Lévy based Cox point processes

Gunnar Hellmund

PhD Thesis

Department of Mathematical Sciences
University of Aarhus
February 2009

To Katharina and Johanna.

Acknowledgments

I would like to thank my family for their love and support during my three years as a PhD student. Especially I want to thank my wife Lene: I am humbled by your love.

I wish to thank my supervisor Eva B. Vedel Jensen for originally encouraging me to initiate the PhD study and introducing me to the theory of point processes. I also wish to thank Rasmus Plenge Waagepetersen and Svend-Erik Graversen for their kind advise and commitment to science.

Preface

This thesis consists of three papers. The first paper *Lévy based Cox point processes* includes background material, basic definitions and other core material. The second paper *Completely random signed measures* answers some fundamental probability theoretical questions regarding Lévy bases, while the third paper *Strong mixing with a view toward spatio-temporal estimating functions* is concerned with estimation theory for (non-stationary) Cox point processes.

Before the actual papers, I provide an informal readers guide as introduction where I put emphasis on the findings I favor.

Contents

Lévy based Cox point processes has been published in Advances in Applied Probability vol. 40 nr. 3, s. 603-629

Completely random signed measures will be published in Statistics and Probability Letters (in press)

Strong mixing with a view toward spatio-temporal estimating functions will appear as a Thiele Research Report and will be submitted to Statistica Sinica for the upcoming special issue on composite likelihood methods.

A readers guide

Below, I give an informal readers guide to the three papers of the thesis. Section number etc. refer to the paper in question. References can be found immediately after the readers guide.

Paper 1: Lévy based Cox point processes

Cox point processes are point processes with random intensity, i.e. given a realization of the generating random field $\Lambda = \lambda$ the resulting point process is a Poisson point process with intensity λ.

The most widely used Cox point process models are log-Gaussian Cox processes (LGCPs) driven by log-Gaussian random fields and shot-noise Cox processes (SNCPs) driven by shot-noise random fields. The introduction of Paper 1 provides a short discussion of the literature and points out the major result that both LGCPs and SNCPs can be constructed through kernel smoothings of so-called Lévy bases.

A Lévy basis L defined on $\mathbb{R}^d$ is a stochastic process indexed by the bounded Borel subsets of $\mathbb{R}^d$ such that

- Its values on disjoint sets are independent
- The values of L are infinitely divisible
- Given any disjoint sequence of bounded Borel sets (A_n) such that $\cup A_n$ is a bounded Borel subset

$$L\left(\cup A_n\right) = \sum L\left(A_n\right), P - a.s.$$

Section 2 provides theoretical background material which also enables the reader to understand integration wrt. Lévy bases. Since the random fields driving (log) Lévy based Cox point processes are kernel smoothings of Lévy bases this is an important Section. Lemma 1 gives sufficient conditions for integrability of a kernel wrt. a Lévy basis. Necessary and sufficient conditions for integrability are given in [2], but their conditions are not as user-friendly as our conditions.

For readers unfamiliar with Cox point processes some key concepts are summarized in Section 3.

The definition of Lévy driven Cox processes is given in Section 4.1. Section 4.2 provides nth order product densities (Proposition 3), and probability generating functional (Proposition 6). A mixing result (Proposition 7) that might be interesting for inference of stationary point processes is given in Section 4.3, while Section 4.4 provides examples of Lévy driven Cox processes incl. the important SNCPs.

The definition of log Lévy driven Cox Processes is given in Section 5.1. As shown in Section 5.4.2. these include log-Gaussian Cox point processes driven by stationary Gaussian random fields - but it also includes the log shot-noise Cox processes and combinations of these two classes as discussed in Section 5.1. A traditional argument for using LGCPs is that the Gaussian random field model random events - events that make the intensity increase or decrease. Using a stationary Gaussian random field will however on average provide equal increases and decreases of the intensity. Log shot-noise random fields can be used to separate (and quantify) the amount of respectively increases and decreases in the intensity generated by random events. The use of a positive shot-noise random field in a log Lévy driven Cox process can describe the positive effect on the intensity by random events and a negative shot-noise random field can be used to describe the negative effect on the intensity by random events. (see figure 3 on page 21 of Paper 1). Section 5.2 provides details on nth order product densities while a mixing result (useful in a stationary setting) is provided by Proposition 11 in Section 5.3.

Finally in Section 6 combinations of Lévy driven Cox processes and log Lévy driven Cox processes are given. Section 7 discusses inhomogeneous (log) Lévy driven Cox processes.

In the concluding discussion in Section 8 spatio-temporal extensions are given in Section 8.2 and in Section 8.3 inference based on summary statistics is mentioned.

Paper 2: Completely random signed measures

A Lévy basis can be seen as a 'generalized' random measure. It is however not a random (signed) measure in the usual sense, since its realizations may not be (signed) Radon measures.

A question to be answered is: How fundamental is the concept of a Lévy basis and how is it related to the usual notion of random measures? These questions are answered in the paper *Completely random signed measures*.

According to Lemma 3.1 the assumption of infinitely divisible values of Lévy bases (that might seem odd at first) is a natural consequence of the independence assumption, if the values of the Lévy basis on bounded Borel sets are assumed to have finite variance or the Lévy basis is dominated by Lebesgue measure. The proof of Lemma 3.1 uses a result in [1] and a translation of the problem to a one dimensional stochastic process setting, a trick that is also used in other proofs.

The question: 'When is a Lévy basis a (signed) random measure?' is answered by Corollary 5.5 (in conjunction with Lemma 4.6 and Definition 5.4).

As a side-effect of these investigations we were able to give a characterization of completely random signed measures that extends known results for completely random measures (Theorem 4.7).

Paper 3: Strong mixing with a view toward spatio-temporal estimating functions

As described in *Lévy based Cox point processes* the Lévy based Cox processes constitute a large and flexible class of point processes driven by random fields. Most often second-order reweighted stationarity or even other kinds of non-stationary models are natural to consider in an applied situation. As noted in Section 1.1 the recent literature on Cox point processes focuses on second-order reweighted stationary processes on the form

$$\Lambda(x) = \exp(\beta \cdot z(x)) \Psi(x), \tag{1}$$

where β is a parameter vector, z is covariate information and Ψ is a stationary field.

The literature on Cox point processes defined on (rectangular) observation windows in $\mathbb{R}^2$ has so far used minimum contrast methods for summary statistics to obtain parameter estimates - most recently in [3]. A major drawback of these methods is that the summary statistics are only well-defined if the Cox point process under investigation is stationary or second-order re-weighted stationary. Instead of using summary statistics, the Bernoulli composite likelihood introduced in [4] can be used in much more general settings (see Section 2.1).

However departing from stationarity introduces difficulties when proving asymptotic results. The concept of strong mixing that is essential for consistency and asymptotic normality (Section 2.2) of parameter estimates in the non-stationary setting is introduced in Section 1.1.

For the first time in the literature on Cox point processes defined on (rectangular) observation windows in $\mathbb{R}^2$ we obtain parameter estimates for a process which is neither stationary nor second-order re-weighted stationary (see the short simulation study in Section 2.3). The fact that parameter estimates can be obtained in a non-stationary setting will make Cox point processes more attractive when modeling the dynamics of natural phenomena.

Furthermore we introduce a flexible class of shot-noise Cox processes based on the (in geostatistical context) popular Whittle-Matern family of covariance functions (Section 1.2). The flexible second-order structure of this family of stationary random shot-noise fields makes it a good choice of the field Ψ in formula (1), when Cox point processes are used to obtain information on the effect of covariates.

References

[1] S. J. Karlin and W. J. Studdun. *Tchebycheff systems: With Applications in Analysis and Statistics.* Pure and applied mathematics. New York: Wiley, 1966.

[2] R.S. Rajput and J. Rosinski. Spectral representation of infinitely divisible processes. *Probab. Theory Relat. Fields*, 82:451–487, 1989.

[3] R. Waagepetersen and Y. Guan. Two-step estimation for inhomogeneous spatial point processes. submitted, 2008.

[4] R. P. Waagepetersen. An estimating function approach to inference for inhomogeneous neyman-scott processes. *Biometrics*, 63(1):252–258, March 2007.

Lévy based Cox point processes

Gunnar Hellmund, Michaela Prokešová and Eva B. Vedel Jensen

T.N. Thiele Centre for Applied Mathematics in Natural Science
Department of Mathematical Sciences
University of Aarhus, Denmark

Abstract

In this paper, we introduce Lévy driven Cox point processes (LCPs) as Cox point processes with driving intensity function Λ defined by a kernel smoothing of a Lévy basis (an independently scattered infinitely divisible random measure). We also consider log Lévy driven Cox point processes (LLCPs) with Λ equal to the exponential of such a kernel smoothing. Special cases are shot noise Cox processes, log Gaussian Cox processes and log shot noise Cox processes. We study the theoretical properties of Lévy based Cox processes, including moment properties described by nth order product densities, mixing properties, specification of inhomogeneity and spatio-temporal extensions.

Keywords: Cox process; infinitely divisible distribution; inhomogeneity; kernel smoothing; Lévy basis; log Gaussian Cox process; mixing; product density; shot noise Cox process

AMS 2000 Subject Classification: Primary 60D05, 60G55; Secondary 60K35

1 Introduction

Cox point processes constitute one of the most important and versatile classes of point process models for clustered point patterns [12, 13, 30, 31, 33, 41]. During the last decades several new classes of Cox point process models have appeared in the literature – e.g. shot noise Cox processes defined by means of generalized gamma measures [5], log Gaussian Cox processes [9, 29] and shot noise Cox processes [27]. These models share some common properties and differ in others, depending on how the driving intensity measure of the Cox process is constructed. One of the aims of this paper is to introduce a unified framework which is able to include all the different models mentioned above thus showing them in new light, investigate their relationships and define further natural extensions of those models.

The starting point for us will be the notion of a Lévy basis L – an independently scattered infinitely divisible random measure. The short terminology of a Lévy basis has been introduced in [2, 3]. Independently scattered infinitely divisible random measures have been studied in detail in [36]. Lévy bases include Poisson random measures, mixed Poisson random measures, Gaussian random measures as well as so-called $G-$measures [5]. Thus having in mind the construction of the shot noise Cox processes the second step in defining the driving intensity Λ of the Cox process should be a kernel smoothing of the Lévy basis

$$\Lambda(\xi) = \int k(\xi, \eta) L(\mathrm{d}\eta),$$

where k is a kernel (weight) function. By this we arrive at the definition of the Lévy driven Cox processes (LCPs) – i.e. Cox processes with the random driving intensity function defined by an integral of a weight function with respect to a Lévy basis. This construction has earlier been discussed by Robert L. Wolpert under the name of Lévy moving average processes [48], see also [49, 50]. It will be shown that LCPs are, under regularity conditions, shot noise Cox processes with additional random noise.

Furthermore, it is also possible to define the driving intensity as the exponential of a kernel smoothing of a Lévy basis (now allowing for non-positive weight functions and non-positive Lévy bases) thus arriving at the log Lévy driven Cox processes (LLCPs). It will be shown that LLCPs have, under regularity conditions, a driving field of the form $\Lambda = \Lambda_1 \cdot \Lambda_2$, where Λ_1 and Λ_2 are independent, Λ_1 is a log Gaussian field and Λ_2 is a log shot noise field. The latter process may describe clustered point patterns with randomly placed empty holes.

Shot noise Cox processes, log Gaussian Cox processes and log shot noise Cox processes will appear as natural building blocks in a modelling framework for Cox processes. Different types of combinations of the building blocks (corresponding to thinning and superposition) will be discussed in the present paper.

Having defined the framework the second aim is to study the theoretical properties of Lévy based Cox processes, including moment properties described by nth order product densities, mixing properties, specification of inhomogeneity and spatio-temporal extensions.

Examples where the new models are needed do already appear in the literature. In [47], an LCP (Cox process with additional random noise) is used in the modelling of tropical rain forest. In [13, pp. 92–100], a point pattern from forestry is described by a shot noise Cox process thinned by a random field, taking unexplained large scale environmental heterogeneity into account. If this field is log Gaussian, the resulting process is one of the combinations of LCPs and LLCPs to be described in the present paper. In the very recent review paper [33], tropical rain forest is modelled by inhomogeneous shot noise Cox processes, obtained by thinning of a homogeneous shot noise Cox process with a log-linear deterministic field depending on explanatory variables. If the deterministic field is substituted with a log Gaussian field, we again arrive at a combination of an LCP and an LLCP.

The present paper is organized as follows. In Section 2 we give a short overview of the theory of Lévy bases and integration with respect to such bases. In Section 3 we recall standard results about Cox processes. In Section 4 we introduce and study the Lévy driven Cox processes and in Section 5 the log Lévy driven Cox processes. Combinations of LCPs and LLCPs are discussed in Section 6, while inhomogeneous LCPs and LLCPs are considered in Section 7. We conclude with a discussion.

2 Lévy bases

This section provides a brief overview of the general theory of Lévy bases, in particular the theory of integration with respect to Lévy bases. For a more detailed exposition, see [3, 36] and references therein.

Let R be a Borel subset of $\mathbb{R}^d$, $\mathcal{B}(\mathrm{R})$ the Borel sets contained in R, and $\mathcal{A}$ the δ-ring of bounded Borel subsets of R.

Following [36], we consider a collection of real-valued random variables $L = \{L(A), A \in \mathcal{A}\}$ with the following properties

- for every sequence $\{A_n\}$ of disjoint sets in $\mathcal{A}$, $L(A_1), \ldots, L(A_n), \ldots$ are independent random variables and $L(\cup_n A_n) = \sum_n L(A_n)$ a.s. provided $\cup_n A_n \in \mathcal{A}$,
- for every A in $\mathcal{A}$, $L(A)$ is infinitely divisible.

If L has these properties, L will be called a *Lévy basis*, cf. [3]. If $L(A) \geq 0$ for all $A \in \mathcal{A}$, L is called a non-negative Lévy basis.

For a random variable X, the logarithm of the characteristic function $\log \mathsf{E}(\mathrm{e}^{ivX})$ will be called the *cumulant function* and will be denoted by $C(v, X)$. This notation has been used in the paper [3] where the terminology of a Lévy basis was introduced. When L is a Lévy basis, the cumulant function of $L(A)$ can by the Lévy-Khintchine representation be written as

$$C(v, L(A)) = iva(A) - \frac{1}{2}v^2 b(A) + \int_{\mathbb{R}} (\mathrm{e}^{ivr} - 1 - ivr\mathbf{1}_{[-1,1]}(r))\, U(\,\mathrm{d}r, A), \quad (1)$$

where a is a $\sigma-$additive set function on $\mathcal{A}$, b is a measure on $\mathcal{B}(\mathrm{R})$, $U(\,\mathrm{d}r, A)$ is a measure on $\mathcal{B}(\mathrm{R})$ for fixed $\mathrm{d}r$ and a Lévy measure on $\mathcal{B}(\mathbb{R})$ for each fixed $A \in \mathcal{B}(\mathrm{R})$, i.e. $U(\{0\}, A) = 0$ and $\int_{\mathbb{R}} (1 \wedge r^2)\, U(\,\mathrm{d}r, A) < \infty$, where $\wedge$ denotes minimum. In fact U is a measure on $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathrm{R})$, cf. [36, Lemma 2.3]. This measure is referred to as the *generalized Lévy measure* and L is said to have *characteristic triplet* (a, b, U). If $b = 0$ then L is called a Lévy jump basis, if $U = 0$ then L is a Gaussian basis, see the examples below. A general Lévy basis L can always be written as a sum of a Gaussian basis and an independent Lévy jump basis. Note that the term $iva(A)$ corresponds to a nonrandom shift of the values of L. The nonrandom shift may be included in the Gaussian component or in the jump component of the Lévy basis or may be shared amongst them.

Let $|a| = a^+ + a^-$. Then, there exists a unique non-negative measure μ on $\mathcal{B}(\mathrm{R})$, satisfying

$$\mu(A) = |a|(A) + b(A) + \int_{\mathbb{R}} (1 \wedge r^2) U(\,\mathrm{d}r, A),$$

for $A \in \mathcal{A}$, cf. [36, Proposition 2.1, Definition 2.2]. We will call μ the *control measure*. In [36, Lemma 2.3] it has been shown that the generalized Lévy measure U factorizes as

$$U(\,\mathrm{d}r, \,\mathrm{d}\eta) = V(\,\mathrm{d}r, \eta)\mu(\,\mathrm{d}\eta), \quad (2)$$

where $V(\,\mathrm{d}r, \eta)$ is a Lévy measure for fixed η. Moreover a and b are absolutely continuous with respect to μ, i.e.

$$a(\,\mathrm{d}\eta) = \tilde{a}(\eta)\mu(\,\mathrm{d}\eta), \qquad b(\,\mathrm{d}\eta) = \tilde{b}(\eta)\mu(\,\mathrm{d}\eta), \quad (3)$$

and obviously $|\tilde{a}|, \tilde{b} \le 1$ μ a.s. .

Let $L'(\eta)$ be a random variable with the cumulant function

$$C(v, L'(\eta)) = iv\tilde{a}(\eta) - \frac{1}{2}v^2\tilde{b}(\eta) + \int_{\mathbb{R}} (\mathrm{e}^{ivr} - 1 - ivr\mathbf{1}_{[-1,1]}(r))\, V(\mathrm{d}r, \eta). \quad (4)$$

Then, we get the representation

$$C(v, L(A)) = \int_A C(v, L'(\eta))\, \mu(\mathrm{d}\eta). \quad (5)$$

The random variables $L'(\eta)$ will play an important role in the following and will be called *spot variables*. Note that $L'(\eta)$ characterizes the behaviour of L at location η. For later use, note that if $\mathsf{E}(L'(\eta))$ and $\mathsf{Var}(L'(\eta))$ exist, then

$$\begin{aligned} \mathsf{E}(L'(\eta)) &= \tilde{a}(\eta) + \int_{[-1,1]^C} rV(\mathrm{d}r, \eta), \\ \mathsf{Var}(L'(\eta)) &= \tilde{b}(\eta) + \int_{\mathbb{R}} r^2 V(\mathrm{d}r, \eta). \end{aligned}$$

By (1) – (3) it is no restriction if we for modelling purposes only consider Lévy bases with characteristic triplet (a, b, U) of the form

$$a(\mathrm{d}\eta) = \tilde{a}_\nu(\eta)\nu(\mathrm{d}\eta) \quad (6)$$
$$b(\mathrm{d}\eta) = \tilde{b}_\nu(\eta)\nu(\mathrm{d}\eta) \quad (7)$$
$$U(\mathrm{d}r, \mathrm{d}\eta) = V_\nu(\mathrm{d}r, \eta)\nu(\mathrm{d}\eta) \quad (8)$$

where ν is a non-negative measure on $\mathcal{B}(\mathrm{R})$, $a_\nu : \mathrm{R} \to \mathbb{R}$ and $b_\nu : \mathrm{R} \to [0, \infty)$ are measurable functions and $V_\nu(\mathrm{d}r, \eta)$ is a Lévy measure for fixed η. The random variable satisfying (5) with μ replaced by ν will be denoted by $L'_\nu(\eta)$. For simplicity, we write $L'_\mu(\eta) = L'(\eta)$, $\tilde{a}_\mu = \tilde{a}$, $\tilde{b}_\mu = \tilde{b}$ and $V_\mu(\mathrm{d}r, \eta) = V(\mathrm{d}r, \eta)$. If $V_\nu(\cdot, \eta)$, $\tilde{a}_\nu(\eta)$ and $\tilde{b}_\nu(\eta)$ do not depend on η neither does the distribution of $L'_\nu(\eta)$ and the Lévy basis L is called *$\nu-$factorizable*. If moreover the measure ν is proportional to the Lebesgue measure, L is called *homogeneous* and all the finite dimensional distributions of L are translation invariant.

Let us now consider integration of a measurable function f on $(\mathrm{R}, \mathcal{B}(\mathrm{R}))$ with respect to a Lévy basis L. The function f is said to be integrable with respect to L, cf. [36], if there exists a sequence of simple functions f_n converging to f $\mu-$a.e. and such that $\int_A f_n \,\mathrm{d}L$ converges in probability, as $n \to \infty$, for all $A \in \mathcal{B}(\mathrm{R})$. The limit is denoted $\int_A f \,\mathrm{d}L$. The integral of a simple function $f_n = \sum_{j=1}^k x_j \mathbf{1}_{A_j^n}$ with respect to L is defined in the obvious manner

$$\int_A f_n \,\mathrm{d}L = \sum_{j=1}^k x_j L(A \cap A_j^n).$$

The following lemma gives conditions for integrability and characterizes the distribution of the resulting integral.

