AALBORG UNIVERSITY

Second order analysis for spatial Hawkes processes

by

Jesper Møller and Giovanni Luca Torrisi

R-2005-20

May 2005

Department of Mathematical Sciences
Aalborg University
Fredrik Bajers Vej 7 G ▪ DK - 9220 Aalborg Øst ▪ Denmark
Phone: +45 96 35 80 80 ▪ Telefax: +45 98 15 81 29
URL: http://www.math.aau.dk

ISSN 1399–2503 ▪ On-line version ISSN 1601–7811

24 May 2005

SECOND ORDER ANALYSIS FOR SPATIAL HAWKES PROCESSES

JESPER MØLLER,* *Aalborg University*

GIOVANNI LUCA TORRISI,** *CNR-Istituto per le Applicazioni del Calcolo "M. Picone"*

Abstract

We derive summary statistics for stationary spatial Hawkes processes which can be considered as spatial versions of classical Hawkes processes. Particularly, we derive the intensity, the pair correlation function and the Bartlett spectrum. Our results for Gaussian fertility rates and the extension to marked Hawkes processes are discussed.

Keywords: Bartlett spectrum; Hawkes process; pair correlation function; spatial point processes; summary statistics; unpredictable marks.

AMS 2000 Subject Classification: Primary 60G55

Secondary 62M30

1. Introduction

Classical Hawkes processes [4, 6, 7, 8, 9] and their extensions to marked Hawkes processes [2, 4, 11, 12, 17, 18] play a fundamental role in the theory of point processes and its applications. In this paper we consider a *spatial Hawkes process* $X \subset \mathbb{R}^d$ $(d \geq 1)$ defined as follows.

(a) The points in X arrive at different times, and each point $\eta \in X$ is of one of two types: an *immigrant* or an *offspring*.

(b) The immigrants constitutes a spatial point process $G_0 \subset \mathbb{R}^d$. Throughout this paper we assume that G_0 is stationary with intensity $\mu_0 \in (0, \infty)$ (where

* Postal address: Department of Mathematical Sciences, Aalborg University Fredrik Bajers Vej 7G, DK-9220 Aalborg, Denmark. Email address: jm@math.auc.dk

** Postal address: CNR-Istituto per le Applicazioni del Calcolo "M. Picone", Viale del Policlinico 137, I-00161 Rome, Italy. Email address: torrisi@iac.rm.cnr.it

"stationarity" and "intensity" are defined in Section 2.1).

(c) If we condition on a point η from X, then independently of the previous history, η generates a Poisson process $\Phi(\eta) \subset X$ of offspring with intensity function $\xi \mapsto \gamma(\xi - \eta)$, where γ is a non-negative locally integrable function called the *fertility rate*.

We can view X as a cluster process with cluster centres at the immigrants and where the cluster associated by an immigrant ξ is the point process consisting of ξ and all offspring generated after some steps by ξ. Figure 1 illustrates the construction of a planar Hawkes process; four clusters are shown, consisting of 1, 1, 2, and 6 points.

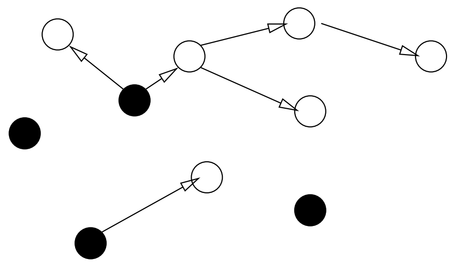

FIGURE 1: Illustration of how a spatial Hawkes process might be generated. Black points are the immigrants and the arrows indicate the offspring produced by each point of the process.

A classical Hawkes process on the line is the special case where $d = 1$, the arrival times are equal to X, G_0 is a Poisson process, and $\gamma(\xi) = 0$ for $\xi < 0$. For $d \geq 1$, we can view our process as a marked Hawkes process, where the points are given by the arrival times in (a) and X defines the marks. However, the arrival times play no important role in this paper and are therefore mostly ignored; they are only used in (c) when describing the independence of the previous history.

It is useful to consider a spatial Hawkes process as a superposition $X = \cup_0^\infty G_n$ where the $(n+1)$-th generation G_{n+1} given the previous generations $G_0, \ldots, G_n$ is a Poisson process on $\mathbb{R}^d$ with intensity function

$$\lambda_{n+1}(\xi) = \sum_{\eta \in G_n} \gamma(\xi - \eta). \tag{1}$$

Note that each G_n is a generalised shot noise Cox process [13], and (1) is a particular case of a model for reproducing individuals studied by Kingman [10] (see also Section

5.5 in [5]). We shall later compare spatial Hawkes processes with a Neyman-Scott process [16]; in a Neyman-Scott process only the immigrants generate offspring (more precisely, G_1 is a Neyman-Scott process if G_0 is a homogeneous Poisson process).

Spatial Hawkes processes may provide natural models for e.g. a population of reproducing individuals or the development of an epidemic. Nevertheless, to the best of our knowledge, spatial Hawkes processes have so far been studied very little in the literature. Brémaud, Massoulié, and Ridolfi [1] consider stationary spatial Hawkes processes with unpredictable marks, and they obtain the Bartlett spectrum (for the point process without the marks), assuming the existence of the Bartlett spectrum of the immigrant process. As explained in Section 3.3, our results easily extend to the case with unpredictable marks, but for ease of presentation we have chosen to concentrate on the unmarked case. Moreover, we are currently preparing a paper on simulation procedures, including perfect simulation, of spatial Hawkes processes.

In this paper we study the first and second order properties of spatial Hawkes processes in line with textbooks in spatial statistics [3, 5, 14] and stochastic geometry [20, 21]. Exploiting (1) we derive the intensity (Section 2.1) and, as our main result, the pair correlation function (Section 2.2). As a corollary, and in a simpler way compared to the exposition in [1], we obtain the Bartlett spectrum (Section 2.3). Our results for the pair correlation function are exemplified in Section 3.1 for Gaussian fertility rates. Other summary statistics such as the F, G, J, and K functions are briefly considered in Section 3.2. In Section 3.3 we discuss how our results easily extends to the case with unpredictable marks. All proofs are deferred to Appendix A–C.

2. Results

Throughout this paper we assume that the mean number of points in an offspring process $\Phi(\xi)$ is strictly less than one, i.e.

$$\nu = \int \gamma(\eta)\,\mathrm{d}\eta < 1. \tag{2}$$

This condition is equivalent to assume that X has finite intensity, cf. Proposition 1 below.

