [10] observed that (5.3) is a discretisation of

$$\tilde{F}(t) = \int_W \left[\frac{\mathbf{1}\{x \in W^{\ominus d(x,Y)}\} \mathbf{1}\{d(x, Y) \leq t\}}{\ell(W^{\ominus d(x,Y)})} \right] d\ell(x) = \int_0^t \frac{h_{d-1}(W^{\ominus s} \cap \partial(Y^{\oplus s}))}{\ell(W^{\ominus s})} ds$$

for all t for which it is defined. Provided W is regular, the integrand is an unbiased estimator of $f(t)$ based on the minus sampling principle, providing us with an alternative interpretation of (5.3).

In general, densities and hazard rates do not exist for G_B . Indeed, the nearest neighbour distance distribution function may be degenerate, for instance for randomly translated grids. However, if the Nguyen–Zessin identity holds, a density does exist.

Proposition 5. *Let Y be a stationary marked point process with finite positive intensity λ for which a regular version of the conditional intensity exists that satisfies the Nguyen–Zessin formula (2.6), and B a Borel mark set with $\nu_M(B) > 0$. Then the nearest neighbour distance $d(0, Y)$ from a point at 0 with mark in B is absolutely continuous with density*

$$g_B(t) = f_F(t) \mathbb{E} \left[\frac{\lambda_B(0; Y)}{\lambda \nu_M(B)} \middle| d(0, Y) = t \right]. \quad (5.5)$$

Proof: Suppose $P\{d(0, Y) \in A\} = 0$ for any Borel subset A of the positive half line, in other words, $\mathbf{1}\{d(0, Y) \in A\} = 0$ P -almost surely. Consequently, $\lambda \nu_M(B) P_B^{10}\{d(0, Y) \in A\} = \mathbb{E}[\lambda_B(0; Y) \mathbf{1}\{d(0, Y) \in A\}] = 0$ by (2.6). Therefore, the nearest neighbour distance distribution with respect to mark set B is absolutely continuous with respect to the distribution of $d(0, Y)$, so, by the Radon–Nikodym theorem

$$P_B^{10}\{d(0, Y) \in A\} = \int_A f_{GF}(a) dF(a)$$

for some measurable, integrable function f_{GF} on the positive half line, and in particular

$$G_B(t) = \int_0^t f_{GF}(s) dF(s) = \int_0^t f_{GF}(s) f_F(s) ds.$$

The Nguyen–Zessin formula further implies that

$$\begin{aligned} \lambda \nu_M(B) G_B(t) &= \mathbb{E}[\lambda_B(0; Y) \mathbf{1}\{Y \cap (B(0, t) \times M) \neq \emptyset\}] \\ &= \mathbb{E}\mathbb{E}[\lambda_B(0; Y) \mathbf{1}\{d(0, Y) \leq t\} | d(0, Y)] \\ &= \int_0^t \mathbb{E}[\lambda_B(0; Y) | d(0, Y) = s] dF(s). \end{aligned}$$

Hence, $\lambda \nu_M(B) f_{GF}(s) = \mathbb{E}[\lambda_B(0; Y) | d(0, Y) = s]$ for almost all s , and (5.5) follows. \square

It is instructive to note that beyond the range of interaction $g_B(t) = f_F(t) \lambda_B(0; \emptyset) / (\lambda \nu_M(B)) = f_F(t) J_B(t)$ so that

$$(h_F(t) - h_{G_B(t)}) J_B(t) = \left[\frac{f_F(t)}{1 - F(t)} - \frac{f_F(t) J_B(t)}{1 - G_B(t)} \right] J_B(t) = \left[\frac{f_F(t)}{1 - F(t)} - \frac{f_F(t)}{1 - F(t)} \right] J_B(t)$$

vanishes as it should.

In the following proposition, we derive Hanisch-style kernel estimators [30] of $f_F(t)$ and $\lambda \nu_M(B) g_B(t)$.

