CHARLES UNIVERSITY IN PRAGUE
FACULTY OF MATHEMATICS AND PHYSICS

DOCTORAL THESIS

Mgr. Blažena Frcalová

Spatio-temporal point processes

Department of Probability and Mathematical Statistics

Supervisor: Prof. RNDr. Viktor Beneš, DrSc.
Field of study: m-4 Probability and Mathematical Statistics

Prague, 2010

It would have been next to impossible to write this thesis without the help and guidance of my supervisor Prof. RNDr. Viktor Beneš, DrSc. I am grateful for all his support, time and friendly attitude throughout my Ph.D studies. I would also like to thank the Fulbright Commission for their financial support that allowed me to spend one year at the Columbia University in New York.

I declare that this thesis has been written by myself and that all used literature is included in the list of references. I agree with using this thesis for study purposes.

Prague, 2 July 2010

Blažena Frcalová

Contents

Title: *Spatio-temporal point processes*
Author: *Blažena Frcalová*
Department: *Department of Probability and Mathematical Statistics*
Author's e-mail address: *frcalova@karlin.mff.cuni.cz*
Supervisor: *Prof. RNDr. Viktor Beneš, DrSc.*
Supervisor's e-mail address: *benesv@karlin.mff.cuni.cz*

Abstract: *The background theory of point processes, spatio-temporal point processes, random measures and random closed sets is given in the beginning of the thesis. Then the special case of spatio-temporal Cox processes constructed from Lévy basis is studied. Formulas for theoretical characteristics are derived using the generating functional. The Cox process on the curve is defined and studied. The analysis of such a process leads to nonlinear filtering methods. Also the methods for model selection are discussed. These methods are used on simulated data, firstly on the simple discrete data and secondly on the continuous data where the curve is a spiral. Then the real data from a neurophysiology experiment is analysed. During the experiment, the spiking activity of a place cell of hippocampus of a rat moving in an arena together with the track of the rat was recorded. The track of the rat and the action potentials (spikes) present the curve and the points on it. At the end of the thesis, other approaches to neurophysiological data are discussed. The first one is an estimation of a conditional intensity of the temporal process of spikes using recursive filtering. In the second one, the track of the rat together with the random driving intensity function of the process of the spikes is viewed as a random marked set.*

Keywords: *Filtering, Cox point process, Spatio-temporal process, Random marked closed set, Random measure, Lévy basis*

Název: *Časoprostorové bodové procesy*
Autor: *Blažena Frcalová*
Katedra: *Katedra pravděpodobnosti a matematické statistiky*
Autorova e-mailová adresa: *frcalova@karlin.mff.cuni.cz*
Vedoucí: *Prof. RNDr. Viktor Beneš, DrSc.*
Vedoucího e-mailová adresa: *benesv@karlin.mff.cuni.cz*

Abstract: *Na začátku práce je přehled základní teorie bodových procesů, časoprostorových bodových procesů, náhodných měr a náhodných uzavřených množin. Dále jsou studovány časoprostorové Coxovy procesy, které jsou konstruovány pomocí Lévyho bází. Za použití vytvořujícího funkcionálu jsou odvozeny základní charekteristiky. Je definován a studován*

Coxův proces na křivce. Analýza takovýchto procesů vede k nelineárním filtrovacím metodám. Jsou diskutovány také metody umožňující výběr modelu. Tyto metody jsou použity na simulovaných datech, nejdříve na jednoduchém diskrétním případě a pak i na spojitém případě se spirálovitou křivkou. Poté je provedena analýza neurofyziologických dat. V průběhu experimentu byla zaznamenávána aktivita neurových buněk z hippocampu u krysy hledající jídlo v omezeném prostoru zároveň s polohou zvířete. Trasa zvířete a akční potenciály (spiky) představují křivku a body na ní. Na konci práce jsou další možné přístupy k neurofyziologickým datům. První je odhad podmíněné intensity časového procesu spiků pomocí rekurzivního filtrování. Ve druhém případě je na trasu krysy spolu s náhodnou řídící funkcí intenzity procesu spiků nahlíženo jako na náhodnou uzavřenou kótovanou množinu.

Klíčová slova: *Filtrování, Coxův bodový proces, Časoprostorový bodový proces, Náhodná uzavřená kótovaná množina, Náhodná míra, Lévyho báze*

Introduction

Point processes and especially spatio-temporal point processes are widely used in many applications to model phenomena from the nature, for example in epidemiology ([7], [21]), forest fires description ([37], [40]), urban development [23], earthquakes modelling ([38], [17]), in weed species growth [12] or in evaluation of spiking activity of neuron [34]. In the last application experiments reveal that the variance of the number of points is higher than their expected value. This suggests the use of the Cox point process model.

The background theory of point processes is reviewed in Chapter 1, together with the theory of random measures and random closed sets. The special case of random closed set, the fibre process, is introduced. At the end of Chapter 1 there is a review of Markov Chain Monte Carlo methods.

The class of the Cox processes with driving intensity function constructed from Lévy basis is studied in Chapter 2. This work was published in our paper [29]. The theory was developed simultaneously, but independently of [30]. Our approach is based on the generating functional which enables us to derive the theoretical characteristics of the process. We study the case when the Lévy basis is a Poisson measure in more details. The ambit set [5] is used in the construction of the model, see Corollary 2.2 and 2.3. This theory is applied to the process on a curve (Theorem 2.4). We consider two models for the intensity function of the Poisson measure. The first one is the model of piecewise constant function and the second model uses the Zernicke polynomials. We derived the formulas for intensity measures and second-order factorial moment measures for both models (Corollary 2.5 and 2.6).

The solution of the nonlinear filtering problem of the driving intensity function of a spatio-temporal Cox point process given its observed events is the conditional expectation. The Bayesian Markov Chain Monte Carlo approach enables to simulate from the target conditional distribution, the relevant methods are developed in Section 2.3. Model selection methods are based on posterior predictive distributions. Summary statistics is computed from the data and compared with the value of that statistics of Cox process with estimated driving intensity function. Residual analysis is based on the innovation process. This method suits well to our model of a point process on the curve and the formula for the variance of scaled innovation, especially for Pearson innovation, is derived (Theorem 2.5).

Chapter 3 starts with a discrete simulation of the model from Section 2.2 on a grid. The curve is represented by a random walk on a grid. The algorithm of the simulation is described and the Fano factor and driving intensity function are estimated. This study contains also an analysis of approximation precision (Lemma 3.1) and it was published in our paper [9]. Then there is a simulation in continuous space and time in Subsection 3.1.2. The real data from the neurophysiological experiment are evaluated in Section 3.2. The experimental animal searching for food was moving in a bounded arena and its track with times and locations of action potentials (spikes) was recorded. The aim is to model the experiment mathematically and also a further hypothesis is studied in Chapter 4. Numerical results are presented graphically, this study was published in our paper [8].