2.1. First order properties

Recall that a point process $Y \subset \mathbb{R}^d$ is stationary if the distribution of Y is invariant under translations in $\mathbb{R}^d$, and its intensity is then given by the mean number of points of Y per unit volume (the intensity may be infinite).

Proposition 1. *Each G_n is stationary with intensity $\rho_n = \mu_0 \nu^n$, and X is stationary with intensity*

$$\rho = \mu_0/(1-\nu). \tag{3}$$

Proof. See Appendix A.

2.2. Second order properties

In this section we find the pair correlation function $g(\xi, \eta)$ for a Hawkes process. Loosely speaking, $\rho^2 g(\xi,\eta)\, \mathrm{d}\xi\, \mathrm{d}\eta$ is the probability for observing a pair of points from X occurring jointly in each of two infinitesimally small balls with centres ξ, η and volumes $\mathrm{d}\xi$, $\mathrm{d}\eta$. The function $\rho^{(2)}(\xi,\eta) = \rho^2 g(\xi,\eta)$ is known as the product density of the second order factorial moment measure, and g is a kind of normalization of $\rho^{(2)}$. For further details, see Appendix B and [14, 20, 21].

2.2.1. *Set up* We need the following terminology and notation.

Consider any Lebesgue integrable functions f and h defined on $\mathbb{R}^d$. Let $f * h$ denote convolution, i.e. the Lebesgue integrable function

$$f * h(\xi) = \int f(\xi - \eta) h(\eta)\, \mathrm{d}\eta, \qquad \xi \in \mathbb{R}^d.$$

Define $\tilde{f}$ by $\tilde{f}(\xi) = f(-\xi)$. Let f^{*n} denote convolution of f with itself $n \geq 1$ times, and set $f^{*0} = \delta$, where δ denotes the Dirac delta function on $\mathbb{R}^d$: $\delta(\xi) = \infty$ if $\xi = 0$, $\delta(\xi) = 0$ if $\xi \in \mathbb{R}^d \setminus \{0\}$, and for any Lebesgue integrable or constant function f,

$$f(\xi) = \int \delta(\xi - \eta) f(\eta)\, \mathrm{d}\eta.$$

Accordingly we set $\delta * f = f * \delta = f$ and $\delta * \delta = \delta$. The *normalized fertility rate* is the density ϕ (with respect to the Lebesgue measure) of an offspring (in the first generation) generated by a point at 0:

$$\phi(\xi) = \gamma(\xi)/\nu, \qquad \xi \in \mathbb{R}^d.$$

Furthermore, let χ denote the mixture density of the densities ϕ^{*n} with geometric weights $(1-\nu)\nu^n$,

$$\chi(\xi) = (1-\nu)\sum_{n=0}^{\infty} \nu^n \phi^{*n}(\xi), \qquad \xi \in \mathbb{R}^d,$$

where in the trivial case $\nu = 0$ we set $\chi = \delta$. Finally, we abuse notation and write e.g. $g_0(\xi, \eta) = g_0(\xi - \eta)$ (which simply means that $g_0(\xi, \eta)$ depends only on (ξ, η) through $\xi - \eta$) for two different functions, however, it will always be clear from the context which function is used.

2.2.2. *Pair correlation function* In the sequel we assume that G_0 has pair correlation function g_0 such that $g_0(\xi, \eta) = g_0(\xi - \eta) = g_0(\eta - \xi)$ for all $\xi, \eta \in \mathbb{R}^d$.

Theorem 1. *We have that*

$$g(\xi, \eta) = g(\kappa) = g(-\kappa), \qquad g(\kappa) = \chi * \tilde{\chi} * \left[g_0 + \frac{1}{\mu_0(1-\nu)}\delta \right](\kappa) \tag{4}$$

whenever $\kappa = \xi - \eta \neq 0$.

Proof. See Appendix B.

Since $\{(\xi, \xi) : \xi \in \mathbb{R}^d\}$ is a nullset with respect to Lebesgue measure, we define arbitrary the value of $g(0)$. The term g_0 in (4) corresponds to the case where ξ and η are not in the same cluster, while the other term $\delta/(\mu_0(1-\nu))$ corresponds to the case where ξ and η are in the same cluster.

From (4) we obtain immediately the following result.

Corollary 1. *If* $g_0 = 1$ *then*

$$g(\kappa) = 1 + \frac{1}{\mu_0(1-\nu)}\chi * \tilde{\chi}(\kappa) \tag{5}$$

for all $\kappa \in \mathbb{R}^d \backslash \{0\}$.

Recall that the pair correlation function for a Poisson process is equal to one. By (5), $g > 1$, which is in agreement with the usual interpretation that this indicates aggregation of the points in X, cf. [14, 20, 21].

2.3. Bartlett spectrum

In this section we derive the Bartlett spectrum of the spatial Hawkes process X.

2.3.1. *Preliminaries* First, recall the notion of the classical Fourier transform: For Lebesgue integrable functions $f : \mathbb{R}^d \mapsto \mathbb{R}$, define the Fourier transform of f by

$$\mathcal{F}f(\omega) := \hat{f}(\omega) = \int \exp(i\ \omega \cdot \xi) f(\xi)\, \mathrm{d}\xi, \qquad \omega \in \mathbb{R}^d,$$

where $\cdot$ is the usual inner product on $\mathbb{R}^d$. The inverse Fourier transform is defined by

$$\mathcal{F}^{-1} f(\xi) := \frac{1}{(2\pi)^d} \int \exp(-i\ \omega \cdot \xi) f(\omega)\, \mathrm{d}\omega, \qquad \xi \in \mathbb{R}^d.$$

We define the Fourier transform of a real constant c as $(2\pi)^d c\delta(\cdot)$, and the inverse Fourier transform of this expression as c. Thus, if f is either Lebesgue integrable or constant, $\mathcal{F}^{-1}\hat{f} = f$, $\widehat{\bar{f}} = \bar{\hat{f}}$, and for any constant or Lebesgue integrable functions f_1 and f_2, we have that $\widehat{f_1 * f_2} = \hat{f}_1 \hat{f}_2$. Here $\bar{z}$ denotes the complex conjugate of $z \in \mathbb{C}$; for later use, let $|z|$ denote the modulus of z.