Proposition 6. *Let Y be a stationary marked point process with finite positive intensity λ for which a regular version of the conditional intensity exists that satisfies the Nguyen–Zessin formula (2.6), B be a Borel mark set with $\nu_M(B) > 0$, and $L = \{x_1, \dots, x_k\}$ a finite set of points in W . Given a symmetric, measurable, non-negative kernel $k_h(\cdot)$ on \mathbb{R} with bandwidth h that integrates to unity, i.e. $k_h(x) = k(x/h)/h$,*

$$\widehat{f_F(t)} = \sum_{x_k \in L} \frac{k_h(t - s_k) \mathbf{1}\{b_k \geq s_k\}}{\#\{i : b_i \geq s_k\}}$$

with the convention $0/0 = 0$ is an unbiased estimator of $\int_0^\infty k_h(t - s) f_F(s) ds$. As before, $s_k = d(x_k, Y)$ and $b_k = d(x_k, \partial W)$. Furthermore,

$$\lambda \nu_M(B) \widehat{g_B(t)} = \sum_{y_k \in Y \cap (W \times B)} \frac{k_h(t - s_k) \mathbf{1}\{b_k \geq s_k\}}{\ell(W^{\ominus s_k})}$$

is an unbiased estimator of $\lambda \nu_M(B) \int_0^\infty k_h(t - s) g_B(s) ds$ for $t \geq 0$ such that $\ell(W^{\ominus t}) > 0$. Here, $y_k = (x_k, m_k)$ and $s_k = d(x_k, Y \setminus \{(x_k, m_k)\})$.

Proof: Note that

$$\begin{aligned} \mathbb{E} \widehat{f_F(t)} &= \sum_{x_k \in L} \mathbb{E} \left[\frac{k_h(t - d(x_k, Y)) \mathbf{1}\{x_k \in W^{\ominus d(x_k, Y)}\}}{\#\{i : x_i \in W^{\ominus d(x_k, Y)}\}} \right] \\ &= \sum_{x_k \in L} \int_0^\infty \left[\frac{k_h(t - s) \mathbf{1}\{x_k \in W^{\ominus s}\}}{\#\{i : x_i \in W^{\ominus s}\}} \right] f_F(s) ds = \int_0^\infty k_h(t - s) f_F(s) ds. \end{aligned}$$

Regarding the nearest-neighbour distance distribution density $g_B(t)$,

$$\begin{aligned} \mathbb{E} \lambda \nu_M(B) \widehat{g_B(t)} &= \lambda \int_W \int_B \mathbb{E}^{!(0, m)} \left[\frac{k_h(t - d(0, Y)) \mathbf{1}\{x \in W^{\ominus d(0, Y)}\}}{\ell(W^{\ominus d(0, Y)})} \right] d\ell(x) d\nu_M(m) \\ &= \lambda \nu_M(B) \int_W \mathbb{E}_B^{!0} \left[\frac{k_h(t - d(0, Y)) \mathbf{1}\{x \in W^{\ominus d(0, Y)}\}}{\ell(W^{\ominus d(0, Y)})} \right] d\ell(x) \\ &= \lambda \nu_M(B) \mathbb{E}_B^{!0} [k_h(t - d(0, Y))] = \lambda \nu_M(B) \int_0^\infty k_h(t - s) g_B(s) ds \end{aligned}$$

by stationarity, the Campbell–Mecke formula and Fubini’s theorem. \square

Since M is a metric space, say with metric $\rho(\cdot, \cdot)$, we may consider the family $\tilde{k}_{\tilde{h}}(m, n) := \tilde{k}(\rho(m, n)/\tilde{h})/\tilde{h}$ of kernels with bandwidth $h > 0$ for some fixed measurable function $\tilde{k} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ that integrates to unity, and apply ideas from kernel estimation theory to both the mark set and the range, as exemplified by the following proposition.

Proposition 7. *Let Y be a stationary marked point process with finite positive intensity λ and marks in the complete, separable metric space (M, ρ) for which a regular version of the conditional intensity exists that satisfies the Nguyen–Zessin formula (2.6). Write $y_k = (x_k, m_k)$, $s_k = d(x_k, Y \setminus \{y_k\})$ and $b_k = d(x_k, \partial W)$. Then, given symmetric kernels k_h with bandwidth h on \mathbb{R} and $\tilde{k}_{\tilde{h}}$ on $M \times M$ that integrate to unity, for $m \in M$ and $t \geq 0$ such that $\ell(W^{\ominus t}) > 0$,*

$$\widehat{h_m(t)} = \frac{\sum_{y_k \in Y \cap (W \times M)} \tilde{k}_{\tilde{h}}(m, m_k) k_h(t - s_k) \mathbf{1}\{b_k \geq s_k\} / \ell(W^{\ominus s_k})}{\sum_{y_k \in Y \cap (W \times M)} \tilde{k}_{\tilde{h}}(m, m_k) \mathbf{1}\{b_k \geq s_k\} \mathbf{1}\{s_k > t\} / \ell(W^{\ominus s_k})}$$