The last Chapter 4 discusses other approaches to the evaluation of neurophysiological. The first one is known from the literature and it is based on sequential filtering of point processes in time. Here, the data are viewed as a temporal point process with spatial marks. The advantage of this method is that it enables us to work with a larger data set since it is computationally faster. However it does not enable to evaluate point process characteristics. In the final approach, the curve is random and together with the random driving intensity function of the process of spikes forms a random marked closed set (RMCS). We show that RMCSs of integer dimension (nonzero and not full) induce weighted random measures. Then we introduce the concept of second-order intensity-reweighted stationary weighted measure and develop a test of the random-field model. In this situation it can be applied to the above inhomogeneous data from the neurophysiological experiment to answer the question about the stochastic independence of the driving intensity on the random track.

The research work was supported by the Grant Agency of the Academy of Sciences of the Czech Republic, Project no. IAA101120604, "Spatio-temporal point processes".

Chapter 1

Theoretical Background

1.1 Point processes

Spatial point processes are basic models in stochastic geometry which can describe various spatial data. They are used in many applications like plant ecology, forestry (positions of trees), computational neurophysiology (spiking activity), zoology (burrows or nests), geography (position of towns), seismology etc.

Definition 1.1: *Let* $(\Omega, \mathcal{F}, \mathsf{P})$ *be a probability space,* $\mathcal{X}$ *a locally compact complete separable metric space with Borel* $\sigma-$*algebra* $\mathcal{B}(\mathcal{X})$*,* $\mathcal{N}$ *the system of locally finite subsets of* $\mathcal{X}$ *with* $\sigma-$*algebra* $\mathfrak{N}$

$$\mathfrak{N} = \sigma\left(N \in \mathcal{N}, \#(N \cap A) = m, A \in \mathcal{B}(\mathcal{X}), m \in \mathbb{N}_0\right). \tag{1.1}$$

A measurable mapping X *from* $(\Omega, \mathcal{A})$ *to* $(\mathcal{N}, \mathfrak{N})$ *is called a point process.*

Distribution of the point process X *is a probability measure* Π *defined by*

$$\Pi(F) = \mathsf{P}(\{\omega; X(\omega) \in F\}) \quad \textit{for } F \in \mathfrak{N}$$

.

In the following text we will assume either $\mathcal{X} = \mathbb{R}^d$ or $\mathcal{X} = S \subset \mathbb{R}^d$. We denote $\mathcal{B}^d = \mathcal{B}(\mathbb{R}^d)$ the Borel $\sigma-$algebra and $\mathcal{B}^d_0 \subset \mathcal{B}^d$ system of bounded sets. From the Definition 1.1 we can see that for each Borel set $A \subset \mathcal{X}$, $X(A) \equiv \#(X \cap A)$ (number of points of X in A) is a random variable. Then $X(.)$ is an integer-valued random measure. Also by X we sometimes mean the image of the mapping X, i.e. a locally finite set of points.

Definition 1.2: *If* $X(\mathcal{X})$ *is almost surely finite then the point process* X *is called finite.*

Definition 1.3: *The intensity measure* Λ *is defined as* $\Lambda(A) = \mathbb{E}(X(A))$ *for each* $A \in \mathcal{B}(\mathcal{X})$*.*

If the intensity measure of a point process X *on* $\mathbb{R}^d$ *is absolutely continuous with respect to the Lebesgue measure, i.e. there is a non-negative function* λ *such that*

$$\Lambda(A) = \int_A \lambda(x)dx \tag{1.2}$$

then the density λ is called an intensity function of X.

Definition 1.4: *Probabilities $\mathsf{P}(\{\omega; X(\omega, A) = 0\})$, $A \in \mathcal{B}(\mathcal{X})$ are called void probabilities.*

Theorem 1.1: *The distribution of a point process X is determined by its void probabilities.*

Proof: See [36]. □

Definition 1.5: *A point process X on $\mathbb{R}^d$ is called stationary if its distribution is invariant under translations, i.e. the distribution of the translated process $(X+x) = \{(y+x), y \in X\}$ is the same as the distribution of X itself for all $x \in \mathbb{R}^d$.*

A point process X is called isotropic if its distribution is invariant under rotation around origin. It means the distribution of the rotated point process $\mathcal{O}X = \{\mathcal{O}x, x \in X\}$ is the same as the distribution of X for all rotations $\mathcal{O}$ around origin $O \in \mathbb{R}^d$.

A point process is called motion invariant if it is both stationary and isotropic.

Lemma 1.1: *If the point process X is stationary with an intensity measure Λ then Λ is a multiple of the Lebesgue measure.*

Proof: It follows since for a stationary point process it must hold that $\Lambda(B) = \Lambda(B+y)$ for all $B \in \mathcal{B}(\mathbb{R}^d)$ and for all $y \in \mathbb{R}^d$ and the Lebesgue measure is the only measure (except the multiple of it) which is invariant under translation. □

From previous lemma 1.1 it follows that a stationary point process has the intensity function and this intensity function is constant. This constant is called for short *intensity*.

Definition 1.6: *The n-order moment measure of a point process X is defined by*

$$M^{(n)}(A) = \mathbb{E} \sum_{X_1,\dots X_n \in X} \mathbf{1}_{[(X_1,\dots,X_n)\in A]}, \quad A \in (\mathcal{B}^d)^n. \tag{1.3}$$

The n-order factorial moment measure of a point process X is defined by

$$\alpha^{(n)}(A) = \mathbb{E} \sum_{X_1,\dots X_n \in X}^{\neq} \mathbf{1}_{[(X_1,\dots,X_n)\in A]}, \quad A \in (\mathcal{B}^d)^n, \tag{1.4}$$

where the symbol $\neq$ means that $X_i \neq X_j$ for $i \neq j$.

Obviously the first moment measure is equal to the first factorial moment measure $M^{(1)} = \alpha^{(1)}$ and it is also the intensity measure Λ.

Definition 1.7: *If the second-order factorial moment measure $\alpha^{(2)}(A)$ can be expressed as*

$$\alpha^{(2)}(A) = \int\int \mathbf{1}_{[(\xi,\nu)\in A]} \lambda^{(2)}(\xi,\nu) d\xi d\nu,$$

where $\lambda^{(2)}$ is a non-negative function, then $\lambda^{(2)}$ is called the second-order product density.