Next, recall the notion of Fourier transform of a tempered distribution: For a Borel measure m on $\mathbb{R}^d$, let

$$< m, \psi > = \int \psi(\xi)\, m(\mathrm{d}\xi), \qquad \psi \in \mathcal{S},$$

where $\mathcal{S}$ is the set of the rapid decreasing functions, see e.g. [4]. By definition, the Fourier transform of the tempered distribution $< m, \cdot >$ is the tempered distribution $< \hat{m}, \cdot >$ such that

$$< \hat{m}, \psi > = < m, \hat{\psi} >, \qquad \psi \in \mathcal{S}.$$

Third, recall that

$$c(\kappa) = \rho^2(g(\kappa) - 1) + \rho\delta(\kappa) \tag{6}$$

is the reduced covariance function of X, since by definition, for Borel sets $A, B \subseteq \mathbb{R}^d$,

$$\mathrm{cov}(X(A), X(B)) = \int \int \mathbf{1}[(\xi, \eta) \in A \times B] c(\xi - \eta)\, \mathrm{d}\xi\, \mathrm{d}\eta$$

where $X(A)$ is the number of points in $X \cap A$ (see e.g. [4]). Denote by $C(\mathrm{d}\xi) = c(\xi)\, \mathrm{d}\xi$ the reduced covariance measure of X. Now, the Bartlett spectrum of X is the Borel measure $\hat{C}$ on $\mathbb{R}^d$ defined by the tempered distribution $< \hat{C}, \cdot >$, see e.g. [4]. Thus $\hat{C}$ is closely related to g.

Fourth, as the reduced covariance function of G_0 is $c_0(\kappa) = \mu_0^2(g_0(\kappa) - 1) + \mu_0\delta(\kappa)$, for all $\psi \in \mathcal{S}$,

$$< C_0, \hat{\psi} > = \int c_0(\xi)\hat{\psi}(\xi)\,\mathrm{d}\xi$$
$$= \mu_0^2 \int g_0(\xi)\hat{\psi}(\xi)\,\mathrm{d}\xi - \mu_0^2 \int \hat{\psi}(\xi)\,\mathrm{d}\xi + \mu_0\hat{\psi}(0). \tag{7}$$

Notice that in the sense of [4], $\mu_0^2 g_0(\kappa) + \mu_0\delta(\kappa)$ is the density of the reduced second moment measure of G_0. Therefore, by Proposition 8.1.II in [4], it is the density of a positive and positive-definite measure, and so by Theorem 8.6.III in [4], there exists a locally finite Borel measure α_0 on $\mathbb{R}^d$ such that

$$\int g_0(\xi)\hat{\psi}(\xi)\,\mathrm{d}\xi = \int \psi(\xi)\,\alpha_0(\mathrm{d}\xi), \quad \psi \in \mathcal{S}. \tag{8}$$

Consequently, by (7) and (8),

$$< C_0, \hat{\psi} > = \mu_0^2 \int \psi(\xi)\,\alpha_0(\mathrm{d}\xi) - \mu_0^2 \int \hat{\psi}(\xi)\,\mathrm{d}\xi + \mu_0\hat{\psi}(0). \tag{9}$$

A straightforward computation shows that the right side in (9) is equal to

$$\int \psi(\xi)\left[\mu_0^2(\alpha_0(\mathrm{d}\xi) - (2\pi)^d\delta(\xi)\,\mathrm{d}\xi) + \mu_0\,\mathrm{d}\xi\right]$$

and so the Bartlett spectrum of G_0 is

$$\hat{C}_0(\mathrm{d}\xi) = \mu_0^2\left[\alpha_0(\mathrm{d}\xi) - (2\pi)^d\delta(\xi)\,\mathrm{d}\xi\right] + \mu_0\,\mathrm{d}\xi. \tag{10}$$

In many applications, g_0 is of the form $g_0 = 1 + f_0$, where f_0 is Lebesgue integrable, in which case $\hat{C}_0$ is absolutely continuous with respect to Lebesgue measure, with density

$$\hat{c_0}(\xi) = \mu_0\left(\mu_0\hat{f}_0(\xi) + 1\right). \tag{11}$$

We refer to $\hat{c_0}$ as the spectral density of G_0.

2.3.2. *The Bartlett spectrum as a corollary to the result for the pair correlation function*
The proof in Appendix C of the following corollary shows how the expression of the Bartlett spectrum follows from Theorem 1 for the pair correlation function.

Corollary 2. *We have that*

$$\hat{C}(\mathrm{d}\xi) = \frac{1}{|1 - \hat{\gamma}(\xi)|^2}\left[\hat{C}_0(\mathrm{d}\xi) + (\mu_0\nu/(1-\nu))\,\mathrm{d}\xi\right] \tag{12}$$

where $\hat{C}_0$ *is given by* (10).

Corollary 2 is in accordance with the result in [1] (see Theorem 5.1 therein). When G_0 has spectral density (11), the spatial Hawkes process has spectral density

$$\hat{c}(\xi) = \frac{\mu_0}{|1-\hat{\gamma}(\xi)|^2}\left[\mu_0 \hat{f}_0(\xi) + \frac{1}{1-\nu}\right] \tag{13}$$

(that is $\hat{C}(B) = \int_B \hat{c}(\xi)\,\mathrm{d}\xi$ for Borel sets $B \subseteq \mathbb{R}^d$).

3. Examples and discussion

3.1. Gaussian fertility rates

Consider a two-dimensional radially symmetric Gaussian fertility rate, i.e. ϕ is the density of $N_2(0, \sigma^2 I)$ where $\sigma^2 > 0$ is the variance. Then $\phi^{*n}(\xi) = \phi^{*n}(r)$ depends only on $r = \|\xi\|$ and is the density of $N_2(0, n\sigma^2 I)$, $n \geq 1$. Further, $\chi = \tilde{\chi}$ and a straightforward calculation shows that

$$\chi * \tilde{\chi}(r) = (1-\nu)^2 \sum_{n=0}^{\infty} (n+1)\nu^n \phi^{*n}(r).$$

Let first $g_0 = 1$ and consider $(g(r)-1)\mu_0$ which by (5) does not depend on the parameter μ_0:

$$(g(r)-1)\mu_0 = (1-\nu) \sum_{n=1}^{\infty} (n+1)\nu^n \phi^{*n}(r), \quad r > 0,$$

which we can calculate by numerical methods using e.g. Maple. The left plot in Figure 2 shows $(g(r)-1)\mu_0$. The effect of increasing σ^2 from 1 to 4 and ν from 0.5 to 0.9 is clearly visible. For comparison we have also shown $(g(r)-1)\mu_0 = \phi^{*2}(r)$ for a Neyman-Scott (or modified Thomas) process when $(\sigma^2, \nu) = (1, 0.9)$ (i.e. when g is the pair correlation function for offspring of the first generation). This curve has to some extent a similar shape as for a spatial Hawkes process, though it is much below the curve for the spatial Hawkes process with $(\sigma^2, \nu) = (1, 0.9)$.