is a ratio-unbiased estimator of

$$\frac{\lambda \int_M \int_0^\infty \tilde{k}_{\tilde{h}}(m, n) k_h(t - s) g_n(s) d\nu_M(n) ds}{\lambda \int_M \int_t^\infty \tilde{k}_{\tilde{h}}(m, n) g_n(s) d\nu_M(n) ds}$$

where $g_n(s) = f_F(s) \mathbb{E}[\lambda((0, n); Y) | d(0, Y) = s] / \lambda$.

Proof: Since

$$\begin{aligned} & \mathbb{E} \left[\sum_{(x,n) \in Y \cap (W \times M)} \frac{\tilde{k}_{\tilde{h}}(m, m_k) k_h(t - d(x, Y \setminus \{(x, n)\})) \mathbf{1}\{x \in W^{\ominus d(x, Y \setminus \{(x, n)\})}\}}{\ell(W^{\ominus d(x, Y \setminus \{(x, n)\})})} \right] = \\ & \lambda \int_W \int_M \mathbb{E}^{!(0,n)} \left[\frac{\tilde{k}_{\tilde{h}}(m, n) k_h(t - d(0, Y)) \mathbf{1}\{x \in W^{\ominus d(0, Y)}\}}{\ell(W^{\ominus d(0, Y)})} \right] d\ell(x) d\nu_M(n) = \\ & \lambda \int_M \tilde{k}_{\tilde{h}}(m, n) \mathbb{E}^{!(0,n)} [k_h(t - d(0, Y))] d\nu_M(n), \end{aligned}$$

by the Campbell–Mecke theorem for stationary processes and Fubini’s theorem. Similarly, one may show the expectation of the denominator is given by

$$\lambda \int_M \tilde{k}_{\tilde{h}}(m, n) \mathbb{E}^{!(0,n)} [\mathbf{1}\{d(0, Y) > t\}] d\nu_M(n).$$

We proceed to show that for any non-negative measurable function h on M , and k on \mathbb{R}^+ ,

$$\int_M h(n) \mathbb{E}^{!(0,n)} [k(d(0, Y))] d\nu_M(n) = \int_M \int_0^\infty h(n) k(s) g_n(s) d\nu_M(n) ds.$$

To do so, follow the route from indicator functions h via step functions by linearity, to general h by approximation and the monotone convergence theorem. Indeed, let $h(n) = \mathbf{1}_B(n)$. If $\nu_M(B) = 0$, the desired identity trivially holds. Otherwise, by Proposition 5

$$\begin{aligned} \int_B \mathbb{E}^{1(0,n)} [k(d(0, Y))] d\nu_M(n) &= \nu_M(B) \mathbb{E}_B^{!0} [k(d(0, Y))] \\ &= \nu_M(B) \int_0^\infty k(s) f_F(s) \mathbb{E} \left[\frac{\lambda_B(0; Y)}{\lambda \nu_M(B)} \middle| d(0, Y) = s \right] ds \\ &= \frac{1}{\lambda} \int_0^\infty k(s) f_F(s) \mathbb{E} [\lambda_B(0; Y) | d(0, Y) = s] ds \\ &= \int_B \int_0^\infty k(s) \frac{f_F(s)}{\lambda} \mathbb{E} [\lambda((0, n); Y) | d(0, Y) = s] d\nu_M(n) ds. \end{aligned}$$

To finish the proof, apply the claim with $h(n) = \tilde{k}_h(m, n)$ and $k(d(0, Y)) = k_h(t - d(0, Y))$. The proof of the second statement follows upon replacement of the kernel $k_h(t - d(0, Y))$ by the indicator function $\mathbf{1}\{d(0, Y) > t\}$. \square

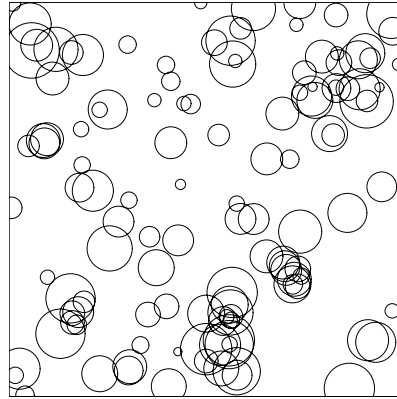


Figure 1: Positions of 126 pine saplings (Kellomaki, Joensuu) within a 10 by 10 meter square. The marks record the height in meter and are represented by a disc.