If both the intensity function λ and the second order product density $\lambda^{(2)}$ exist, the pair correlation function is defined by

$$\rho(\xi,\nu) = \frac{\lambda^{(2)}(\xi,\nu)}{\lambda(\xi)\lambda(\nu)}, \quad \xi,\nu \in \mathcal{X}. \tag{1.5}$$

The interpretation of $\lambda^{(2)}(\xi,\nu)d\xi d\nu$ is as probability of observing a pair of points in two infinitesimally small balls with centres ξ and ν.

A measure μ on $\mathcal{X}$ is called a diffusion measure if $\Lambda(\{\xi\}) = 0$ for $\forall \xi \in \mathcal{X}$.

Definition 1.8: *Let Λ be a locally finite diffusion measure on $\mathcal{X}$. The point process X such that*

1. *random variable $X(B)$ has the Poisson distribution with the parameter $\Lambda(B)$, $B \in \mathcal{B}_0^d$;*

2. *$X(B_1), X(B_2), \dots X(B_n)$ are independent random variables for each $n \in \mathbb{N}$ and each $B_1, B_2 \dots, B_n \in \mathcal{B}_0^d$ disjoint Borel sets;*

is called the Poisson point process (on $\mathcal{X}$ with intensity measure Λ). Such a process we will denote Poisson process $(\mathcal{X}, \Lambda)$.
If X is a Poisson process on $S \subseteq \mathbb{R}^d$ with constant intensity function then it is called a homogeneous Poisson process. Otherwise it is called inhomogeneous.
The homogeneous Poisson process with the intensity measure equal to the Lebesgue measure is called the unit Poisson process.
The random measure $X(.)$ of a Poisson point process X is called a Poisson random measure.

Definition 1.9: *Let μ be a locally finite diffusion measure on $\mathbb{R}^d$ and $B \in \mathcal{B}^d$ such that $0 < \mu(B) < \infty$. Let $n \in \mathbb{N}$, $X_1, X_2, \dots, X_n$ be independent identically distributed d-dimensional random vectors such that:*

$$\mathsf{P}(X_i \in A) = \frac{\mu(A)}{\mu(B)}$$

for each $A \subset B, A \in \mathcal{B}^d$. Then $X = (X_1, X_2, \dots, X_n)$ is called binomial point process on B.

Theorem 1.2: *Let X be a Poisson point process on $B \in \mathcal{B}^d$ with intensity measure Λ such that $\Lambda(B) < \infty$. Then, conditionally on $X(B) = n$, X is binomial process on B with $\mu = \Lambda$.*

Theorem 1.3: *Let Λ be a locally finite diffusion measure on $\mathbb{R}^d$. Then Poisson process with intensity measure Λ exists and it is uniquely determined by Λ.*

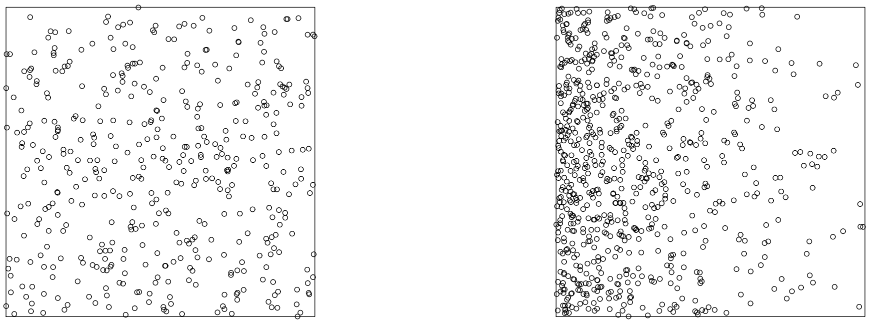

Figure 1.1: Poisson point process. Left a) - Homogeneous Poisson process with intensity 5 in observation window $[0, 10] \times [0, 10]$, Right b) - Inhomogeneous Poisson process with intensity function $\lambda(x, y) = 25 \exp(-x/3)$ in observation window $[0, 10] \times [0, 10]$.

Definition 1.10: *Suppose that $Z = \left\{Z(\xi), \xi \in \mathbb{R}^d\right\}$ is a nonnegative random field so that with probability one $\xi \to Z(\xi)$ is a locally integrable function. If the conditional distribution of a point process X given $Z = z$ is that of a Poisson process on $\mathbb{R}^d$ with the intensity function z, then X is called a Cox process driven by Z. Z is called a driving intensity function and $\Lambda(B) = \int_B Z(\xi)d\xi$, $B \in \mathcal{B}$ is a driving (intensity) measure of the Cox process.*

The Cox process is a generalization of the Poisson process. It is also called a doubly stochastic Poisson point process (see [16]).

Example 1.1: [Mixed Poisson process] If any realization of the random field Z from definition 1.10 is constant in $\mathbb{R}^d$ then the Cox process driven by random variable Z is called a *mixed Poisson process*. You can see three realizations of a mixed Poisson process in Figure 1.2.

Definition 1.11: *$Y : \Omega \times \mathbb{R}^d \to \mathbb{R}$ is a Gaussian random field if all of its finite-dimensional distributions are multidimensional normal distributions. Let X be a Cox process on $\mathbb{R}^d$ driven by $Z = exp(Y)$ where Y is a Gaussian random field. Then X is called a log-Gaussian Cox process (LGCP).*

Definition 1.12: *Let X_g be a point process and $\mathcal{W}$ be a complete separable metric space. Let a random mark $w_i \in \mathcal{W}$ be attached to each point $x_i \in X_g$ then*

$$X_m = \{(x_i, w_i)\}$$

is a marked point process with points in $\mathbb{R}^d$ and marks in $\mathcal{W}$. The process X_g is called a ground process of the marked point process X_m. If the marks k_i are identically distributed with distribution Q on $B(\mathcal{W})$, then Q is called the mark distribution.

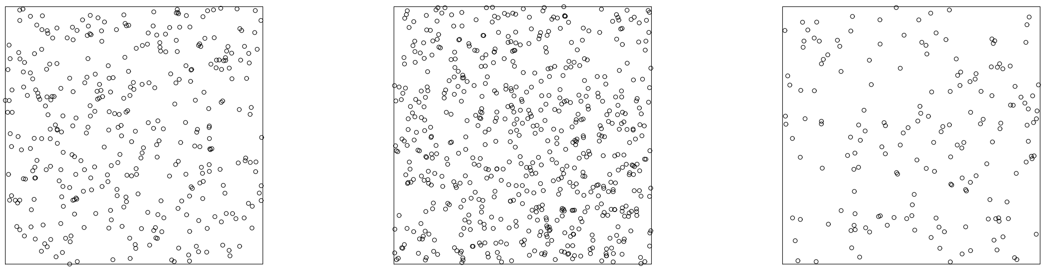

Figure 1.2: A mixed Poisson point process in observation window $[0, 10] \times [0, 10] \subset \mathbb{R}^2$. The realization of driving intensity function is 3.44 (left), 6.42 (middle) and 1.67 (right).