Lemma 1 *Let f be a measurable function on $(\mathrm{R}, \mathcal{B}(\mathrm{R}))$ and L a Lévy basis on R with characteristic triplet (a, b, U). If the following conditions* *(i)* $\int_{\mathrm{R}} |f(\eta)|\, |a|(\,\mathrm{d}\eta) < \infty$

(ii) $\int_{\mathrm{R}} f(\eta)^2\, b(\,\mathrm{d}\eta) < \infty$

(iii) $\int_{\mathrm{R}} \int_{\mathbb{R}} |f(\eta) r|\, V(\,\mathrm{d}r, \eta)\, \mu(\,\mathrm{d}\eta) < \infty$

are satisfied, then the function f is integrable with respect to L and $\int_{\mathrm{R}} f \,\mathrm{d}L$ is a well defined random variable with the cumulant function

$$
\begin{aligned}
C\left(v, \int_{\mathrm{R}} f \,\mathrm{d}L\right) \quad &= \quad iv \int_{\mathrm{R}} f(\eta) a(\,\mathrm{d}\eta) - \frac{1}{2} v^2 \int_{\mathrm{R}} f(\eta)^2 b(\,\mathrm{d}\eta) \\
&\quad + \int_{\mathrm{R}} \int_{\mathbb{R}} (e^{if(\eta)vr} - 1 - if(\eta) vr \mathbf{1}_{[-1,1]}(r))\, V(\,\mathrm{d}r, \eta)\, \mu(\,\mathrm{d}\eta).
\end{aligned} \tag{9}
$$

Proof. It suffices to check that the regularity conditions of [36, Theorem 2.7] are satisfied under the assumptions of Lemma 1. More specifically we need to check that

(a) $\int_{\mathrm{R}} |h(f(\eta), \eta)| \mu(\,\mathrm{d}\eta) < \infty$,

(b) $\int_{\mathrm{R}} f(\eta)^2 \tilde{b}(\eta) \mu(\,\mathrm{d}\eta) < \infty$,

(c) $\int_{\mathrm{R}} \left(\int_{\mathbb{R}} \min\{1, (rf(\eta))^2\} V(\,\mathrm{d}r, \eta) \right) \mu(\,\mathrm{d}\eta) < \infty$,

where

$$
h(u, \eta) = u \tilde{a}_\tau(\eta) + \int_{\mathbb{R}} (\tau(ru) - u\tau(r)) V(\,\mathrm{d}r, \eta).
$$

Here,

$$
\tau(r) = r \mathbf{1}_{[-1,1]}(r) + \frac{r}{|r|} \mathbf{1}_{[-1,1]^C}(r)
$$

and

$$
\tilde{a}_\tau(\eta) = \tilde{a}(\eta) + \int_{[-1,1]^C} \frac{r}{|r|} V(\,\mathrm{d}r, \eta).
$$

To proof (a), note that $|\tau(ru) \leq |ur|$. Therefore,

$$
|h(f(\eta), \eta)| \leq |f(\eta) \tilde{a}_\tau(\eta)| + 2 \int_{\mathbb{R}} |f(\eta) r| V(\,\mathrm{d}r, \eta).
$$

Using (i) and (iii) of Lemma 1, it follows that

$$
\int_{\mathrm{R}} |h(f(\eta), \eta)| \mu(\,\mathrm{d}\eta) \leq \int_{\mathrm{R}} |f(\eta) \tilde{a}(\eta)| \mu(\,\mathrm{d}\eta) + 3 \int_{\mathrm{R}} \int_{\mathbb{R}} |f(\eta) r| V(\,\mathrm{d}r, \eta) \mu(\,\mathrm{d}\eta) < \infty.
$$

Condition (b) is the same as (ii) and (c) follows from (iii) and

$$
\min\{1, (rf(\eta))^2\} \leq |rf(\eta)|.
$$

□

The conclusions of Lemma 1 hold under weaker assumptions, see [20, Proposition 5.6] or [36, Theorem 2.7]. The assumptions in Lemma 1 are simple to check and suffice for our purposes. The master thesis [20] also contains new selfcontained proofs of a number of other results concerning integration with respect to a Lévy basis.

Using equation (4) we can rewrite (9) as

$$
C\left(v, \int_{\mathrm{R}} f \,\mathrm{d}L\right) = \int_{\mathrm{R}} C(vf(\eta), L'(\eta))\, \mu(\,\mathrm{d}\eta). \tag{10}
$$

The logarithm of the Laplace transform of a random variable X will be called the *kumulant function* and denoted by $K(v,X) = \log \mathsf{E}(e^{-vX})$ for $v \in \mathbb{R}$, in accordance with the notation used in [3]. If the kumulant function of the integral $\int_{\mathrm{R}} f \,\mathrm{d}L$ exists, then

$$K\left(v, \int_{\mathrm{R}} f \,\mathrm{d}L\right) = \int_{\mathrm{R}} K(vf(\eta), L'(\eta))\, \mu(\mathrm{d}\eta). \tag{11}$$

Example 1 (Gaussian Lévy basis). If L is a Gaussian Lévy basis with characteristic triplet $(a, b, 0)$, then $L(A)$ is $N(a(A), b(A))$ distributed for each set $A \in \mathcal{A}$. If (6) and (7) hold, we obtain $L'_\nu(\eta) \sim N(\tilde{a}_\nu(\eta), \tilde{b}_\nu(\eta))$. Furthermore,

$$C\left(v, \int_{\mathrm{R}} f \,\mathrm{d}L\right) = iv \int_{\mathrm{R}} f(\eta)\, a(\mathrm{d}\eta) - \frac{1}{2}v^2 \int_{\mathrm{R}} f(\eta)^2\, b(\mathrm{d}\eta).$$

It follows that

$$\int_{\mathrm{R}} f \,\mathrm{d}L \sim N\Big(\int_{\mathrm{R}} f(\eta)\, a(\mathrm{d}\eta), \int_{\mathrm{R}} f(\eta)^2\, b(\mathrm{d}\eta)\Big).$$

The basis is $\nu-$factorizable when $\tilde{a}_\nu$ and $\tilde{b}_\nu$ are constant. A concrete example of a Gaussian Lévy basis is obtained by attaching independent Gaussian random variables $\{X_i\}$ to a locally finite sequence $\{\eta_i\}$ of fixed points and let

$$L(A) = \sum_{\eta_i \in A} X_i, \quad A \in \mathcal{A}.$$

Another example of a Gaussian Lévy basis is the white noise process, cf. e.g. [24, Section 1.3]. □

Example 2 (Poisson Lévy basis). The simplest Lévy jump basis is the Poisson basis for which $L(A) \sim Po(\nu(A))$, where ν is a non-negative measure on $\mathcal{B}(\mathrm{R})$. Clearly, L is a non-negative Lévy basis. This basis has characteristic triplet $(\nu, 0, \delta_1(\mathrm{d}r)\nu(\mathrm{d}\eta))$, where δ_c denotes the Dirac measure concentrated at c. Note that $\tilde{a}_\nu(\eta) \equiv 1$ and $V_\nu(\mathrm{d}r, \eta) = \delta_1(\mathrm{d}r)$. This basis is always $\nu-$factorizable. The random variable $L'_\nu(\eta)$ has a $Po(1)$ distribution. □

Example 3 (generalized G-Lévy basis). A broad and versatile class of (non-negative) Lévy jump bases are the so-called generalized G–Lévy bases with characteristic triplet of the form $(a, 0, U)$ depending on a non-negative measure ν on $\mathcal{B}(\mathrm{R})$. The measures a and U satisfy (6) and (8) with

$$V_\nu(\mathrm{d}r, \eta) = \mathbf{1}_{\mathbb{R}_+}(r)\frac{r^{-\alpha-1}}{\Gamma(1-\alpha)}\,\mathrm{e}^{-\theta(\eta)r}\,\mathrm{d}r \quad \text{and} \quad \tilde{a}_\nu(\eta) = \int_0^1 \frac{r^{-\alpha}}{\Gamma(1-\alpha)}\mathrm{e}^{-\theta(\eta)r}\,\mathrm{d}r,$$

where $\alpha \in (-\infty, 1)$ and $\theta : \mathrm{R} \to (0, \infty)$ is a measurable function. Γ denotes the gamma function. The class includes two important special cases – the gamma Lévy basis for $\alpha = 0$ with $L'_\nu(\eta) \sim \Gamma(1, \theta(\eta))$, and the inverse Gaussian Lévy basis for $\alpha = \frac{1}{2}$ with $L'_\nu(\eta) \sim IG(\sqrt{2}, \sqrt{2\theta(\eta)})$. In case the function θ is constant $\theta(\eta) = \theta$ we get that $L(A) \sim G(\alpha, \nu(A), \theta)$, i.e. L is a G-measure as defined in [5, Section 2]. □

The following theorem is a special case of the Lévy-Ito decomposition. This theorem will play a crucial role for the interpretation of some of the Lévy driven Cox processes to be considered in the subsequent sections.

Theorem 2 *Suppose that the Lévy basis L has no Gaussian part ($b = 0$) and its generalized Lévy measure U satisfies the following conditions*

- $U(\{(r,\eta)\}) = 0$ *for all* $(r,\eta) \in \mathbb{R} \times \mathrm{R}$ *(U is diffuse),*
- $\int_{[-1,1]\times A} |r|\, U(\mathrm{d}r, \mathrm{d}\eta) < \infty$ *for all* $A \in \mathcal{A}$.

Then

$$L(A) = a_0(A) + \int_{\mathbb{R}} rN(\mathrm{d}r, A), \quad A \in \mathcal{A}, \tag{12}$$

where

$$a_0(A) = a(A) - \int_{[-1,1]} rU(\mathrm{d}r, A), \quad A \in \mathcal{A},$$

and N is a Poisson measure on $\mathbb{R} \times \mathrm{R}$ *with intensity measure U.*

The conditions of Theorem 2 are satisfied for a Poisson Lévy basis and a generalized G-Lévy basis if ν is a diffuse locally finite measure on $\mathcal{B}(\mathrm{R})$.

3 Cox processes

Let S be a Borel subset of $\mathbb{R}^d$ and suppose that $\{\Lambda(\xi) : \xi \in \mathrm{S}\}$ is a non-negative random field which is almost surely integrable (with respect to the Lebesgue measure) on bounded Borel subsets of S. A point process X on S is a *Cox process with the driving field* Λ, if conditionally on Λ, X is a Poisson process with intensity Λ, cf. [11, 12, 31]. The driving measure Λ_M of the Cox process X is defined by

$$\Lambda_M(B) = \int_B \Lambda(\xi)\,\mathrm{d}\xi, \quad B \in \mathcal{B}_b(\mathrm{S}),$$

where $\mathcal{B}_b(\mathrm{S})$ is the bounded Borel subsets of S.

In the following, the intensity function of X will be denoted by $\rho(\xi)$ and, more generally, $\rho^{(n)}(\xi)$ is the nth order product density of X. It follows from the conditional structure of X that $\rho^{(n)}$ can be computed from Λ by

$$\rho^{(n)}(\xi_1, \ldots, \xi_n) = \mathsf{E} \prod_{i=1}^{n} \Lambda(\xi_i), \qquad \xi_i \in \mathrm{S}. \tag{13}$$

(for a proof, using moment measures, see e.g. [12]). A useful characteristic of a point process is the pair correlation function defined by

$$g(\xi_1, \xi_2) = \frac{\rho^{(2)}(\xi_1, \xi_2)}{\rho^{(1)}(\xi_1)\,\rho^{(1)}(\xi_2)}, \qquad \xi_1, \xi_2 \in \mathrm{S},$$

provided that $\rho^{(1)}(\xi_i) > 0$ for $i = 1, 2$. (We let $g(\xi_1, \xi_2) = 0$, if $\rho^{(1)}(\xi_1)\,\rho^{(1)}(\xi_2) = 0$, cf. [31, p. 31].) Note that for a Cox process, the pair correlation function can be calculated as

$$g(\xi_1, \xi_2) = \frac{\mathsf{E}\,\Lambda(\xi_1, \xi_2)}{\mathsf{E}\,\Lambda(\xi_1)\,\mathsf{E}\,\Lambda(\xi_2)}.$$

It can be shown that a Cox process is overdispersed relative to the Poisson process, i.e.

$$\mathsf{Var}(X(B)) \geq \mathsf{E}\,X(B),$$

where $X(B)$ denotes the number of points from X falling in B.

Examples of Cox processes include shot noise Cox processes (SNCPs, see [5, 27, 49]) with driving field of the form

$$\Lambda(\xi) = \sum_{(r,\eta)\in\Phi} r\, k(\xi,\eta),$$

where k is a probability kernel ($k(\cdot,\eta)$ is a probability density) and Φ is the atoms of a Poisson measure on $\mathbb{R}_+ \times \mathrm{R}$, say. Concrete examples of probability kernels are the uniform kernel

$$k(\xi,\eta) = \frac{1}{\omega_d R^d}\mathbf{1}_{[0,R]}(\|\xi-\eta\|),$$

where $R > 0$ and $\omega_d = \pi^{d/2}/\Gamma(1+d/2)$ is the volume of the unit ball in $\mathbb{R}^d$, and the Gaussian kernel

$$k(\xi,\eta) = \frac{1}{(2\pi\sigma^2)^{d/2}}\exp(-\|\xi-\eta\|^2/2\sigma^2), \tag{14}$$

where $\sigma^2 > 0$. Another important class of Cox processes are the log Gaussian Cox processes (LGCPs, see [29]) driven by the exponential of a Gaussian field Ψ

$$\Lambda(\xi) = \exp(\Psi(\xi)).$$

4 Lévy driven Cox processes (LCPs)

4.1 Definition

A point process X on S is called a *Lévy driven Cox process (LCP)* if X is a Cox process with a driving field of the form

$$\Lambda(\xi) = \int_{\mathcal{R}} k(\xi,\eta)L(\mathrm{d}\eta), \qquad \xi \in \mathrm{S}, \tag{15}$$

where L is a non-negative Lévy basis on R. Furthermore, k is a non-negative function on $\mathrm{S}\times\mathrm{R}$ such that $k(\xi,\cdot)$ is integrable with respect to L for each $\xi\in\mathrm{S}$ and $k(\cdot,\eta)$ is integrable with respect to the Lebesgue measure on S for each $\eta\in\mathrm{R}$.

Note that it is always possible for each pair (k, L) to construct an associated pair $(\tilde{k},\tilde{L})$ generating the same driving field Λ where now $\tilde{k}(\cdot,\eta)$ is a probability kernel. We may simply let

$$\tilde{k}(\xi,\eta) = k(\xi,\eta)/\alpha(\eta),$$
$$\tilde{L}(\mathrm{d}\eta) = \alpha(\eta)L(\mathrm{d}\eta)$$

where

$$\alpha(\eta) = \int_{\mathrm{S}} k(\xi,\eta)\,\mathrm{d}\xi$$

is assumed to be strictly positive. In the formulation and analysis of the models it is however convenient not always to restrict to probability kernels.

It is important to note that from the non-negativity of the Lévy basis L and [12, Theorem 6.1.VI], we get that L is equivalent to a random measure on

R. Thus, the measurability of Λ defined in (15) follows from measurability of k as a function of η and ξ and Tonelli's theorem. Therefore, Λ is a well-defined random field and (under the condition of local integrability - see below) the driving measure $\int_B \Lambda(\xi)\,\mathrm{d}\xi$, $B \in \mathcal{B}_b(\mathrm{S})$, is also a well-defined random measure determined by the finite-dimensional distributions of L.

It will be assumed that the function k and the Lévy basis L have been chosen such that Λ is almost surely locally integrable, i.e. $\int_B \Lambda(\xi)\,\mathrm{d}\xi < \infty$ with probability 1 for $B \in \mathcal{B}_b(\mathrm{S})$. A sufficient condition for the last property is that, cf. [31, Remark 5.1],

$$\int_B \mathsf{E}\,\Lambda(\xi)\,\mathrm{d}\xi < \infty, \qquad B \in \mathcal{B}_b(\mathcal{S}). \tag{16}$$

If L is factorizable, then (16) is satisfied if the following conditions hold

$$\begin{aligned} &\int_1^\infty rV(\,\mathrm{d}r) < \infty, \\ &\int_B \int_{\mathrm{R}} k(\xi,\eta)\mu(\,\mathrm{d}\eta)\,\mathrm{d}\xi < \infty, \qquad B \in \mathcal{B}_b(\mathrm{S}). \end{aligned}$$

4.2 The nth order product densities of an LCP

It is possible to derive a number of properties of LCPs, using the theory of Lévy bases presented in Section 2. Below, the nth order product densities, the generating functional and the void probabilities of an LCP are considered. In the proposition below, (complete) Bell polynomials, well known in combinatorics, are used, see [10].

Proposition 3 *Suppose that*

$$\mathsf{E}\left(\int_{\mathrm{R}} k(\xi,\eta)L(\mathrm{d}\eta)\right)^n < \infty$$

and

$$\int_{\mathrm{R}}\int_{\mathbb{R}_+} (k(\xi,\eta)r)^n V(\mathrm{d}r,\eta)\mu(\mathrm{d}\eta) < \infty,$$

for all $\xi \in \mathrm{S}$. Then, the nth order product density of an LCP is given by

$$\rho^{(n)}(\xi_1,\ldots,\xi_n) = \frac{1}{2^n n!}\sum_{t\in T_n}\left(\prod_{j=1}^n t_j\right) B_n(\kappa_1(t),\ldots,\kappa_n(t)),$$

$\xi_1,\ldots,\xi_n \in \mathrm{S}$, *where T_n denotes the set of all functions from $\{1,\ldots,n\}$ to $\{-1,1\}^n$, B_n is the nth complete Bell polynomial evaluated at*

$$\kappa_j(t) = \int_{\mathrm{R}}\left(\sum_{i=1}^n t_i k(\xi_i,\eta)\right)^j \kappa_j(L'(\eta))\mu(\mathrm{d}\eta), \quad j=1,\ldots,n,$$

and $\kappa_j(L'(\eta))$ is the jth cumulant moment of the spot variable $L'(\eta)$.

Proof. First we rewrite $\rho^{(n)}(\xi_1,\ldots,\xi_n) = \mathsf{E}\prod_{i=1}^n \Lambda(\xi_i)$, using the polarization formula, cf. [15, p. 43],

$$\mathsf{E}\prod_{i=1}^{n}\Lambda(\xi_i) = \frac{1}{2^n n!}\sum_{t\in T_n}\left(\prod_{i=1}^{n} t_i\right)\mathsf{E}\left(\sum_{i=1}^{n} t_i\Lambda(\xi_i)\right)^n. \tag{17}$$

The terms

$$\mathsf{E}\left(\sum_{i=1}^{n} t_i\Lambda(\xi_i)\right)^n$$

can be computed by evaluating the nth complete Bell polynomial in the first n cumulants of $\sum_{i=1}^n t_i\Lambda(\xi_i) = \int_{\mathrm{R}}\sum_{i=1}^n t_i k(\xi_i,\eta)\,L(\mathrm{d}\eta)$. Thus, we have

$$\mathsf{E}\left(\sum_{i=1}^{n} t_i\Lambda(\xi_i)\right)^n = B_n(\kappa_1(t),\ldots,\kappa_n(t)),$$

where $\kappa_j(t)$ is the jth cumulant of

$$\int_{\mathrm{R}}\sum_{i=1}^{n} t_i k(\xi_i,\eta)L(\mathrm{d}\eta).$$

Under the assumptions of the proposition, $\kappa_j(t)$ can be calculated by differentiating (11) j times with $f(\eta) = \sum_{i=1}^n t_i k(\xi_i,\eta)$. We get

$$\kappa_j(t) = \int_{\mathrm{R}}\left(\sum_{i=1}^{n} t_i k(\xi_i,\eta)\right)^j \kappa_j(L'(\eta))\mu(\mathrm{d}\eta).$$

Note that

$$\mathsf{E}\left(\int_{\mathrm{R}}\sum_{i=1}^{n} t_i k(\xi_i,\eta)L(\mathrm{d}\eta)\right)^j < \infty$$

and

$$\int_{\mathrm{R}}\int_{\mathbb{R}_+}(\sum_{i=1}^{n} t_i k(\xi_i,\eta))^j V(\mathrm{d}r,\eta)\mu(\mathrm{d}\eta) < \infty,$$

$j = 1,\ldots,n$, under the assumptions of the proposition. □

Corollary 4 *Suppose that $k(\xi,\cdot)$ satisfies the assumptions of Lemma 1 for each $\xi\in\mathrm{S}$. Then, the intensity function of the LCP exists and is given by*

$$\rho(\xi) = \int_{\mathrm{R}} k(\xi,\eta)\,\mathsf{E}(L'(\eta))\mu(\mathrm{d}\eta) \tag{18}$$

for all $\xi\in\mathrm{S}$. Furthermore, if

$$\mathsf{E}\left(\int_{\mathrm{R}} k(\xi,\eta)L(\mathrm{d}\eta)\right)^2 < \infty, \tag{19}$$

and

$$\int_{\mathrm{R}}\int_{\mathbb{R}}(k(\xi,\eta)r)^2\,V(\mathrm{d}r,\eta)\,\mu(\mathrm{d}\eta) < \infty, \tag{20}$$

for each $\xi\in\mathrm{S}$, the pair correlation function of the process exists and is given by

$$g(\xi,\zeta) = 1 + \frac{\int_{\mathrm{R}} k(\xi,\eta)k(\zeta,\eta)\,\mathsf{Var}(L'(\eta))\mu(\mathrm{d}\eta)}{\rho(\xi)\rho(\zeta)}, \tag{21}$$

for all $\xi,\zeta\in\mathrm{S}$.

Proof. The result follows from Proposition 3, using that the first and second complete Bell polynomials are given by $B_1(x) = x$, $B_2(x_1, x_2) = x_1^2 + x_2$. Also recall that $\kappa_1(L'(\eta)) = \mathsf{E}(L'(\eta))$ and $\kappa_2(L'(\eta)) = \mathsf{Var}(L'(\eta))$. $\square$

Corollary 5 (Stationary LCP) *Let* $\mathrm{S} = \mathrm{R} = \mathbb{R}^d$ *and assume that* k *is a homogeneous kernel in the sense that*

$$k(\xi, \eta) = k(\xi - \eta) \qquad \textit{for all} \quad \xi, \eta \in \mathbb{R}^d. \tag{22}$$

Let $\int k(\eta)\,\mathrm{d}\eta = \alpha$. *Assume that* L *is a homogenous Lévy basis with control measure* $\mu(\mathrm{d}\eta) = c\,\mathrm{d}\eta$ *for some* $c > 0$. *Then, (18) and (21) take the following simplified form*

$$\begin{aligned} \rho =& c\,\mathsf{E}\,L'\alpha \\ g(\xi, \zeta) =& 1 + \frac{\mathsf{Var}\,L'}{(\mathsf{E}\,L')^2}\frac{I_k(\zeta - \xi)}{c}, \end{aligned}$$

where I_k *only depend on the kernel* k

$$I_k(\zeta - \xi) = \int_{\mathbb{R}^d} \frac{k(\zeta - \xi + \eta)k(\eta)}{\alpha^2}\,\mathrm{d}\eta.$$

Note that the fraction $\frac{\mathsf{Var}\,L'}{(\mathsf{E}\,L')^2}$ is equal to $\frac{1}{\mathsf{E}\,L'}$, 1 and $\mathsf{E}\,L'$ for the Poisson, gamma and inverse Gaussian basis, respectively. The choice of the Lévy basis changes substantially the correlations in the LCP and the overall variability in the point pattern even when the corresponding LCPs are stationary and all other parameters of the model are the same. As an illustration, Figure 1 shows 3 stationary LCPs observed on a $[0, 100] \times [0, 200]$ window with $c = 0.003$, $\mathsf{E}\,L' = 2$ and a Gaussian kernel obtained as 10 times the kernel (14) with $\sigma = 4$. The spot variable L' is distributed as $\mathsf{E}\,L'$ times a $Po(1)$-distributed variable, as a $\Gamma(1, \mathsf{E}\,L')$-distributed variable and a $IG(1, 1/\,\mathsf{E}\,L')$-distributed variable, respectively. From left to right, an increasing irregularity is clearly visible.