6. FORESTRY EXAMPLE

Below, we illustrate the use of the marked J -function by means of a data set of pine saplings in Finland (cf. Figure 1) collected by Professor Seppo Kellomaki from Joensuu, and kindly provided by Professor Antti Penttinen. The observation window W is the square $[-5, 5] \times$

$[-8, 2]$, but note that the data were originally recorded in a larger circular plot with polar coordinates. After transformation to Euclidean coordinates, to get rid of alignments at larger distances, some random rounding was done as a result of which there are a few close pairs of neighbours. For each of the 126 pines, the height and diameter at breast height were measured. The marks are strongly positively correlated. Moreover, a number of trees are broken resulting in zero diameter at breast height. For these reasons, we base our analysis on the height marks only, i.e. take $M = \mathbb{R}^+$.

We begin our analysis with first order characteristics. The intensity estimator is $\hat{\lambda} = 1.26$, and the sample mean of the mark distribution $\hat{\mu} = 2.83$. A histogram and kernel estimator are plotted in Figure 2. Note that the histogram counts in bin B are unbiased estimators of $\lambda \nu_M(B) \ell(W)$, the kernel estimator is ratio-unbiased due to the fact that the number of points is random.

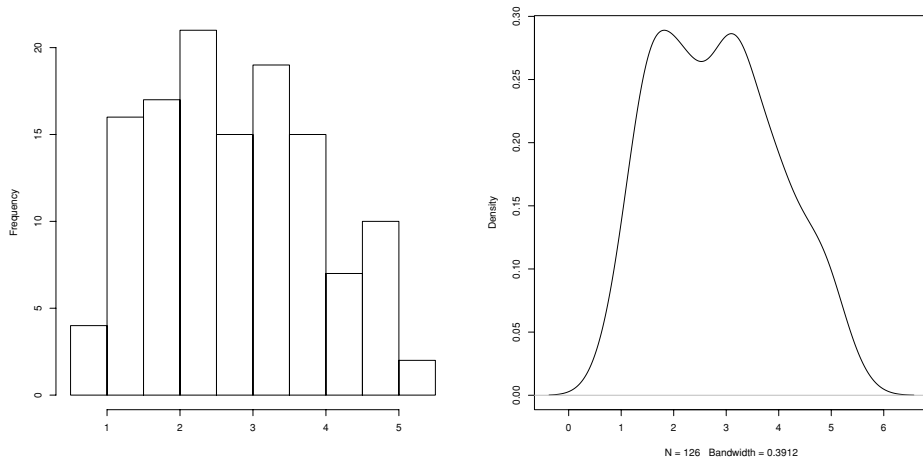


Figure 2: Histogram and Gaussian kernel estimator of the mark distribution for the pine saplings data.

The Hanisch estimators for the nearest neighbour distance distribution function and empty space function have been implemented in R using the package `spatstat` [4]. The values t at which we evaluated the estimators were separated by 0.025. For the lattice L , we used a regular 100 by 100 grid girting the boundary. The J -function of the point process of locations is less than 1 which suggests a clustering of trees over the considered range. Note that since there are many small saplings in the field, one would not expect to see a hard core effect that is typically observed in older more established forests, a feature reinforced by the rounding involved in transforming the data from polar to Euclidean coordinates. As for the difference in hazard rates $h_F(t) - h_G(t)$, its graph is monotonically decreasing up to about $t = 0.2$, then increases.

