Blažena Frcalová; Viktor Beneš

Spatio-temporal modelling of a Cox point process sampled by a curve, filtering and Inference

Kybernetika, Vol. 45 (2009), No. 6, 912--930

Persistent URL: http://dml.cz/dmlcz/140029

KYBERNETIKA — VOLUME *45* (2009), NUMBER 6, PAGES 912 – 930

SPATIO–TEMPORAL MODELLING OF A COX POINT PROCESS SAMPLED BY A CURVE, FILTERING AND INFERENCE

Blažena Frcalová and Viktor Beneš

The paper deals with Cox point processes in time and space with Lévy based driving intensity. Using the generating functional, formulas for theoretical characteristics are available. Because of potential applications in biology a Cox process sampled by a curve is discussed in detail. The filtering of the driving intensity based on observed point process events is developed in space and time for a parametric model with a background driving compound Poisson field delimited by special test sets. A hierarchical Bayesian model with point process densities yields the posterior. Markov chain Monte Carlo "Metropolis within Gibbs" algorithm enables simultaneous filtering and parameter estimation. Posterior predictive distributions are used for model selection and a numerical example is presented. The new approach to filtering is related to the residual analysis of spatio-temporal point processes.

Keywords: Cox point process, filtering, spatio-temporal process

AMS Subject Classification: 60G55, 60D05, 62M30

1. INTRODUCTION

Spatio-temporal point processes (see [10, 26]) are of great interest in applications. In seismological data studies they help to solve problems in the prediction of large earthquakes with clusters of aftershocks [20]. The local random nature of forest fire ignitions as well as its dynamics in time enable to idealize the occurrence of fires as a space-time point process [19, 22]. In epidemiology records of spatial locations of incident cases are naturally developing in time, cf. [7] motivated by gastrointestinal infections. Experience dependent changes and dynamic representations of biological signals are characteristic features of neural systems. E.g. in [11] the evaluation of spike trains from rat hippocampus enables to detect a temporal evolution (caused by adaptation) of locations of firing fields. A classical approach to stochastic modelling is to consider a temporal point process with spatial marks [10, 11, 20, 22] and to employ the conditional intensity. In the present paper we develop another approach which consists in the modelling of spatio-temporal events, cf. [6, 7, 13, 19]. Specially, it is devoted to Cox processes which are suitable in a variety of situations when overdispersion and clustering takes place. A parametric model based on Lévy jump basis [13] is used for the driving intensity. Such tools are applied in finance [2, 9], physics (turbulence) [3], agriculture [6]. In hierarchical models they may serve as priors [8].

The nonlinear filtering problem for Cox point processes consists in the inference of random driving intensity based on observed events of the process. It was studied for temporal point processes by many authors, early solutions [5, 15, 25] were based on stochastic differential equations which led in practice to serious numerical difficulties. These references also contain statistical techniques concerning the parameter estimation and model testing.

First attempts to filtering of spatio-temporal point processes [12] were still based on stochastic differential equations. Modern approaches use stochastic simulations, either sequential Monte Carlo [11] or MCMC, typically in the Bayesian paradigm. In [7] the log-Gaussian spatio-temporal Cox point process is investigated and the values of driving intensity evaluated on a grid using the Markov property. Filtering and transition together enable prediction. The hierarchical Bayesian approach to filtering was used for temporal point processes with known parameters in [17]. Representation of the driving intensity by finite point processes and the use of their density with respect to Poisson process in Markov chain Monte Carlo (MCMC) leads to simple Metropolis birth-death algorithm, cf. [18, 24]. The spatio-temporal modeling developed below is more complex. Moreover, simultaneously the parameters of the model are estimated within MCMC. The method developed is directed to a biological application [11, 16], where a special case of a spatio-temporal Cox process, sampled by a curve, has to be investigated. We consider first a fixed curve and than develop the case of a random curve. The filtering algorithm is presented in a simulation study. It arises that in three-dimensions (space and time) with edge effect corrections the algorithm is computationally satisfactorily fast. Thanks to ergodicity properties of the chain the estimation based on posterior mean leads to acceptable results. The posterior predictive distributions enable to quantify the model selection. The residual analysis of spatio-temporal point processes was developed in [20], based on the conditional intensity which is not available in a closed form for Cox processes. By repeating the MCMC it is also possible to perform the residual analysis in our approach despite the fact that our model is based on the driving intensity rather than on the conditional intensity.

Section 2 of the paper is devoted to the theoretical background of Lévy based Cox processes of the type investigated. In Section 3 a Cox process on a curve in space and time is developed. This theory is extended to sampling of a spatio-temporal Cox process by means of a random curve in Section 4 together with the filtering algorithm. In Section 5 the model selection and parameter estimation is discussed and demonstrated in a synthetic example. Finally in Section 6 the residual analysis is described and brief conclusions end the paper.

2. BACKGROUND

Consider $\mathbb{R}^d$ with the Borel σ-algebra $\mathcal{B}^d = \mathcal{B}$. Let $Z = \{Z(A);\ A \in \mathcal{B}\}$ be an independently scattered random measure. That means for every sequence $\{A_n\}$

of disjoint sets in $\mathcal{B}$, the random vectors $Z(A_n)$ are mutually independent and $Z(\cup_n A_n) = \sum_n Z(A_n)$ almost surely. Assume that $Z(A)$ is moreover infinitely divisible for all $A \in \mathcal{B}$, in this case Z is called a Lévy basis.

The following background comes from [23]. A Lévy measure χ on $\mathcal{B}^1$ is defined by conditions $\chi(\{0\}) = 0$ and $\int_{\mathbb{R}}(|x|^2 \wedge 1)\chi(\mathrm{d}x) < \infty$. Let $(a, 0, \nu)$ be a generating triplet of a Lévy jump basis [13]. Here a is a signed measure, $\nu(\mathrm{d}x, A)$ is a Lévy measure for fixed $A \in \mathcal{B}$ and a measure on $\mathcal{B}$ in the second variable. It is interpreted as the mean number of jumps in A with sizes in $\mathrm{d}x$. Zero in the second place of the triplet implies that the cumulant transform defined as $\mathcal{C}\{\zeta \ddagger Z(A)\} = \log \mathrm{E}(e^{i\zeta Z(A)})$ is

$$\mathcal{C}\{\zeta \ddagger Z(A)\} = i\zeta a(A) + \int_{\mathbb{R}} \{e^{i\zeta x} - 1 - i\zeta x \mathbf{1}_{|x|\leq 1}\}\nu(\mathrm{d}x, A),$$

$\zeta \in \mathbb{R}$. It is important that ν can be factorized as

$$\nu(\mathrm{d}x, \mathrm{d}\xi) = \mu(\mathrm{d}x, \xi)U(\mathrm{d}\xi), \tag{1}$$

where $\mu(\mathrm{d}x, \xi)$ is a Lévy measure on $\mathbb{R}$ for fixed $\xi \in \mathbb{R}^d$ and $U(\mathrm{d}\xi)$ is a measure on $\mathcal{B}$. Then assuming that the density a' exists, $a(\mathrm{d}\eta) = a'(\eta)U(\mathrm{d}\eta)$, $\eta \in \mathbb{R}^d$, we can write

$$\mathcal{C}\{\zeta \ddagger Z'(\eta)\} = i\zeta a'(\eta) + \int_{\mathbb{R}} \{e^{i\zeta x} - 1 - i\zeta x \mathbf{1}_{|x|\leq 1}\}\mu(\mathrm{d}x, \eta), \tag{2}$$

$\zeta \in \mathbb{R}$, for an additive process $Z'(\eta)$. For a fine discussion about the correspondence of Z and Z' see [21].