Next, the upper curve in Figure 2 is $(g(r)-1)\mu_0$ when the immigrant pair correlation function is that of the Thomas process above, i.e. when $g_0 = 1 + \phi^{*2}/\mu_0$ and

$$(g(r)-1)\mu_0 = (1-\nu) \sum_{n=1}^{\infty} (n+1)\nu^n \left[\phi^{*n}(r) + (1-\nu)\phi^{*(n+2)}(r)\right], \quad r > 0,$$

and where again $(\sigma^2, \nu) = (1, 0.9)$. This curve is only slightly above the corresponding curve where $g_0 = 1$.

The right plot in Figure 2 shows spectral densities (13) on a logarithmic scale when still $(\sigma^2, \nu) = (1, 0.9)$ and $\mu_0 = 1$ or $\mu_0 = 10$ and $g_0 = 1$ or $g_0 = 1 + \phi^{*2}/\mu_0$. For a low immigrant intensity, it is hard to distinguish between the two cases of spectral densities.

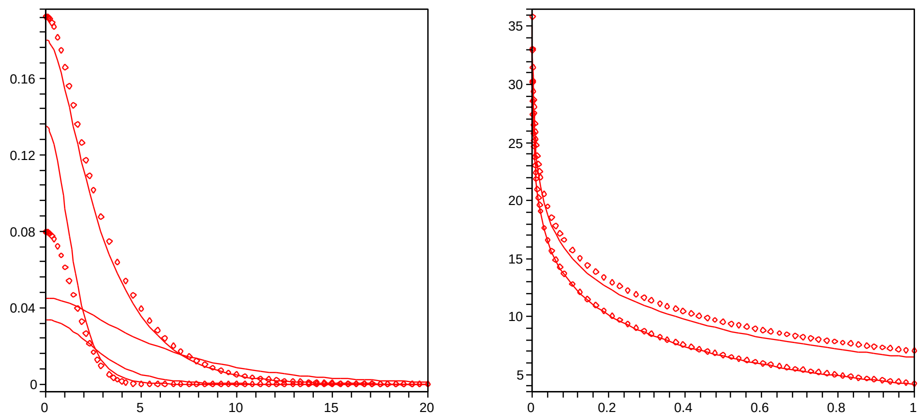

FIGURE 2: *Left:* Transformed pair correlation functions $(g(r)-1)\mu_0$ with $d = 2$ and a radially symmetric Gaussian density ϕ with variance σ^2. Full lines: for spatial Hawkes processes with $(\sigma^2, \nu) = (1, 0.9), (1, 0.5), (4, 0.9), (4, 0.5)$ (from top to bottom at $r = 0$) and $g_0 = 1$. Dotted lines: for a modified Thomas process (bottom) and for a spatial Hawkes process when g_0 is the pair correlation of the modified Thomas process (top), with $(\sigma^2, \nu) = (1, 0.9)$ in both cases. *Right:* Log spectral densities when $(\sigma^2, \nu) = (1, 0.9)$ and $\mu_0 = 1$ (bottom) or $\mu_0 = 10$ (top) and $g_0 = 1$ (full lines) or $g_0 = 1 + \phi^{*2}/\mu_0$ (dotted lines).

3.2. Summary statistics

The pair correlation function is frequently used in spatial statistics as a characteristic of the second order properties of a spatial point process, and from g we obtain Ripley's K-function [3, 5, 14, 19, 20, 21]. The use of the Bartlett spectrum has played a minor role in spatial statistics [15], possibly because g is easier to interpret. However, in light of the much simpler expression (13) compared to (4), using the spectral density as a second order characteristic for spatial Hawkes processes seems appealing. The empty space function F, the nearest-neighbour distribution function G, and the related J

function, which are all widely used summary statistics (see [14] and the references therein), seem intractable for spatial Hawkes processes. If G_0 is Poisson, then by [20], since X is a Poisson cluster process, $J \leq 1$.

3.3. Unpredictable marks

The definition in Section 1 extends immediately to a marked Hawkes process with points ξ which are still either immigrants or offspring and unpredictable marks Z_ξ: The immigrant process is as before and offspring are generated in a similar way except that the fertility rate $\xi \to \gamma(\xi - \eta, Z_\eta)$ associated to a marked point (η, Z_η) may depend on the mark. Furthermore, the marks are unpredictable in the sense that each mark Z_ξ follows the same probability distribution, which is independent of the point ξ and the previous history of the marked point process. It is easily seen that this property of the marks implies that the results in Section 2 are still valid for the spatial point process given by the points of the marked Hawkes process, provided we replace $\gamma(\cdot)$ by $\mathrm{E}\gamma(\cdot, Z)$ where the expectation is with respect to a generic mark Z. The Bartlett spectrum so obtained agrees with Theorem 5.1 in [1].

Appendix A

Proof of Proposition 1: It follows immediately from the definition of a spatial Hawkes process that stationarity of G_0 implies stationarity of $(G_0, G_1, \ldots)$ and hence also stationarity of $X = \cup_0^\infty G_n$. We show by induction that, for any $n \geq 0$, G_n has intensity $\rho_n = \mu_0 \nu^n$. The case $n = 0$ is satisfied by assumption. Using first (1), next Campbell's theorem (see e.g. [20]), and third the induction hypothesis, we obtain for $n \geq 1$,

$$\rho_n = \mathrm{E}\lambda_n(\xi) = \rho_{n-1} \int \gamma(\xi - \eta)\, \mathrm{d}\eta = \mu_0 \nu^{n-1} \nu = \mu_0 \nu^n.$$

Hence $X = \cup_0^\infty G_n$ has intensity $\rho = \sum_{n \geq 0} \rho_n = \mu_0/(1 - \nu)$.