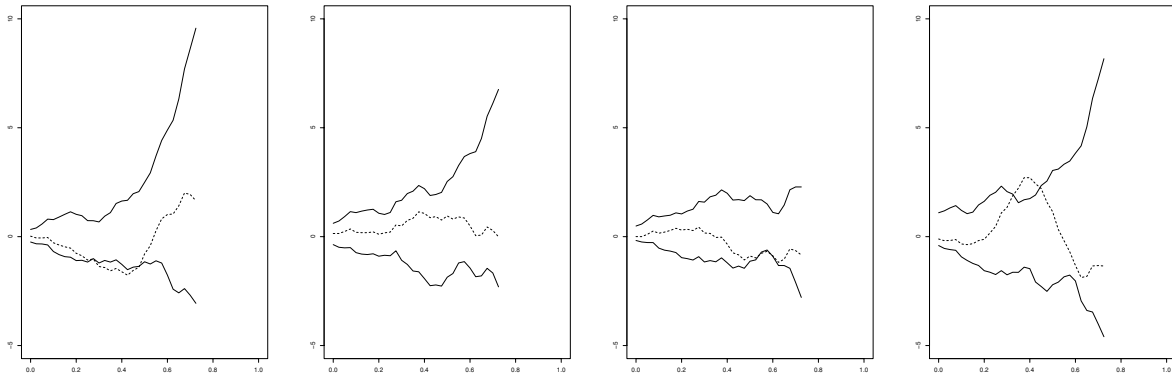


Figure 3: Difference of hazard rates for pine saplings data for $m = (1.5, 2.5, 3.5, 4.5)$.

To investigate whether random labelling is a tenable hypothesis, we condition on the locations, and sample the labels without replacement. Within the scope of this paper we have restricted ourselves to exploratory data analysis. Of course a more formal test could be designed quite easily [6]. Empirical evidence [5, 8, 36] suggests that the power of tests based on a J -function is comparable to that of the more powerful of the alternatives based on F or nearest neighbour distances. Figure 3 presents graphs of $\widehat{h_F(t)} - \widehat{h_m(t)}$ based on box kernels for the mark with $\tilde{h} = 0.5$, and Gaussian kernels with $h = 0.1$ for the range. We also plotted upper and lower envelopes based on 19 independent resamplings of the marks without replacement. It can be seen that the deviation from the null hypothesis is significant for large and small marks, while the graph keeps within the envelopes for the intermediate heights. Further insight perhaps is gained from the plots of the densities obtained by Hanisch kernel estimation with a Gaussian kernel of bandwidth 0.10 depicted in Figure 4. For $B = (0, 2)$, the estimated curve is bimodal up to $t = 0.8$, which is picked up by the hazard rate statistic. For $B = (4, \infty)$, the two local peaks at smaller range are reflected in the hazard rate statistic, but the more pronounced local peak around $t = 0.8$ is not significant because of the increased variance indicated by the widening of the envelope. In both cases, the shape of $\widehat{g_B(t)}$ is quite different from that of $\widehat{g_M(t)}$ (the dotted line in Figure 4), its estimated null hypothesis value.

7. DISCUSSION AND CONCLUSION

In this paper, we introduced a summary statistic for stationary marked point patterns based on comparing the distance to the nearest (other) point of the pattern seen from an arbitrarily chosen origin to that from a typical marked point. This J -statistic captures both the type and strength of interaction, and reduces to a simple form under random labelling. We derived representations in terms of both Papangelou conditional intensity and correlation functions, thus relating the J -function to fundamental concepts in marked point process theory.

Further variations on the theme are possible. For example, the balls $B(0, t)$ used in the definition of $J_A(t)$ may be replaced by any bounded Borel set. This would be particularly useful in a directional analysis of a non-isotropic marked point process. In another vein, only distances to points with a certain type of mark may be considered. More precisely,

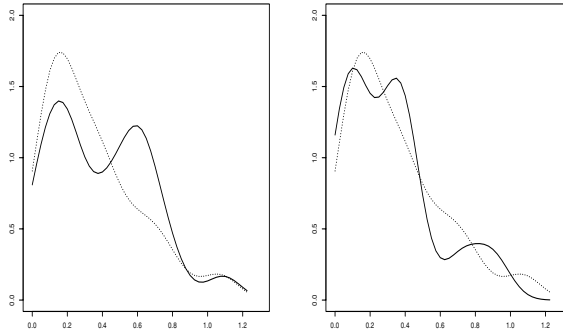


Figure 4: Hanisch estimators of g_B with Gaussian kernel for pine saplings data and $B = (0, 2)$ and $(4, \infty)$. The dotted line is the corresponding estimator \hat{g}_M .

if $F_B(t) = P\{Y \cap (B(0, t) \times B) \neq \emptyset\}$ is the empty space function of $Y \cap (B \times M)$ and $G_{AB}(t) = P_A^{!0}\{Y \cap (B(0, t) \times B) \neq \emptyset\}$ the cross nearest neighbour distance distribution function from points with mark in A to those with mark in B , an A -to- B cross J -function for marked point processes is given by

$$J_{AB}(t) = \frac{1 - G_{AB}(t)}{1 - F_B(t)}$$

for Borel subsets A and B of the mark space M with $\nu_M(A) > 0$, and all $t \geq 0$ for which $F_B(t) < 1$. The function $J_{AB}(t)$ measures the influence of the presence of a point with mark in the set A on the presence of points with a mark in B within distance t compared to the same event seen from an arbitrary origin. Values $J_{AB} > 1$ can be interpreted as indicating inhibition of points with mark in B by those with mark in A . Similarly, values less than 1 mean that the presence of a point with a mark in the set A decreases the probability of finding points with a mark value in B nearby.