An integral of a deterministic function f with respect to a Lévy basis is defined as a limit (in probability) of integrals of simple functions $f_n \to f$. Necessary and sufficient conditions for the existence are known [23].

Lemma 2.1. Assuming that the following integrals exist for a measurable function f, it holds

$$\mathcal{C}\{\zeta \ddagger \int_{\mathbb{R}^d} f \,\mathrm{d}Z\} = \int_{\mathbb{R}^d} \mathcal{C}\{\zeta f(\xi) \ddagger Z'(\xi)\}U(\mathrm{d}\xi). \tag{3}$$

We will apply Lévy bases to the theory of simple point processes in $\mathbb{R}^d$ [18] and specially in space and time [10]. Consider a Lévy basis Z on $\mathbb{R}^d$ with triplet $(a, 0, \nu)$ and assume that a nonnegative locally integrable random field is obtained as

$$\Lambda(\xi) = \int_{\mathbb{R}^d} g(\xi, \eta)\, Z(\mathrm{d}\eta),\ \xi \in \mathbb{R}^d, \tag{4}$$

where g is a measurable function on $\mathbb{R}^{2d}$. For a compound Poisson process Z' a sufficient condition for local integrability follows from the Campbell theorem [18]: the mean jump size has to be finite and $h(\xi) = \int g(\xi, \eta)\, U(\mathrm{d}\eta)$ should be an integrable function of ξ on each bounded set.

A Cox point process X with driving (random) measure Λ_m is a point process such that conditionally on $\Lambda_m = \lambda_m$ it is the Poisson process with intensity measure λ_m.

We will assume that the density Λ of Λ_m with respect to Lebesgue measure exists, it is called the driving intensity function. The generating functional of a point process is defined as

$$G(u) = \mathrm{E}\left(\prod_{i=1}^{N} u(x_i)\right),$$

for measurable functions $u : \mathbb{R}^d \mapsto [0,1]$ with bounded support, where x_i are the N events of the point process observed within the support of u. For a Cox process X with random driving intensity function $\Lambda(s),\ s \in \mathbb{R}^d$, the generating functional has form

$$G(u) = \mathrm{E}\exp\left(-\int_{\mathbb{R}^d} (1-u(\sigma))\Lambda(\sigma)\,\mathrm{d}\sigma\right).$$

Theorem 2.2. Consider a Lévy jump basis Z on $\mathbb{R}^d$ with triplet $(a, 0, \nu)$ and a nonnegative locally integrable random field (4). Then the generating functional of a Cox point process X driven by Λ is

$$G(u) = \exp\left[-\int_{\mathbb{R}^d} f(\xi)a'(\xi)\,U(\mathrm{d}\xi) \right. \tag{5}$$

$$\left. + \int_{\mathbb{R}^d}\int_{\mathbb{R}} \left(e^{-rf(\xi)} - 1 + rf(\xi)\mathbf{1}_{[-1,1]}(r)\right)\,\mu(\mathrm{d}r,\xi)\,U(\mathrm{d}\xi)\right],$$

where

$$f(\xi) = \int_{\mathbb{R}^d} (1-u(\sigma))g(\sigma,\xi)\,\mathrm{d}\sigma. \tag{6}$$

Proof. A direct consequence of Lemma 1, see [16]. $\square$

Corollary 2.3. Specially for $a'(\xi) = \int_{-1}^{1} r\mu(\mathrm{d}r,\xi)$ (zero drift) it holds

$$G(u) = \exp\left[\int_{\mathbb{R}^d}\int_{\mathbb{R}} (e^{-rf(\xi)} - 1)\mu(\mathrm{d}r,\xi)\,U(\mathrm{d}\xi)\right]. \tag{7}$$

The distribution of a point process is determined by void probabilities. The void probabilities

$$\mathrm{P}(X(D) = 0) = G(1-\mathbf{1}_D) = \mathrm{E}e^{-\Lambda(D)},\ D \in \mathcal{B},$$

have under the assumptions of Theorem 1 form (5) with $u = 1 - \mathbf{1}_D$, i.e.

$$f(\xi) = \int_D g(\sigma,\xi)\,\mathrm{d}\sigma.$$

Moment characteristics of a point process are obtained by means of differentiation of the generating functional, the intensity measure

$$M(D) = \mathrm{E}X(D) = -\frac{\partial}{\partial z} G(1 - z\mathbf{1}_D)\,|_{z=0}, \quad D \in \mathcal{B}^d, \tag{8}$$

and the factorial second moment measure

$$\alpha^{(2)}(C) = \mathrm{E} \sum_{\xi,\eta \in X}^{\neq} \mathbf{1}_{[(\xi,\eta)\in C]}, \quad C \subset \mathbb{R}^{2d}, \tag{9}$$

as

$$\alpha^{(2)}(D_1, D_2) = \frac{\partial}{\partial z_1}\frac{\partial}{\partial z_2} G(1 - z_1\mathbf{1}_{D_1} - z_2\mathbf{1}_{D_2})\,|_{z_1=z_2=0}, \tag{10}$$

$D_1, D_2 \in \mathcal{B}^d$.

By [3] positive Lévy bases have Lévy–Itô representation

$$Z(D) = \bar{a}(D) + \int_{\mathbb{R}_+} x\Phi(\mathrm{d}x, D),$$

where $\bar{a}$ is a diffuse measure on $\mathbb{R}^d$ and Φ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d$. This leads to an expression

$$\Lambda(\xi) = \int_{\mathbb{R}^d} g(\xi,\sigma)\left(\bar{a}(\mathrm{d}\sigma) + \int_{\mathbb{R}_+} r\Phi(\mathrm{d}r, \mathrm{d}\sigma)\right) \tag{11}$$

and a connection with the class of shot-noise Cox processes (SNCP), cf. [18].

The class of non-Gaussian Ornstein–Uhlenbeck processes was extended in [2] by means of superpositions to achieve possibly a long range dependence. For spatio-temporal Cox processes this property (still in temporal sense) can be studied by means of second order characteristics. Superposition for driving intensities

$$\Lambda = \Lambda_1 + \Lambda_2,$$

where Λ_i is driven according to (4) by Z_i, $i = 1, 2$ independent, respectively, leads to the corresponding relation

$$G(u) = G_1(u)G_2(u)$$

for Cox process generating functionals. Using (10) we obtain for $u = 1 - z_1\mathbf{1}_A - z_2\mathbf{1}_B$

$$\begin{aligned} \alpha_\Lambda^{(2)}(A, B) &= \alpha_{\Lambda_1}^{(2)}(A, B) + \alpha_{\Lambda_2}^{(2)}(A, B) \\ &+ \left[\frac{\partial}{\partial z_1}G_1(u)\frac{\partial}{\partial z_2}G_2(u) + \frac{\partial}{\partial z_1}G_2(u)\frac{\partial}{\partial z_2}G_1(u)\right]_{z_1=z_2=0}. \end{aligned} \tag{12}$$

In the following we will mainly study a special case of the model (4) suggested for the purpose of spatio-temporal modelling by [3]. They define an Ornstein–Uhlenbeck (OU) type process $\Lambda(t,\sigma)$, $t \in \mathbb{R}$ (time), $\sigma \in \mathbb{R}^d$ (space) by

$$\Lambda(t,\sigma) = \int_{-\infty}^{t} e^{\gamma(s-t)} Z(B_{s-t}(\sigma) \times \mathrm{d}s),\ \sigma \in \mathbb{R}^d,\ t \in \mathbb{R}, \tag{13}$$

$\gamma > 0$ a parameter, where Z is a Lévy basis and $\{B_s(\sigma)\},\ s \le 0$ is a family of subsets on $\mathbb{R}^d$ which we will assume to be of the form

$$B_s(\sigma) = \{\rho \in \mathbb{R}^d;\ \chi(\rho, \sigma) \le -\omega s\}$$

for a metric χ on $\mathbb{R}^d$, $\omega > 0$ is a parameter. A spatio-temporal Cox process driven by nonnegative locally integrable Ornstein–Uhlenbeck type process is denoted OUCP. Further Leb denotes the Lebesgue measure in $\mathbb{R}^d$.