Appendix B

For integers $m, n \geq 0$ and Borel sets $C \subseteq \mathbb{R}^d \times \mathbb{R}^d$, define the measure

$$\alpha_{m,n}(C) = \mathrm{E} \sum_{\xi \in G_m, \eta \in G_n :\, \xi \neq \eta} \mathbf{1}[(\xi, \eta) \in C]$$

where $\mathbf{1}[\cdot]$ is the indicator function. In the terminology of [14], $\alpha_{m,n}$ is the cross moment measure of (G_m, G_n) if $m \neq n$, and $\alpha_{n,n}$ is the second order factorial moment measure of G_n. Since the point processes $G_0, G_1, \ldots$ are almost surely pairwise disjoint, the second order factorial moment measure of $X = \cup_0^\infty G_n$ is given by

$$\alpha^{(2)}(C) = \sum_{m,n\geq 0} \alpha_{m,n}(C). \tag{14}$$

If $\alpha^{(2)}$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d \times \mathbb{R}^d$ with density $\rho^{(2)}$, then $\rho^{(2)}$ is called the second order product density, and the pair correlation function of X is given by $g(\xi, \eta) = \rho^{(2)}(\xi, \eta)/\rho^2$. By (14), if $\alpha_{m,n}$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d \times \mathbb{R}^d$ with density $\rho_{m,n}$, we can take $\rho^{(2)} = \sum_{m,n} \rho_{m,n}$. As in Theorem 1 we assume that G_0 has pair correlation function $g_0(\xi, \eta) = g_0(\xi - \eta) = g_0(\eta - \xi)$, and write $\rho_0^{(2)}(\xi) = g_0(\xi)\mu_0^2$ for its second order product density. We show in Lemma 1 below that $\rho_{m,n}$ exists. In particular, $g_n = \rho_{n,n}/\rho_n^2$ is the pair correlation function of G_n, where $\rho_n = \mu_0 \nu^n$, cf. Proposition 1.

Lemma 1. *We have that $\alpha_{m,n}$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d \times \mathbb{R}^d$, and for all $m, n \geq 0$ and $\xi \neq \eta$,*

$$\rho_{m,n}(\xi, \eta) = \rho_{m,n}(\xi - \eta) = \gamma^{*m} * \tilde{\gamma}^{*n} * \rho_0^{(2)}(\xi - \eta) + \sum_{k=0}^{\min\{m,n\}} \mu_0 \nu^k \gamma^{*(m-k)} * \tilde{\gamma}^{*(n-k)}(\xi - \eta). \tag{15}$$

Proof. The result is trivially true when $m = n = 0$.

Let $m < n$, and note that (G_m, G_n) is determined by the marked point process obtained by attaching to each point $\xi \in G_m$ a mark M_ξ given by the point process of all those points $\eta - \xi$ such that $\eta \in G_n$ and η is an offspring generated by ξ in $n - m$ steps. The point processes M_ξ, $\xi \in G_m$, are i.i.d. and independent of $G_0, \ldots, G_m$. Furthermore, if $m > 0$, conditional on $G_0, \ldots, G_{m-1}$, we have that G_m has intensity function λ_m. Hence by Campbell's theorem,

$$\mathrm{E}\left[\sum_{\xi \in G_m} f(\xi, M_\xi) \,\middle|\, G_0, \ldots, G_{m-1}\right] = \mathrm{E}\left[\int \lambda_m(\xi) f(\xi, M_\xi)\,\mathrm{d}\xi \,\middle|\, G_0, \ldots, G_{m-1}\right] \tag{16}$$

for non-negative measurable functions f, where we set $\lambda_0 = \mu_0$ and condition on nothing if $m = 0$. Thereby, since with probability one the translated point processes

$M_\xi + \xi,\ \xi \in G_m$, are pairwise disjoint and their union is equal to G_n,

$$\alpha_{m,n}(C) = \mathrm{E}\int \lambda_m(\xi) \sum_{\eta\in G_n} \mathbf{1}[(\xi,\eta)\in C]\,\mathrm{d}\xi = \mathrm{E}\int \lambda_n(\eta)\int \lambda_m(\xi)\mathbf{1}[(\xi,\eta)\in C]\,\mathrm{d}\xi\,\mathrm{d}\eta \tag{17}$$

where to obtain the first equality we consider $\mathrm{E}[\cdots] = \mathrm{E}\mathrm{E}[\cdots|G_0,\ldots,G_{m-1}]$ and to obtain the second equality we consider $\mathrm{E}[\cdots] = \mathrm{E}\mathrm{E}[\cdots|G_0,\ldots,G_{n-1}]$.

Similarly, for $m > n$, we obtain (17). For $m = n > 0$, since G_n conditional on $G_0,\ldots,G_{n-1}$ is a Poisson process with intensity function λ_n, Slivnyak-Mecke's theorem (see e.g. [14]) and considering again $\mathrm{E}[\cdots] = \mathrm{E}\mathrm{E}[\cdots|G_0,\ldots,G_{n-1}]$ imply that

$$\alpha_{n,n}(C) = \mathrm{E}\int \lambda_n(\xi)\,\mathrm{E}\left[\sum_{\eta\in G_n} \mathbf{1}[(\xi,\eta)\in C]\,\middle|\, G_0,\ldots,G_{n-1}\right]\mathrm{d}\xi,$$

and so by Campbell's theorem we obtain (17) with $m = n$.

Therefore, by (17) and Fubini's theorem, for all $m,n \geq 0$, $\alpha_{m,n}$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d \times \mathbb{R}^d$, with density

$$\rho_{m,n}(\xi,\eta) = \mathrm{E}[\lambda_m(\xi)\lambda_n(\eta)]. \tag{18}$$

Consequently,

$$\rho_{m,n}(\xi,\eta) = \rho_{n,m}(\eta,\xi) \tag{19}$$

for all $\xi,\eta \in \mathbb{R}^d$ and $m,n \geq 0$.

In the remainder of this proof, let $\xi \neq \eta$.