The distribution of a point process X on S can be characterized by the probability generating functional G_X. This functional is defined by

$$G_X(u) = \mathsf{E}\prod_{\xi \in X} u(\xi),$$

for functions $u : \mathrm{S} \to [0, 1]$ with $\{\xi \in \mathrm{S} : u(\xi) < 1\}$ bounded. As proved e.g. in [12] the probability generating functional of a Cox process can be computed by

$$G_X(u) = \mathsf{E}\exp\left(-\int_{\mathrm{S}}(1 - u(\xi))\Lambda(\xi)\,\mathrm{d}\xi\right). \tag{23}$$

Void probabilities can be calculated as

$$v(B) := P(X \cap B = \emptyset) = \mathsf{E}\exp\left(-\int_B \Lambda(\xi)\,\mathrm{d}\xi\right), \quad B \in \mathcal{B}_b(\mathrm{S}).$$

Below, we give the expressions for G_X and v for an LCP.

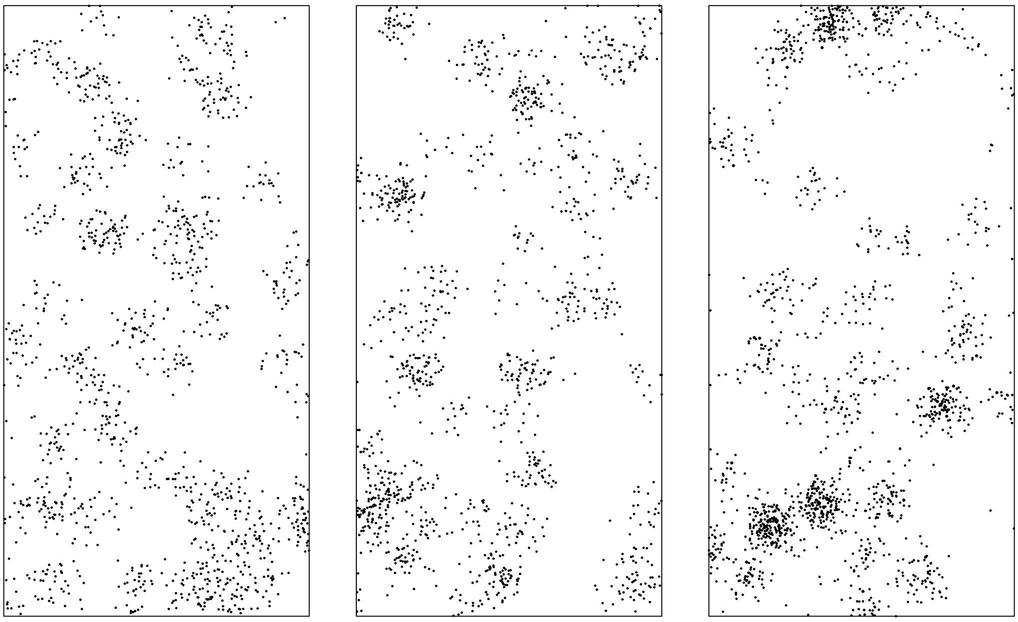

Figure 1: Examples of realizations of homogeneous LCPs with Poisson (left), gamma (middle) and inverse Gaussian (right) Lévy bases. For details, see the text.

Proposition 6 *The probability generating functional of an LCP has the following form*

$$\begin{aligned} G_X(u) &= \exp\Bigg(- \int_{\mathbb{R}_+} \int_{\mathrm{R}} \left[1 - \exp\left(- \int_{\mathcal{S}} (1 - u(\xi)) k(\xi, \eta) r \, \mathrm{d}\xi \right) \right] U(\mathrm{d}r, \mathrm{d}\eta) \\ &\qquad - \int_{\mathrm{R}} \int_{\mathrm{S}} (1 - u(\xi)) k(\xi, \eta) \, \mathrm{d}\xi a_0(\mathrm{d}\eta) \Bigg), \end{aligned}$$

while the void probabilities are given by

$$\begin{aligned} v(B) &= \exp\Bigg(- \int_{\mathbb{R}_+} \int_{\mathrm{R}} \left[1 - \exp(-r \int_B k(\xi, \eta) \, \mathrm{d}\xi) \right] U(\mathrm{d}r, \mathrm{d}\eta) \\ &\qquad - \int_{\mathrm{R}} \int_B k(\xi, \eta) \, \mathrm{d}\xi a_0(\mathrm{d}\eta) \Bigg), \quad B \in \mathcal{B}_b(\mathrm{S}). \end{aligned}$$

Proof. Since $\Lambda(\xi)$ is almost surely locally integrable,

$$\int_{\mathrm{S}} (1 - u(\xi)) \Lambda(\xi) \, \mathrm{d}\xi \le \int_{\mathrm{S}} \mathbf{1}_{\mathrm{supp}(1-u)}(\xi) \Lambda(\xi) \, \mathrm{d}\xi < \infty \tag{24}$$

is a well-defined non-negative random variable and its kumulant transform exists. (In (24), the support of the function $1 - u$ is denoted $\mathrm{supp}(1 - u)$.)

Using the key relation (11) for the kumulant function, we get

$$
\begin{aligned}
\log\left(G_X(u)\right) =& \log\left(\mathsf{E}\exp(-\int_{\mathrm{S}}(1-u(\xi))\Lambda(\xi)\,\mathrm{d}\xi)\right)\\
=& K\left(1, \int_{\mathrm{S}}(1-u(\xi))\Lambda(\xi)\,\mathrm{d}\xi\right)\\
=& K\left(1, \int_{\mathrm{S}}(1-u(\xi))\int_{\mathrm{R}}k(\xi,\eta)L(\mathrm{d}\eta)\,\mathrm{d}\xi\right)\\
=& K\left(1, \int_{\mathrm{R}}\left(\int_{\mathrm{S}}(1-u(\xi))k(\xi,\eta)\,\mathrm{d}\xi\right)L(\mathrm{d}\eta)\right)\\
=& \int_{\mathrm{R}}K\left(\int_{\mathrm{S}}(1-u(\xi))k(\xi,\eta)\,\mathrm{d}\xi,\ L'(\eta)\right)\mu(\mathrm{d}\eta)\\
=& -\int_{\mathrm{R}}\int_{\mathrm{S}}(1-u(\xi))k(\xi,\eta)\,\mathrm{d}\xi a_0(\mathrm{d}\eta)\\
&+\int_{\mathrm{R}}\int_{\mathbb{R}_+}\left(\exp\left(-\int_{\mathrm{S}}(1-u(\xi))k(\xi,\eta)r\,\mathrm{d}\xi\right)-1\right)V(\mathrm{d}r,\eta)\mu(\mathrm{d}\eta).
\end{aligned}
$$

The result for the void probabilities is obtained by choosing $u(\xi) = \mathbf{1}_{B^c}(\xi)$. $\square$

4.3 Mixing properties

The following proposition gives conditions for stationarity and mixing of an LCP. Mixing and ergodicity are important e.g. for establishing the consistency of model parameter estimates, including nonparametric estimates of the n-th order product density $\rho^{(n)}$ and the pair correlation function g. Mixing [12, Definition 10.3.I] implies ergodicity [12, p. 341]. The case of an LCP with G-Lévy basis has been treated in [5, Proposition 2.2].

Proposition 7 *Let* $\mathrm{S} = \mathrm{R} = \mathbb{R}^d$ *and assume that the Lévy basis* L *and the kernel* k *are homogeneous. Then, an LCP with driving field* Λ *of the form (15) is stationary and mixing.*

Proof. Note that a Cox process is stationary/mixing if and only if the driving field of the Cox process has the same property [12, Proposition 10.3.VII]. Using the assumptions of the proposition it is easily seen that $\{\Lambda(\xi + x) : \xi \in \mathbb{R}^d\}$ has the same distribution as $\{\Lambda(\xi) : \xi \in \mathbb{R}^d\}$ for all $x \in \mathbb{R}^d$.

According to [12, Proposition 10.3.VI(a)], Λ is mixing if and only if

$$\mathsf{L}_\Lambda[h_1 + T_x h_2] \to \mathsf{L}_\Lambda[h_1]\mathsf{L}_\Lambda[h_2],$$

as $\|x\| \to \infty$. Here, h_1 and h_2 are arbitrary non-negative bounded functions on $\mathbb{R}^d$ of bounded support, and L_Λ is the Laplace functional defined by

$$\mathsf{L}_\Lambda[h] = \mathsf{E}\exp\left(-\int_{\mathbb{R}^d}h(\xi)\Lambda(\xi)\,\mathrm{d}\xi\right),$$

and $T_x h(\xi) = h(\xi + x)$, $\xi, x \in \mathbb{R}^d$. We get

$$
\begin{aligned}
&\mathsf{L}_\Lambda[h_1 + T_x h_2] \\
&= \mathsf{E} \exp\left(-\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(h_1(\xi) + h_2(\xi + x))k(\xi - \eta)L(\mathrm{d}\eta)\,\mathrm{d}\xi\right) \\
&= \mathsf{E} \exp\left(-\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} h_1(\xi)k(\xi - \eta)\,\mathrm{d}\xi + \int_{\mathbb{R}^d} h_2(\xi)k(\xi - \eta - x)\,\mathrm{d}\xi\right) L(\mathrm{d}\eta)\right) \\
&= \mathsf{E}\left[\exp\left(-\int_{\mathbb{R}^d} \tilde{h}_1(\eta)L(\mathrm{d}\eta)\right) \cdot \exp\left(-\int_{\mathbb{R}^d} \tilde{h}_2(\eta + x)L(\mathrm{d}\eta)\right)\right],
\end{aligned}
$$

where

$$\tilde{h}_i(\eta) = \int_{\mathbb{R}^d} h_i(\xi)k(\xi - \eta)\,\mathrm{d}\xi.$$

If k has bounded support, then we can find a $C > 0$ such that for $\|x\| > C$

$$\{\eta \in \mathbb{R}^d : \tilde{h}_1(\eta) > 0\} \cap \{\eta \in \mathbb{R}^d : \tilde{h}_2(\eta + x) > 0\} = \emptyset.$$

It follows that for $\|x\| > C$

$$
\begin{aligned}
\mathsf{L}_\Lambda[h_1 + T_x h_2] &= \mathsf{E} \exp\left(-\int_{\mathbb{R}^d} \tilde{h}_1(\eta)L(\mathrm{d}\eta)\right) \cdot \mathsf{E} \exp\left(-\int_{\mathbb{R}^d} \tilde{h}_2(\eta + x)L(\mathrm{d}\eta)\right) \\
&= \mathsf{L}_\Lambda[h_1]\mathsf{L}_\Lambda[h_2],
\end{aligned}
$$

since L is independently scattered. If k does not have bounded support, we define a series of functions with bounded support

$$k_n(\xi - \eta) = k(\xi - \eta)\mathbf{1}_{[0,n)}(\|\xi - \eta\|), \quad n = 1, 2, \ldots$$

that converges monotonically from below to k. It follows that $\tilde{h}_{i,n}$ defined by

$$\tilde{h}_{i,n}(\eta) = \int_{\mathbb{R}^d} h_i(\xi)k_n(\xi - \eta)\,\mathrm{d}\xi$$

converges monotonically from below to $\tilde{h}_i(\eta)$ and for fixed n we can find C_n such that for $\|x\| > C_n$

$$
\begin{aligned}
&\mathsf{E}\left[\exp\left(-\int_{\mathbb{R}^d} \tilde{h}_{1,n}(\eta)L(\mathrm{d}\eta)\right) \cdot \exp\left(-\int_{\mathbb{R}^d} \tilde{h}_{2,n}(\eta + x)L(\mathrm{d}\eta)\right)\right] \\
&= \mathsf{E} \exp\left(-\int_{\mathbb{R}^d} \tilde{h}_{1,n}(\eta)L(\mathrm{d}\eta)\right) \cdot \mathsf{E} \exp\left(-\int_{\mathbb{R}^d} \tilde{h}_{2,n}(\eta + x)L(\mathrm{d}\eta)\right)
\end{aligned}
$$

Using the reasoning just after [12, Proposition 10.3.VI], it follows that

$$\mathsf{L}_\Lambda[h_1 + T_x h_2] \to \mathsf{L}_\Lambda[h_1]\mathsf{L}_\Lambda[h_2],$$

for the original functions h_1 and h_2. □

4.4 Examples of LCPs

4.4.1 Shot noise Cox processes (SNCPs) with random noise

Under the assumptions of Theorem 2, the driving field of an LCP takes the form

$$\Lambda(\xi) = \int_{\mathrm{R}} k(\xi, \eta)a_0(\mathrm{d}\eta) + \sum_{(r,\eta)\in\Phi} rk(\xi, \eta), \tag{25}$$

where Φ is the atoms of a Poisson measure on $\mathbb{R}_+ \times \mathrm{R}$ with intensity measure U. An LCP X with such a driving field is distributed as a superposition $X_1 \cup X_2$ where X_1 and X_2 are independent, X_1 is a Poisson point process with intensity function

$$\rho_1(\xi) = \int_{\mathrm{R}} k(\xi, \eta) a_0(\mathrm{d}\eta)$$

and X_2 is a shot noise Cox process as defined in [27] with driving field

$$\Lambda_2(\xi) = \sum_{(r,\eta)\in\Phi} r k(\xi, \eta).$$

An LCP with driving field Λ of the form (25) is therefore an SNCP with additional random noise. Simulation of the associated Lévy basis can be performed, using the algorithm introduced in [16], if L is factorizable, otherwise the algorithm developed in [50] may be used, see also [46]. A third option is the method used in [27]. An overview of available methods of simulating Lévy processes can be found in [39].

For $a_0 \equiv 0$, we get the familiar SNCPs. In [27], three specific examples of stationary SNCPs are considered. Using the notion of a Lévy basis, they are specified by $U(\mathrm{d}r, \mathrm{d}\eta) = V(\mathrm{d}r, \eta)\nu(\mathrm{d}\eta)$, where $\nu(\mathrm{d}\eta) \propto \mathrm{d}\eta$ and

- V is concentrated in a single point $c > 0$, i.e. $V(\mathrm{d}r) = \delta_c(\mathrm{d}r)$. If $c = 1$, the corresponding Lévy basis is Poisson. If $c \neq 1$, $L(A) \sim cPo(\nu(A))$. LCPs of this type are the well-known Matérn cluster process [25] and the Thomas process [42].
- $V((0, \infty)) < \infty$. In this case, Φ can be represented as a marked Poisson point process. Examples of LCPs with such a Lévy basis are the Neyman-Scott processes, cf. [34].
- $V(\mathrm{d}r) = \mathbf{1}_{\mathbb{R}_+}(r)\frac{r^{-\alpha-1}}{\Gamma(1-\alpha)}\mathrm{e}^{-\theta r}\,\mathrm{d}r$ corresponding to a G-Lévy basis. The resulting LCP is a so-called shot noise G Cox process [5].

In Figure 2, we show an example of an SNCP with a homogeneous Poisson process (a_0 is proportional to Lebesgue measure) as additional random noise. More precisely, the process $X = X_1 \cup X_2$ is defined on $[0, 200] \times [0, 100]$, X_1 is a Poisson process with intensity 0.01 and X_2 is an SNCP with Gaussian kernel (14) with $\sigma = 2$ and an intensity measure U of the form $U(\mathrm{d}r, \mathrm{d}\eta) = \delta_{25}(r) \cdot 0.0025\,\mathrm{d}\eta$. The process X_2 is thereby a Thomas process.

4.4.2 LCPs driven by smoothed discrete random fields

We suppose that $\{\eta_i\}$ is a locally finite sequence of fixed points and let

$$L(A) = \sum_{\eta_i \in A} X_i,$$

where $\{X_i\}$ is a sequence of independent and identically distributed non-negative random variables with infinitely divisible distribution. If, for instance, X_i is gamma or inverse Gaussian distributed, then L is a special case of a gamma or inverse Gaussian Lévy basis, respectively. The driving intensity of the associated LCP will take the form

$$\Lambda(\xi) = \sum_{\eta_i} k(\xi, \eta_i) X_i.$$

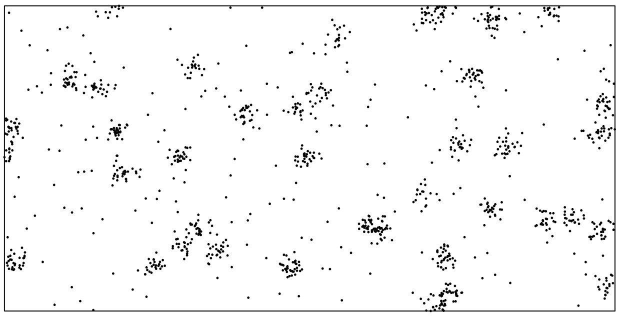

Figure 2: Example of a shot noise Cox process with extra noise. For details, see the text.

5 Log Lévy driven Cox processes (LLCPs)

5.1 Definition

A point process X on S is called a *log Lévy driven Cox process (LLCP)* if X is a Cox process with intensity field of the form

$$\Lambda(\xi) = \exp\left(\int_{\mathrm{R}} k(\xi,\eta) L(\mathrm{d}\eta)\right), \tag{26}$$

where L is a Lévy basis and k is a kernel such that $k(\xi,\cdot)$ is integrable with respect to L for each $\xi \in \mathrm{S}$, $k(\cdot,\eta)$ is integrable with respect to Lebesgue measure on S for each $\eta \in \mathrm{R}$ and Λ is almost surely locally integrable.

Since the driving intensity field of an LLCP is always non-negative because of the exponential function, we can generally use kernels and Lévy bases which also have negative values. Moreover, using the Lévy-Khintchine representation (1), we see that each Lévy basis L is equal to a sum of two independent parts – a Lévy jump part (let us denote it by L_J) and a Gaussian part (let us denote it by L_G). Thus we can represent the driving intensity of an LLCP as a product of two independent driving fields

$$\Lambda(\xi) = \exp\left(\int_{\mathrm{R}} k(\xi,\eta) L_J(\mathrm{d}\eta)\right) \exp\left(\int_{\mathrm{R}} k(\xi,\eta) L_G(\mathrm{d}\eta)\right) = \Lambda_J(\xi)\Lambda_G(\xi). \tag{27}$$

If $L_J \equiv 0$, Λ is the driving field of a log Gaussian Cox process (LCP) [9, 29]; if $L_G \equiv 0$, Λ is under regularity conditions the driving field of a log shot noise Cox process, see the examples in Section 5.4 below.

Because of the exponential function in the definition of $\Lambda(\xi)$, stronger conditions on k and L are needed in order to ensure that Λ is almost surely locally integrable. A sufficient condition is that the kumulant transform $K(-k(\xi,\eta), L'(\eta))$ exists for all $\xi \in \mathrm{S}$ and $\eta \in \mathrm{R}$, and that

$$\int_B \exp\left(\int_{\mathrm{R}} K(-k(\xi,\eta), L'(\eta))\,\mu(\mathrm{d}\eta)\right) \mathrm{d}\xi < \infty, \qquad \text{for all } B \in \mathcal{B}_b(\mathrm{S}). \tag{28}$$

This result follows from the definition of the kumulant function and from the key relation (11) for the kumulant transform. In particular, we use that

$$\mathsf{E}\,\Lambda(\xi) = \exp\left(K(1, -\int_{\mathrm{R}} k(\xi,\eta) L(\mathrm{d}\eta))\right) = \exp\left(\int_{\mathrm{R}} K(-k(\xi,\eta), L'(\eta))\,\mu(\mathrm{d}\eta)\right).$$

Note that

$$\begin{aligned} K(-k(\xi,\eta),L'(\eta)) &= k(\xi,\eta)\tilde{a}(\eta)+\frac{1}{2}k(\xi,\eta)^2\tilde{b}(\eta) \\ &\quad + \int_{\mathbb{R}}(e^{k(\xi,\eta)r}-1-k(\xi,\eta)r\mathbf{1}_{[-1,1]}(r))V(\mathrm{d}r,\eta). \end{aligned}$$

If L is factorizable, then (28) is satisfied if *either* there exist $B>0$, $C>0$ and $D>0$ such that

$$|k(\xi,\eta)| \le C \text{ for all } \xi \in \mathrm{S}, \eta \in \mathrm{R} \tag{29}$$

$$\int_{\mathrm{R}} |k(\xi,\eta)|^i \mu(\mathrm{d}\eta) < B\cdot D^i, \quad i=1,2,\dots, \quad \xi\in\mathrm{S} \tag{30}$$

$$\int_{\mathbb{R}} \left(e^{(C\vee D)|r|}-1-(C\vee D)|r|\mathbf{1}_{[-1,1]}(r)\right)V(\mathrm{d}r)<\infty, \tag{31}$$

or there exist $C>0$ and $R>0$ such that

$$|k(\xi,\eta)| \le C \text{ for all } \xi \in \mathrm{S}, \eta \in \mathrm{R} \tag{32}$$

$$k(\xi,\eta)=0 \text{ for } \|\xi-\eta\|>R \tag{33}$$

$$\mu \text{ is locally finite} \tag{34}$$

$$\int_{\mathbb{R}} \left(e^{C|r|}-1-C|r|\mathbf{1}_{[-1,1]}(r)\right)V(\mathrm{d}r)<\infty. \tag{35}$$

Note that (29) and (30) are satisfied for the Gaussian kernel if μ is Lebesgue measure, while (32) and (33) hold for the uniform kernel. In the case of a purely Gaussian basis, (30) is only needed for $i=2$ and conditions (31) and (35) are trivially fulfilled since $V\equiv 0$.