The definition of J_{AB} is not symmetric in the mark sets, which is sometimes an advantage as pointed out in [22] and illustrated in practice by [15]. Note that if $\nu_M(B) = 0$, the expected number of points with mark in B is zero. The cross J -function has properties similar to those of the multivariate J -function [22]. For further details, see the Appendix.

From a statistical perspective, we discussed Hanisch style kernel estimators of densities and hazard rates of J_A that suppress the variance explosion at larger range encountered by plug-in estimators of J_A itself. Finally, the new statistic was used to explore the spatial structure of a forestry data set.

ACKNOWLEDGEMENTS

The author is grateful to A. Penttinen for the data and helpful discussions, to C. Beisbart for posing a question that motivated this work, and to A. Baddeley and R. Turner for access to ‘spatstat’.

References

1. Baddeley, A.J. and Gill, R.D. (1994). The empty space hazard of a spatial pattern. Research Report 1994/3, Department of Mathematics, The University of Western Australia.
2. Baddeley, A.J. and Gill, R.D. (1997). Kaplan–Meier estimators of distance distributions for spatial point processes. *Ann. Statist.*, 25:263–292.
3. Baddeley, A.J. and Møller, J. (1989). Nearest-neighbour Markov point processes and random sets. *Int. Statist. Rev.*, 57:89–121.
4. Baddeley, A.J. and Turner, R. Spatstat: An R library for spatial statistics. www.cran.r-project.org.
5. Baddeley, A.J., Kerscher, M., Schladitz, K. and Scott, B.T. (2000). Estimating the J function without edge correction. *Statist. Neerlandica*, 54:315–328.
6. Besag, J.E. and Diggle, P.J. (1977). Simple Monte Carlo tests for spatial pattern. *Appl. Statist.*, 26:327–333.
7. Bedford, T. and Berg, J. van den (1997). A remark on the Van Lieshout and Baddeley J -function for point processes. *Adv. in Appl. Probab.*, 29:19–25.
8. Chen, J. (2003). *Summary statistics in point patterns and their applications*. PhD thesis, Curtin University of Technology.
9. Chen, J., Baddeley, A.J. and Nair, G. (2001). Uncorrected estimators of J -functions in multivariate point patterns. *Paper presented at 11th International Workshop on Stereology, Stochastic Geometry and Related Fields*, Perth, Australia.
10. Chiu, S.N. and Stoyan, D. (1998). Estimators of distance distributions for spatial patterns. *Statist. Neerlandica*, 52:239–246.
11. Cox, D.R. and Lewis, P.A.W. (1972). Multivariate point processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, volume 3, pages 401–448. University of California Press.
12. Cressie, N.A.C. (1991). *Statistics for spatial data*. John Wiley and Sons, New York.