We now prove by induction that (15) is satisfied when $m = n$. For $m = n = 0$, this is trivially satisfied. Assume that it is satisfied for a fixed $m = n \geq 0$. By (1) and (18),

$$\rho_{n+1,n+1}(\xi,\eta) = \mathrm{E}\sum_{x_1,x_2\in G_n:\,x_1\neq x_2} \gamma(\xi - x_1)\gamma(\eta - x_2) + \mathrm{E}\sum_{x\in G_n} \gamma(\xi - x)\gamma(\eta - x).$$

Applying the definition of $\rho_{n,n}$ in the first term and Campbell's theorem in the second term above,

$$\rho_{n+1,n+1}(\xi,\eta) = \int\int \gamma(\xi - x_1)\gamma(\eta - x_2)\rho_{n,n}(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 + \int \gamma(\xi - x)\gamma(\eta - x)\rho_n\,\mathrm{d}x$$

which after a straightforward computation reduces to

$$\rho_{n+1,n+1}(\xi,\eta) = \gamma * \tilde{\gamma} * (\rho_n\delta + \rho_{n,n})(\xi - \eta).$$

Therefore, by the induction hypothesis,

$$\begin{aligned}\rho_{n+1,n+1}(\xi,\eta) &= \mu_0\nu^n\gamma*\tilde{\gamma}(\xi-\eta)+\gamma^{*(n+1)}*\tilde{\gamma}^{*(n+1)}*\rho_0^{(2)}(\xi-\eta)\\ &+ \sum_{k=0}^{n-1}\mu_0\nu^k\gamma^{*(n+1-k)}*\tilde{\gamma}^{*(n+1-k)}(\xi-\eta)\end{aligned}$$

whereby the induction proof is completed, since $\delta(\xi-\eta)=0$. Note that from this and (18) we obtain that

$$\rho_{n,n}(\xi,\eta)=\rho_{n,n}(\xi-\eta)=\rho_{n,n}(\eta-\xi). \tag{20}$$

Next, let $m<n$. Then

$$\begin{aligned}\rho_{m,n}(\xi,\eta) &= \mathrm{E}[\lambda_m(\xi)\lambda_n(\eta)]\\ &= \mathrm{E}\sum_{x_1\in G_{n-1}}\gamma(\eta-x_1)\lambda_m(\xi)\\ &= \mathrm{E}\mathrm{E}\left[\sum_{x_1\in G_{n-1}}\gamma(\eta-x_1)\lambda_m(\xi)\,\middle|\,G_0,\dots,G_{n-2}\right]\\ &= \mathrm{E}\int\gamma(\eta-x_1)\lambda_m(\xi)\lambda_{n-1}(x_1)\,\mathrm{d}x_1\\ &= \int\gamma(\eta-x_1)\rho_{m,n-1}(\xi,x_1)\,\mathrm{d}x_1\end{aligned}$$

where we have used (18) in the first and last equalities, (1) in the second equality, and Campbell's theorem in the fourth equality. Iterating this calculation we obtain

$$\begin{aligned}\rho_{m,n}(\xi,\eta) &= \int\gamma(\eta-x_1)\int\gamma(x_1-x_2)\cdots\int\gamma(x_{n-m-1}-x_{n-m})\rho_{m,m}(\xi,x_{n-m})\\ &\quad \mathrm{d}x_{n-m}\cdots\mathrm{d}x_2\,\mathrm{d}x_1\\ &= \tilde{\gamma}^{*(n-m)}*\rho_{m,m}(\xi-\eta)\end{aligned}$$

using (20). Since $\rho_{m,m}$ satisfies (15), we obtain that $\rho_{m,n}$ satisfies (15).

Third, let $m>n$. Observe that if $h=f_1*f_2$ then $\tilde{h}=\tilde{f}_1*\tilde{f}_2$. Combining this with (15) (for the case so far verified), (19), and the fact that $\rho_0^{(2)}(\cdot)$ is symmetric, we obtain that

$$\begin{aligned}\rho_{m,n}(\xi,\eta) &= \gamma^{*n}*\tilde{\gamma}^{*m}*\rho_0^{(2)}(\eta-\xi)+\sum_{k=0}^{n}\mu_0\nu^k\gamma^{*(n-k)}*\tilde{\gamma}^{*(m-k)}(\eta-\xi)\\ &= \gamma^{*m}*\tilde{\gamma}^{*n}*\rho_0^{(2)}(\xi-\eta)+\sum_{k=0}^{n}\mu_0\nu^k\gamma^{*(m-k)}*\tilde{\gamma}^{*(n-k)}(\xi-\eta).\end{aligned}$$

Thereby (15) is verified for all $m,n\geq 0$.

Proof of Theorem 1: By (15), $\rho^{(2)}(\xi,\eta)=\rho^{(2)}(\xi-\eta)$ whenever $\xi\neq\eta$, where for $\kappa\neq 0$,

$$\begin{aligned}\rho^{(2)}(\kappa) &= \sum_{m\geq 0,\, n\geq 0}\rho_{m,n}(\kappa)\\ &= \left(\sum_{m\geq 0}\gamma^{*m}\right)*\left(\sum_{n\geq 0}\tilde{\gamma}^{*n}\right)*\rho_0^{(2)}(\kappa)\\ &+ \sum_{k,m,n:\, m\geq k,\, n\geq k,\, k\geq 0}\mu_0\nu^k\gamma^{*(m-k)}*\tilde{\gamma}^{*(n-k)}(\kappa)\\ &= \left(\sum_{m\geq 0}\gamma^{*m}\right)*\left(\sum_{n\geq 0}\tilde{\gamma}^{*n}\right)*\left(\rho_0^{(2)}+\rho\delta\right)(\kappa) \qquad (21)\end{aligned}$$

using (3). By (3) and (21), $g(\xi,\eta)=g(\xi-\eta)=\rho^{(2)}(\xi-\eta)/\rho^2$ is easily seen to be given by the last expression in (4). Finally, it follows that $g(\cdot)$ is symmetric.

Appendix C

Proof of Corollary 2: Consider the Borel measure on $\mathbb{R}^d$ $G(\mathrm{d}\xi)=g(\xi)\,\mathrm{d}\xi$. The linear functional

$$< G,\psi > = \int g(\xi)\psi(\xi)\,\mathrm{d}\xi, \qquad \psi\in\mathcal{S}$$

is a tempered distribution. Indeed, by (4),

$$< G,\psi > = \int \chi*\tilde{\chi}*g_0(\xi)\psi(\xi)\,\mathrm{d}\xi+\frac{1}{\mu_0(1-\nu)}\int\chi*\tilde{\chi}(\xi)\psi(\xi)\,\mathrm{d}\xi, \qquad \psi\in\mathcal{S}$$

and both the functions $\chi*\tilde{\chi}*g_0$ and $\chi*\tilde{\chi}$ are locally integrable (the first one is the density of a p.p.d. measure, the second one is a probability density).