5.2 The nth order product densities of an LLCP

The nth order product densities of LLCPs are easily derived, using Lévy theory.

Proposition 8 *The nth order product density is given by*

$$\rho^{(n)}(\xi_1,\dots,\xi_n)=\exp\left(\int_{\mathrm{R}} K(-\sum_{i=1}^n k(\xi_i,\eta),L'(\eta))\mu(\mathrm{d}\eta)\right), \tag{36}$$

$\xi_1,\dots,\xi_n\in\mathrm{S}$*, provided the right-hand side exists.*

Proof. The formula follows directly from the definition of the kumulant function and from the key relation (11). We get

$$\begin{aligned} \rho^{(n)}(\xi_1,\dots,\xi_n) &= \mathsf{E}\prod_{i=1}^n \Lambda(\xi_i) = \mathsf{E}\exp\left(\sum_{i=1}^n \int_{\mathrm{R}} k(\xi_i,\eta)L(\mathrm{d}\eta)\right) \\ &= \exp\left(K(1,-\sum_{i=1}^n \int_{\mathrm{R}} k(\xi_i,\eta)L(\mathrm{d}\eta))\right) \\ &= \exp\left(\int_{\mathrm{R}} K(-\sum_{i=1}^n k(\xi_i,\eta),L'(\eta))\mu(\mathrm{d}\eta)\right). \end{aligned}$$

□

Corollary 9 *The intensity function of an LLCP X is given by*

$$\rho(\xi) = \exp\left(\int_{\mathrm{R}} K(-k(\xi,\eta), L'(\eta))\,\mu(\mathrm{d}\eta)\right), \tag{37}$$

provided the right-hand side exists. When the second order product density exists, the pair correlation function of an LLCP takes the following form

$$\begin{aligned}
g(\xi,\zeta) &= \exp\Bigg(\int_{\mathrm{R}}[K(-k(\xi,\eta)-k(\zeta,\eta), L'(\eta)) \\
&\qquad -K(-k(\xi,\eta), L'(\eta)) - K(-k(\zeta,\eta), L'(\eta))]\mu(\mathrm{d}\eta)\Bigg) \\
&= \exp\Bigg(\int_{\mathrm{R}} k(\xi,\eta)k(\zeta,\eta)b(\mathrm{d}\eta) \\
&\qquad + \int_{\mathrm{R}}\int_{\mathbb{R}}\left[e^{(k(\xi,\eta)+k(\zeta,\eta))r} - e^{k(\xi,\eta)r} - e^{k(\zeta,\eta)r} + 1\right] V(\mathrm{d}r,\eta)\,\mu(\mathrm{d}\eta)\Bigg).
\end{aligned}$$

Corollary 10 (Stationary LLCP) *Let* $\mathrm{S} = \mathrm{R} = \mathbb{R}^d$*. Assume that k is a homogeneous kernel and L a homogeneous Lévy basis with $\mu(\mathrm{d}\eta) = c\,\mathrm{d}\eta$ for some $c > 0$. Then,*

$$\rho = \exp\left(c\int_{\mathbb{R}^d} K(-k(\eta), L')\,\mathrm{d}\eta\right)$$

and

$$\begin{aligned}
g(\xi,\zeta) &= \exp\Bigg(\tilde{b}c\int_{\mathbb{R}^d} k(\xi-\zeta+\eta)k(\eta)\,\mathrm{d}\eta \\
&\qquad + c\int_{\mathbb{R}^d}\int_{\mathbb{R}}(e^{(k(\xi-\zeta+\eta)+k(\eta))r} - e^{k(\xi-\zeta+\eta)r} - e^{k(\eta)r} + 1)V(\mathrm{d}r)\,\mathrm{d}\eta\Bigg).
\end{aligned}$$

5.3 Mixing properties

Proposition 11 *Let* $\mathrm{S} = \mathrm{R} = \mathbb{R}^d$ *and assume that the Lévy basis L is homogeneous and the kernel k is homogeneous. Then, an LLCP with driving field of the form (26) is stationary and mixing.*

Proof. As in the proof of Proposition 7, we immediately get the stationarity. The method of proving mixing has to be modified compared to the one used in Proposition 7. First, rewrite

$$\begin{aligned}
&\mathsf{L}_\Lambda[h_1 + T_x h_2] \\
&= \mathsf{E}\Bigg[\exp\left(-\int_{\mathbb{R}^d} h_1(\xi)\exp\left(\int_{\mathbb{R}^d} k(\xi-\eta)L(\mathrm{d}\eta)\right)\mathrm{d}\xi\right) \\
&\qquad \cdot\exp\left(-\int_{\mathbb{R}^d} h_2(\xi+x)\exp\left(\int_{\mathbb{R}^d} k(\xi-\eta)L(\mathrm{d}\eta)\right)\mathrm{d}\xi\right)\Bigg] \\
&= \mathsf{E}[A\cdot B_x],
\end{aligned}$$

say. If k has bounded support, A and B_x will be independent if $\|x\|$ is large enough. If k does not have bounded support, we use a series of functions k_n

with bounded support that converges to k. To be precise, let as in Proposition 7

$$k_n(u) = k(u)\mathbf{1}_{[0,n)}(\|u\|), \quad u \in \mathbb{R}^d,$$

$n = 1, 2, \ldots$. We have $k_n \to k$ and $|k_n| \leq |k|$. Now, let

$$A_n = \exp\left(-\int_{\mathbb{R}^d} h_1(\xi) \exp\left(\int_{\mathbb{R}^d} k_n(\xi - \eta) L(\mathrm{d}\eta)\right) \mathrm{d}\xi\right)$$

and

$$B_{x,n} = \exp\left(-\int_{\mathbb{R}^d} h_2(\xi + x) \exp\left(\int_{\mathbb{R}^d} k_n(\xi - \eta) L(\mathrm{d}\eta)\right) \mathrm{d}\xi\right).$$

Note that $0 \leq A, A_n, B_x, B_{x,n} \leq 1$. Now, consider the following inequality

$$\begin{aligned}
&|\mathsf{E}[A \cdot B_x] - \mathsf{E}\, A \cdot \mathsf{E}\, B_x| \\
&\leq |\mathsf{E}[A \cdot B_x] - \mathsf{E}[A \cdot B_{x,n}]| + |\mathsf{E}[A \cdot B_{x,n}] - \mathsf{E}[A_n \cdot B_{x,n}]| \\
&\quad + |\mathsf{E}[A_n \cdot B_{x,n}] - \mathsf{E}\, A_n \cdot \mathsf{E}\, B_{x,n}| + |\mathsf{E}\, A_n \cdot \mathsf{E}\, B_{x,n} - \mathsf{E}\, A_n \cdot \mathsf{E}\, B_x| \\
&\quad + |\mathsf{E}\, A_n \cdot \mathsf{E}\, B_x - \mathsf{E}\, A \cdot \mathsf{E}\, B_x|. \\
&= \delta_{1xn} + \delta_{2xn} + \delta_{3xn} + \delta_{4xn} + \delta_{5xn},
\end{aligned}$$

say. Let us evaluate each of these five terms. Using that $0 \leq A \leq 1$ and that L is homogeneous, we get

$$\begin{aligned}
\delta_{1xn} &= |\mathsf{E}[A \cdot (B_x - B_{x,n})]| \\
&\leq \mathsf{E}[A \cdot |B_x - B_{x,n}|] \\
&\leq \mathsf{E}\,|B_x - B_{x,n}| \\
&= \mathsf{E}\,|B_0 - B_{0,n}|.
\end{aligned}$$

Now, since $k_n \to k$ and $|k_n| \leq |k|$, where k is $L-$integrable, it follows that

$$\int_{\mathbb{R}^d} k_n(\xi - \eta) L(\mathrm{d}\eta) \to \int_{\mathbb{R}^d} k(\xi - \eta) L(\mathrm{d}\eta),$$

almost surely. We can therefore find n_1 (not dependent on x) such that for $n \geq n_1$, $\delta_{1xn} \leq \epsilon$, say. Using the same type of arguments, we can find n_2, n_4, n_5 such that for $n \geq n_i$, $\delta_{ixn} \leq \epsilon$, $i = 2, 4, 5$. Now choose a fixed $n \geq \max(n_1, n_2, n_4, n_5)$ and consider

$$\delta_{3xn} = |\mathsf{E}[A_n \cdot B_{x,n}] - \mathsf{E}\, A_n \cdot \mathsf{E}\, B_{x,n}|.$$

Using the previous results for bounded functions of bounded support, we finally find a constant $C > 0$ such that for x with $\|x\| > C$ we have $\delta_{3xn} \leq \epsilon$. This completes the proof. $\square$

5.4 Examples of LLCPs

5.4.1 Log shot noise Cox processes (LSNCPs)

Under the assumptions of Theorem 2, the driving field of an LLCP takes the form

$$\Lambda(\xi) = \exp\left(d(\xi) + \sum_{(r,\eta) \in \Phi} r k(\xi, \eta)\right), \tag{38}$$

where $d(\xi)$ is a deterministic function and Φ is the atoms of a Poisson measure on $\mathbb{R} \times \mathrm{R}$ with intensity measure U. Such a process is called a *log shot noise Cox process (LSNCP)*.

It is important to realize that SNCPs and LSNCPs are quite different model classes. An SNCP X with driving field of the form

$$\Lambda(\xi) = \sum_{(r,\eta)\in\Phi} rk(\xi,\eta)$$

is a superposition of independent Poisson processes $X_{(r,\eta)}$, $(r,\eta) \in \Phi$, where $X_{(r,\eta)}$ has intensity function $rk(\cdot,\eta)$. (The process $\{\eta : (r,\eta) \in \Phi\}$ is usually called the centre process (although it is not necessarily locally finite) while $X_{(r,\eta)}$ is called a cluster around η.) The presence of a particular cluster $X_{(r,\eta)}$ will not affect the presence of the other clusters.

In contrast to this, the driving field of an LSNCP takes the form

$$\Lambda(\xi) = \exp(d(\xi)) \prod_{(r,\eta)\in\Phi} \exp(rk(\xi,\eta)).$$

A cluster $X_{(r,\eta)}$ with negative, numerically large values of $rk(\cdot,\eta)$ will very likely contain 0 points and moreover, wipe out points from other clusters in the neighbourhood of η. In the resulting point pattern, empty holes may therefore be present. Examples of such point patterns are shown in Figure 3. Here, $\{\eta\}$ is a homogeneous Poisson process on $[0,100]\times[0,200]$ with intensity $c = 0.003$, $V(\mathrm{d}r) = \frac{1}{3}\delta_1(r) + \frac{2}{3}\delta_{-1}(r)$ and the kernel is (left) $k(\xi) = 1(|\xi| \le R)$ and (right) $k(\xi) = (1 - \frac{|\xi|^3}{R^3})1(|\xi| \le R)$, respectively, with $R = 10$.

5.4.2 Log Gaussian Cox processes (LGCPs)

In this subsection, we consider LLCPs with driving field of the form

$$\Lambda(\xi) = \exp\left(\int_{\mathrm{R}} k(\xi,\eta) L(\mathrm{d}\eta)\right), \tag{39}$$

where L is a Gaussian Lévy basis.

Clearly, the resulting process is an LGCP [9, 29]. If k and L are homogeneous, the process is stationary. In this case, the random intensity function $\Lambda(\xi)$ is well defined for all $\xi \in \mathbb{R}^d$ and almost surely integrable if

$$k(\xi) \le C, \xi \in \mathbb{R}^d, \text{ and } \int_{\mathbb{R}^d} k(\xi)^2 \, \mathrm{d}\xi < \infty. \tag{40}$$

The covariance function of the Gaussian field

$$\Psi(\xi) = \int_{\mathbb{R}^d} k(\xi - \eta) L(\mathrm{d}\eta)$$

takes the form

$$\mathsf{Cov}(\Psi(\xi_1), \Psi(\xi_2)) = \int_{\mathbb{R}^d} k(\xi_1 - \xi_2 + \eta) k(\eta) \, \mathrm{d}\eta = c(\xi_1 - \xi_2),$$

say. Note that under (40) c is integrable. Under the mild additional assumption that the set of discontinuity points of k has Lebesgue measure 0, c is also

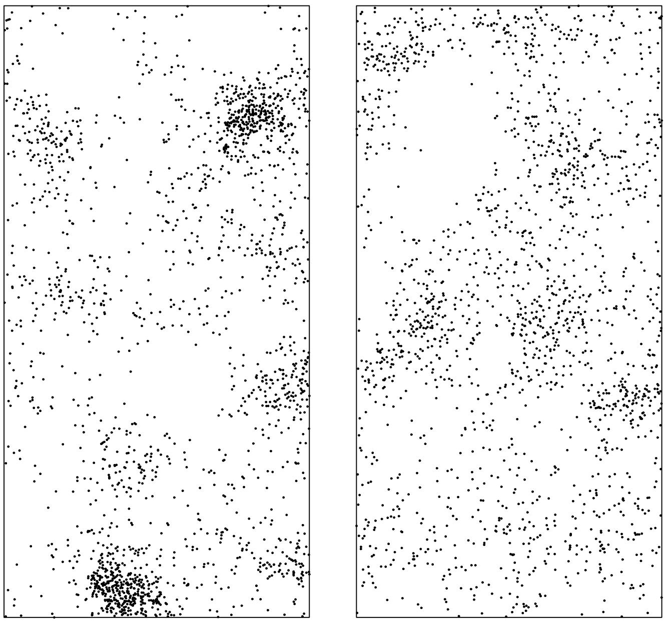

Figure 3: Examples of log shot noise Cox processes. Notice the circular empty holes in the point patterns. For details, see the text.

continuous. In the proposition below, we show that any stationary LGCP with a continuous and integrable covariance function can indeed be obtained as a kernel smoothing (39) of a Gaussian Lévy basis. The proposition is a generalization of a result mentioned in [21].

Proposition 12 *Any stationary Gaussian random field with continuous and integrable covariance function can be generated by a kernel smoothing of a homogeneous Lévy basis.*

Proof. Let $\{\Psi(\xi) : \xi \in \mathbb{R}^d\}$ be an arbitrary stationary zero mean Gaussian field. Let $c(\xi_1, \xi_2) = c(\xi_1 - \xi_2)$ denote its covariance function which is a function of $\xi = \xi_1 - \xi_2$ due to the stationarity. Since c is continuous and positive definite, it follows from Bochner's Theorem that

$$c(\xi) = \int_{\mathbb{R}^d} e^{i\xi\cdot\eta} \tau(\mathrm{d}\eta)$$

for some non-negative measure τ. Since c is integrable and symmetric, τ has a symmetric density f, which can be found using the inverse Fourier-transform. $\sqrt{f}$ is continuous and a member of $L^2(\mathbb{R}^d)$. Note: For a symmetric function defined on $\mathbb{R}^d$ the Fourier transform and its inverse are the same up to multiplication/division with the constant $(2\pi)^{d/2}$.

By the convolution theorem for the Fourier(-Plancerel) transform we get

$$\left(\widehat{\hat{\sqrt{f}} * \hat{\sqrt{f}}}\right)^{-1} = \widehat{\hat{\sqrt{f}}}^{-1} \cdot \widehat{\hat{\sqrt{f}}}^{-1} = f,$$

thus

$$\hat{\sqrt{f}} * \hat{\sqrt{f}}(\xi) = c(\xi).$$

Put $k = \hat{\sqrt{f}}$ and let L denote a homogeneous Lévy basis, with characteristic triplet $(0, 1, 0)$. Then, since the covariance function for $\int k dL$ is equal to $k * k$, our proof is complete. □

In [29, Theorem 3], conditions for ergodicity is given in the special case of a stationary LGCP. Note that under (40) $c(\xi) \to 0$ for $\|\xi\| \to \infty$ and the conditions for ergodicity stated in [29, Theorem 3b] are satisfied.

6 Combinations of LCPs and LLCPs

The driving field of an LLCP has the form

$$\Lambda(\xi) = \exp\left(\int_{\mathrm{R}} k(\xi,\eta) L_J(\mathrm{d}\eta)\right) \exp\left(\int_{\mathrm{R}} k(\xi,\eta) L_G(\mathrm{d}\eta)\right) = \Lambda_J(\xi)\Lambda_G(\xi).$$

It seems natural to extend the model such that the kernels used in the jump part and the Gaussian part do not need to be the same. We thereby arrive at Cox processes with a driving field of the form

$$\Lambda(\xi) = \exp\left(\int_{\mathrm{R}} k(\xi,\eta) L_J(\mathrm{d}\eta)\right) \Lambda_G(\xi) \tag{41}$$

with Λ_G an arbitrary log Gaussian random field.

If L_J satisfies the regularity conditions of Theorem 2, we get

$$\Lambda(\xi) = \exp\left(d(\xi) + \sum_{(r,\eta)\in\Phi} r k(\xi,\eta) + Y(\xi)\right),$$

where $d(\xi)$ is a deterministic function, Φ is the atoms of a Poisson measure with intensity measure U and Y is an independent Gaussian field.

A related model can be found in [37] for modelling the positions of offsprings in a long-leaf pine forest given the positions of the parents and information about the topography. The model is in [37] formulated conditional on the positions η of the parents.

There are, of course, other possibilities for combining shot noise components and log Gaussian components in the driving field than the one suggested above. For instance, if L_J is a non-negative Lévy jump basis, we may consider Cox processes driven by

$$\begin{aligned}
\Lambda(\xi) &= \left(\int_{\mathrm{R}} k(\xi,\eta) L_J(\mathrm{d}\eta)\right) \Lambda_G(\xi) \\
&= \left(\int_{\mathrm{R}} k(\xi,\eta) a_0(\mathrm{d}\eta) + \sum_{(r,\eta)\in N} r k(\xi,\eta)\right) \Lambda_G(\xi),
\end{aligned} \tag{42}$$

cf. [45]. In [13, pp. 92-100], a Cox process model of the type described in (42) has been considered but now with the Gaussian field replaced by a Boolean field. Such a model will be able to produce shot noise point patterns with empty holes generated by the Boolean field.

7 Inhomogeneous LCPs and LLCPs

In [33], it has recently been suggested to introduce inhomogeneity into a Cox process such that the resulting process becomes second-order intensity reweighted stationary, see [1] for details. In this section, we describe four types of inhomogeneity. Only Type 3 leads to second-order intensity reweighted stationary processes.

We concentrate on SNCPs with $a_0 \equiv 0$, cf. Section 4.4.1. The interpretation of the type of inhomogeneity introduced may be facilitated by using the cluster representation of a shot noise Cox process X. It is not needed that the process of cluster centres (mothers) is locally finite in order to use this interpretation.

Example 4 (Type 1). The kernel is assumed to be homogeneous $k(\xi, \eta) = k(\xi - \eta)$ while the Lévy basis satisfies $V(\mathrm{d}r, \eta) = V(\mathrm{d}r)$, $\nu(\mathrm{d}\eta) = cf(\eta)\,\mathrm{d}\eta$. If the function f is non-constant, mothers will be unevenly distributed (according to ν) but the distribution of the clusters will not depend on the location in the sense that the distribution of $X_{(r,\eta)} - \eta$ does not depend on η. □

Example 5 (Type 2). The kernel is assumed to be homogeneous $k(\xi, \eta) = k(\xi-\eta)$ while the Lévy basis satisfies $V(\mathrm{d}r, \eta) = V(\mathrm{d}(\frac{r}{f(\eta)}))$ and $\nu(\mathrm{d}\eta) = c\,\mathrm{d}\eta$. In this case, the mothers will be evenly distributed while the distribution of the clusters may be location dependent. A model with (k, V) replaced by $k(\xi, \eta) = k(\xi - \eta)f(\eta)$ and $V(\mathrm{d}r, \eta) = V(\mathrm{d}r)$ will result in the same type of LCP. □

Example 6 (Type 3). The kernel is inhomogeneous of the form $k(\xi, \eta) = k(\xi - \eta)f(\xi)$ while the Lévy basis is homogeneous $V(\mathrm{d}r, \eta) = V(\mathrm{d}r)$ and $\nu(\mathrm{d}\eta) = c\,\mathrm{d}\eta$. The resulting LCP will be a location dependent thinning of a stationary LCP. This option has been discussed in [33, 44] with the following log-linear specification of the function f

$$f(\xi) = \exp(z(\xi) \cdot \beta).$$

Here, $z(\xi)$ is a list of explanatory variables and β a parameter vector. Note that Type 2 and 3 inhomogeneity will typically have a similar appearance. The reason is that they can be regarded as only differing in the specification of the kernel as either of the form

$$k(\xi, \eta) = k(\xi - \eta)f(\eta) \text{ (Type 2)}$$

or

$$k(\xi, \eta) = k(\xi - \eta)f(\xi) \text{ (Type 3)},$$

and $k(\xi - \eta)(f(\eta) - f(\xi))$ is only non-negligible if ξ and η are close enough so that $k(\xi - \eta)$ is non-negligible and at the same time there is an abrupt change in f between ξ and η. □

Example 7 (Type 4). Inhomogeneity may also be introduced into the process by a local scaling mechanism [17, 18]. Here, the kernel is inhomogeneous

$$k(\xi, \eta) = k\left(\frac{\xi - \eta}{f(\eta)}\right) \frac{1}{f(\eta)^d}.$$

while $V(\mathrm{d}r, \eta) = V(\mathrm{d}r)$ and $\nu(\mathrm{d}\eta) = c\,\mathrm{d}\eta/f(\eta)^d$. The inhomogeneity of the resulting point process can be explained by local scaling. $\square$

In Figure 4, examples of inhomogeneous LCPs of Type 1, 2 and 4 are given on $\mathrm{S} = \mathrm{R} = [0, 100] \times [0, 200]$. Here, k is the Gaussian kernel (14) with $\sigma = 2$, $c = 1/200$ and V is concentrated in $r = 18$. The inhomogeneity function f is linear in all three cases, $f(x, y) = y/100$.