Corollary 2.4. On $\mathbb{R}^d \times \mathbb{R}$ consider a Cox process X with driving intensity (13). Then the generating functional has form (5) with

$$f(s, \rho) = \int_s^\infty \int_{B_{s-t}(\rho)} (1 - u(t, \sigma)) e^{\gamma(s-t)} \mathrm{d}\sigma\, \mathrm{d}t. \tag{14}$$

Denote $D_t = \{\sigma \in \mathbb{R}^d;\ (t, \sigma) \in D\},\ t \in \mathbb{R}$. Void probabilities of X have form $G(1 - \mathbf{1}_D)$ in (5) with

$$f(s, \rho) = \int_s^\infty \mathrm{Leb}(B_{s-t}(\rho) \cap D_t) e^{\gamma(s-t)} \mathrm{d}t.$$

P r o o f. (13) is of type (4) with

$$g(\xi, \eta) = g((t, \sigma), (s, \rho)) = \mathbf{1}_{[-\infty, t]}(s) \mathbf{1}_{B_{s-t}(\sigma)}(\rho) e^{\gamma(s-t)} \tag{15}$$

and so

$$f(s, \rho) = \int_s^\infty \int_{\mathbb{R}^d} (1 - u(\sigma, t)) \mathbf{1}_{B_{s-t}(\sigma)}(\rho) e^{\gamma(s-t)} \mathrm{d}\sigma\, \mathrm{d}t$$

and using the properties of $B_s(\sigma)$ we obtain the result. □

Corollary 2.5. Let

$$\Lambda_j(t, \sigma) = \int_{-\infty}^t e^{\gamma_j(s-t)} Z_j(B_{s-t}(\sigma) \times \mathrm{d}s),\ j = 1, 2,$$

Z_j be independent identically distributed. Under the conditions (7) and $\nu(\mathrm{d}x, \mathrm{d}\xi) = \mu(\mathrm{d}x)\, \mathrm{d}\xi$ for the superposition $\Lambda = \Lambda_1 + \Lambda_2$ it holds

$$\alpha^{(2)}_\Lambda(A, B) = \alpha^{(2)}_{\Lambda_1}(A, B) + \alpha^{(2)}_{\Lambda_2}(A, B) + m_1^2 [F_1(A) F_2(B) + F_1(B) F_2(A)],$$

where $m_1 = \int_\mathbb{R} x \mu(\mathrm{d}x)$ and for $C = C_1 \times C_2$, $C_1 \subset \mathbb{R}$

$$F_j(C) = \int \int \int_{C_1 \cap [s, \infty]} e^{\gamma_j(s-t)} \mathrm{Leb}(B_{s-t}(\phi) \cap C_2)\, \mathrm{d}t\, \mathrm{d}\phi\, \mathrm{d}s.$$

P r o o f. Use (12), (7) and (14). □

3. SPATIO–TEMPORAL COX POINT PROCESS ON A CURVE

Consider a continuous map $y : [0,T] \mapsto \mathbb{R}^d$, where $[0,T] \subset \mathbb{R}$ is a compact interval. Denote

$$Y = \{(t, y_t), t \in [0,T]\}$$

the curve in $\mathbb{R}^{d+1}$. Further consider a nonnegative locally integrable random function

$$\Lambda = \{\Lambda(t,u), u \in \mathbb{R}^d, t \in [0,T]\} \tag{16}$$

of form (4), i.e. $\Lambda = \int g \,\mathrm{d}Z$. We define a spatio-temporal Cox point process X_Y with events on Y so that conditionally on a realization $\Lambda = \lambda$ the number of points in $Y \cap B,\ B \in \mathcal{B}$ within $0 \le t_1 < t_2 \le T$ is Poisson distributed with mean

$$\int_{t_1}^{t_2} \mathbf{1}_B(y_t)\lambda(t, y_t)\,\mathrm{d}t.$$

That means X_Y is a Cox point process with random driving measure

$$\Lambda_Y([t_1,t_2] \times B) = \int_{t_1}^{t_2} \mathbf{1}_B(y_t)\Lambda(t, y_t)\,\mathrm{d}t. \tag{17}$$

Denote the intensity measure $M(\cdot) = \mathrm{E}X_Y(\cdot) = \mathrm{E}\Lambda_Y(\cdot)$. We will assume in the following zero drift condition (7) for the Lévy jump basis Z and a special form of (1)

$$\nu(\mathrm{d}x, \mathrm{d}\xi) = \mu(\mathrm{d}x)\rho(\xi)\,\mathrm{d}\xi, \tag{18}$$

where the Lévy measure μ is finite which corresponds to the compound Poisson process Z'. We can normalize the right hand side of (18) so that μ is the jump size distribution and ρ the spatio-temporal intensity (density of the measure U). Denote m_j the jth moment of μ, i.e.

$$m_j = \int x^j \mu(\mathrm{d}x),\ j = 1, 2, \ldots.$$

We will use in the following product sets $C_1 \times C_2$ where $C_1 \subset [0,T]$ is a temporal set (typically an interval) and $C_2 \subset \mathbb{R}^d$ is a bounded spatial set.

Theorem 3.1. Denote

$$f_C(\xi) = \int_{C_1} \mathbf{1}_{C_2}(y_t) g((t, y_t), \xi)\,\mathrm{d}t, \quad \xi \in \mathbb{R}^{d+1} \tag{19}$$

$C = C_1 \times C_2$, similarly $f_D,\ D = D_1 \times D_2$. It holds

$$M(C) = m_1 \int f_C(\xi)\rho(\xi)\,\mathrm{d}\xi, \tag{20}$$

and the factorial second moment measure of X_Y

$$\alpha^{(2)}(C,\, D) = M(C)M(D) + m_2 \int f_C(\xi) f_D(\xi)\rho(\xi)\,\mathrm{d}\xi. \tag{21}$$

Proof. Using the formula for the generating functional of a Cox process we have

$$G(1 - z\mathbf{1}_C) = \mathrm{E}\exp\left(-z\int_{C_1} \mathbf{1}_{C_2}(y_t)\lambda(t, y_t)\,\mathrm{d}t\right).$$

Using (4) and Fubini theorem we obtain

$$G(1 - z\mathbf{1}_C) = \exp\left(\mathcal{C}\left\{iz \ddagger \int f_C\,\mathrm{d}Z\right\}\right)$$

and from Lemma 1 we have

$$G(1 - z\mathbf{1}_C) = \exp\left\{\int\int (e^{-zrf_C(\xi)} - 1)\,\mu(\mathrm{d}r)\rho(\xi)\,\mathrm{d}\xi\right\}.$$