- Revised reprint 1993.
13. Daley, D.J. and Vere-Jones, D. (1988). *An introduction to the theory of point processes*. Springer Verlag, New York. Second edition Volume I, *Elementary theory and methods*, 2003.
 14. Diggle, P.J. (1983). *Statistical analysis of spatial point patterns*. Academic Press, London.
 15. Foxall, R. and Baddeley, A.J. (2002). Nonparametric measures of association between a spatial point process and a random set, with geological applications. *J. Roy. Statist. Soc. Ser. C*, 51:165–182.
 16. Hanisch, K.-H. (1984). Some remarks on estimators of the distribution function of nearest neighbour distance in stationary spatial point patterns. *Math. Operationsforsch. Statist., Ser. statist.*, 15:409–412.
 17. Hansen, M.B., Gill, R.D. and Baddeley, A.J. (1996). Kaplan–Meier type estimators for linear contact distributions. *Scand. J. Statist.*, 23:129–155.
 18. Kerscher, M. (1998). Regularity in the distribution of superclusters? *Astron. Astrophys.*, 336:29–34.
 19. Kerscher, M., Schmalzing, J., Buchert, T. and Wagner, H. (1998). Fluctuations in the IRAS 1.2 Jy catalogue. *Astron. Astrophys.*, 333:1–12.
 20. Kerscher, M., Pons–Bordería, M.J., Schmalzing, J., Trasarti–Battistoni, R., Buchert, T., Martínez, V.J. and Valdarnini, R. (1999). A global descriptor of spatial pattern interaction in the galaxy distribution. *Astrophysical Journal*, 513:543–548.
 21. Lieshout, M.N.M. van and Baddeley, A.J. (1996). A nonparametric measure of spatial interaction in point patterns. *Statist. Neerl.*, 50:344–361.
 22. Lieshout, M.N.M. van and Baddeley, A.J. (1999). Indices of dependence between types in multivariate point patterns. *Scand. J. Statist.*, 26:511–532.
 23. Nguyen, X.-X. and Zessin, H. (1979). Integral and differential characterizations of the Gibbs process. *Math. Nachr.*, 88:105–115.
 24. Ogata, Y. and Tanemura, M. (1989). Likelihood estimation of soft-core interaction potentials for Gibbsian point patterns. *Ann. Inst. Statist. Math.*, 41:583–600.
 25. Paulo, M.J. (2002). *Statistical sampling and modelling for cork oak and eucalyptus stands* PhD thesis, Wageningen University.
 26. Penttinen, A. and Stoyan, D. (1989). Statistical analysis for a class of line segment processes. *Scand. J. Statist.*, 16:153–168.
 27. Ripley, B.D. (1988). *Statistical inference for spatial processes*. Cambridge University Press.
 28. Ripley, B. D. (1989). Gibbsian interaction models. In *Spatial statistics: past, present and future*, D.A. Griffiths (Ed.) Image, New York.
 29. Ripley, B.D. and Kelly, F.P. (1977). Markov point processes. *J. London Math. Soc.*, 15:188–192.
 30. Silverman, B.W. (1986). *Density estimation for statistics and data analysis*. Chapman

and Hall, London.

31. Stein, A., Lieshout, M.N.M. van and Booltink, H.W.G. (2001). Spatial interaction of methylene blue stained soil pores. *Geoderma*, 102:101–121.
32. Stoyan, D. (1984). On correlations of marked point processes. *Math. Nachr.*, 116:197–207.
33. Stoyan, D., Kendall, W.S. and Mecke, J. (1987). *Stochastic geometry and its applications*. Akademie-Verlag, Berlin, 1987. Second edition 1995.
34. Stoyan, D. and Stoyan, H. (1994). *Fractals, random shapes and point fields. Methods of geometrical statistics*. John Wiley and Sons, Chichester. Translated from the 1992 German original.
35. Stoyan, D. and Stoyan, H. (2000). Improving ratio estimators of second order point process characteristics. *Scand. J. Statist.*, 27:641–656.
36. Thónnnes, E. and Lieshout, M.N.M. van (1999). A comparative study on the power of Van Lieshout and Baddeley's J -function. *Biom. J.*, 41:721–734.
37. White, S.D.M. (1979). The hierarchy of correlation functions and its relation to other measures of galaxy clustering. *Mon. Not. R. astr. Soc.*, 186:145–154.

APPENDIX

Below, some basic properties of the cross J -function are briefly discussed.

Write Y_A for the thinning of Y with retention probability of a marked point (x, m) equal to $\mathbf{1}\{m \in A\}$. If A_1, \dots, A_n form a partition of M , an analysis based on the cross J -function amounts to an analysis of the multivariate point pattern $Y = (X_1, \dots, X_n)$ with $X_i = Y_{A_i}$ as in [22] or the thesis by Chen [8]. In general, though, A and B need not be disjoint. Indeed, for $B = M$, $G_{AM} \equiv G_A$, and we regain the J -function with respect to the mark set A .

Independence If $A \cap B = \emptyset$ and Y_A , the locations of points marked A , and Y_B are independent, then $J_{AB}(t) \equiv 1$ for all $t \geq 0$ for which $F_B(t) < 1$.

If Y_A is independent of Y_{A^c} , then $1 - F(t) = (1 - F_A(t))(1 - F_{A^c}(t))$ and $1 - G_A(t) = (1 - G_{AA}(t))(1 - F_{A^c}(t))$, which implies that $J_A(t)$ reduces to $J_{AA}(t)$, the marginal J function of Y_A , for all t for which $F(t) < 1$.