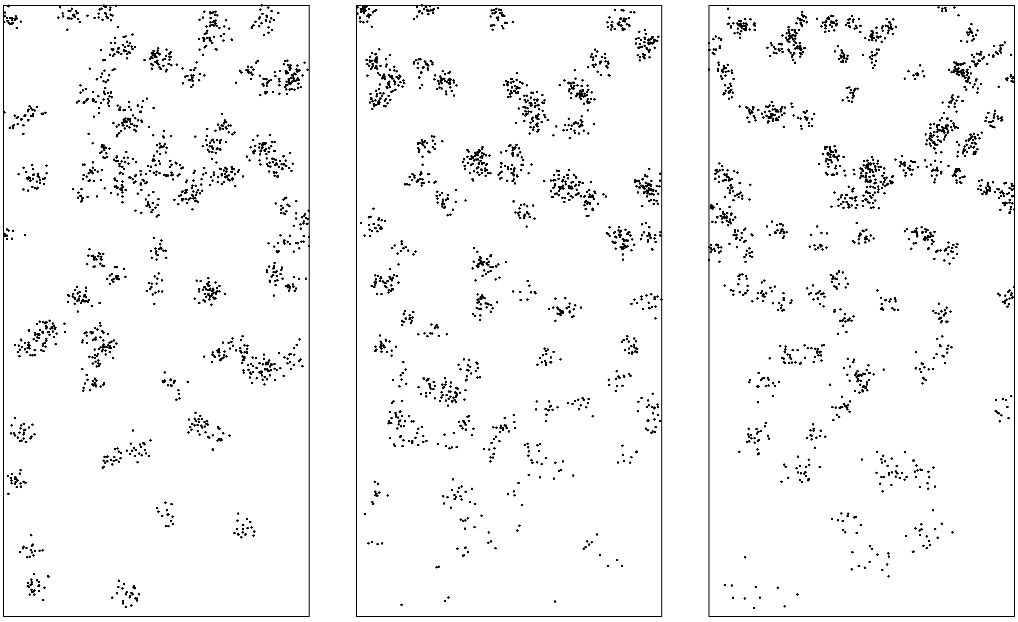

Figure 4: Examples of realizations of inhomogeneous LCPs. From left to right, Type 1, 2 and 4, respectively. For details, see the text.

Inhomogeneity may be introduced into an LSNCP by changing L or k as indicated in the examples above. Compared to LCPs, the effects are now multiplicative.

8 Discussion

During the last years, there has been some debate concerning which one of the two model classes (SNCP or LGCP) are most appropriate [28, 32, 38, 49]. The modelling framework described in the present paper provides the possibility for using models involving both SNCP and LGCP components and subsequently test whether it is appropriate to reduce the model to a pure SNCP model or a pure LGCP model. Figure 5 summarizes the most important model classes treated in the present paper. Below, we discuss a few additional issues.

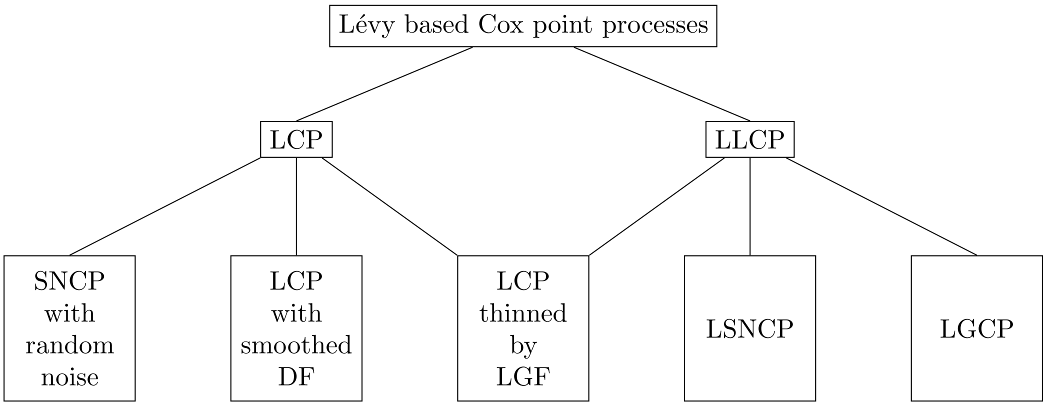

Figure 5: Overview of Lévy based Cox point processes. The abbreviations DF and LGF are used for a discrete random field and a log Gaussian random field, respectively. For further details, see the text.

8.1 Probability densities of LCPs and LLCPs

It is possible to derive an expression for the density of an LCP or an LLCP, using the methodology of Lévy bases. For instance, in the case of an LCP with $a_0 \equiv 0$, the density of X_B for $B \in \mathcal{B}_b(\mathrm{S})$ can be written as an expansion involving complete Bell polynomials evaluated at certain cumulants. The derivation of this result utilizes (17). Unfortunately, the expansion seems to be too complicated to be of practical use for inference. The same type of conclusion has been reached for likelihood inference in G-shot-noise Cox processes and in log Gaussian Cox processes, see [5, Section 4.2.1] and [29]. Closed form expressions for densities of other types of Cox processes are available, see [26].

8.2 Spatio-temporal extensions

The LCPs and LLCPs extend easily to spatio-temporal Cox processes. The set S on which the process is living is now a Borel subset of $\mathbb{R}^d \times \mathbb{R}$ where the last copy of $\mathbb{R}$ indicates time. The dependency on the past at time t and position x may be modelled using an ambit set

$$A_t(x), \quad x \in \mathbb{R}^d, t \in \mathbb{R},$$

satisfying

$$\begin{aligned} &(x,t) \in A_t(x) \\ &A_t(x) \subseteq \mathbb{R}^d \times (-\infty, t] \end{aligned}$$

A spatio-temporal LCP is then defined by a driving intensity of the form

$$\Lambda(x,t) = \int_{A_t(x)} k((x,t),(y,s)) L(\mathrm{d}(y,s)),$$

where L is a non-negative Lévy basis on $\mathrm{R} \subseteq \mathbb{R}^d \times \mathbb{R}$ and k is a non-negative weight function. Likewise, a spatio-temporal LLCP has driving field of the form

$$\Lambda(x,t) = \exp\left(\int_{A_t(x)} k((x,t),(y,s)) L(\mathrm{d}(y,s)) \right),$$

where L and k do not need to be non-negative anymore. Using the tools of Lévy theory, it is possible to derive moment relations as shown in the present paper for the purely spatial case [35]. This approach to spatio-temporal modelling is expected to be very flexible and has earlier been used with success in growth modelling [23], see also [22]. It will be interesting to investigate how it performs compared to earlier methods described in [6, 7, 8, 14].

8.3 Statistical inference

Statistical inference for Cox processes has been studied earlier in a number of papers, including [4, 19, 28, 32, 33, 43]. It remains to investigate to what degree known procedures, based on summary statistics, likelihood or Bayesian reasoning, can be adjusted to deal with LCPs and LLCPs. For a stationary LCP with $\Lambda = \rho_1 + \Lambda_2$, it is easy to determine the summary statistics F, G and J in terms of the corresponding characteristics F_2, G_2 and J_2 of the shot noise component with intensity field Λ_2. Thus,

$$\begin{aligned}
1 - F(r) &= \exp(-\rho_1|B(0,r)|)(1 - F_2(r)), \\
1 - G(r) &= \exp(-\rho_1|B(0,r)|)\left(\frac{\rho_1}{\rho_1+\rho_2}(1 - F_2(r)) + \frac{\rho_2}{\rho_1+\rho_2}(1 - G_2(r))\right), \\
J(r) &= \frac{\rho_1}{\rho_1+\rho_2} + \frac{\rho_2}{\rho_1+\rho_2}J_2(r).
\end{aligned}$$

However, in general simple expressions for G_2 and J_2 in terms of model parameters are not available. Likewise, it does not seem to be possible to derive general closed form expressions for F, G and J in the case of an LLCP.

In order to evaluate whether both a jump part and a Gaussian part is needed in an LLCP, we may consider a third order summary statistic, suggested in the paper [29] (for stationary point processes)

$$z(t) = \frac{1}{\pi^2 t^4}\int_{\|\xi\|\le t}\int_{\|\zeta\|\le t}\frac{\rho^{(3)}(\xi,\zeta)}{(\rho^{(1)})^3\, g(\xi)\, g(\zeta)\, g(\xi-\zeta)}\,\mathrm{d}\xi\,\mathrm{d}\zeta, \qquad t > 0, \tag{43}$$

where the following abbreviated notation is used due to the stationarity

$$\begin{aligned}
g(\xi_1,\xi_2) &= g(\xi_2 - \xi_1), \\
\rho^{(3)}(\xi_1,\xi_2,\xi_3) &= \rho^{(3)}(\xi_2-\xi_1, \xi_3-\xi_1).
\end{aligned}$$

When computing the integrand in $z(t)$ for an LLCP we obtain

$$\begin{aligned}
&\frac{\rho^{(3)}(\xi_1,\xi_2,\xi_3)}{(\rho^{(1)})^3\, g(\xi_1,\xi_2)\, g(\xi_2,\xi_3)\, g(\xi_1,\xi_3)} \\
&= \frac{\mathsf{E}(\prod_{i=1}^3 \Lambda_J(\xi_i))\,(\prod_{i=1}^3 \mathsf{E}\,\Lambda_J(\xi_i))}{\mathsf{E}(\Lambda_J(\xi_1)\Lambda_J(\xi_2))\ \mathsf{E}(\Lambda_J(\xi_2)\Lambda_J(\xi_3))\ \mathsf{E}(\Lambda_J(\xi_1)\Lambda_J(\xi_3))},
\end{aligned} \tag{44}$$

where $\Lambda_J(\xi) = \exp\left(\int_{\mathrm{R}} k(\xi,\eta)L_J(\mathrm{d}\eta)\right)$ is the part of the driving intensity originating from the pure jump part of the Lévy basis. Thus, this characteristic of X is not influenced by the Gaussian part of the model. In particular, $z \equiv 1$ for log Gaussian Cox processes. A non-parametric unbiased estimator of $z(t)$ has been derived in [29, Theorem 2].

Assessment of the full potential of the new modelling framework described in the present paper will also require more detailed studies of inhomogeneity and practical experience with concrete applications of the models.

Acknowledgements

We are grateful to Ole E. Barndorff-Nielsen for sharing his thoughts with us. This research has been supported by grants from the Carlsberg Foundation and the Danish Natural Science Research Council.

References

[1] Baddeley, A. J.; Moller, J.; Waagepetersen, R. (2000) Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statist. Neerlandica* **54**, 329–350.

[2] Barndorff-Nielsen, O.E., Blæsild, P. and Schmiegel, J. (2004). A parsimonious and universal description of turbulent velocity increments. *Eur. Phys. J. B* **41**, 345 –363.

[3] Barndorff-Nielsen, O.E. and Schmiegel, J. (2004). Lévy based tempo-spatial modelling; with applications to turbulence. *Uspekhi Mat. Nauk* **159**, 63–90.

[4] Best, N.G., Ickstadt, K. and Wolpert, R.L. (2000). Spatial Poisson regression for health and exposure data measured at disparate resolutions. *J. Amer. Statist. Assoc.* **95**, 1076–1088.

[5] Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. Appl. Prob.* **31**, 929–953.

[6] Brix, A. and Chadoeuf, J. (2002). Spatio-temporal modeling of weeds by shot-noise G Cox processes. *Biom. J.* **44**, 83–99.

[7] Brix, A. and Diggle, P.J. (2001). Spatio-temporal prediction for log-Gaussian Cox processes. *J. Roy. Statist. Soc. B* **63**, 823–841.

[8] Brix, A. and Møller, J. (2001). Space-time multitype log Gaussian Cox processes with a view to modelling weed data. *Scand. J. Statist.* **28**, 471–488.

[9] Coles, P. and Jones, B. (1991). A lognormal model for the cosmological mass distribution. *Monthly Notes of the Royal Astronomical Society* **248**, 1–13.

[10] Comtet, L. (1970). *Analyse Combinatoire. Tomes I, II.* Presses Universitaires de France, Paris.

[11] Cox, D.R. (1955). Some statistical models related with series of events. *J. Roy. Statist. Soc. B* **17**, 129–164.

[12] Daley, D.J. and Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes.* Springer Verlag, New York.

[13] Diggle, P.J. (2003). *Statistical Analysis of Spatial Point Patterns.* Second Edition. Hodder Arnold, London.

[14] Diggle, P., Rowlingson, B. and Su, T. (2005). Point process methodology for on-line spatio-temporal disease surveillance. *Environmetrics* **16**, 423–434.

[15] Federer, H. (1969). *Geometric Measure Theory.* Springer Verlag, Berlin.

[16] Ferguson, T.S. and Klass, M.J. (1972). A representation of independent increment processes without Gaussian components. *Ann. Math. Statist.* **43**, 1634–43.

[17] Hahn, U. (2007). *Global and Local Scaling in the Statistics of Spatial Point Processes.* Habilitationsschrift, Institut für Mathematik, Universität Augsburg.

[18] Hahn, U., Jensen, E.B.V., van Lieshout, M.N.M. and Nielsen, L.S. (2003). Inhomogeneous spatial point processes by location dependent scaling. *Adv. Appl. Prob. (SGSA)* **35**, 319–336.

[19] Heikkinen, J. and Arjas, E. (1998). Non-parametric Bayesian estimation of a spatial Poisson intensity. *Scand. J. Statist.* **25**, 435–450.

[20] Hellmund, G. (2005). *Lévy driven Cox processes with a View towards Modelling Tropical Rain Forest.* Master Thesis, Department of Mathematical Sciences, University of Aarhus, Denmark.

[21] Higdon, D. (2002). Space and space-time modeling using process convolutions. In *Quantitative Methods for Current Environmental Issues*, eds. C.W. Anderson, V. Barnett, P.C. Chatwin and A.H. El-Shaarawi, Springer, London, pp. 37–56.

[22] Jensen, E.B.V., Jónsdóttir, K.Y., Schmiegel, J. and Barndorff-Nielsen, O.E. (2007). Spatio-temporal modelling – with a view to biological growth. In *Statistics of Spatio-Temporal Systems*, eds. B.F. Finkenstadt, L. Held and V. Isham. Monographs on Statistics and Applied Probability, Chapman & Hall/CRC, pp. 47-75.

[23] Jónsdóttir, K.Y., Schmiegel, J. and Jensen, E.B.V. (2008). Lévy-based growth models. *Bernoulli* **14**, 62–90.

[24] Khoshnevisan, D. (2002). *Multiparameter Processes. An Introduction to Random Fields.* Springer Monographs in Mathematics. Springer-Verlag, New York.

[25] Matérn, B. (1960). *Spatial Variation.* Meddelanden **49**, Statens Skogsforskningsinstitut, Stockholm. Second edition. Lectures Notes Statist. **36**, Springer, Berlin, 1986.

[26] McCullagh, P. and Møller, J. (2006). The permanental process. *Adv. Appl. Prob. (SGSA)* **38**, 873–888.

[27] Møller, J. (2003a). Shot noise Cox processes. *Adv. Appl. Prob. (SGSA)* **35**, 614–640.

[28] Møller, J. (2003b). A comparison of spatial point process models in epidemiological applications. In *Highly Structured Stochastic Systems*, eds. P.J. Green, N.L. Hjort and S. Richardson. Oxford University Press, Oxford, pp. 264–268.

[29] Møller, J., Syversveen, A.R. and Waagepetersen, R.P. (1998). Log Gaussian Cox processes. *Scand. J. Statist.* **25**, 451–482.

[30] Møller, J. and Torrisi, G.L. (2005). Generalised shot noise Cox processes. *Adv. Appl. Prob. (SGSA)* **37**, 48–74.

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

[32] Møller, J. and Waagepetersen, R.P. (2002). Statistical inference for Cox processes. In *Spatial Cluster Modelling*, eds A.B. Lawson and D. Denison. Chapman & Hall/CRC, Boca Raton, pp. 37–60.

[33] Møller, J. and Waagepetersen, R.P. (2007). Modern statistics for spatial point processes. *Scand. J. Statist.* **34**, 643–684.

[34] Neyman, J. and Scott, E.L. (1958). Statistical approach to problems of cosmology. *J. R. Statist. Soc. B* **20**, 1-43.

[35] Prokešová, M., Hellmund, G. and Jensen, E.B.V. (2006). On spatio-temporal Lévy based Cox processes. In *Proceedings S^4G, International Conference on Stereology, Spatial Statistics and Stochastic Geometry, Prague, June 26 to 29 2006*, eds. R. Lechnerová, I. Saxl and V. Beneš. Union Czech Mathematicians and Physicists, Prague, pp. 111–116.

[36] Rajput, R.S. and Rosinski, J. (1989). Spectral representation of infinitely divisible processes. *Probab. Theory Relat. Fields* **82**, 451–487.

[37] Rathburn, S.R. and Cressie, N. (1994). A space-time survival point process for longleaf pine forest in southern Georgia. *J. Amer. Statist. Assoc.* **89**, 1164–1174.

[38] Richardson, S. (2003). Spatial models in epidemiological applications. In *Highly Structured Stochastic Systems*, eds. P.J. Green, N.L. Hjort and S. Richardson. Oxford University Press, Oxford, pp. 237–259.

[39] Rosinski, J. (2007). Simulation of Lévy processes. To appear in *Encyclopedia of Statistics in Quality and Reliability: Computational Intensive Methods and Simulation.* Wiley.

[40] Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University press, Cambridge.

[41] Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and its Applications.* Second edition. Wiley, Chichester.

[42] Thomas, M. (1949). A generalization of Poisson's binomial limit for use in ecology. *Biometrika* **36**, 18–25.

[43] Van Lieshout, M.N.M. and Baddeley, A.J. (1996). A nonparametric measure of spatial interaction in point patterns. *Statist. Neerlandica* **50**, 344–361.

[44] Waagepetersen, R.P. (2005). Discussion of the paper by Baddeley, Turner, Møller & Hazelton (2005). *J. Roy. Statist. Soc. B* **67**, 662.

[45] Waagepetersen, R.P. (2007). Personal communication.

[46] Walker, S. and Damien, P. (2000). Representations of Lévy processes without Gaussian components. *Biometrika* **87**(2), 477–483.

[47] Wiegand, T., Gunatilleke, S., Gunatilleke, N. and Okuda, T. (2007). Analyzing the spatial structure of a Sri Lankan tree species with multiple scales of clustering. *Ecology* **88**, 3088–3102.

[48] Wolpert, R.L. (2001). Lévy moving averages and spatial statistics. Lecture. Slides at `//citeseer.ist.psu.edu/wolpert01leacutevy.html`.

[49] Wolpert, R.L. and Ickstadt, K. (1998a). Poisson/gamma random fields models for spatial statistics. *Biometrika* **85**, 251–267.

[50] Wolpert, R.L. and Ickstadt, K. (1998b). Poisson/gamma random fields models for spatial statistics. Simulation of Lévy random fields. In *Practical Nonparametric and Semiparametric Bayesian Statistics*, eds. D. Dey, P. Müller and D. Sinha. Springer Lecture Notes in Statistics Vol. 133, Springer, New York, pp. 227-242.

Completely random signed measures

Gunnar Hellmund

T.N. Thiele Centre for Applied Mathematics in Natural Science
Department of Mathematical Sciences
University of Aarhus, Denmark

Abstract

Completely random signed measures are defined, characterized and related to Lévy random measures and Lévy bases.

1 Introduction

Completely random measures were defined in Kingman (1993). As described in Kingman (1993); Karr (1991) and more recently in Daley and Vere-Jones (2003, 2008) completely random measures are related to point process models, in particular Poisson cluster point processes. We make a natural extension of completely random measures to completely random signed measures and give a characterization of this class of signed random measures. It is shown that the class of Lévy random measures, introduced and used in Lévy adaptive regression kernel models Tu et al. (2006), and the class of Lévy bases, defined in Barndorff-Nielsen and Schmiegel (2004) and used in spatio-temporal modeling in Barndorff-Nielsen and Schmiegel (2004); Hellmund et al. (2008); Jónsdóttir et al. (2008), are natural extensions of completely random signed measures. Furthermore we show the assumption of infinitely divisibility in the definition of Lévy random measures and Lévy bases can be replaced by other very mild assumptions. The most basic concept involved in the definition of Lévy random measures and Lévy bases is thus independence.

2 Signed random measures

We let $\mathcal{X}$ denote a Borel subset of $\mathbb{R}^d$ for some $d \geq 1$ and $\mathcal{B} = \mathcal{B}(\mathcal{X})$ denote the trace of the Borel sigma algebra on $\mathcal{X}$. By $\mathcal{B}_b$ we denote the set of bounded Borel subsets of $\mathcal{X}$. A subset of $\mathcal{X}$ is bounded, if the closure of the set is compact.