By differentiating the result for intensity follows. Analogously we obtain

$$G(1 - z\mathbf{1}_C - v\mathbf{1}_D) = \exp\left\{\int\int (e^{-r(zf_C(\xi)+vf_D(\xi))} - 1)\,\mu(\mathrm{d}r)\rho(\xi)\,\mathrm{d}\xi\right\}$$

and by differentiating the factorial second moment measure. □

Since the measure μ is finite we get from (4) a representation

$$\Lambda(\xi) = \sum_j w_j g(\xi, \eta_j) \tag{22}$$

where η_j are events of a Poisson process with intensity function ρ and w_j are jump sizes. In fact formula (20) follows then from the Campbell theorem

$$\mathrm{E}\Lambda(\xi) = m_1 \int g(\xi, \eta)\rho(\eta)\,\mathrm{d}\eta.$$

We can extend the definition of OUCP to a Cox process on a curve Y by using an Ornstein–Uhlenbeck type process Λ in (16). Specially we have

Corollary 3.2. Consider the random function Λ from (13) and an OUCP X_Y. For the intensity measure $M(\cdot) = \mathrm{E}X_Y(\cdot)$ of a product set $C = C_1 \times C_2$ and for the factorial second moment measure formulas (20), (21) of Theorem 2 hold, respectively, with

$$f_C(s, \sigma) = \int_{C_1 \cap [s,\infty)} e^{\gamma(s-t)}\mathbf{1}_{C_2 \cap B_{s-t}(\sigma)}(y_t)\,\mathrm{d}t. \tag{23}$$

Proof. Put (15) into (19). □

For the model in $\mathbb{R}^3$ of a piecewise constant ρ

$$\rho(\xi) = \sum_{ijk} \rho_{ijk}\mathbf{1}_{A_{ijk}}(\xi), \tag{24}$$

where $A_{ijk} = A_i \times A_{jk}$, A_i a temporal interval, $A_{jk} \subset \mathbb{R}^2$ we obtain specially

Corollary 3.3. Under the assumptions of Corollary 4 and with the model (24) it holds

$$M(C) = m_1 \int_{C_1} \mathbf{1}_{C_2}(y_t) \sum_{ijk} \rho_{ijk} \int_{(-\infty,t]\cap A_i} e^{\gamma(s-t)} \mathrm{Leb}(B_{s-t}(y_t) \cap A_{jk})\, \mathrm{d}s\, \mathrm{d}t \quad (25)$$

and

$$\alpha^{(2)}(C,D) = m_2 \int_{C_1} \mathbf{1}_{C_2}(y_t) \int_{D_1} \mathbf{1}_{D_2}(y_u) \sum_{ijk} \rho_{ijk} \quad (26)$$

$$\times \int_{A_i\cap[-\infty,\min(u,t)]} e^{\gamma(2s-t-u)} \mathrm{Leb}(A_{jk} \cap B_{s-t}(y_t) \cap B_{s-u}(y_u))\, \mathrm{d}s\, \mathrm{d}u\, \mathrm{d}t + M(C)M(D).$$

Proof. Formula (25) follows putting (24) and (23) in (20) and similarly using Fubini theorem we obtain (26). □

4. FILTERING AND BAYESIAN MCMC

One of important questions in the analysis of Cox point processes is the inference on the driving intensity Λ and its characteristics. A rigorous approach to this problem is the filtering, see [12, 17]. Filtering and transition together yield prediction, cf. [7] for a log-Gaussian spatio-temporal point process. Transition density is available for the OU processes which are Markov (in time), e. g. for $\Lambda(t,\sigma)$ in (13).

Generally given a realization of a spatio-temporal Cox point process X driven by Λ, the solution of the nonlinear filtering problem is the conditional expectation $\mathrm{E}[\Lambda|X]$. Typically conditioning up to a real time t has been considered, we will get back to this situation later in Section 6. Here we develop another approach based on filtering global point processes. Since $\mathrm{E}[\Lambda|X]$ is not explicitly available the Bayes formula for probability densities enters:

$$f(\lambda|x) \propto f(x|\lambda) f(\lambda),$$

and from the definition of the Cox process $f(x|\lambda)$ is a density of an inhomogeneous Poisson process with intensity λ. The aim is to simulate samples from the density $f(\lambda|x)$ which enables to solve the filtering problem and estimate empirically any characteristics of Λ. Simulation is possible using Markov chain Monte Carlo (MCMC) techniques.

For a spatio-temporal Cox process on a curve, given a realization X and given a curve Y, the solution of the nonlinear filtering problem is the conditional expectation $\mathrm{E}[\Lambda|X,Y]$. Here

$$f(\lambda|x,y) \propto f(x|\lambda,y) f(\lambda|y), \quad (27)$$

and $f(x|\lambda,y)$ is a density of an inhomogeneous Poisson process with intensity measure λ_Y, cf. (17), given $\Lambda = \lambda$.

We will consider a more general situation than in the previous section here, namely that the curve Y is random, but independent of Λ. The presented model may be useful in neurophysiology [4, 16] where an experimental rat is moving in an arena along a random track and electrical impulses (spikes) of its brain are measured in time and space (location of the rat). Thus Y is a random element (with distribution P_Y) in the space of curves in a bounded region $\mathcal{A} \in \mathcal{B}$, with positive time derivative on $[0, T]$. The velocity is a covariate which is not considered here. Thanks to independence of Λ and Y formulas for generating functionals and moment measures are obtained from those in the previous section (which are conditional given $Y = y$) by averaging with respect to P_Y. In this situation we speak about a Cox process sampled by a curve rather than a Cox process on a curve, since an unknown spatio-temporal intensity is sampled along a random curve.

An approach to filtering based on the point process densities with respect to the unit Poisson process is available. Simultaneously the parameters in a parametric model of Z can be estimated as posterior means. Let $W = \mathcal{A} \times [0, T]$, $\mathcal{A} \in \mathcal{B}$, be a bounded window where the data $x = \{\tau_j\}$, a realization of the Cox process driven by (22), and the curve Y are observed. Each τ_j reflects time and location of an event on Y. We have now

$$f(\psi, b|x) \propto f(x|\psi, b, y) f(\psi|b) f(b), \tag{28}$$

where $\psi = \{\eta_j, w_j\}$ represents the compound Poisson process Z' and b is a vector of unknown parameters, e.g. those of a model for the intensity function ρ in (18), of a model for the jump size distribution, etc. Because of the independence of Λ and Y conditioning on y appears in formula (28) in the term $f(x|\psi, b, y)$ only.

In the spatio-temporal situation $\psi = \{t_j, z_j, w_j\}$, where z_j are locations and t_j times of events of Z'. Since the Cox process is conditionally Poisson we have the likelihood

$$f(x|\psi, b, y) \propto \exp\left(-\int_0^T \lambda(t, y_t)\,\mathrm{d}t\right) \prod_{\tau_i \in x} \lambda(\tau_i),$$

which corresponds to a density (w.r.t. a unit Poisson process on the time axis) of the inhomogeneous spatio-temporal Poisson process on Y given $\Lambda = \lambda$. There are two point process densities competing in formula (28).

The second one is

$$f(\psi|b) \propto \exp\left(-\int_W \rho(v)\,\mathrm{d}v\right) \prod_{(t_j, z_j, w_j) \in \psi} \rho(t_j, z_j) h(w_j),$$

where h is the probability density of jump size. Finally $f(b)$ is a prior distribution of parameters.