Example 1.2: [Marked Poisson process] If X is a Poisson process $(\mathcal{X}, \Lambda)$ and conditionally on X, the marks w_i attached to each point $x_i \in X$ are mutually independent. Then $X_m = \{(x_i, w_i)\,; x_i \in X\}$ is called a marked Poisson process.

Definition 1.13: *Let X be a point process with intensity measure Λ. Then the Campbell measure of X is defined as*

$$C(A) = \mathbb{E} \sum_{x \in X} \mathbf{1}_{[(x,X) \in A]}, \qquad A \in \mathcal{B}^d \times \mathfrak{N}. \tag{1.6}$$

The Campbell measure can be also characterized by the property

$$C(B \times U) = \mathbb{E} X(B) \mathbf{1}_{[X \in U]}, \quad U \in \mathfrak{N},\ B \in \mathcal{B}^d \tag{1.7}$$

Definition 1.14: *Let $(S, \mathcal{S})$ and $(T, \mathcal{T})$ are two measurable spaces. Then the function $P : S \times \mathcal{T} \mapsto [0, 1]$ is called Markov kernel from $(S, \mathcal{S})$ to $(T, \mathcal{T})$ if*

1. *$P(., A)$ is non-negative measurable function for all $A \in \mathcal{T}$,*
2. *$P(x, .)$ is a probability measure for all $x \in S$.*

Lemma 1.2: *Let X be a point process with a σ-finite intensity measure Λ. Then there exists a Markov kernel P from $(\mathbb{R}^d, \mathcal{B}^d)$ to $(\mathcal{N}, \mathfrak{N})$ such that*

$$C(B \times U) = \int_B P(x, U) \Lambda(dx) \qquad B \in \mathcal{B}^d,\ U \in \mathfrak{N}. \tag{1.8}$$

The Markov kernel $P(x, U) = P_x(U)$ is called a Palm distribution at the point $x \in \mathbb{R}^d$.

Proof: See [18]. □

Definition 1.15: *Let X be a stationary point process with intensity λ. Then the reduced second moment measure $\mathcal{K}$ is defined as*

$$\lambda\mathcal{K}(B) = \int_{\mathcal{N}} X(B \setminus \{0\})P_0(dX) \quad B \in \mathcal{B}^d.$$

If X is stationary and isotropic define the K-function as

$$K(r) = \mathcal{K}(b(0,r)), \quad r > 0,$$

where $b(0,r)$ is a ball centred in 0 with radius r.

1.1.1 Finite point processes with density

Let X_P be a finite Poisson point process on $\mathbb{R}^d$ with a diffusion intensity measure Λ_P. Then $\Lambda_P(\mathbb{R}^d) < +\infty$ and the distribution of X_P can be expressed [36] as

$$\Pi(U) = \mathsf{P}(X_P \in U) = \sum_{n=0}^{\infty} \mathsf{P}\left(X_P(\mathbb{R}^d) = n\right) \mathsf{P}\left(X_P \in U | X_P(\mathbb{R}^d) = n\right) =$$

$$= \sum_{n=0}^{\infty} \frac{\Lambda_P(\mathbb{R}^d)^n}{n!} \exp\{-\Lambda_P(\mathbb{R}^d)\} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \mathbf{1}_{[(x_1,x_2,\dots x_n)\in U]} \frac{\Lambda_P(dx_1)}{\Lambda_P(\mathbb{R}^d)} \cdots \frac{\Lambda_P(dx_n)}{\Lambda_P(\mathbb{R}^d)} =$$

$$= \exp\{-\Lambda_P(\mathbb{R}^d)\} \left[\mathbf{1}_{[\emptyset\in U]} + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \mathbf{1}_{[(x_1,x_2,\dots x_n)\in U]} \Lambda_P(dx_1) \dots \Lambda_P(dx_n)\right]$$

for $U \in \mathfrak{N}$.

Suppose that the distribution of a finite point process X can be expressed by

$$\mathsf{P}(X \in U) = \exp\{-\Lambda_P(\mathbb{R}^d)\}\times$$

$$\times \left[p(\emptyset)\mathbf{1}_{[\emptyset\in U]} + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \mathbf{1}_{[(x_1,x_2,\dots x_n)\in U]} p(x_1,\dots x_n) \Lambda_P(dx_1) \dots \Lambda_P(dx_n)\right]$$

$$= \int_U p(\varphi)\Pi(d\varphi), \tag{1.9}$$

for some non-negative function p.

Definition 1.16: *Function $p : \mathcal{N} \to \mathbb{R}_+$ in (1.9) is called the density of X with respect to X_P.*

Specially, if the intensity measure Λ_P is the restriction of Lebesgue measure on some bounded subset $B \in \mathbb{R}^d_0$, then X is a point process on B with the density p with respect to the unit Poisson process.

Example 1.3: Let X be the Poisson point process on $B \in \mathcal{B}^d$ with intensity function λ. Then

$$\mathsf{P}(X \in U) =$$

$$= \exp\left\{-\int_B \lambda(x)dx\right\}\left[\mathbf{1}_{[\emptyset \in U]} + \sum_{n=1}^{\infty}\frac{1}{n!}\int_B \dots \int_B \mathbf{1}_{[(x_1,x_2,\dots x_n)\in U]}\lambda(x_1)\dots\lambda(x_n)dx_1\dots dx_n\right] =$$

$$= \exp\left\{-\int_B 1dx\right\}\exp\left\{-\int_B (\lambda(x)-1)\,dx\right\}\mathbf{1}_{[\emptyset\in U]} +$$

$$\exp\left\{-\int_B 1dx\right\}\sum_{n=1}^{\infty}\frac{1}{n!}\int_B \dots \int_B \mathbf{1}_{[(x_1,x_2,\dots x_n)\in U]}\exp\left\{-\int_B (\lambda(x)-1)\,dx\right\}\prod_{i=1}^{n}\lambda(x_i)dx_1\dots dx_n$$

for $U \in \mathfrak{N}$. So the density of X with respect to unit Poisson process is

$$p(\varphi) = \exp\left\{-\int_B (\lambda(x)-1)\,dx\right\}\prod_{x_i\in\varphi}\lambda(x_i), \quad \varphi \in \mathcal{N}. \tag{1.10}$$

1.2 Spatio-temporal point processes

Spatio-temporal point process can be considered as a point process in $\mathbb{R}\times\mathbb{R}^d$ or $\mathbb{R}_+\times\mathbb{R}^d$. Each point represents time and location of some event (see [44]). Locations and times of earthquakes in some region are example of such a process (see [38]). The important feature of spatio-temporal point processes is the fact that there is the temporal coordinate and we can see the evolutionary progress. Also in some cases the spatial coordinates can be seen as d-dimensional marks in temporal point process and spatio-temporal point process is a marked temporal process. The following background is from [46].