Let $\mathcal{M}$ denote the set of signed Radon measures on $\mathcal{B}$, i.e. an element in $\mathcal{M}$ is a σ-additive set function, which takes finite values on compact sets, in particular on all bounded Borel subsets of $\mathcal{X}$. We let $\mathcal{M}^+$ denote the subset of $\mathcal{M}$ consisting of positive Radon measures. There are several different definitions of signed Radon measures in the literature, we use, what we believe is the most general, see Ash (1972).

We define a random signed measure M as a measurable mapping from a probability space $(\Omega, \mathcal{E}, P)$ into $(\mathcal{M}, \mathcal{F})$ where

$$\mathcal{F} = \sigma \{\pi_f | f \in C_c(\mathcal{X})\}, \quad \pi_f : \mathcal{M} \to \mathbb{R} : \mu \to \int_{\mathcal{X}} f(x) \mu(dx),$$

and $C_c(\mathcal{X})$ is the set of continuous functions on $\mathcal{X}$ with compact support.

Let $\mathcal{F}^+$ denote the trace of $\mathcal{F}$ on $\mathcal{M}^+$, then $\mathcal{F}^+ = \mathcal{B}(\mathcal{M}^+)$. A random measure is defined as a measurable mapping from a probability space into $(\mathcal{M}^+, \mathcal{F}^+)$.

The lemma below, used in the sequel, concerns the fixed atoms of a signed random measure. We say $x \in \mathcal{X}$ is a fixed atom of M if and only if

$$P(|M(\{x\})| > 0) > 0.$$

Lemma 2.1. *A signed random measure has at most countable many fixed atoms.*

Proof. Assume there are more than countable many fixed atoms, then there exist $\{x_n | n \geq 1\}$ contained in a bounded set, such that

$$\exists b > 0, a > 0 \forall n \geq 1 : P(|M(x_n)| > b) > a.$$

Thus $P(\limsup_n \{|M(x_n)| > b\}) \geq a$ and $\sum_n M(\{x_n\})$ cannot be convergent, which is a contradiction. □

3 A result on infinitely divisibility

Lemma 3.1. *Let M denote a stochastic process with index set $\mathcal{B}_b$ such that $(M(B_n))_{n\geq 1}$ are independent and*

$$M(\cup B_n) = \sum M(B_n)$$

P-a.s. for disjoint sets $(B_n)_{n\geq 1} \subset \mathcal{B}_b$ and $\cup B_n \in \mathcal{B}_b$.

Then $M(B)$ is infinitely divisible for any $B \in \mathcal{B}_b$, if the cumulant function of $M(A), A \in \mathcal{B}_B$ is of the form

$$\mathcal{C}\{M(A) \ddagger t\} = \log E\left[e^{itM(A)}\right] = \int_A f_{t,B}(x) \lambda_B(dx) \tag{1}$$

for some measurable function $f_{t,B} : \mathcal{X} \to \mathbb{C}$ for all $t \in \mathbb{R} \backslash \{0\}$ and an atom-less finite measure λ_B on $(B, \mathcal{B}_B)$, where $\mathcal{B}_B = \mathcal{B}(B)$.

Remark 3.2. For a given $t \in \mathbb{R}$ and $B \in \mathcal{B}_b$, because of independence, the cumulant transform defines a complex measure on $(B, \mathcal{B}_B)$. If $M(A)$ is zero with probability one on all sets in $\mathcal{B}_B$ with Lebesgue measure zero, then Lebesgue measure dominates the complex measure generated by the cumulant transform for all $t \in \mathbb{R}$ and thus, by Radon-Nikodym, condition (3.1) in the above lemma is fulfilled.

Proof. Let $B \in \mathcal{B}_b$ be given. Since $\lambda_B(B)$ is finite using Lemma 12.2, p.268 in Karlin and Studdun (1966) we can choose $(B_s)_{0\leq s\leq 1}$ such that $B_0 = \emptyset$, $B_{s'} \subseteq B_s, s' \leq s$, $B_1 = B$ and $\lambda_B(B_s) = s \cdot \lambda_B(B)$ for $s \in [0,1]$. It is clear the stochastic process $(M_s) = M(B_{s\wedge 1})_{s\geq 0}$ has independent increments and P-a.s. $M(B_0) = 0$. Furthermore (M_s) is stochastic continuous: Let $0 \leq s < 1$ be given, then for any $s' \in (s,1)$:

$$M_{s'} - M_s = M(B_{s'} \backslash B_s)$$

Since

$$\lambda_B(B_{s'} \backslash B_s) = (s' - s)\lambda_B(B) \to 0, s' \downarrow s,$$

we see

$$M(B_{s'} \backslash B_s) \tilde{\to} 0.$$

Thus in probability $M(B_{s'} \backslash B_s) \to 0$ as s' goes to s. Left continuity for $0 < s \leq 1$ is proved similar.

By definition 1.6 in Sato (2005) (M_s) is an additive process in law. Therefore $M_1 = M(B)$ is infinitely divisible, see Theorem 9.1 Sato (2005). □

4 Completely random signed measures

Definition 4.1. A completely random (signed) measure is a random (signed) measure M with independent values on disjoint sets, i.e. $\{M(A_n)\}$ are independent, if $\{A_n\}$ is a family of disjoint sets.

Corollary 4.2. *A completely random signed measure with cumulant transform satisfying condition (1) in Lemma 3.1 has infinitely divisible values.*

In the sequel we use the definition of a Poisson point process found in Kingman (1993).

Definition 4.3. A Poisson point process Φ on $\mathcal{Y}$, a Borel subset of $\mathbb{R}^l$ for some $l \geq 1$, is a random countable subset of $\mathcal{Y}$, such that

- The number of points $N(A)$ in a Borel subset A of $\mathcal{Y}$ is Poisson distributed with mean value $\mu(A)$, where μ is a measure on $\mathcal{B}(\mathcal{Y})$ (μ may be infinite on bounded sets!).
- Given disjoint sets $A_1, \ldots, A_n$ the random variables $N(A_1), \ldots, N(A_n)$ are independent.

The theorem below is stated in Kingman (1993), which also provides a sketch of a proof. We give a detailed proof, since we use important elements of the proof in the sequel.

Theorem 4.4. *Given a completely random measure M fulfilling condition (1) in Lemma 4.2 there exists a Radon measure m, an at most countable set of fixed atoms $\{x_i\}_{i\in I} \subset \mathcal{X}$, independent positive random variables $\{W_i\}$ and a Poisson point process Φ on $\mathcal{X} \times \mathbb{R}_+$, such for any $B \in \mathcal{B}_b$*

$$M(B) \sim m(B) + \sum_i W_i \cdot 1_B(x_i) + \sum_{(X_j,U_j)\in\Phi} U_j \cdot 1_B(X_j). \quad (2)$$

Proof. Using Lemma 2.1 we note the set of fixed atoms is at most countable. In the remaining part of the proof we assume M has no fixed atoms.

By Lemma 3.1 $M(B)$ is infinitely divisible for any $B \in \mathcal{B}_b$.

For any B in $\mathcal{B}_b$ there exists a constant $m(B) \in \mathbb{R}_+$ and a measure ν_B on $\mathbb{R}_+$, such that $\int (|x| \wedge 1)\, \nu_B(dx)$ is finite and for any $t \in \mathbb{R}_+$:

$$\lambda_t(B) = \log E\left[e^{itM(B)}\right] = m(B) \cdot it + \int_{(0,\infty]} (e^{itz} - 1)\nu_B(dt),$$

see Exc. 11, Chap. 15 in Kallenberg (2002).

By the properties of λ_t, $(A, B) \to \nu_B(A)$ is a bi-measure. There exists a unique σ-finite measure ν on $(\mathcal{X} \times \mathbb{R}_+, \mathcal{B} \otimes B([0,\infty)))$ satisfying

$$\nu(B \times C) = \nu_B(C)$$

for all $B \in \mathcal{B}_b$ and $C \in B((0,\infty])$, see (9.17) in Sato (2005). We see m defines a Radon measure on $\mathcal{B}$. Assume without loss of generality, that $m \equiv 0$.

Let Φ denote a Poisson point process on $\mathcal{X} \times \mathbb{R}_+$ with mean measure ν. Notice the number of points from Φ in a bounded set need not be finite. Define

$$M'(B) = \sum_{(X_j, U_j) \in \Phi} U_j \cdot 1_B(X_j)$$

for any $B \in \mathcal{B}_b$. Then for any $B \in \mathcal{B}_b$, using ν is σ-finite and applying Campbell's Theorem Kingman (1993) we get

$$\begin{aligned} E\left[\exp\left(itM'(B)\right)\right] &= \exp\left(\int_{B\times(0,\infty]} \left(e^{itz} - 1\right) \nu(dx \times dz)\right) \\ &= \exp\left(\int_{(0,\infty]} \left(e^{itz} - 1\right) \nu_B(dz)\right) \end{aligned}$$

Assuming without loss of generality, that M' is a random measure, M and M' are equal in distribution. ☐

We were not able to find a reference to the lemma below, concerning additive processes, thus a short proof is provided.

Lemma 4.5. *Given a continuous additive process (X_s), $s \geq 0$ with paths of bounded variation, $X_0 = 0$ and with characteristic triplet on the form $(A_s, 0, \gamma(s))$ we have $A_s \equiv 0$. i.e. the process has no Gaussian part and is deterministic, see Sato (2005) for notation.*

Proof. The total variation process $\left(V_s^X\right)$ of (X_s) is an increasing, continuous additive process. Set

$$Y_s = \frac{e^{-V_s^X}}{E\left[e^{-V_s^X}\right]}, s \geq 0$$

Then Y_s is a continuous martingale of bounded variation and thus constant, therefore V_s^X is deterministic, implying (X_s) is integrable and $\gamma(s)$ is of bounded variation.

$X_s - \gamma(s)$ therefore defines a continuous martingale of bounded variation, thus $X_s = \gamma(s)$ ☐

Lemma 4.6. *Let M denote a completely random signed measure and suppose the condition on the cumulant transform in Lemma 3.1 is fulfilled, then M has no Gaussian part.*

Proof. Let $B \in \mathcal{B}_b$ be given and let (M_s) denote the additive process in law constructed in the proof of Lemma 3.1 In the proof below, all references are to Sato (2005).

By Theorem 11.5 we can choose a *cadlag* modification of (M_s). Given a *cadlag* modification $\tilde{M}$ of M, $n \geq 1$ and the partition $0 = s_0 < s_1 = 1/n < \cdots < s_{n-1} = (n-1)/n < s_n = 1$ of the interval $[0,1]$ we have P-a.s.

$$\sum_i \left| \tilde{M}_{s_i} - \tilde{M}_{s_{i-1}} \right| = \sum_i \left| M_{s_i} - M_{s_{i-1}} \right| \leq |M|(B) < \infty,$$

since M is a random signed measure. Without loss of generality we assume M is an additive process of bounded variation (see Lemma 21.8 (i)).

Using Theorem 9.8 the law of M is uniquely determined by a characteristic triplet $(A_s, \nu_s, \gamma(s))$ each component satisfying some conditions given in the theorem. For every $s \geq 0$, ν_s is a Lévy measure on $\mathbb{R}$. Define $H = (0, \infty) \times \mathbb{R} \setminus \{0\}$ and let $\mathcal{B}(H)$ denote the Borel subsets of H.

By (19.1) we define

$$J(D, \omega) = \# \{ s > 0 | (s, M_s - M_{s-}) \in D \in \mathcal{B}(H) \}$$

Because of bounded variation we have (Lemma 21.8) for any $s > 0$

$$\int_{(0,s] \times \mathbb{R} \setminus \{0\}} |x| \, J(d(t,x), \omega) < \infty$$

as shown page 141_3-142^5 $\int_{\{|x| \leq 1\}} |x| \, \nu_s(dx) < \infty$ for all $s > 0$.

Using Theorem 19.3 and Lemma 21.8 there exist processes M^J and M^G, such that $M = M^J + M^G$ and M^G is P-a.s. an additive process, continuous in s with characteristic triplet $\left(A_s, 0, \gamma(s) - \int_{\{|x| \leq 1\}} x \nu_s(dx)\right)$ and of bounded variation, thus $A_s \equiv 0$ (see 4.5). $\square$

Theorem 4.7. *Given a completely random signed measure M fulfilling condition (1) in Lemma 3.1. Then for all $B \in \mathcal{B}_b$:*

$$M(B) \sim \mu^+(B) - \mu^-(B) + \sum_{i=1}^{k} W_i \cdot 1_B(x_i) + \sum_{(X_j, U_j) \in \Phi} U_j \cdot 1_B(X_j)$$

where Φ is a Poisson point process on $\mathcal{X} \times \mathbb{R}$, W_i are independent random variables, the $x_i, i \in I$ is an at most countable set of points in $\mathcal{X}$ and μ^+, μ^- are Radon measures.

Proof. Assume without loss of generality M has no fixed atoms. Applying Rajput and Rosinski (1989) Proposition 2.1 (see this reference for notation) for every $B \in \mathcal{B}_b$:

$$\mathcal{C}\{M(B) \ddagger t\} = it \cdot \left(a(B) - \int_{\{|z| \leq 1\}} z U_B(dz) \right) + \int_{\mathbb{R}} \left(e^{itz} - 1 \right) U_B(dz)$$

From the proof of the previous lemma, we have that

$$\int_{\{|z|\leq 1\}} (|z|)\, U_B\,(dz) < \infty. \tag{3}$$

Thus we can apply Campbell's Theorem Kingman (1993). In an argument, similar to the one found in the proof of Theorem 4.4, we can construct a Poisson point process Φ on $\mathcal{X} \times \mathbb{R}$, such that

$$\mathcal{C}\Big\{\sum_{(X_j,U_j)\in\Phi} U_j \cdot 1_B\,(X_j) \ddagger t\Big\} = \int_{\mathbb{R}} \big(e^{itz} - 1\big)\, U_B\,(dz)$$

(see Lemma 2.3 in Rajput and Rosinski (1989) for existence of a (mean) measure on $\mathcal{X} \times \mathbb{R}$ with the acquired properties)

It remains to note that $a\,(B) - \int_{\{|z|\leq 1\}} zU_B\,(dz)$ is finite for all bounded sets B. $\square$

5 Lévy bases

Definition 5.1. A stochastic process L indexed by $\mathcal{B}_b$ is called a Lévy basis, if $L\,(B)$ is infinitely divisible for all B in $\mathcal{B}_b$ and $L\,(B_n)$,$n \geq 1$ are independent and

$$L\,(\cup_n B_n) = \sum_n L\,(B_n)$$

P-a.s. for any sequence of disjoint sets $(B_n)_{n\geq 1} \subseteq \mathcal{B}_b$, $\cup B_n \in \mathcal{B}_b$.

Remark 5.2. The condition of infinite divisibility can be left out, if condition (1) in Lemma 3.1 is fulfilled.

It is proved in Rajput and Rosinski (1989) Lemma 2.3 that the cumulant transform of a Lévy basis L can be written as

$$\mathcal{C}\{L\,(dx) \ddagger t\} = \Big\{ita\,(x) - \tfrac{1}{2}t^2 b\,(x) + \int_{\mathbb{R}} \big(e^{itz} - 1 - itz \cdot 1_{[-1,1]}\,(z)\big)\, \rho\,(x, dz)\Big\}\lambda\,(dx) \tag{4}$$

λ is called the control measure and is σ-finite, a is a Borel measurable mapping into the real numbers and b is a density wrt. λ of a measure, $\rho\,(x,\cdot)$ is a Lévy measure for given x.

A Lévy basis such that $a \equiv 0$ and $\rho \equiv 0$ is called a purely Gaussian Lévy basis.

Theorem 5.3. *Let L denote a Lévy basis. L has the same distribution as the sum of a purely Gaussian Lévy basis and a completely random signed measure restricted to $\mathcal{B}_b$ (all terms being independent) if and only if for any $B \in \mathcal{B}_b$:*

$$\int_B \int_{\{|z|\leq 1\}} |z|\, \rho\,(x, dz)\, \lambda\,(dx) < \infty$$

Proof. From the proof of Lemma 4.6 we see the condition is necessary.

Assume the condition is fulfilled. Using the representation of the cumulant transform of the Lévy we can make the following rearrangements:

$$\mathcal{C}\{dx \ddagger t\} = \Big\{it\Big(a\left(x\right) - \int_{\mathbb{R}} \left(z \cdot 1_{[-1,1]}\left(z\right)\right) \rho\left(x, dz\right)\Big) \\ - \frac{1}{2}t^2 b\left(x\right) + \int_{\mathbb{R}} \left(e^{itz} - 1\right) \rho\left(x, dz\right)\Big\}\lambda\left(dx\right) \quad (5)$$

The non-Gaussian part of L has cumulant transform

$$\{it\left(a\left(x\right) - \int_{\mathbb{R}} \left(z \cdot 1_{[-1,1]}\left(z\right)\right) \rho\left(x, dz\right)\right) + \int_{\mathbb{R}} \left(e^{itz} - 1\right) \rho\left(x, dz\right)\}\lambda\left(dx\right) \quad (6)$$

Following the proof of Theorem 4.7 there is a completely random signed measure with cumulant transform (6). □

Definition 5.4. Lévy random measures are Lévy bases with no Gaussian part.

Corollary 5.5. *A Lévy random measure is a completely random signed measure if and only if for any $B \in \mathcal{B}_b$:*

$$\int_B \int_{\{|z|\leq 1\}} |z|\, \rho\left(x, dz\right) \lambda\left(dx\right) < \infty.$$

References

Ash, R. B., 1972. Real Analysis and Probability. Probability and Mathematical Statistics. New York: Academic Press.

Barndorff-Nielsen, O., Schmiegel, J., 2004. Levy-based spatial-temporal modelling, with applications to turbulence. Uspekhi Matematicheskikh Nauk 59 (1), 63–90.

Daley, D., Vere-Jones, D., 2003. An Introduction to the Theory of Point Processes, Vol. I. New York: Springer-Verlag.

Daley, D. J., Vere-Jones, D., 2008. An Introduction to the Theory of Point Processes, Vol. II. New York: Springer-Verlag.

Hellmund, G., Prokesova, M., Vedel, E., September 2008. Lévy based cox point processes. Advances in Applied Probability 40 (3), to appear.

Jónsdóttir, K., Schmiegel, J., Jensen, E., 2008. Lévy-based growth models. Bernoulli 14 (1), 62–90.

Kallenberg, O., 2002. Foundations of Modern Probability, second edition Edition. Probability and Its Applications. New York: Springer-Verlag.

Karlin, S. J., Studdun, W. J., 1966. Tchebycheff systems: With Applications in Analysis and Statistics. Pure and applied mathematics. New York: Wiley.

Karr, A. F., 1991. Point Processes and Their Statistical Inference. Bora Raton: CRC Press.

Kingman, J. F. C., 1993. Poisson Processes. Oxford Studies in Probability. Oxford: Oxford University Press.

Rajput, B., Rosinski, J., 1989. Spectral representations of infinitely divisible processes. Probability Theory and Related Fields 82, 451–487.

Sato, K., 2005. Lévy Processes and Infinitely Divisible Distributions. Cambridge studies in advanced mathematics. Cambridge: Cambridge University Press.

Tu, C., Clyde, M., Wolpert, R., 2006. Lévy adaptive regression kernels. Discussion Paper 2006-08, Duke University Department of Statistical Science.
URL `ftp://ftp.stat.duke.edu/pub/WorkingPapers/06-08.html`

Strong mixing with a view toward spatio-temporal estimating functions

Gunnar Hellmund, Rasmus Plenge Waagepetersen

T.N. Thiele Centre for Applied Mathematics in Natural Science
Department of Mathematical Sciences
University of Aarhus, Denmark

Department of Mathematical Sciences
University of Aalborg, Denmark

Abstract

We provide conditions for strong mixing of point processes on $\mathbb{R}^d, d \geq 1$ driven by shot-noise random fields. We discuss how the results provided can be readily applied in estimation of spatio-temporal models.

Introduction

Mixing is often an essential prerequisite for establishing theoretical properties of inference procedures for spatial processes. In this paper we establish results regarding strong mixing of spatial point processes. We give particular attention to point processes driven by shot-noise random fields.

For a particular class of estimating functions we discuss how the strong mixing results can be applied to establish consistency and asymptotic normality for the estimating function parameter estimates. This class of estimating functions e.g. contains estimating functions obtained from certain composite likelihood functions [Møller and Waagepetersen, 2007] which provide computationally easier alternatives to maximum likelihood estimation for spatial point processes.

The discussion of mixing and estimating functions is focused on the spatial case. Generalizations to spatio-temporal settings are however obvious.

So far, inference for point processes driven by random fields has mainly been restricted to stationary or second-order re-weighted stationary processes [Diggle, 2002, Møller and Waagepetersen, 2007, Waagepetersen and Guan, 2008]. However, our mixing results are applicable in a wider context of e.g. shot-noise processes with an inhomogeneous cluster center process. Such processes are neither stationary nor second-order re-weighted stationary. We conclude the paper by considering a specific example of estimation in a non-second-order reweighted stationary model, motivated by a point pattern dataset from the Yasuni tropical rain forest plot in Ecuador.

1 Point processes driven by random fields

A point process driven by a random field $\mathbf{X}$ on a subset of $\mathbb{R}^d$ is defined in terms of a random intensity function Λ, such that given $\Lambda = \lambda$, $\mathbf{X}$ is a Poisson point process with intensity λ. Realizations of these point processes are most often clustered point patterns. The intensity m and the second-order product density m^2 for a point process with random intensity Λ are given by

$$m(x) = E[\Lambda(x)] \text{ and } m^2(x,y) = E[\Lambda(x)\Lambda(y)], \quad x, y \in \mathbb{R}^d.$$

A sufficient condition for second-order re-weighted stationarity Baddeley et al. [2000] is that the pair correlation function $g(x,y) = m^2(x,y)/(m(x)m(y))$ is translation invariant, i.e. $g(x,y)$ only depends on $x-y$.