The "Metropolis within Gibbs" method can be used to simulate an MCMC chain $(\psi, b)^{(l)}$, $l = 0, \ldots, J$, which tends in distribution to the desired conditional distribution (28). The birth-death algorithm [18] is available for variable ψ with e.g. a uniform proposal distribution for both birth and death of a point. The real parameters are updated by a Gaussian random walk or a Langevin–Hastings algorithm. Geometric ergodicity of the chain follows under mild conditions, in temporal case cf. [14].

Example 1. Consider a compound Poisson process Z' and the model (13) (with $d = 2$). The driving intensity function has from (22) a representation

$$\Lambda(t,v) = \sum_{t_j \le t} w_j e^{\gamma(t_j - t)} \mathbf{1}_{B_{t_j-t}(v)}(z_j). \tag{29}$$

In practice this formula is an approximation, theoretically unbounded domain of ρ is substituted by some W_0 bounded, $W \subset W_0$, containing also events at negative times t_j. Let the jumps have an exponential distribution with density $h(a) = \frac{1}{\alpha}\exp(-\frac{a}{\alpha})$, $a \ge 0$, where $\alpha > 0$ is a parameter. Further let

$$B_s(x_1, x_2) = [x_1 + \omega s,\ x_2 - \omega s] \times [x_1 + \omega s,\ x_2 - \omega s],\ s \le 0.$$

Consider a cubic subdivision of W_0, denote the cubes $A_{ijk} = A_i \times A_{jk}$, A_i is a time interval. For the model (24) the vector of parameters is

$$b = (\alpha, \omega, \gamma, \{\rho_{ijk},\ i,j,k = 1,\dots,n\}).$$

The prior distributions are also chosen one-dimensional exponential with fixed hyperparameters $l_\alpha, l_\omega, l_\gamma, l_{ijk} \gg 0$ for random $\alpha, \omega, \gamma, \rho_{ijk}$, respectively.

Under these assumptions, denoting N_ψ the number of events of ψ in W_0, we can rewrite (28) as

$$\begin{aligned} f(\psi, b|x) \propto{} & \exp\left(-\int_0^T \sum_{t_j \le t} w_j e^{\gamma(t_j-t)} \mathbf{1}_{B_{t_j-t}(y_t)}(z_j)\,\mathrm{d}t\right) \\ & \times \prod_{\tau_i \in x} \lambda(\tau_i) \prod_{ijk} \exp(-\rho_{ijk}\mathrm{Leb}(A_{ijk} \cap W_0)) \alpha^{-N_\psi} \exp\left(-\sum_i \frac{w_i}{\alpha}\right) \\ & \times \left(\prod_{(t,z)\in\psi} \sum_{jlk} \rho_{jlk} \mathbf{1}_{A_{jlk}}(t,z)\right) l_\alpha^{-1} e^{-\frac{\alpha}{l_\alpha}} l_u^{-1} e^{-\frac{u}{l_u}} l_\gamma^{-1} e^{-\frac{\gamma}{l_\gamma}} \prod_{jk} l_{ijk}^{-1} e^{-\frac{\rho_{ijk}}{l_{ijk}}}. \end{aligned} \tag{30}$$

The full-conditional distributions for the Gibbs sampler are then

$$\begin{aligned} f(\psi|b,x) \propto{} & \exp\left(-\int_0^T \sum_{t_j \le t} w_j e^{\gamma(t_j-t)} \mathbf{1}_{B_{t_j-t}(y_t)}(z_j)\,\mathrm{d}t\right) \prod_{\tau_i \in x} \lambda(\tau_i) \\ & \times \alpha^{-N_\psi} \exp\left(-\sum_i \frac{w_i}{\alpha}\right) \left(\prod_{(t,z)\in\psi} \sum_{jlk} \rho_{jlk} \mathbf{1}_{A_{jlk}}(t,z)\right), \\ f(\rho_{ijk}|\psi,\omega,\alpha,\gamma,x) \propto{} & \exp(-\rho_{ijk}\mathrm{Leb}(A_{ijk} \cap W_0)) \left(\prod_{(t,z)\in\psi} \sum_{jlm} \rho_{jlm} \mathbf{1}_{A_{jlm}}(t,z)\right) \\ & \times e^{-\frac{\rho_{ijk}}{l_{ijk}}}, \quad i,j,k = 1,\dots,n, \\ f(\omega|\psi,\rho,\alpha,\gamma,x) \propto{} & \exp\left(-\int_0^T \sum_{t_j \le t} w_j e^{\gamma(t_j-t)} \mathbf{1}_{B_{t_j-t}(y_t)}(z_j)\,\mathrm{d}t\right) \prod_{\tau_i \in x} \lambda(\tau_i) e^{-\frac{u}{l_u}}, \end{aligned}$$

$$f(\gamma|\psi,\rho,\alpha,\omega,x) \propto \exp\left(-\int_0^T \sum_{t_j\le t} w_j e^{\gamma(t_j-t)} \mathbf{1}_{B_{t_j-t(y_t)}}(z_j)\,\mathrm{d}t\right) \prod_{\tau_i\in x} \lambda(\tau_i) e^{-\frac{\gamma}{l_\gamma}},$$

$$f(\alpha|\psi,\rho,\omega,\gamma,x) \propto \alpha^{-N_\psi} \exp\left(-\sum_i \frac{w_i}{\alpha}\right) e^{-\frac{\alpha}{l_\alpha}}.$$

To draw from these densities we use Metropolis–Hastings steps, i. e. in each iteration proposal distributions yield new candidates, we evaluate Hastings ratios H and the proposals are accepted with probability equal to $\min\{1, H\}$ each, respectively.

5. ESTIMATION AND MODEL SELECTION

Using ergodicity properties of the MCMC chain we can estimate statistical characteristics of Λ. Denote $\Lambda^{(l)}(t,v)$ from (29) the lth iteration of the intensity (conditioned on a realization of x, y) of the MCMC chain. J is the number of iterations, $K,\ 0 < K < J$, the burn-in of the chain, put $k = J - K$. The filtered conditional expectation of Λ is estimated by the average value

$$\hat{\Lambda}(t,v) = \frac{1}{k} \sum_{l=K+1}^{J} \Lambda^{(l)}(t,v), \tag{31}$$

analogously we get estimators of higher moments and conditional variance of Λ.

In the Bayesian framework there exist several tools for model selection including Bayes factors, posterior predictive distributions or an extended Bayesian analysis. We restrict attention to the consideration of posterior predictive distributions. Consider a summary statistics $V(x,y)$ computed from the data and compare it with $V(X,y)$ where X is a Cox process with the estimated driving intensity.

We use summary statistics corresponding to the first order and the second order characteristics of the spatio-temporal point process. Those of the first order are the counts, i. e. numbers of points $N(C_j)$ of X in subregions $C_j \subset W,\ j = 1,\dots,k$ hitting Y. A measure of discrepancy of the model is e. g.

$$\sum_{j=1}^{k} (M(C_j) - N(C_j))^2.$$

For the second order analysis usually estimates of K or L-functions have been used which enable graphical tests based on repeated simulations of the model with estimated parameters. For inhomogeneous processes these estimators are based on the assumptions of the second-order intensity reweighted stationarity [16], p. 32. Since our study is directed to applications where this assumption is not fulfilled and also since these estimators are not unbiased we proceed another way which does not yield a straightforward graphical presentation, however it is based on unbiased estimation and follows the theory for the Cox processes on a curve developed in Theorem 2. We will evaluate the factorial second moment measure $\alpha^{(2)}$ for pairs of subsets of the window and compare it with the estimator

$$\hat{\alpha}^{(2)}(C,D) = \sum_{\xi,\eta\in X}^{\neq} \mathbf{1}_{[\xi\in C,\eta\in D]}, \quad C,\ D\subset\mathbb{R}^2 \tag{32}$$

unbiased from (9). The statistics

$$\sum_{i\neq j}(\alpha^{(2)}(C_i,C_j)-\hat{\alpha}^{(2)}(C_i,C_j))^2$$

can be compared for various models of ρ.