1.2.1 Conditional intensity

In the following we will consider a spatio-temporal process $X = \{(t_i, \xi_i)\}$ on $\mathbb{R}_+ \times \mathcal{X}$, $\mathcal{X}$ bounded subset of $\mathbb{R}^d$ with positive Lebesgue measure $|\mathcal{X}|$ and such that its projections on $\mathcal{X}$ (space) and on $\mathbb{R}_+$ (time) are point processes. Suppose that

$$t_1 < t_2 < \dots < t_n < \dots$$

and process

$$X_t = \{(t_i, \xi_i) \in X : t_i \le t\} \text{ on } S_t = [0, t] \times \mathcal{X}$$

has a density g_{X_t} with respect to the unit Poisson process. Define two families of conditional probability densities,

$$\left\{p_n(t|x_{t_{n-1}}) : n \in \mathbb{N}\right\} \tag{1.11}$$

and

$$\left\{f_n(\xi|x_{t_{n-1}}, t_n) : n \in \mathbb{N}\right\} \tag{1.12}$$

with respect to the Lebesgue measure. The density p_n is the density of the time of $n-$th point given the history up to time t_{n-1} and the density f_n is the density of the location of $n-$th point given the history up to time t_{n-1} and the time of $n-$th point.

Theorem 1.4: *The density of point process X_t on $[0,t]\times\mathcal{X}$ can be expressed as*

$$g_{X_t}(x) = \exp(t|\mathcal{X}|)\prod_{i=1}^{n} p_i(t_i|x_{t_{i-1}})f_i(\xi_i|x_{t_{i-1}},t_i)S_{n+1}(t|x_{t_n}).$$

where

$$S_{n+1}(t|x_{t_n}) = \int_t^{\infty} p_{n+1}(t|x_{t_n})dt, \quad t > t_n$$

is the survival function of $p_{n+1}(.|x_{t_n})$.

Proof:

$$\mathsf{P}(X_t \in U) = \int_U g_{X_t}(z)\Pi(dz) = \exp\{-t|\mathcal{X}|\}\, g_{X_t}(\emptyset)\mathbf{1}_{[\emptyset\in U]}+$$

$$+\sum_{n=1}^{\infty}\exp\{-t|\mathcal{X}|\}\frac{(t|\mathcal{X}|)^n}{n!}\int_{\mathcal{X}}\cdots\int_{\mathcal{X}}\int_0^t\cdots\int_0^t \mathbf{1}_{[z\in U]}\frac{1}{(t|\mathcal{X}|)^n}g_{X_t}(z)dz =$$

$$= \exp\{-t|\mathcal{X}|\}\, g_{X_t}(\emptyset)\mathbf{1}_{[\emptyset\in U]}+ \tag{1.13}$$

$$+\sum_{n=1}^{\infty}\exp\{-t|\mathcal{X}|\}\int_{\mathcal{X}}\cdots\int_{\mathcal{X}}\int_0^t\int_{t_1}^t\cdots\int_{t_{n-1}}^t \mathbf{1}_{[\{t_{(n)},\xi_{(n)}\}\in U]}g_{X_t}(t_{(n)},\xi_{(n)})d\xi_1\ldots d\xi_n dt_1\ldots dt_n$$

and

$$\mathsf{P}(X_t\in U) = \sum_{n=0}^{\infty}\mathsf{P}\left(X_t\in U, X_t([0,t]\times\mathcal{X}) = n\right) =$$

$$= \sum_{n=1}^{\infty}\int_{\mathbb{R}_+\times\mathcal{X}}\cdots\int_{\mathbb{R}_+\times\mathcal{X}}\mathbf{1}_{[\{t_{(n)},\xi_{(n)}\}\in U]}\mathbf{1}_{[t_{n-1}>t]}\times$$

$$\times\prod_{i=1}^{n+1}p_i\left(t_i|t_{(i-1)},\xi_{(i-1)}\right)f_i\left(\xi_i|t_{(i-1)},\xi_{(i-1)},t_i\right)d\xi_1\ldots d\xi_{n+1}dt_1\ldots dt_{n+1} =$$

$$= \sum_{n=1}^{\infty}\int_{\mathcal{X}}\cdots\int_{\mathcal{X}}\int_0^t\cdots\int_{t_{n-1}}^t\mathbf{1}_{[\{t_{(n)},\xi_{(n)}\}\in U]}\times \tag{1.14}$$

$$\times\prod_{i=1}^{n}p_i\left(t_i|t_{(i-1)},\xi_{(i-1)}\right)f_i(\xi_i|x_{t_{i-1}},t_i)S_{n+1}(t|t_{(n)},\xi_{(n)})d\xi_1\ldots d\xi_n dt_1\ldots dt_n$$

By comparing equations (1.13) and (1.14), we will get

$$g_{X_t}(x) = \exp(t|\mathcal{X}|)\prod_{i=1}^{n} p_i(t_i|x_{t_{i-1}})f_i(\xi_i|x_{t_{i-1}},t_i)S_{n+1}(t|x_{t_n}).$$

$\square$

Definition 1.17: *For the spatio-temporal point process the conditional intensity function is*

$$\lambda^*(t,\xi|\mathcal{H}_t) = \lambda_g(t) f^*(\xi|t), \quad \text{if } t_{n-1} < t \leq t_n, \tag{1.15}$$

where

$$\lambda_g(t) = \frac{p_n(t|X_{t_{n-1}})}{S_n(t|X_{t_{n-1}})}, \quad \text{if } t_{n-1} < t \leq t_n,$$

$$f^*(\xi|t) = f_n(\xi|X_{t_{n-1}}, t), \quad \text{if } t_{n-1} < t \leq t_n,$$

and $\mathcal{H}_t = \sigma\{X_s, s < t\}$ *denotes the* σ*-algebra of the process until time* t*.*