1.1 Mixing for point processes

For a point process $\mathbf{X}$ and given $W_0 \subseteq \mathbb{R}^d, \tau > 0$ define

$$W_\tau = \left\{x \in \mathbb{R}^d | \forall y \in W_0 : \|x-y\| \geq \tau \right\}$$

and

$$\mathcal{F}_0 = \sigma\left(\mathbf{X} \cap W_0\right), \quad \mathcal{F}_\tau = \sigma\left(\mathbf{X} \cap W_\tau\right).$$

The strong mixing coefficient between the σ-algebras $\mathcal{F}_0$ and $\mathcal{F}_\tau$ is given as

$$\alpha\left(\tau\right) = \sup\left\{\left|P\left(A_\tau \cap A_0\right) - P\left(A_\tau\right)P\left(A_0\right)\right| | A_\tau \in \mathcal{F}_\tau, A_0 \in \mathcal{F}_0\right\}.$$

The following theorem states that the mixing coefficient of a point process driven by a random field is bounded above by the mixing coefficient of the driving random field Λ.

Theorem 1.1. *For a given point process, let Λ denote the underlying random intensity. Set*

$$\alpha_\Lambda\left(\tau\right) = \sup\left\{\left|P\left(A_\tau \cap A_0\right) - P\left(A_\tau\right)P\left(A_0\right)\right| | A_\tau \in \mathcal{G}_\tau, A_0 \in \mathcal{G}_0\right\},$$

where $\mathcal{G}_\tau = \sigma(\Lambda\left(x\right), x \in W_\tau)$, and $\mathcal{G}_0 = \sigma(\Lambda\left(x\right), x \in W_0)$. Then

$$\alpha\left(\tau\right) \leq \alpha_\Lambda\left(\tau\right).$$

Proof. Let $A_\tau \in \mathcal{F}_\tau, A_0 \in \mathcal{F}_0$ be given. Let $\Lambda|_{W_\tau}$ and $\Lambda|_{W_0}$ denote the restrictions of Λ to W_τ and W_0. Then because $\mathbf{X}$ given Λ is a Poisson point process:

$$\begin{aligned}
&\left|P\left(A_\tau \cap A_0\right) - P\left(A_\tau\right)P\left(A_0\right)\right| = \left|E\left[P\left(A_\tau \cap A_0|\Lambda\right)\right] - E\left[P\left(A_\tau|\Lambda\right)\right]E\left[P\left(A_0|\Lambda\right)\right]\right| \\
&= \left|E\left[P\left(A_\tau|\Lambda\right)P\left(A_0|\Lambda\right)\right] - E\left[P\left(A_\tau|\Lambda\right)\right]E\left[P\left(A_0|\Lambda\right)\right]\right| \leq \alpha_\Lambda\left(\tau\right),
\end{aligned}$$

using Doukhan [(1') p. 3 1994], since $P\left(A_\tau|\Lambda\right) \leq 1$ is measurable wrt. $\mathcal{G}_\tau$, and $P\left(A_0|\Lambda\right) \leq 1$ is measurable wrt. $\mathcal{G}_0$. □

Frequently used random fields are log-Gaussian random fields and shot-noise random fields. Recently the literature on point processes with random intensity has focused on second-order reweighted stationary point processes with random intensity of the form

$$\Lambda(x) = \exp(\beta \cdot z(x))\Psi(x),$$

where $z(x)$ denotes covariate information, β is a vector of parameters and Ψ is a stationary random field. In the case of log-Gaussian random fields, popular choices of $\log(\Psi)$ are Gaussian random fields with exponential or Gaussian covariance function. In the case of the inhomogeneous Thomas point process [Waagepetersen, 2007]

$$\Psi(x) = \sum_{c \in \Phi} \frac{1}{2\pi\sigma^2} \exp(-\frac{|x-c|^2}{2\sigma^2}),$$

where Φ is a homogeneous Poisson point process. The random field Ψ for an inhomogeneous Thomas point process has Gaussian covariance function.

According to [Belyaev, 1959] random fields with Gaussian covariance function have smooth realizations and are not strong mixing: Using a simple power-series expansion it can easily be seen that the σ-algebras generated by the restriction of the random field to any open ball is identical to the σ-algebra generated by the field restricted to any set containing an open ball. However, as shown in the sequel, in the case of point processes driven by shot-noise random fields an analytic random field does not rule out strong mixing of the resulting point process.

1.2 The Whittle-Matérn class

Guttorp and Gneiting [2006] and Stein [1999] discusses the widely used and very flexible Whittle-Matérn class of isotropic covariance functions for random fields on $\mathbb{R}^d, d \geq 1$. An often used parametrization of this family is

$$c(r) = \frac{\sigma}{2^{\nu-1}\Gamma(\nu)}(\frac{2\nu^{1/2}r}{\rho})^{\nu}\mathcal{K}_{\nu}(\frac{2\nu^{1/2}r}{\rho}); \rho, \sigma, \nu > 0,$$

where $\mathcal{K}_\nu$ denotes the second modified Bessel function of order ν. For fixed ρ and σ the limiting covariance function of this family, as ν goes to infinity, is a Gaussian covariance function of the form

$$c(r) = \sigma \exp(-\frac{r^2}{\rho^2}),$$

For $\nu = \frac{1}{2}$ we obtain the exponential covariance function. In Matérn [1960] it is noted (p.17) that the Whittle-Matérn covariance function with $\nu \geq \frac{d}{2}$ can be seen as the correct generalization to d dimensions of type III distributions. Furthermore using the results in Matérn [1960] we see that point processes on $\mathbb{R}^d$ driven by shot-noise random fields on the form

$$\Psi(x) = \sum_{c \in \Phi} \varphi\gamma(\kappa|x-c|)^{\nu}\mathcal{K}_{\nu}(\kappa|x-c|); \varphi, \nu, \kappa > 0,$$

where

$$\gamma^{-1} = 2^{\nu+d}\pi^{d/2}\kappa^{-1}\Gamma(\nu + d/2),$$

and Φ is a homogeneous Poisson point process, have covariance functions belonging to the Whittle-Matérn family. The intensity function of these processes is equal to $\varphi \cdot \alpha$, where α denotes the intensity of Φ. Furthermore for $r = |x - y|$

$$\frac{Cov(\Psi(x), \Psi(y))}{\varphi^2\gamma^2\alpha} = \frac{2^{1-2\nu-d/2}\sqrt{\pi}}{\Gamma(2\nu + (d+1)/2)\kappa^{2\nu+d/2}} r^{2\nu+d/2}\mathcal{K}_{2\nu+d/2}(\kappa r).$$

Since $\mathcal{K}_\nu$ is exponential decaying it is a consequence of Theorem 1.2 below that this flexible class of shot-noise random fields can be used to generate point processes that are strong mixing.

1.3 Cluster processes

In this section we consider cluster point processes $\mathbf{X}$ with cluster centers generated from a Poisson process. Specifically we assume $\mathbf{X}$ is a point process with random intensity function

$$x \to \sum_{(\xi,\zeta)\in\Phi} k(\xi, \zeta, x), \tag{1.1}$$

where k is a positive function and Φ is a Poisson point process on $\mathbb{R}^d \times \mathbb{R}^l$ with intensity function

$$\chi : \mathbb{R}^d \times \mathbb{R}^l \to \mathbb{R}_+, l \geq 1.$$

In order to have a well-defined point process it is necessary that the mean intensity function exists and is locally integrable, i.e. for all $x \in \mathbb{R}^d$

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^l} k(\xi, \zeta, x)\,\chi(\xi, \zeta)\,d\zeta d\xi < \infty,$$

and for all bounded Borel subsets $B \subset \mathbb{R}^d$

$$\int_B\int_{\mathbb{R}^d}\int_{\mathbb{R}^l} k(\xi, \zeta, x)\,\chi(\xi, \zeta)\,d\zeta d\xi dx < \infty.$$

For introductions to cluster processes see [Daley and Vere-Jones, 2003, 2008, Møller, 2003, Møller and Torrisi, 2005]. The class of shot-noise Cox processes Møller [2003] is obtained with $l = 1$ and $k(\xi, \zeta, x)$ of the form $\zeta f(\xi, x)$. The generalized shot-noise Cox processes in Møller and Torrisi [2005] also have random intensity functions of the form (1.1) with $l = 2$ and $k(\xi, \zeta, x) = \zeta_1 f(\xi/\zeta_2, x/\zeta_2)/\zeta_2^d$ for some kernel f but Φ is not restricted to be Poisson.

Recall the notation for mixing coefficients and define

$$W_{c,\tau} = \left\{x \in \mathbb{R}^d | \forall y \in W_0 : \|x - y\| > \frac{\tau}{2}\right\}$$

and

$$W_{c,0} = \mathbb{R}^d \backslash W_{c,\tau}.$$

The following theorem generalizes a result in Waagepetersen and Guan [2008]. A proof is given in the Appendix.

Theorem 1.2.

$$\begin{aligned}\alpha(\tau) \leq & 4\int_{W_0}\int_{W_{c,\tau}\times\mathbb{R}^l} k(\xi,\zeta,x)\,\chi(\xi,\zeta)\,d(\zeta,\xi)\,dx \\ & + 4\int_{W_\tau}\int_{W_{c,0}\times\mathbb{R}^l} k(\xi,\zeta,x)\,\chi(\xi,\zeta)\,d(\zeta,\xi)\,dx.\end{aligned}$$

In particular, if $h(z)$ is $O\left(e^{-\gamma\cdot z}\right)$ for some positive constant γ and

$$\int_{\mathbb{R}^l} k(\xi,\zeta,x)\cdot\chi(\xi,\zeta)\,d\zeta \leq h(\|\xi - x\|),$$

then $\alpha(\tau)$ is $O\left(e^{-\nu\cdot\tau}\right)$ for some positive constant ν if W_0 is bounded.

The condition in the theorem trivially holds if k has compact support. Another example is a shot-noise Cox process with $k(\xi,\zeta,x) = \zeta\cdot f(x)\cdot f_k(\|x-\xi\|)$ where f is bounded, f_k is a Laplace or Gaussian kernel and $\int\zeta\cdot\chi(\xi,\zeta)\,d\zeta < \infty$. Certain strong mixing results [Guyon, 1995] only require that $\alpha(\tau) \in O(\tau^{-p})$.

2 Estimating functions

Maximum likelihood estimation is in general difficult for point processes with random intensity due to intractable likelihood functions. Instead parameter estimates are often obtained using minimum contrast estimation based on the K-function, which is well-defined in the stationary or second-order re-weighted stationary case [see Diggle, 2002, Møller and Waagepetersen, 2007, 2004, Waagepetersen and Guan, 2008]. In the case of cluster processes, which are not second-order re-weighted stationary, minimum contrast estimation based on the K-function is not possible and instead one may use certain composite likelihood functions as discussed below. Furthermore we discuss how strong mixing results become useful for establishing asymptotic results for a general class of estimating functions including those obtained by differentiating log composite likelihood functions.

2.1 Composite likelihoods

We assume the distribution $P_{\mathbf{X}}$ of $\mathbf{X}$ is defined in terms of a vector of parameters $\theta \in \Theta \subseteq \mathbb{R}^q$ for some $q \geq 1$. The intensity m and second-order product density m^2 is related to probabilities of occurrence of points or pairs of points and this leads to composite likelihoods of the form

$$\ell(\theta) = \frac{1}{|W|}\left(\sum_{x\in W}\ln m^1(x) - \int_W m^1(x)\,d(x)\right). \tag{2.1}$$

[Schoenberg, 2005, Waagepetersen, 2007] and

$$\ell_{pairwise}(\theta) = \frac{1}{|W|}\left(\sum_{\substack{x,y\in\mathbf{X}\cap W:\\ \|x-y\|<r}}^{\neq}\ln m^2(x,y) - \Big(\sum_{\substack{x,y\in\mathbf{X}\cap W:\\ \|x-y\|<r}}^{\neq}1\Big)\ln\int_{\substack{W\times W:\\ \|x-y\|<r}} m^2(x,y)d(x,y)\right) \tag{2.2}$$

[Guan, 2006]. Guan [2006] considered the stationary case but (2.2) may be applied in a much wider context of non stationary point processes. In (2.2), pairs of points with large inter point distance tend to add more noise than information and that is the reason for using only pairs of points whose inter point distance is less than the user-specified parameter r.

In Waagepetersen [2007] the Bernoulli composite log likelihood was introduced,

$$\ell_{Bernoulli}(\theta) = \frac{1}{|W|}\left(\sum_{\substack{x,y\in W:\\ \|x-y\|<r}}^{\neq} \ln m^2(x,y) - \int_{\substack{W\times W:\\ \|x-y\|<r}} m^2(x,y)\, d(x,y)\right). \quad (2.3)$$

Estimating functions are given by the gradients of the composite likelihoods. The estimating functions $\nabla\ell_{pairwise}$ and $\nabla\ell_{Bernouilli}$ are closely related, since $\ell_{Bernouilli}$ is obtained from $\nabla\ell_{pairwise}$ by replacing the last random term in $\nabla\ell_{pairwise}$ by its expectation.

Consider a disjoint partioning of $\mathbb{R}^d$ into equally sized squares ($d = 2$) or boxes ($d \geq 2$) S_p, $p \in \mathbb{Z}^d$. The above composite log likelihood functions and many other types of inference functions based e.g. on method of moments or on product densities m^k, $k > 2$, may be written as a sum

$$\frac{1}{a}\sum_p \ell_p(\theta), \quad (2.4)$$

where $\ell_p(\theta)$ only depends on $\mathbf{X}$ through $\mathbf{X}\cap C_p$ for a bounded neighborhood C_p of S_p (see figure (2.1)). The scaling factor a is often the size $|W|$ of W or the number of points p in W.

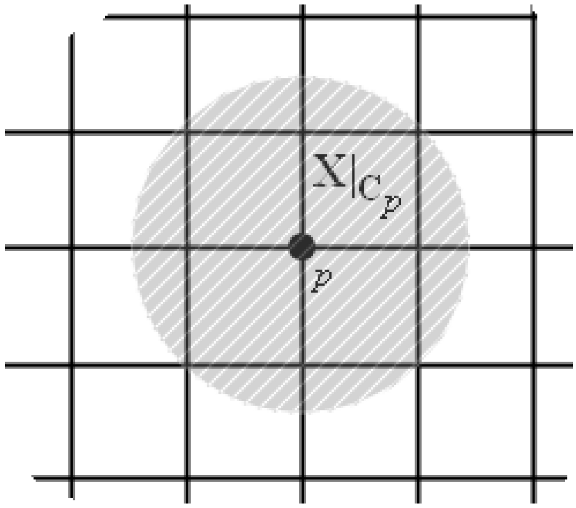

Figure 2.1: In a partition of $\mathbb{R}^d$ into squares, p marks the center of a cell S_p and C_p marks a neighborhood.

Regarding (2.3) for example,

$$\ell_p(\theta) = \sum_{\substack{x\in\mathbf{X}\cap S_p, y\in\mathbf{X}:\\ \|x-y\|<r}}^{\neq} \ln m^2(x,y) - \int_{\substack{S_p\cap W\times W:\\ \|x-y\|<r}} m^2(x,y)d(x,y)$$

and

$$C_p = \{y \in \mathbb{R} | \exists x \in S_p : \|x - y\| \leq r\}.$$

2.2 Mixing and asymptotic results

In this section we present a set of theorems useful for establishing asymptotic properties of estimates obtained e.g. from composite likelihoods as discussed in the previous section. Consider an increasing sequence of observation windows W_n and a sequence of estimating functions $u(n, \theta)$ of the form

$$u(n, \theta) = \frac{1}{a_n} \sum_p u_p(\theta) \tag{2.5}$$

using the notation of the previous section, where $u_p(\theta)$ only depends on $\mathbf{X}$ through $\mathbf{X} \cap C_p$ and C_p is a bounded neighborhood of S_p. The estimating function $u(n, \theta)$ might for instance be given by the gradient of one of the composite likelihoods defined in the previous section. An estimate $\hat{\theta}_n$ is obtained by solving $u(n, \theta) = 0$.

The primary application of strong mixing is to obtain a central limit theorem for a suitably scaled version of (2.5). Define the mixing coefficient

$$\alpha_{2,\infty}(m) = \sup_{A,B} \{\alpha(\sigma(u_p(\theta), p \in A), \sigma(u_p(\theta), p \in B))\},$$

where $A, B \subseteq \mathbb{Z}^d, |A| \leq 2, dist(A, B) \geq m$,

$$\alpha(\sigma(u_p(\theta), p \in A), \sigma(u_p(\theta), p \in B)) = \sup_{M_A, M_B} |P(M_A \cap M_B) - P(M_A) P(M_B)|$$

and

$$M_A \in \sigma(u_p(\theta), p \in A), M_B \in \sigma(u_p(\theta), p \in B).$$

Assume for all $\exists \delta > 0 \forall \theta \in \Theta$:

$$\sup_p E\left[\|u_p(\theta)\|^{2+\delta}\right] < \infty \tag{2.6}$$

and

$$\sum_{m \geq 1} m^{d-1} \alpha_{2,\infty}(m)^{\delta/2+\delta} < \infty. \tag{2.7}$$

It then follows from Theorem 3.3.1 in Guyon [1995] that the variance of $a_n^{1/2} u(n, \theta)$ is bounded and that $a_n^{1/2} u(n, \theta)$ is asymptotically normal, provided the variance matrix of $a_n^{1/2} u(n, \theta)$ stays positive definite as n tends to infinity.

In applications $u_p(\theta)$ is a measurable mapping of $\mathbf{X}|_{C_p}$, thus

$$\alpha(\sigma(u_p, p \in A), \sigma(u_p, p \in B)) \leq \alpha\left(\sigma\left(\mathbf{X}|_{\cup C_p}, p \in A\right), \sigma\left(\mathbf{X}|_{\cup C_p}, p \in B\right)\right),$$

where

$$\alpha\left(\sigma\left(\mathbf{X}|_{\cup C_p}, p \in A\right), \sigma\left(\mathbf{X}|_{\cup C_p}, p \in B\right)\right) = \sup_{M_A^{\mathbf{X}}, M_B^{\mathbf{X}}} \left|P\left(M_A^{\mathbf{X}} \cap M_B^{\mathbf{X}}\right) - P\left(M_A^{\mathbf{X}}\right) P\left(M_B^{\mathbf{X}}\right)\right|$$

and

$$M_A^{\mathbf{X}} \in \sigma\left(\mathbf{X}|_{\cup C_p}, p \in A\right), M_B^{\mathbf{X}} \in \sigma\left(\mathbf{X}|_{\cup C_p}, p \in B\right).$$

Mixing results for the process $\mathbf{X}$ are therefore crucial for establishing mixing of $u_p(\theta), p \in \mathbb{Z}^d$.

The following theorem is an example of how mixing of $u_p(\theta), p \in \mathbb{Z}^d$ may be used to establish consistency of roots $\hat{\theta}_n$ of u_n. Define

$$R = \left\{\theta \in \Theta | \lim_{n\to\infty} \|E_{\theta_0}[u(n,\theta)]\| = 0\right\}$$

and

$$R_\epsilon = \{\theta \in \Theta | \exists \theta_R \in R : \|\theta - \theta_R\| < \epsilon\},$$

where $\|\cdot\|$ is the Euclidean norm. We will assume $R = \{\theta_0\}$, i.e. asymptotically there is at most one root.

Theorem 2.1. *Suppose $u(n,\theta)$ is continuous and Θ is compact, given* (1)-(3) *below, $\hat{\theta}_n$ is consistent, i.e. in probability,*

$$\lim_{n\to\infty} \hat{\theta}_n = \theta_0.$$

if $a_n \uparrow \infty$ as $n \uparrow \infty$.

1. $\forall \epsilon > 0 \exists c > 0, N \geq 1 \forall n \geq N : \inf_{\Theta \setminus R_\epsilon} \|E_{\theta_0}[u(n,\theta)]\| \geq c$
2. *the conditions* (2.6) *and* (2.7) *are satisfied*
3. $\exists h : x \downarrow 0 \Rightarrow h(x) \downarrow 0$ *and* $\exists B_n = O_P(1)$*:*

$$\|u(n,\theta) - u(n,\theta')\| \leq B_n \cdot h(\|\theta - \theta'\|), P-a.s.,$$

where $B_n = O_P(1)$ means that B_n is a sequence of stochastic variables not depending on θ and bounded in probability, i.e.

$$\forall \epsilon > 0 \exists \delta > 0 : \limsup_{n\to\infty} P(B_n \geq \delta) \leq \epsilon.$$

Proof. By Theorem 3.1 in Crowder [1986] it suffices to prove that

$$\sup_{\Theta \setminus R_\epsilon} \|u(n,\theta) - E_{\theta_0}[u(n,\theta)]\| \to_{n\to\infty} 0. \tag{2.8}$$

Note that

$$Var\left(u(n,\theta)_i\right) = Var\left(\frac{1}{a_n}\sum_{p\in W_n} u_p(\theta)_i\right) = \frac{1}{a_n^2}\left(\sum_p \sigma_p^2 + 2\sum_{p_1,p_2}^{\neq} \sigma_{p_1 p_2}\right),$$

where i refers to the $i'th$ estimating function and

$$\sigma_p^2 = Var\left(u_p(\theta)_i\right), \sigma_{p_1p_2} = Cov\left(u_{p_1}(\theta)_i, u_{p_2}(\theta)_i\right).$$

Using condition (2) and the first part of Theorem 3.3.1 in [Guyon, 1995] and the accompanying Remark (1) we obtain for all θ and $n \geq 1$

$$a_n Var\left(u\left(n,\theta\right)_i\right) < \infty.$$

Thus in L^2 and therefore in probability

$$\left\|u\left(n,\theta\right) - E_{\theta_0}\left[u\left(n,\theta\right)\right]\right\| \to 0, n \to \infty. \tag{2.9}$$

We therefore obtain (2.8) from condition (3), since this condition by Theorem 21.10 (i) in Davidson [1994] assures stochastic equicontinuity, which combined with (2.9) and Theorem 21.9 in [Davidson, 1994] provides uniform convergence $\square$

Theorem 3.3 in [Crowder, 1986] provides conditions, for which

$$V_n\left(\theta_0\right)^{-1/2} M_n\left(\theta_0\right)\left(\hat{\theta}_n - \theta_0\right)$$

has the same asymptotic distribution as

$$-V_n\left(\theta_0\right)^{-1/2} u\left(n,\theta_0\right) \tag{2.10}$$

where

$$\begin{aligned} V_n\left(\theta\right) &= Var_{\theta_0}\left(u\left(n,\theta\right)\right) \\ M_n\left(\theta\right) &= E_{\theta_0}\left[\frac{\partial}{\partial\theta^T} u\left(n,\theta\right)\right]. \end{aligned}$$

Asymptotic normality of $\hat{\theta}_n$ then follows from asymptotic normality of $u(n,\theta)$.