Also we can apply Monte–Carlo tests for these posterior predictive distributions. Using estimated parameters we simulate 19 realizations of the Cox process model and evaluate lower and upper value of summary statistics (counts and factorial second moment measure) for each subset. Then we observe how the quantities obtained from data ($N(C_j)$ and $\hat{\alpha}^{(2)}(C_i,C_j)$) fit within these bounds, respectively.

There are two ways of numerical evaluation of model selection procedure using posterior predictive distribution which we demonstrate on counts. The first one is based on formula (17) and $M(\cdot)=\mathrm{E}\Lambda_Y(\cdot)$. We approximate the mean value $\mathrm{E}\Lambda_Y(\cdot)$ as

$$\hat{\Lambda}_Y([t,s]\times B)\approx \triangle\sum_{p=1}^{m}\mathbf{1}_B(y_{t_p})\hat{\Lambda}(t_p,y_{t_p}), \tag{33}$$

where $t_p=t+p\triangle$, $\triangle=(s-t)/m$, where $\hat{\Lambda}(t_p,y_{t_p})$ is evaluated from (31). Thus we obtain an estimate of $M([t,s]\times B)$ based on the auxiliary process iterations since Λ comes from (29).

The second way is to involve also parameter estimators. For the numerical evaluation of (25) under the discretization $t_p=p\triangle,\ s_q=q\triangle,\ p,q$ integers, $\triangle>0$, we can use for $D=D_1\times D_2,\ D_1\subset\mathbb{R}$ an approximation

$$M(D)\approx m_1\triangle^2\sum_{lmn}\rho_{lmn}\sum_{q:s_q\in A_l}\sum_{\substack{p:t_p\in D_1\\ t_p\geq s_q}} e^{\gamma(s_q-t_p)}\mathbf{1}_{D_2}(y_{t_p})\,\mathrm{Leb}(A_{mn}\cap B_{s_q-t_p}(y_{t_p})).$$

To enumerate (26) we can use either an analogous discretization (there is one more integration) or we can evaluate integrals (23) separately on a grid of points and then integrate numerically directly in (21).

Example 2. Let $\mathcal{A}$ be a circle $b(0,r)\subset\mathbb{R}^2$, $W=\mathcal{A}\times[0,T]$, consider a deterministic spiral curve

$$Y=\{t,t\cos(\beta t),t\sin(\beta t)\},\ 0<t<r, \tag{34}$$

$\beta>0$ a parameter. Numerical results are presented in Figure 1. A simulated point process realization on Y is filtered using MCMC, the model from Example 1 is used. The posterior predictive distributions in space and time were evaluated using formula (33). The empirical and theoretical counts are compared in planar subregions and in time. We observe a reasonably good fit in both space and time for given data.

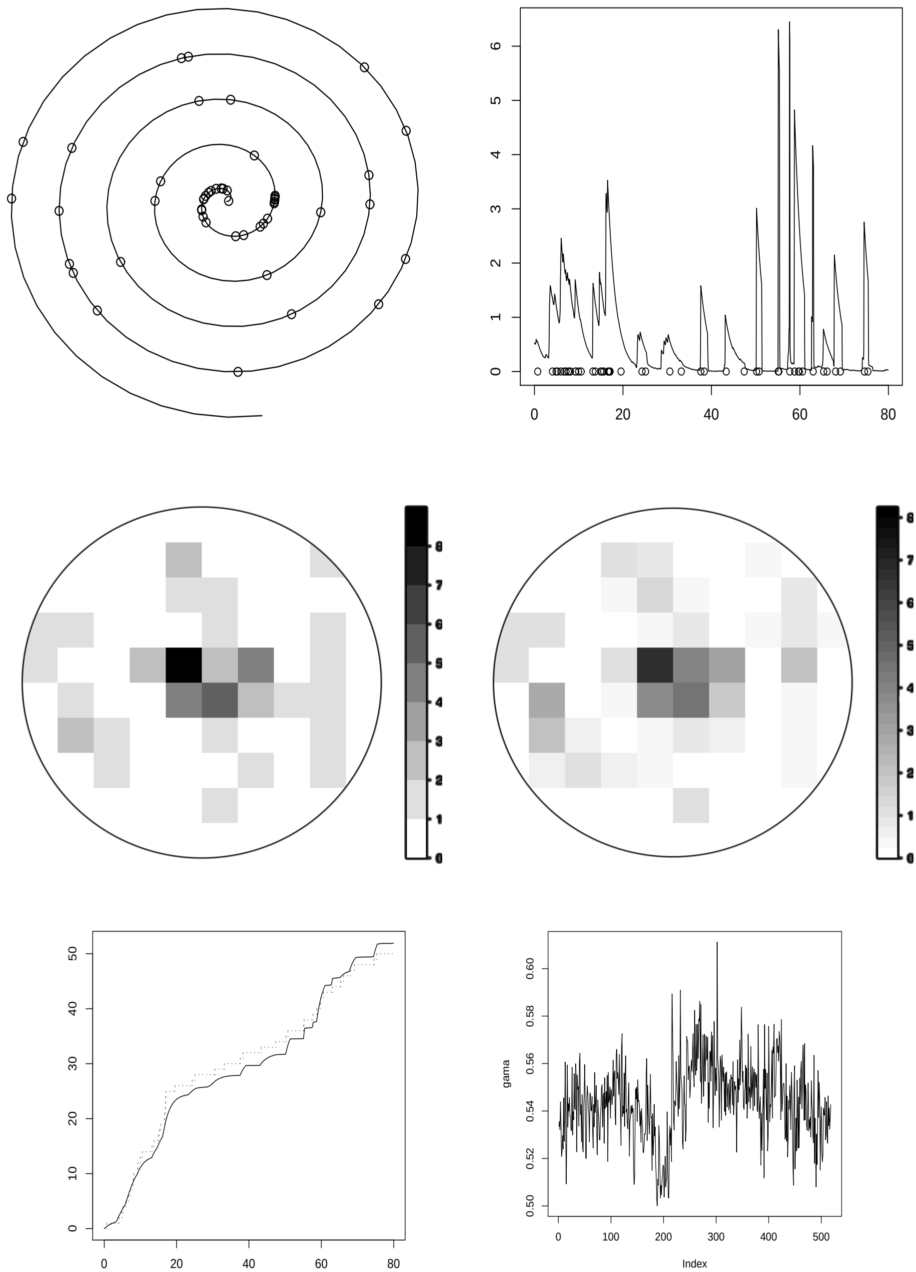

Fig. 1. MCMC computations, Example 2. Upper left (data) Simulated realization with 51 events (circles) of a point process on the curve (34) which is projected onto the plane. Right (filtering) The graph of $\hat{\Lambda}(t, y_t)$, $t \in [0, T]$, cf. (31) after 4×10^6 iterations of MCMC. Middle (model selection) left – the counts $N(D)$, right – evaluation of $M(D)$ from (33), here $D = [0, T] \times A_{ij}$, $\{A_{ij}\}$ is 10×10 planar grid restricted by circular $\mathcal{A}$. Lower left (model selection) $N(C)$ (grey dotted curve) and $M(C)$ (black curve) from (33) for $C = \mathcal{A} \times [0, t]$ with increasing time t (horizontal axis). Right (parameter estimation) 5×10^5 iterations (with step 1000) of γ. The chain is well mixing.