The interpretation of $\lambda^*(t,\xi)d\xi dt$ is as the probability of observing a point at the location (t,ξ) conditioned by the history up to time t. It can be shown that the density of X_t can be written as

$$g_{X_t}(x) = \exp\left(-\int_{[0,t)\times\mathcal{X}} [\lambda^*(s,\xi)|\mathcal{H}_s) - 1]\, d(s,\xi)\right) \prod_{i=1}^{n} \lambda^*(t_i,\xi_i|\mathcal{H}_{t_i}),$$

where

$$x = \{(t_1,\xi_1),\ldots,(t_n,\xi_n)\}, \quad t_1 < \ldots < t_n.$$

Example 1.4: [Poisson process] If the process X is a Poisson point process then the conditional intensity function λ^* is non-random and equal to the intensity function λ and thus the density of X_t is

$$g_{X_t}(x) = \exp\left(-\int_{S_t} [\lambda(u) - 1]\, du\right) \prod_{i=1}^{n} \lambda(t_i,\xi_i), \tag{1.16}$$

(cf. (1.10)). The densities (1.11) and (1.12) are given by

$$p_n(t|X_{t_{n-1}}) = \lambda_g(t) \exp\left(-\int_{t_{n-1}}^{t} \lambda_g(s)ds\right), \quad t > t_{n-1}$$

and

$$f_n(\xi|X_{t_{n-1}}, t_n) = \frac{\lambda(t_n,\xi)}{\lambda_g(t_n)}, \quad \xi \in \mathcal{X}$$

where

$$\lambda_g(t) = \int_{\mathcal{X}} \lambda(t,\xi)d\xi.$$

These results hold under the assumption that

$$\int_t^{\infty} \lambda_g(u)du = \infty \;\text{ for all } t \geq 0$$

Denote the time of arrival of $n-$th point by T_n. Then, given (ξ_{n-1}, t_{n-1}),

$$\int_{t_{n-1}}^{T_n} \lambda_g(s)ds$$

is exponentially distributed with parameter 1.

1.2.2 Cox processes

A spatio-temporal Cox process X is a Cox point process (cf. Definition 1.10) with the intensity function given by $\lambda(t,\xi) = \mathbb{E}(\Lambda(t,\xi)),\ t \in \mathbb{R}_+,\ \xi \in \mathcal{X}$, where Λ is the random driving intensity function. Because

$$M^{(2)}(C \times D) = \mathbb{E} \sum_{X_1,X_2 \in X} \mathbf{1}_{[X_1,X_2 \in C \times D]} = \mathbb{E}\left[\mathbb{E} \sum_{X_1,X_2 \in X} \mathbf{1}_{[X_1,X_2 \in C \times D]} | \Lambda\right] = \mathbb{E}\Lambda_m(C)\Lambda_m(D),$$

where $\Lambda_m(.) = \int \Lambda(x)dx$, the pair correlation is given by

$$\rho\left((t,\xi),(s,\eta)\right) = \frac{\mathbb{E}(\Lambda(t,\xi)\Lambda(s,\eta))}{\mathbb{E}(\Lambda(t,\xi))\mathbb{E}(\Lambda(s,\eta))}.$$

Two basic types of spatio-temporal Cox processes follow namely the shot noise Cox process and the log-Gaussian Cox process.

Definition 1.18: *Shot noise Cox process is a Cox process X driven by*

$$\Lambda(t,\xi) = \sum_{(u,c,\gamma) \in \Phi} \gamma k\left((u,c),(t,\xi)\right) \tag{1.17}$$

where $k(.,.)$ is a kernel function, i.e. $\left(\int k(x,y)dy = 1\right)$ and Φ is a marked Poisson process with points in $\mathbb{R}_+ \times \mathcal{X}$ and marks in $\mathbb{R}$.

Log-Gaussian Cox process was defined in Section 1.1. In spatio-temporal applications some variants were used.

In the analysis of weed data, Brix and Moller (see [12]) considered the following model of the driving intensity

$$\Lambda(t,\xi) = m(t,\xi)\exp(W(\xi)),$$

where m is a mean function satisfying

$$m(t',\xi) \le m(t,\xi) \quad \text{for } t' \le t,\ \xi \in \mathcal{X},$$

and W is a zero mean Gaussian process on $\mathcal{X}$.

Brix and Diggle (see [11]) considered a model

$$\Lambda(t,\xi) = \lambda(t,\xi)\exp(S(t,\xi)),$$

where the mean of a Gaussian process $S(t,\xi)$ on $\mathbb{R}_+ \times \mathcal{X}$ is a constant. In the stationary case, a convenient parametrisation is to set $\mathbb{E}\left[S(t,\xi)\right] = -0.5\sigma^2$, where $\sigma^2 = \mathrm{Var}\left[S(t,\xi)\right]$. This gives $\mathbb{E}\left[\exp\left\{S(t,\xi)\right\}\right] = 1$, and hence $\lambda(t,\xi)$ is the unconditional space-time intensity. $S(t,\xi)$ has mean $-0.5\sigma^2$ and covariance function $\sigma^2\rho(t,t',\xi,\xi')$. If we denote $u = (\xi,\xi') \in \mathcal{X}^2$ and $v = (t,t') \in \mathbb{R}_+^2$

$$\rho(v,u) = r(u)\exp(-v/\beta),$$

where r is a function of a temporal coordinate.

1.2.3 Mechanistic modelling

While the previous modelling is called empirical (see [21]), in the present subsection we consider so called mechanistic modelling in which we would like to explain how the evolution of the process depends on its past history. This leads to the investigation of the conditional intensity λ^*, (see (1.15)). Given a parametric model of the conditional intensity of the point process X on $[0,T] \times \mathcal{X}$ with parameter $\theta \in \mathbb{R}^d$ and given data $(t_i, \xi_i) \in [0,T] \times \mathcal{X} : i = 1, \ldots, n$, with $t_1 < t_2, \ldots < t_n$, the log-likelihood is

$$L(\theta) = \sum_{i=1}^{n} \log(\lambda^*(t_i, \xi_i)|\mathcal{H}_{t_i-}) - \int_0^T \int_{\mathcal{X}} \lambda^*(t, \xi|\mathcal{H}_t) d\xi dt. \tag{1.18}$$

Since the integral in (1.18) is difficult to compute Diggle ([21]) suggested instead maximizing a partial likelihood for the inference. Let

$$p_i = \frac{\lambda^*(\xi_i, t_i|\mathcal{H}_{t_i})}{\sum_{j=i}^{n} \lambda^*(x_j, y_i|\mathcal{H}_{t_i})}. \tag{1.19}$$

Then, the partial log-likelihood is

$$L_p(\theta) = \sum_{i=1}^{n} \log p_i. \tag{1.20}$$

This method was used with data of 2001 foot-and-mouth epidemic in England.