An explicit expression for J_{AB} under the random labelling may be obtained from [21, Theorem 3], but is much less pleasing than the expression for $J_A(t)$.

Representation in terms of Papangelou conditional intensity Let Y be a stationary marked point process with finite positive intensity λ for which a regular version of the conditional intensity $\lambda((x, m); Y)$ exists that satisfies (2.6). Then, by a similar appeal to the Nguyen–Zessin formula as used in the proof for $J_B(t)$, $G_{AB}(t) < 1$ implies $F_B(t) < 1$, and for such $t \geq 0$,

$$J_{AB}(t) = \mathbb{E} \left[\frac{\lambda_A(0; Y)}{\lambda \nu_M(A)} \middle| Y \cap (B(0, t) \times B) = \emptyset \right] = \left(\mathbb{E}_A^{!0} \left[\frac{\lambda \nu_M(A)}{\lambda_A(0; Y)} \middle| Y \cap (B(0, t) \times B) = \emptyset \right] \right)^{-1}.$$

It should be noted that in case $\nu_M(B) = 0$, the conditioning event has probability 1, hence $J_{AB} \equiv 1$ is non-informative.

Representation in terms of product densities Let Y be a stationary marked point process. Suppose that all order factorial moment measures exist as locally finite measures, and have a Radon–Nikodym derivative $\rho^{(n)}$ with respect to the n -fold product of $\ell \times \nu_M$, $n \in \mathbb{N}$. Then

$$J_{AB}(t) = \frac{1}{\nu_M(A)} \sum_{n=0}^{\infty} \frac{(-\lambda)^n}{n!} J_n^{AB}(t),$$

where

$$J_n^{AB}(t) = \int_A \int_{B(0,t) \times B} \dots \int_{B(0,t) \times B} \xi_{n+1}((0, m), y_1, \dots, y_n) d\nu_M(m) d\ell \times \nu_M(y_1) \dots d\ell \times \nu_M(y_n).$$

The first order approximation is

$$\begin{aligned} J_{AB}(t) &\approx 1 - \frac{1}{\lambda \nu_M(A)} \int_A \left[\int_{B(0,t) \times B} \left(\rho^{(2)}((0, m), y) - \lambda^2 \right) d\ell \times \nu_M(y) \right] d\nu_M(m) \\ &= 1 - \lambda \int_{B(0,t)} \left[\frac{M^{0x}(A \times B)}{\nu_M(A)} g(0, x) - \nu_M(B) \right] d\ell(x) \\ &= 1 - \lambda \nu_M(B) \int_{B(0,t)} [k_f(0, x) g(0, x) - 1] d\ell(x) \end{aligned}$$

where g is the pair correlation function of the unmarked point process associated with Y , and $k_f(0, x) = \int_M \int_M f(m_1, m_2) dM^{0,x}(m_1, m_2)$ for $f(m_1, m_2) = \mathbf{1}_A(m_1) \mathbf{1}_B(m_2) / (\nu_M(A) \nu_M(B))$. Thus, a first order approximation of the J -function resembles a second order analysis based on the pair correlation function and mark correlation function $\mathbf{1}_A(m) \mathbf{1}_B(n)$ as recommended in Stoyan and Stoyan [34]. Note that the first order approximation is additive in B , but that this property is not shared by a truncation at larger n .

Estimation The Hanisch idea applies equally to cross versions of the nearest neighbour distance distribution function. Indeed, if s_k is replaced by $d(x_k, Y \cap (W \times A) \setminus \{x_k, m_k\})$ in (5.1), an unbiased estimator of $\lambda \nu_M(B) G_{BA}(t)$ is obtained.

If Y is a stationary marked point process with finite positive intensity λ for which a regular version of the conditional intensity exists that satisfies the Nguyen–Zessin formula (2.6), the cross nearest neighbour distance distribution is absolutely continuous with density $g_{BA}(t)$ as in (5.5) but with f_F replaced by a density of the empty space function of the locations in Y_A , and with $d(0, Y_A) = t$ for the conditioning event. Density estimators of F_A and $\lambda \nu_M(B) G_{BA}(t)$ can then be defined as in Proposition 6, if we replace Y by Y_A in the definition of s_k .