6. RESIDUAL ANALYSIS

The model selection procedures from the previous section enable to compare various models but they do not present a proper goodness-of-fit test. The Monte–Carlo test for summary statistics, when not rejected, does not guarantee the validity of the model on a given significance level. The residual analysis does a better job in this direction. For temporal and spatio-temporal point processes it is well developed, see [17], based on the conditional intensity and martingale theory in time. The purely spatial case is more complicated and the Papangelou conditional intensity is recommended as the basic tool by [1]. The authors note that spatial Cox processes are hard to analyze since with the exceptions when the density w.r.t. unit Poisson process exists in a closed form, the Papangelou conditional intensity is not computationally tractable.

For a Cox point process X (either temporal, spatial or spatio-temporal) with driving intensity measure Λ_m we can define an innovation process generally as

$$I(B) = X(B) - \mathrm{E}[\Lambda_m(B) \mid X], \quad B \in \mathcal{B}. \tag{35}$$

It holds

$$\mathrm{E}I(B) = 0.$$

Given a model for Λ_m depending on a parameter $\theta \in \mathbb{R}^p$ we obtain its estimator $\hat{\theta}$ and we can observe how the residual process

$$R_{\hat{\theta}}(B) = X(B) - \mathrm{E}_{\hat{\theta}}[\Lambda_m(B) \mid X] \tag{36}$$

oscillates around zero. A possibility to perform a statistical test depends on the way how exactly the conditioning in (35),(36) is defined. In the temporal case denoting $N_t, t \geq 0$ the counting process corresponding to X and Λ the density of Λ_m, assuming that the conditional intensity λ^* exists and

$$\begin{aligned}
\lambda_t^* &= \lim_{\triangle t \downarrow 0} \frac{1}{\triangle t} \mathrm{E}[N_{t+\triangle t} - N_t \mid N_s,\ s < t] \\
&= \lim_{\triangle t \downarrow 0} \frac{1}{\triangle t} \mathrm{E}[\mathrm{E}[N_{t+\triangle t} - N_t \mid N_s,\ s < t;\ \Lambda_p,\ t \leq p < t + \triangle t] \mid N_s,\ s < t] \\
&= \lim_{\triangle t \downarrow 0} \frac{1}{\triangle t} \mathrm{E}\left[\int_t^{t+\triangle t} \Lambda(s)\,\mathrm{d}s \mid N_s,\ s < t\right] = \mathrm{E}[\Lambda(t) \mid N_s,\ s < t]
\end{aligned}$$

the innovation process $N_t - \int_0^t \lambda_s^*\,\mathrm{d}s$ is a martingale [12]. In the spatio-temporal case denote $N_s(C) = \mathrm{card}\{x \in X;\ x \in [0,s] \times C\},\ \ C \in \mathcal{B}^d$. Analogously the conditional intensity λ^*

$$\lambda^*(t,\xi)\,\mathrm{d}t\,\mathrm{d}\xi = \mathrm{E}[N(\mathrm{d}t \times \mathrm{d}\xi) \mid N_s(C),\ s < t,\ C \in \mathcal{B}^d]$$

of a Cox process corresponds to

$$\mathrm{E}[\Lambda(t,\xi) \mid N_s(C),\ s < t,\ C \in \mathcal{B}^d] \tag{37}$$

and

$$N_t(C) - \int_0^t \int_C \lambda^*(s,\xi)\,\mathrm{d}s\,\mathrm{d}\xi$$

is a martingale with mean zero, $C \in \mathcal{B}$. Scaled innovations

$$V_h = \int_{\mathbb{R}\times\mathbb{R}^d} H(t,\xi)[N(\mathrm{d}t \times \mathrm{d}\xi) - \lambda^*(t,\xi)\,\mathrm{d}t\,\mathrm{d}\xi],$$

where H is a predictable process, are investigated.

For the Cox process on a curve studied in this paper we have an analogous argument. Define

$$\lambda_s^* = \mathrm{E}[\Lambda(s,y_s)|N_u, u < s],$$

$N_t - \int_0^t \lambda_s^*\,\mathrm{d}s$ is a martingale with mean zero. For $C \in \mathcal{B},\ C \subset \mathcal{A}$ and a random process $\{H(t),\, t \in [0,T]\}$ the scaled innovation V_C is defined as

$$V_C = \int_0^T \mathbf{1}_C(y_t)H(t)[N(\mathrm{d}t) - \lambda_t^*\,\mathrm{d}t].$$

Theorem 6.1. For a nonnegative predictable process $\{H(t),\, t \in [0,T]\}$ the scaled innovation has variance

$$\operatorname{var} V_C = \mathrm{E}\left[\int_0^T \mathbf{1}_C(y_t)H^2(t)\lambda_t^*\,\mathrm{d}t\right].$$

P r o o f. Denote $G(t) = \mathbf{1}_C(y_t)H(t)$, $\{G(t),\ t \in [0,T]\}$ is a predictable process. Since by [12], Theorem 4.6.1

$$\mathrm{E}\left(\int_0^T G(t)[N(\mathrm{d}t) - \lambda_t^*\,\mathrm{d}t]\right) = 0$$

we have (integral limits $0,T$ are omitted)

$$\begin{aligned} \operatorname{var} V_C &= \mathrm{E}\left(\left[\int G(t)N(\mathrm{d}t)\right]^2\right) \\ &\quad + \mathrm{E}\left(\left[\int G(t)\lambda_t^*\,\mathrm{d}t\right]^2\right) - 2\mathrm{E}\left[\int G(t)N(\mathrm{d}t)\int G(t)\lambda_t^*\,\mathrm{d}t\right]. \end{aligned}$$

Using Fubini and Theorem 1 from [25] we have

$$\begin{aligned} \operatorname{var} V_C &= \mathrm{E}\int G^2(t)\lambda_t^*\,\mathrm{d}t \\ &\quad -2\mathrm{E}\left[\int G(s)\lambda_s^* ds \int G(t)[N(\mathrm{d}t) - \lambda_t^*\,\mathrm{d}t]\right] \end{aligned}$$

and the second term vanishes again by [12], Theorem 4.6.1. □

The choice

$$H(t) = \mathbf{1}_D(t)(\lambda_t^*)^{-\frac{1}{2}},\ D \in \mathcal{B}^1$$

leads to the Pearson innovation

$$V_p = \int_D \mathbf{1}_C(y_t)[(\lambda_t^*)^{-\frac{1}{2}} N(\mathrm{d}t) - (\lambda_t^*)^{\frac{1}{2}}\,\mathrm{d}t] \tag{38}$$

with

$$\operatorname{var} V_p = \mathrm{Leb}\{t \in D;\, y_t \in C\}.$$

The residual data analysis based on a realization of the Cox process on the curve

$$x = \{\tau_j\} = \{s_j, \eta_j\}_{j=1,\dots,k},\ s_j \in \mathbb{R}, \eta_j \in \mathbb{R}^2$$

follows. Denote $\hat{\Lambda}(s)$ the MCMC estimator of λ_s^*. The Pearson residual corresponding to (38), time t and a measurable set $C \subset \mathcal{A}$ is then

$$R_{\hat{\theta}}(t, C) = \sum_{\substack{(s_l,\eta_l)\\ s_l \le t,\, \eta_l \in C}} \hat{\Lambda}(s_l)^{-\frac{1}{2}} - \int_0^t \mathbf{1}_C(y_s)[\hat{\Lambda}(s)]^{\frac{1}{2}}\,\mathrm{d}s. \tag{39}$$

Evaluation of the sum desires k MCMC chains conditioned up to time $s_j,\ j = 1, \dots, k$. A problem is the integral approximation in (39) which desires either more chains (computationally demanding) or the approximation of values of $\hat{\Lambda}(s)$ from chains conditioned at times larger than the argument s.