Define the function

$$\beta=(1-\nu)\sum_{n=1}^{\infty}\nu^n\phi^{*n}$$

and note that $\chi=(1-\nu)\delta+\beta$. From (4), for all $\xi\in\mathbb{R}^d\setminus\{0\}$,

$$g(\xi)=\chi*\tilde{\chi}*g_0(\xi)+\frac{1}{\mu_0}[\tilde{\beta}(\xi)+\beta(\xi)+(1-\nu)^{-1}\beta*\tilde{\beta}(\xi)].$$

Therefore,

$$\begin{aligned}< G,\hat{\psi} > &= \int\chi*\tilde{\chi}*g_0(\xi)\hat{\psi}(\xi)\,\mathrm{d}\xi+\frac{1}{\mu_0}\int(\tilde{\beta}(\xi)+\beta(\xi))\hat{\psi}(\xi)\,\mathrm{d}\xi\\ &+\frac{1}{\mu_0(1-\nu)}\int(\beta*\tilde{\beta})(\xi)\hat{\psi}(\xi)\,\mathrm{d}\xi, \qquad \psi\in\mathcal{S}. \qquad (22)\end{aligned}$$

Since the functions $\tilde{\beta}$ and β are Lebesgue integrable,

$$\int (\tilde{\beta}(\xi) + \beta(\xi))\hat{\psi}(\xi)\,\mathrm{d}\xi = \int \psi(\xi)(\hat{\tilde{\beta}}(\xi) + \hat{\beta}(\xi))\,\mathrm{d}\xi$$
$$= (1-\nu)\int \psi(\xi)\left(\frac{1}{1-\overline{\hat{\gamma}(\xi)}} + \frac{1}{1-\hat{\gamma}(\xi)} - 2\right)\mathrm{d}\xi \tag{23}$$

and

$$\int \beta * \tilde{\beta}(\xi)\hat{\psi}(\xi)\,\mathrm{d}\xi = \int \psi(\xi)\hat{\beta}(\xi)\hat{\tilde{\beta}}(\xi))\,\mathrm{d}\xi$$
$$= (1-\nu)^2\int \psi(\xi)\left(\frac{1}{1-\hat{\gamma}(\xi)} - 1\right)\left(\frac{1}{1-\overline{\hat{\gamma}(\xi)}} - 1\right)\mathrm{d}\xi. \tag{24}$$

We notice that

$$\int \chi * \tilde{\chi} * g_0(\xi)\hat{\psi}(\xi)\,\mathrm{d}\xi = \int\int \chi * \tilde{\chi}(\xi-\eta)g_0(\eta)\hat{\psi}(\xi)\,\mathrm{d}\eta\,\mathrm{d}\xi$$
$$= \int g_0(\eta)\,\mathrm{d}\eta \int \chi * \tilde{\chi}(\xi-\eta)\hat{\psi}(\xi)\,\mathrm{d}\xi \tag{25}$$
$$= \int \chi * \tilde{\chi} * \hat{\tilde{\psi}}(-\eta)g_0(\eta)\,\mathrm{d}\eta. \tag{26}$$

In (25) we can exchange the order of integration, since the function $(\xi,\eta) \mapsto \chi * \tilde{\chi}(\xi - \eta)g_0(\eta)\hat{\psi}(\xi)$ is Lebesgue integrable. Indeed, since g_0 is the density of a p.p.d. measure, $g_0 \le g_0(0)$, and the function $(\xi,\eta) \mapsto \chi * \tilde{\chi}(\xi-\eta)\hat{\psi}(\xi)$ is Lebesgue integrable, since $\chi * \tilde{\chi}$ is a probability density and $\hat{\psi}$ is a rapid decreasing function. Notice also that the function $\xi \mapsto \chi * \tilde{\chi} * \widehat{\tilde{\psi}}(-\xi)$ is Lebesgue integrable and

$$\mathcal{F}^{-1}\chi * \tilde{\chi} * \widehat{\tilde{\psi}}(-\xi) = \frac{1}{(2\pi)^d}\mathcal{F}\chi * \tilde{\chi} * \widehat{\tilde{\psi}}(\xi)$$
$$= \hat{\chi}(\xi)\hat{\tilde{\chi}}(\xi)\psi(\xi) = \frac{(1-\nu)^2\psi(\xi)}{|1-\hat{\gamma}(\xi)|^2}. \tag{27}$$

By (27), the function $\xi \mapsto \mathcal{F}^{-1}\chi * \tilde{\chi} * \widehat{\tilde{\psi}}(-\xi)$ is Lebesgue integrable. Indeed,

$$|\mathcal{F}^{-1}\chi * \tilde{\chi} * \widehat{\tilde{\psi}}(-\xi)| = \frac{(1-\nu)^2|\psi(\xi)|}{|1-\hat{\gamma}(\xi)|^2} \le \frac{(1-\nu)^2|\psi(\xi)|}{(1-|\hat{\gamma}(\xi)|)^2} \le |\psi(\xi)|$$

and ψ is Lebesgue integrable as it is rapid decreasing. Therefore, the Fourier transform of $\xi \mapsto \mathcal{F}^{-1}\chi * \tilde{\chi} * \widehat{\tilde{\psi}}(-\xi)$ is well-defined, and by taking the Fourier transform in (27), we obtain that

$$\chi * \tilde{\chi} * \widehat{\tilde{\psi}}(-\xi) = (1-\nu)^2\mathcal{F}\frac{\psi(\xi)}{|1-\hat{\gamma}(\xi)|^2}. \tag{28}$$

Thus, by (26) and (28),

$$\int \chi * \tilde{\chi} * g_0(\xi)\hat{\psi}(\xi)\,\mathrm{d}\xi = (1-\nu)^2 \int g_0(\xi)\mathcal{F}\frac{\psi(\xi)}{|1-\hat{\gamma}(\xi)|^2}\,\mathrm{d}\xi. \tag{29}$$

By applying the extended Parseval relation, see equation 8.6.10 in [4], it follows that

$$\int g_0(\xi)\mathcal{F}\frac{\psi(\xi)}{|1-\hat{\gamma}(\xi)|^2}\,\mathrm{d}\xi = \int \frac{\psi(\xi)}{|1-\hat{\gamma}(\xi)|^2}\,\alpha_0(\mathrm{d}\xi), \tag{30}$$

where α_0 is the Borel measure defined by (8). Therefore, by (29) and (30),

$$\int \chi * \tilde{\chi} * g_0(\xi)\hat{\psi}(\xi)\,\mathrm{d}\xi = (1-\nu)^2 \int \frac{\psi(\xi)}{|1-\hat{\gamma}(\xi)|^2}\,\alpha_0(\mathrm{d}\xi). \tag{31}$$