1.3 Random measures

A measure μ on $(\mathbb{R}^d, \mathcal{B}^d)$ is said to be a locally finite measure if it is finite on bounded Borel sets. By $\mathcal{M} \equiv \mathcal{M}(\mathbb{R}^d)$ we denote the set of all locally finite measures on $(\mathbb{R}^d, \mathcal{B}^d)$.

Let $\mathfrak{M}$ be the smallest σ-algebra on $\mathcal{M}$ with respect to which the function $\mu \mapsto \mu(B)$ is measurable for all $B \in \mathcal{B}^d$.

Definition 1.19: *A random measure on $\mathbb{R}^d$ is a measurable mapping*

$$\Psi : (\Omega, \mathcal{F}, \mathsf{P}) \to (\mathcal{M}, \mathfrak{M}).$$

The probability measure $\mathsf{P}\Psi^{-1}$ is the distribution of the random measure Ψ and the measure $\Lambda(\cdot) = \mathbb{E}\Psi(\cdot)$ is called the intensity measure of Ψ.

Definition 1.20: *Let Ψ be a random measure on $\mathbb{R}^d$ with distribution P and $k \in \mathbb{N}$. The measure*

$$M_k(\cdot) = \mathbb{E}\Psi^k(\cdot) = \int \mu^k(\cdot)\, P(d\mu)$$

on $(\mathbb{R}^{dk}, \mathcal{B}^{dk})$ is called the moment measure of k-th order of Ψ.

Specially $M_1 \equiv \Lambda$ is the intensity measure of Ψ.

Definition 1.21: *Let Ψ be a random measure on $\mathbb{R}^d$ with distribution P and intensity measure Λ. The Campbell measure C corresponding to Ψ is a measure on $\mathbb{R}^d \times \mathcal{M}$ defined by*

$$\int_{\mathbb{R}^d \times \mathcal{M}} f(x, \mu) C(d(x, \mu)) = \int_{\mathcal{M}} \int_{\mathbb{R}^d} f(x, \mu) \mu(dx) P(d\mu),$$

where f is an arbitrary nonnegative measurable function on $\mathbb{R}^d \times \mathcal{M}$ (cf. (1.6)).

Note that the Campbell measure C can also be characterized by the property (cf. (1.7))

$$C(A \times \mathcal{U}) = \mathbb{E}\Psi(A) \mathbf{1}_{\mathcal{U}}(\Psi),$$

where $A \in \mathcal{B}_0^d$ and $\mathcal{U}$ a measurable subset of $\mathcal{M}$.

Definition 1.22: *Let Ψ be a random measure on $\mathbb{R}^d$ with distribution P and a locally finite intensity measure Λ. Then there exists a probability kernel $x \mapsto P_x$ from $(\mathbb{R}^d, \mathcal{B}^d)$ to $(\mathcal{M}, \mathfrak{M})$ such that*

$$\int_{\mathcal{M}} \int_{\mathbb{R}^d} f(x, \mu) \mu(dx) P(d\mu) = \int_{\mathbb{R}^d} \int_{\mathcal{M}} f(x, \mu) P_x(d\mu) \Lambda(dx) \tag{1.21}$$

for an arbitrary nonnegative measurable function f on $\mathbb{R}^d \times \mathcal{M}$. The distribution P_x is called the Palm distribution of the random measure Ψ at the point $x \in \mathbb{R}^d$ (cf. (1.8)).

If $(P'_x : x \in X)$ is another probability kernel satisfying (1.21) then for any measurable set $\mathcal{U} \subset \mathcal{M}$,

$$P_x(\mathcal{U}) = P'_x(\mathcal{U}) \text{ for } \Lambda \text{ almost all } x \in X.$$

For $z \in \mathbb{R}^d$, let t_z denote the corresponding shift operator on $\mathcal{M}$ defined by

$$t_z \mu(B) = \mu(B - z), \quad B \in \mathcal{B}^d.$$

Definition 1.23: *The random measure Ψ is called stationary if its distribution is shift invariant, i.e., if $t_z \Psi$ has the same distribution as Ψ for any $z \in \mathbb{R}^d$.*

A stationary random measure Ψ on $\mathbb{R}^d$ with intensity $\lambda > 0$ has Palm distributions

$$P_x(\mathcal{U}) = P_0(t_x^{-1} \mathcal{U}), \quad x \in \mathbb{R}^d,\ \mathcal{U} \in \mathcal{M}.$$

Definition 1.24: *The reduced second moment measure $\mathcal{K}$ of a stationary random measure is defined by*

$$\mathcal{K}(B) = \lambda^{-1} \int \mu(B \setminus \{0\}) P_0(d\mu), \quad B \in \mathcal{B}_0^d. \tag{1.22}$$

The K-function is defined by

$$K(r) = \mathcal{K}(b(0, r)), \quad r > 0. \tag{1.23}$$

We can also write

$$\mathcal{K}(B) = \frac{1}{\lambda^2|A|}\mathbb{E}\int_A \Psi(B-x)\Psi(dx), \tag{1.24}$$

with an arbitrary bounded Borel set A of positive Lebesgue measure.

We can introduce also an inhomogeneous reduced second moment measure $\mathcal{K}_{inhom}$ of a random measure Ψ with intensity function $\lambda > 0$.

Definition 1.25: *Let the measure*

$$M(C,B) = \mathbb{E}\left[\int_C \frac{\Psi(dy)}{\lambda(y)} \int_B \frac{\Psi(dx)}{\lambda(x)}\right], \quad C, B \in \mathcal{B}_0^d,$$

be finite. Ψ is a second-order intensity-reweighted stationary (SOIRS) random measure if it holds

$$M(C,B) = M(C+x, B+x), \quad x \in \mathbb{R}^d.$$

Under SOIRS assumption we define (cf. [1] for point processes)

$$\mathcal{K}_{inhom}(B) = \frac{1}{|A|}\mathbb{E}\int_A \int \frac{1_B(x+y)}{\lambda(x)\lambda(y)}\Psi(dy)\Psi(dx)$$

independently of the choice of A. If there exists a second-order product density $\lambda^{(2)}$ such that

$$M(C,B) = \int_C \int_B \frac{\lambda^{(2)}(u,v)}{\lambda(u)\lambda(v)} dudv,$$

then $\rho(u,v) = \frac{\lambda^{(2)}(u,v)}{\lambda(u)\lambda(v)}$ is the pair correlation function, cf. (1.5).