Finally Pearson residuals can be plotted at times $0 < t_1 < \dots < t_n = T$ with bounds $2\sigma_i$ at t_i,

$$\sigma_i = [\mathrm{Leb}\{t \le t_i;\, y_t \in C\}]^{\frac{1}{2}}.$$

7. CONCLUSIONS

In the paper a model of a spatio-temporal Cox point process with driving intensity based on background driving Lévy process is investigated. We derive basic properties of the model by means of the closed form of the generating functional. An important case where the events of the process lie on a curve, is studied in detail, under the assumption that the curve is independent of the driving intensity. Then the nonlinear filtering problem is solved using the Bayesian approach. Densities with respect to Poisson process of both the Cox process and the background driving compound Poisson process are involved. Markov chain Monte Carlo enables to draw approximately from the posterior distribution and make the desired inference. Posterior predictive distributions evaluate the model selection.

There have been essentially two approaches to modeling in spatio-temporal point processes which have played different roles. The use of conditional intensity enables the residual analysis but less formulas for basic characteristics analysis. Models not based on conditioning serve in the opposite way. In our paper within the Lévy based modeling and stochastic simulations we tried to achieve both as is shown in the final section. Filtering and statistics can be done simultaneously, the procedures which have been mainly solely investigated in previous studies. The Bayesian MCMC of space-time realizations is fast enough for filtering while the residual analysis is computationally demanding in our setting. This procedure is more easily performed using sequential Monte Carlo methods [18] within a conditional intensity model.

ACKNOWLEDGEMENT

The authors thank to the referees for valuable comments and suggestions. The research was supported by grants of the Grant Agency of the Academy of Sciences of the Czech Republic (Grant 101120604), the Ministry of Education, Youth and Sport of the Czech Republic (Project MSM 0021620839) and the Czech Science Foundation (Grant 201/05/H007).

(Received April 7, 2008.)

REFERENCES

[1] A. Baddeley, R. Turner, J. Møller, and M. Hazelton: Residual analysis for spatial point processes (with discussion). J. Royal Stat. Soc. *B 67* (2005), 617–666.

[2] O. Barndorff-Nielsen and N. Shephard: Non-Gaussian Ornstein–Uhlenbeck based models and some of their uses in financial economics. J. Royal Stat. Soc. *B 63* (2001), 167–241.

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

[4] V. Beneš and B. Frcalová: Modelling and simulation of a neurophysiological experiment by spatio-temporal point processes. Image Anal. Stereol. *1* (2008), 27, 47–52.

[5] P. Brémaud: Point Process and Queues. Springer, New York 1981.

[6] A. Brix and J. Møller: Space-time multi type log Gaussian Cox processes with a view to modelling weed data. Scand. J. Statist. *28* (2002), 471–488.

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

[8] M. A. Clyde and R. L. Wolpert: Nonparametric function estimation using overcomplete dictionaries. Bayesian Statistics *8* (2007), 1–24.

[9] R. Cont and P. Tankov: Financial Modelling with Jump Processes. Chapman and Hall/CRC, Boca Raton 2004.

[10] D. Daley and D. Vere-Jones: An Introduction to the Theory of Point Processes I, II. Springer, New York 2003, 2008.

[11] A. Ergun, R. Barbieri, U. T. Eden, M. A. Wilson, and E. N. Brown: Construction of point process adaptive filter algorithms for neural systems using sequential Monte Carlo methods. IEEE Trans. Biomed. Engrg. *54* (2007), 3, 307–326.

[12] P. M. Fishman and D. Snyder: The statistical analysis of space-time point processes. IEEE Trans. Inform. Theory *22* (1976), 257–274.

[13] G. Hellmund, M. Prokešová, and E. Vedel Jensen: Lévy based Cox point processes. Adv. Appl. Probab. *40* (2008), 3, 603–629.

[14] M. Jacobsen: Point Processes Theory and Applications. Marked Point and Piecewise Deterministic Processes. Birkhäuser, Boston 2006.

[15] A. F. Karr: Point Processes and Their Statistical Inference. Marcel Dekker, New York 1985.

[16] P. Lánský and J. Vaillant: Stochastic model of the overdispersion in the place cell discharge. BioSystems *58* (2000), 27–32.

[17] R. Lechnerová, K. Helisová, and V. Beneš: Cox point processes driven by Ornstein–Uhlenbeck type processes. Method. Comp. Appl. Probab. *10* (2008), 3, 315–336.

[18] J. Møller and R. Waagepetersen: Statistics and Simulations of Spatial Point Processes. World Sci., Singapore 2003.

[19] J. Møller and C. Diaz-Avalos: Structured spatio-temporal shot-noise Cox point process models, with a view to modelling forest fires. Scand. J. Statist. (2009), to appear.

[20] Y. Ogata: Space-time point process models for eartquake occurences. Ann. Inst. Statist. Math. *50* (1998), 379–402.

[21] J. Pedersen: The Lévy–Ito Decomposition of Independently Scattered Random Measure. Res. Report 2, MaPhySto, University of Aarhus 2003.

[22] R. D. Peng, F. P. Schoenberg, and J. Woods: A space-time conditional intensity model for evaluating a wildfire hazard risk. J. Amer. Statist. Assoc. *100* (2005), 469, 26–35.

[23] B. S. Rajput and J. Rosinski: Spectral representations of infinitely divisible processes. Probab. Theory Related Fields *82* (1989), 451–487.

[24] G. Roberts, O. Papaspiliopoulos, and P. Dellaportas: Bayesian inference for non-Gaussian Ornstein–Uhlenbeck stochastic volatility processes. J. Royal Statist. Soc. *B 66* (2004), 369–393.

[25] D. L. Snyder: Filtering and detection for doubly stochastic Poisson processes. IEEE Trans. Inform. Theory *18* (1972), 91–102.

[26] F. P. Schoenberg, D. R. Brillinger, and P. M. Guttorp: Point processes, spatial-temporal. In: Encycl. of Environmetrics (A. El. Shaarawi and W. Piegorsch, eds.), Vol. 3, Wiley, New York 2003, pp. 1573–1577.

[27] J. Zhuang: Second-order residual analysis of spatiotemporal point processes and applications in model evaluation. J. Royal Statist. Soc. *B 68* (2006), 635–653.

Blažena Frcalová and Viktor Beneš, Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics – Charles University in Prague, Sokolovská 83, 186 75 Praha 8. Czech Republic.

e-mails: frcalova@karlin.mff.cuni.cz, benesv@karlin.mff.cuni.cz