Asymptotic existence of $\hat{\theta}_n$ is ensured by the following condition

$$\forall\epsilon > 0 \exists\delta > 0, N \geq 1 \forall n \geq N : \inf_{\|\theta-\theta_0\|=\epsilon} \left(\theta - \theta_0\right)^T E_{\theta_0}\left[u\left(n,\theta\right)\right] > \delta$$

[Theorem 3.2 Crowder, 1986]. A related route to existence of a consistent sequence of roots is given by Theorem 2 in [Waagepetersen and Guan, 2008]. This approach also relies on a central limit theorem for $a_n^{1/2} u(n,\theta)$.

2.3 Estimation in non-stationary cluster point processes

In this example we provide some details for estimation in shot-noise Cox point processes $\mathbf{X}$ with random intensity on the form:

$$\Lambda(x) = \sum_{c\in\Phi} k(\|c - x\|),$$

where k is an Epanechnikov type kernel, i.e.

$$k(\|c - x\|) = \begin{cases} \frac{2\delta\gamma}{\pi}(1 - \gamma\|c - x\|^2), & \text{if } \|c - x\| < \frac{1}{\sqrt{\gamma}}; \\ 0, & \text{otherwise.} \end{cases}$$

where $\delta, \gamma > 0$ and Φ is a Poisson point process on $\mathbb{R}^2$ with intensity function

$$\lambda(c) = \exp(\beta_0 + \beta \cdot z(c)),$$

$\beta_0 \in \mathbb{R}, \beta \in \mathbb{R}^l$ and $z(c)$ is a set of l covariate fields.

The intensity function is given as

$$m^1(x) = \frac{2\delta\gamma}{\pi} \int_{\|c-x\|<\frac{1}{\gamma}} e^{\beta_0+\beta\cdot z(c)}(1 - \gamma \|c - x\|^2)dc; x \in \mathbb{R}^2.$$

This cluster point process is neither stationary nor second-order re-weighted stationary. Strong mixing of the point process is trivial since the kernel has bounded support, thus using a first order Bernoulli composite log likelihood (2.1) we can estimate the parameters β, γ and the quantity $\delta \cdot \exp(\beta_0)$.

We assume an enlarged observation window $\overline{W}$ contains covariate information.

2.3.1 Data

As a brief practical example we will look at positions of trees of the species *Rinorea lindeniana* in the Yasuni tropical forest plot in the Ecuadorian rain forest, data are kindly provided by professor Henrik Balslev, Department of Biology, University of Aarhus, Denmark.

The distribution of species in the $500m \times 500m$ plot in relation to habitat information was discussed in Valencia et al. [2004]. Specifically the observation window was divided into six different habitats as shown in figure 2.2. We will use an indicator of each of the six different habitats as covariate information. Furthermore since we need covariate information in an enlarged observation window, we will only look at trees within the $[60, 440] \times [60, 440]$ sub window - see the right hand side of figure 2.2. In this specific case β becomes a four-dimensional vector - the sixth habitat is ignored, since it has only a minor effect due to it's small size and location outside W, and the fifth habitat is used as set off. The estimated values were: $\beta = (-3.98, -0.35, -0.05, -1.87)$, $\gamma = \exp(-6.27)$ and $\delta \cdot \exp(\beta_0) = \exp(-3.53)$.

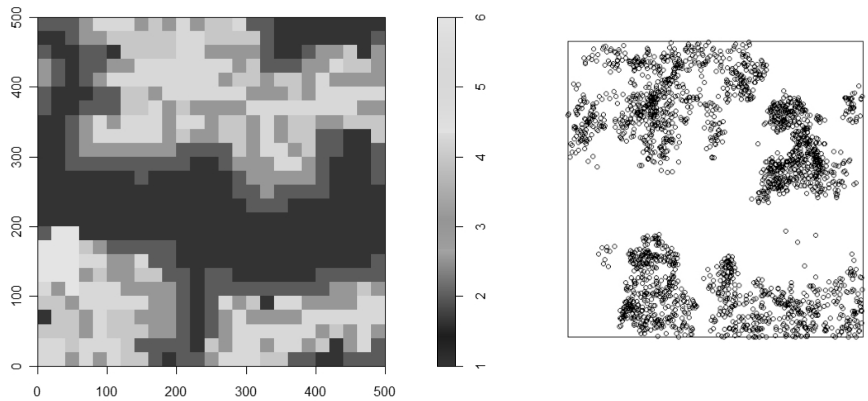

Figure 2.2: Left: Habitats in the Yasuni tropical forest plot. Right: Positions of Rinorea lindeniana in the $[60, 440] \times [60, 440]$ sub window.

2.3.2 Simulation

We simulated 100 point patterns of the above type using habitat covariate information from the Yasuni plot and parameters $\beta = (-3.98, -0.35, -0.05, -1.87)$, $\gamma = \exp(-6.27)$, $\delta = \exp(6.5 - 3.53)$ and $\beta_0 = -6.5$. See figure 2.3 for plots of four simulated point patterns. For each point pattern we estimated the parameters β and γ and the quantity $\delta \cdot \exp(\beta_0)$. As expected estimates are well-behaved (see figure 2.4). The empirical mean values obtained from these simulations were $\overline{\beta} = (-4.04, -0.39, -0.08, -1.78)$, $\overline{\ln(\gamma)} = -5.48$, $\overline{\ln(\delta \cdot \exp(\beta_0))} = -3.54$. Although the estimates of γ seem to be significantly biased, they are within a reasonable distance from the true value.

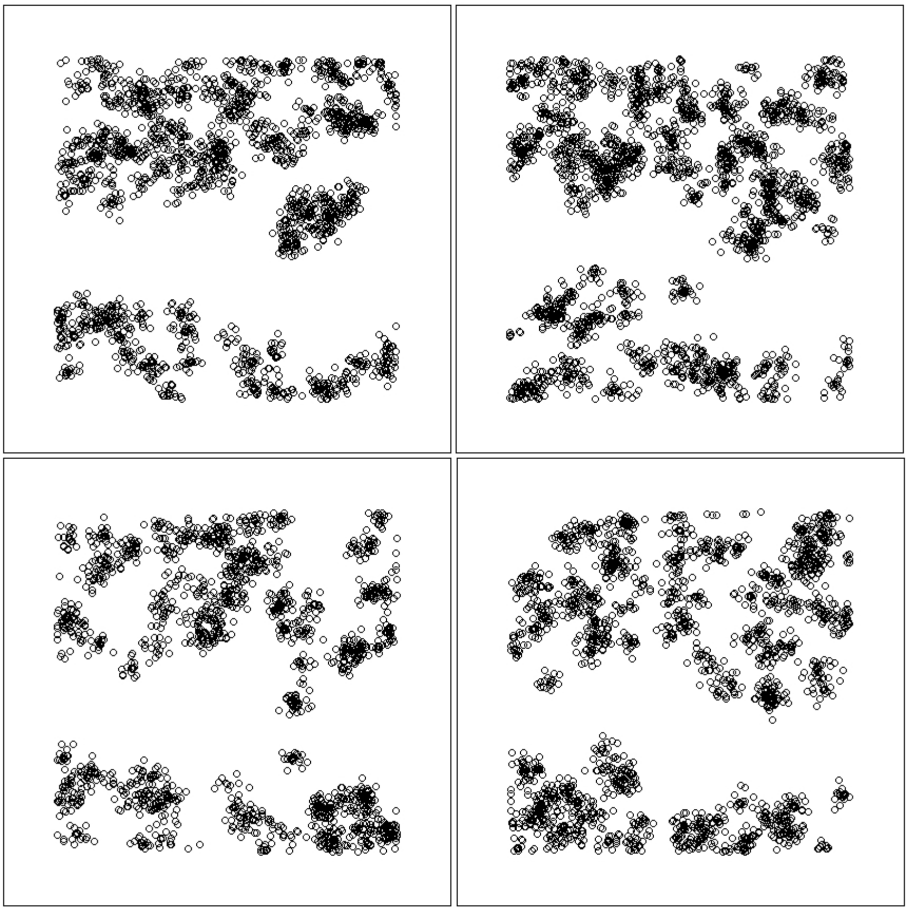

Figure 2.3: Simulated point patterns, using parameters $\beta = (-3.98, -0.35, -0.05, -1.87)$, $\gamma = \exp(-6.27)$, $\delta = \exp(6.5 - 3.53)$ and $\beta_0 = -6.5$.

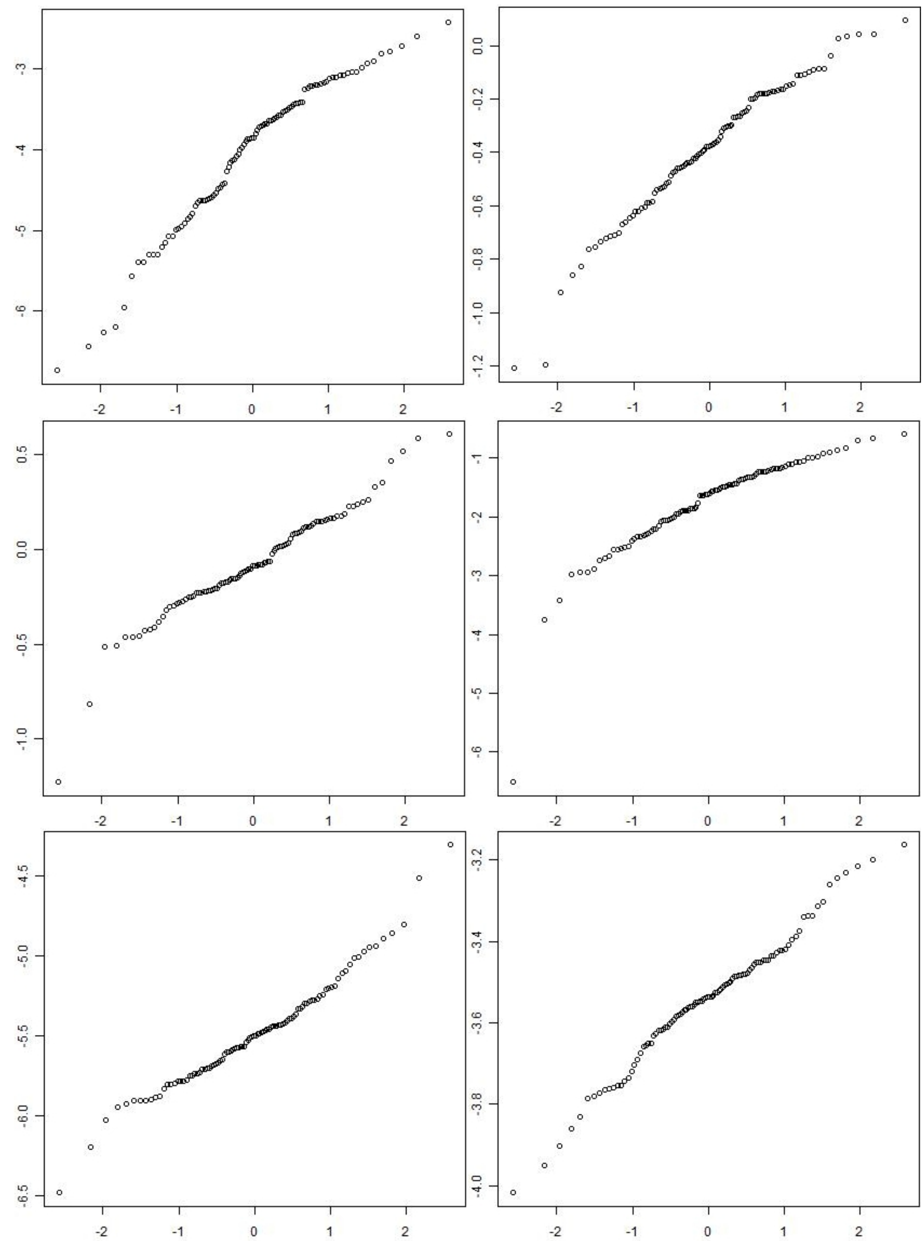

Figure 2.4: QQ-plots of parameter estimates obtained using first order Bernoulli composite likelihood. From above left: β the 4-dimensional covariate parameter vector, $\ln(\gamma), \ln(\delta \cdot \exp(\beta_0))$.

Discussion

We have derived useful conditions for strong mixing of cluster point processes.

A very important application of the mixing results was given in section 2, through a discussion of asymptotic results for estimating functions for spatial data. The results in section 2 may be further generalized and adapted to other dependence structures using Guyon [1995] and Davidson [1994].

The asymptotic results found above gives a theoretically justification for method of moments and composite likelihood estimation methods in (non-stationary) point processes driven by random fields. The shot-noise random fields introduced in section 1.2 may provide a flexible class for use in descriptive statistics based on observed point patterns and covariate information, while models that are not second-order re-weighted stationary may occur when trying to describe the dynamics behind some clustered point patterns. Shot-noise random fields are fairly simple and very fast to simulate. Fast estimation procedures and fast simulation routines provide evident opportunities for simulation based inference. Subsampling methods as described in Heagerty and Lumley [2000] can easily be adopted.

Throughout our discussion we focused on spatial estimating functions in order to provide a clear presentation. However, the setup can obviously be generalized to spatio-temporal settings if time is interpreted as one of the spatial dimensions.

Appendix

Proof of theorem 1.2: It is well known that $\mathbf{X}$ can be seen as a union of independent Poisson point processes

$$\mathbf{X} = \cup_{(\xi,\zeta)\in\Phi}\mathbf{X}_{(\xi,\zeta)},$$

where $\mathbf{X}_{(\xi,\zeta)}$ has intensity function $k(\xi,\zeta,x)$.

Let $\mathbf{X}_\tau$ and $\mathbf{X}_0$ denote Poisson cluster processes, such that

$$\mathbf{X}_\tau = \cup_{(\xi,\zeta)\in\Phi|_{W_{c,\tau}\times\mathbb{R}^l}}\mathbf{X}_{(\xi,\zeta)}$$

and

$$\mathbf{X}_0 = \cup_{(\xi,\zeta)\in\Phi|_{W_{c,0}\times\mathbb{R}^l}}\mathbf{X}_{(\xi,\zeta)}.$$

Then $\mathbf{X} = \mathbf{X}_\tau \cup \mathbf{X}_0$, furthermore $\mathbf{X}_\tau$ and $\mathbf{X}_0$ are independent.

Let G_τ and G_0 denote elements in the sigma-algebra $\mathcal{N}_{lf}$, such for given $A_\tau \in \mathcal{F}_\tau$ and $A_0 \in \mathcal{F}_0$

$$A_\tau = \{\mathbf{X}\cap W_\tau \in G_\tau\}$$

and

$$A_0 = \{\mathbf{X}\cap W_0 \in G_0\}.$$

Where

$$\mathcal{N}_{lf} = \sigma\left\{M \subset \mathbb{R}^d | \forall B \in \mathcal{B}_b\left(\mathbb{R}^d\right) \exists m \geq 0 : \#\left(M \cap B\right) = m\right\},$$

and $\mathcal{B}_b\left(\mathbb{R}^d\right)$ are the bounded Borel subsets of $\mathbb{R}^d$.

Define

$$A_\tau^* = \{\mathbf{X}_\tau \cap W_\tau \in G_\tau\}, A_0^* = \{\mathbf{X}_0 \cap W_0 \in G_0\}$$

and

$$B_\tau = \{\mathbf{X}_\tau \cap W_0 = \emptyset\}, B_0 = \{\mathbf{X}_0 \cap W_\tau = \emptyset\}$$

Because of independence between $\mathbf{X}_\tau$ and $\mathbf{X}_0$ we see, that

$$\begin{aligned} P(A_\tau \cap A_0 \cap B_\tau \cap B_0) &= P(A_\tau^* \cap B_\tau) P(A_0^* \cap B_0) \\ P(A_\tau \cap B_\tau \cap B_0) &= P(A_\tau^* \cap B_\tau) P(B_0) \\ P(A_0 \cap B_0 \cap B_\tau) &= P(A_0^* \cap B_0) P(B_\tau). \end{aligned}$$

Thus

$$\begin{aligned} & |P(A_\tau \cap A_0) - P(A_\tau) P(A_0)| \\ = & |P(A_\tau^* \cap B_\tau) P(A_0^* \cap B_0) + P\left(A_\tau \cap A_0 \cap \left(B_\tau^C \cup B_0^C\right)\right) - P(A_\tau) P(A_0)| \\ \leq & P\left(B_\tau^C\right) + P\left(B_0^C\right) + P(A_\tau^* \cap B_\tau) P(A_0^* \cap B_0)(1 - P(B_0) P(B_\tau)) \\ & + P\left(A_\tau \cap \left(B_\tau^C \cup B_0^C\right)\right) + P\left(A_0 \cap \left(B_\tau^C \cup B_0^C\right)\right) \\ \leq & 4P\left(B_\tau^C\right) + 4P\left(B_0^C\right), \end{aligned}$$

since

$$1 - P(B_0) P(B_\tau) = P\left(B_0^C \cup B_\tau^C\right).$$

Using Markov's Inequality

$$\begin{aligned} P\left(B_\tau^C\right) &= P(\#(\mathbf{X}_\tau \cap W_0) \geq 1) \leq E[\#(\mathbf{X}_\tau \cap W_0)] \\ &= \int_{W_0} \int_{W_{c,\tau} \times \mathbb{R}^l} k(\xi, \zeta, x) \chi(\xi, \zeta) \, d(\zeta, \xi, x) \end{aligned}$$

and

$$P\left(B_0^C\right) \leq \int_{W_\tau} \int_{W_{c,0} \times \mathbb{R}^l} k(\xi, \zeta, x) \chi(\xi, \zeta) \, d(\zeta, \xi, x)$$

References

A. J. Baddeley, J. Møller, and R. Waagepetersen. Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, 54:329–350, 2000.

Yu. K. Belyaev. Analytic random processes. *Theory of Probability and its Applications*, 4(4):402–409, 1959.

M. Crowder. On consistency and inconsistency of estimating equations. *Econometric Theory*, 2(3):305–330, 1986.

D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes Vol I+II.* Springer (New York), 2003, 2008.

J. Davidson. *Stochastic limit theory.* Oxford University Press (Oxford), 1994.

P.J. Diggle. *Statistical Analysis of Spatial Point Patterns.* Hodder Arnold (London), 2002.

P. Doukhan. *Mixing: Properties and examples.* Springer-Verlag (New York), 1994.

Y. T. Guan. A composite likelihood approach in fitting spatial point process models. *Journal Of The American Statistical Association*, 101(476):1502–1512, 2006.

P. Guttorp and T. Gneiting. Studies in the history of probability and statistics xlix on the matérn correlation family. *Biometrika*, 93(4):989–995, 2006.

X. Guyon. *Random Fields on a Network: Modeling, Statistics, and Applications.* Springer-Verlag (New York), 1995.

P. J. Heagerty and T. Lumley. Window subsampling of estimating functions with application to regression models. *Journal of the American Statistical Association*, 95:197–211, 2000.

B. Matérn. *Spatial Variation.* Statens Skogsforskningsinstitut (Stockholm), 1960.

J. Møller. Shot noise Cox processes. *Advances In Applied Probability*, 35(3):614–640, 2003.

J. Møller and G. L. Torrisi. Generalised shot noise Cox processes. *Advances In Applied Probability*, 37(1):48–74, 2005.

J. Møller and R. P. Waagepetersen. Modern statistics for spatial point processes. *Scandinavian Journal Of Statistics*, 34(4):643–684, 2007.

J Møller and R.P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes.* Chapman & Hall/CRC (Boca Raton), 2004.

F. P. Schoenberg. Consistent parametric estimation of the intensity of a spatial-temporal point process. *Journal Of Statistical Planning And Inference*, 128(1): 79–93, 2005.

M.L. Stein. *Interpolation of Spatial Data: Some Theory for Kriging.* Springer-Verlag (New York), 1999.

R. Valencia, R. B. Foster, G. Villa, R. Condit, J. C. Svenning, C. Hernandez, K. Romoleroux, E. Losos, E. Magard, and H. Balslev. Tree species distributions and local habitat variation in the amazon: large forest plot in eastern ecuador. *Journal Of Ecology*, 92(2):214–229, 2004.

R. Waagepetersen and Y. Guan. Two-step estimation for inhomogeneous spatial point processes. submitted, 2008.

R. P. Waagepetersen. An estimating function approach to inference for inhomogeneous neyman-scott processes. *Biometrics*, 63(1):252–258, 2007.