Combining (22), (23), (24) and (31), we obtain that

$$< G, \hat{\psi} >= \int \psi(\xi)\left[\frac{(1-\nu)^2}{|1-\hat{\gamma}(\xi)|^2}\,\alpha_0(\mathrm{d}\xi) + \frac{1-\nu}{\mu_0|1-\hat{\gamma}(\xi)|^2}\,\mathrm{d}\xi - \frac{1-\nu}{\mu_0}\,\mathrm{d}\xi\right] \tag{32}$$

Arguing as for (7) we have, for all $\psi \in \mathcal{S}$,

$$< C, \hat{\psi} > = \rho^2 \int g(\xi)\hat{\psi}(\xi)\,\mathrm{d}\xi - \rho^2 \int \hat{\psi}(\xi)\,\mathrm{d}\xi + \rho\hat{\psi}(0). \tag{33}$$

Therefore, by (32) and (33)

$$\begin{aligned}
< C, \hat{\psi} > &= \rho^2 \int \psi(\xi)\left[\frac{(1-\nu)^2}{|1-\hat{\gamma}(\xi)|^2}\,\alpha_0(\mathrm{d}\xi) + \frac{1-\nu}{\mu_0|1-\hat{\gamma}(\xi)|^2}\,\mathrm{d}\xi - \frac{1-\nu}{\mu_0}\,\mathrm{d}\xi\right] \\
&\quad - \rho^2 \int \hat{\psi}(\xi)\,\mathrm{d}\xi + \rho\hat{\psi}(0) \\
&= \int \psi(\xi)\left[\frac{\mu_0^2}{|1-\hat{\gamma}(\xi)|^2}\,\alpha_0(\mathrm{d}\xi) + \frac{\rho}{|1-\hat{\gamma}(\xi)|^2}\,\mathrm{d}\xi - (2\pi)^d\rho^2\delta(\xi)\,\mathrm{d}\xi\right] \\
&= \int \psi(\xi)\left[\frac{\mu_0^2}{|1-\hat{\gamma}(\xi)|^2}\,\alpha_0(\mathrm{d}\xi) + \frac{\rho}{|1-\hat{\gamma}(\xi)|^2}\,\mathrm{d}\xi - \frac{(2\pi)^d\mu_0^2}{|1-\hat{\gamma}(\xi)|^2}\,\delta(\xi)\,\mathrm{d}\xi\right] \\
&= \int \psi(\xi)\left[\frac{\mu_0^2}{|1-\hat{\gamma}(\xi)|^2}\,[\alpha_0(\mathrm{d}\xi) - (2\pi)^d\delta(\xi)\,\mathrm{d}\xi] + \frac{\mu_0}{|1-\hat{\gamma}(\xi)|^2}\,\mathrm{d}\xi\right] \\
&\quad + \int \psi(\xi)\frac{\nu\mu_0/(1-\nu)}{|1-\hat{\gamma}(\xi)|^2}\,\mathrm{d}\xi \\
&= \int \psi(\xi)\frac{\hat{C}_0(\mathrm{d}\xi)}{|1-\hat{\gamma}(\xi)|^2} + \int \psi(\xi)\frac{\nu\mu_0/(1-\nu)}{|1-\hat{\gamma}(\xi)|^2}\,\mathrm{d}\xi, \qquad (34) \\
&= < \hat{C}, \psi >
\end{aligned}$$

where the equality in (34) follows by (10).

References

[1] Brémaud, P., Massoulié, L. and Ridolfi, A. (2004). Power spectra of random spike fields and related processes. Manuscript available at http://lcavwww.epfl.ch/∼bremaud.

[2] Brémaud, P., Nappo, G. and Torrisi, G. L. (2002). Rate of convergence to equilibrium of marked Hawkes processes. *Journal of Applied Probability* **39,** 123–136.

[3] Cressie, N. A. C. (1993). *Statistics for Spatial Data* second ed. Wiley, New York.

[4] Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes, Volume I: Elementary Theory and Methods* second ed. Springer, New York.

[5] Diggle, P. J. (2003). *Statistical Analysis of Spatial Point Patterns* second ed. Arnold, London.

[6] Hawkes, A. G. (1971). Point spectra of some mutually exciting point processes. *Journal of Royal Statistical Society Series B* **33,** 438–443.

[7] Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58,** 83–90.

[8] Hawkes, A. G. (1972). Spectra of some mutually exciting point processes with associated variables. In *Stochastic Point Processes.* ed. P. A. W. Lewis. Wiley, New York, pp. 261–271.

[9] Hawkes, A. G. and Oakes, D. (1974). A cluster representation of a self-exciting process. *Journal of Applied Probability* **11,** 493–503.

[10] Kingman, J. F. C. (1977). Remarks on the spatial distribution of a reproducing population. *Journal of Applied Probability* **14,** 577–583.

[11] Møller, J. and Rasmussen, J. G. (2004). Approximate simulation of Hawkes processes. *Technical report* R-2004-28. Department of Mathematical Sciences, Aalborg University.

[12] Møller, J. and Rasmussen, J. G. (2005). Perfect simulation of Hawkes processes. *Advances in Applied Probability* **37,** to appear.

[13] Møller, J. and Torrisi, G. L. (2005). Generalised shot noise Cox processes. *Advances in Applied Probability* **37,** 48–74.

[14] Møller, J. and Waagepetersen, R. P. (2003). *Statistical Inference and Simulation for Spatial Point Processes.* Chapman & Hall, Boca Raton, Florida.

[15] Mugglestone, M. A. and Renshaw, E. (1996). A practical guide to the spectral analysis of spatial point processes. *Computational Statistics and Data Analysis* **21,** 43–65.

[16] Neyman, J. and Scott, E. L. (1958). Statistical approach to problems of cosmology. *Journal of the Royal Statistical Society Series B* **20,** 1–43.

[17] Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statististical Association* **83,** 9–27.

[18] Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics* **50,** 379–402.

[19] Ripley, B. D. (1977). Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society Series B* **39,** 172–212.

[20] Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and Its Applications* second ed. Wiley, Chichester.

[21] Stoyan, D. and Stoyan, H. (1995). *Fractals, Random Shapes and Point Fields.* Wiley, Chichester.