Definition 1.26: *Let Ψ be a random measure in $\mathbb{R}^d$, let C be its Campbell measure and let W be a locally compact space. Let w be a measurable mapping (weight function)*

$$w : suppC \to W$$

(we consider the natural product σ-algebra on $suppC \subset \mathbb{R}^d \times \mathcal{M}$). Then, we call the pair (Ψ, w) a weighted random measure in $\mathbb{R}^d$ with weight space W.

Note that a weighted random measure induces a random measure $\tilde{\Psi}$ on the product space $\mathbb{R}^d \times W$:

$$\tilde{\Psi}(B \times D) = \Psi\{x \in B : w(x,\Psi) \in D\}, \tag{1.25}$$

$B \in \mathcal{B}^d$, $D \in \mathcal{B}(W)$.

Definition 1.27: *We say that the weighted random measure (Ψ, w) is stationary if Ψ is stationary and the weight function is translation covariant, i.e., $w(x,\mu) = w(x+z, t_z\mu)$ for any $(x,\mu) \in suppC$ and $z \in \mathbb{R}^d$.*

1.4 Random closed sets and fibre processes

The following theory is from [6]. Denote ω_k the volume of a unit ball in $\mathbb{R}^k$.

Definition 1.28: *Let $k \in \{0, 1, \ldots, d\}$ be fixed. The Hausdorff measure $\mathcal{H}^k$ of order k in $\mathbb{R}^d$ is defined as*

$$\mathcal{H}^k(A) = \lim_{\delta \to 0_+} \inf_{\substack{A \subset \bigcup_i G_i \\ diam\ G_i \leq \delta}} \sum_i \omega_k \left(\frac{diam\ G_i}{2} \right)^k,$$

where $diam\, G_i$ denotes the diameter of G_i and the infimum is taken over all at most countable coverings of A with sets of diameters less or equal to δ.

A mapping f is Lipschitz if there exits a constant M such that $||f(x) - f(y)|| \leq M||x - y||$ for any x, y from the domain of f.

Definition 1.29: *We call a subset $A \subset \mathbb{R}^d$ k-rectifiable if it is a Lipschitz image of a bounded subset of $\mathbb{R}^k$. A is $(\mathcal{H}^k, k)-$rectifiable if*

1. *A is $\mathcal{H}^k-$measurable,*
2. *$\mathcal{H}^k(A) < \infty$,*
3. *$A = \cup_{i=0}^{\infty} W_i$ with $\mathcal{H}^k(W_0) = 0$ and W_i $k-$rectifiable for $i \geq 1$.*

Finally A is $\mathcal{H}^k-$rectifiable if $A \cap K$ is $(\mathcal{H}^k, k)-$rectifiable for any $K \subset \mathbb{R}^d$ compact.

Basic class of random elements used in stochastic geometry are random closed sets.

Definition 1.30: *Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space, $\mathcal{C}^d$ be a system of all closed sets from $\mathbb{R}^d$ and $\mathfrak{C}^d = \sigma\left\{C^K :\ K \text{ is a compact subset in } \mathbb{R}^d\right\}$ where $C^K = \left\{D \in \mathcal{C}^d;\ D \cap K \neq \emptyset\right\}$. Then a random closed set X in $\mathbb{R}^d$ is measurable mapping from $(\Omega, \mathcal{F})$ to $\left\{\mathcal{C}^d, \mathfrak{C}^d\right\}$.*

To formalize the meaning of a $k-$dimensional random closed set, we use the general concept due to Zähle [47] who introduced random $\mathcal{H}^k-$sets as random closed sets in $\mathbb{R}^d$ which are $\mathcal{H}^k-$rectifiable. It is shown in [47] that the space

$$\mathcal{C}_k^d := \left\{F \in \mathcal{C}^d : F \text{ is } \mathcal{H}^k - \text{rectifiable}\right\}$$

is an $\mathfrak{C}^d-$measurable subsystem of $\mathcal{C}^d$.

For particular dimension, a random $\mathcal{H}^1-$set will be called a (random) fibre system and a random $\mathcal{H}^{d-1}-$set a (random) surface system.

Here we use also a definition of a random fibre process from [45] based on differential geometry. It is less general, however suitable for us since we will need a curve parametrization by time in Chapter 2.

Definition 1.31: *Fibre y in $\mathbb{R}^d$ is a subset of $\mathbb{R}^d$ which is image of a curve $y(t) = (y_1(t), \ldots, y_d(t))$ such that*

1. $y : [0, T] \mapsto \mathbb{R}^d$ *is once continuously differentiable,*

2. $||y'(t)||^2 = \sum_{i=1}^{d} |y_i'(t)|^2 > 0$ *for* $\forall t \in [0, T]$*, where* y' *is the derivative,*

3. the mapping y *is one-to-one, so that a fibre does not intersect itself.*

We can define a length measure

$$y(B) = \int_0^T \mathbf{1}_B(y(t)) \sqrt{\sum_{i=1}^{d} |y_i'(t)|^2 dt}$$

for any set $B \in \mathcal{B}^d$.

Definition 1.32: *The fibre system* Y *in* $\mathbb{R}^d$ *is a closed subset of* $\mathbb{R}^d$ *which can be represented as a union of at most countably many fibres* $y^{(i)}$ *in* $\mathbb{R}^d$ *with the property that any compact set* K *is intersected by a finite number of the fibres, and such that distinct fibres have only endpoints in common:*

$$y^{(i)}((0, T)) \cap y^{(j)}((0, T)) = \emptyset \ \text{if} \ i \neq j.$$

The length measure of the fibre system Y *is a measure defined as*

$$Y(B) = \sum_{y^{(i)} \in Y} y^{(i)}(B), \quad B \in \mathcal{B}^d.$$

Denote the family of all fibre systems by $\mathbb{D}$ and generate the σ-algebra

$$\mathcal{D} = \sigma \left\{ Y \in \mathbb{D} : \ Y(B) < x, \ B \in \mathcal{B}^d, x \in \mathbb{R} \right\}.$$

This σ-algebra is in fact the trace σ-algebra $\mathcal{D} \cap \mathfrak{C}^d$ on the space of closed sets.

Definition 1.33: *A measurable mapping* $\mathcal{Y}$ *from* $(\Omega, \mathcal{F})$ *to* $(\mathbb{D}, \mathcal{D})$ *is called a fibre process. Distribution of the fibre process* $\mathcal{Y}$ *is a probability measure* Π *defined by*

$$\Pi(F) = \mathsf{P}(\{\omega; \mathcal{Y}(\omega) \in F\}) \ \text{for} \ F \in \mathcal{D}.$$

The length measure $\mathcal{Y}(.)$ of a fibre process is a random measure in the sense of Section 1.3. Thus we obtain the notion of the intensity measure $\Lambda(B) = \mathbb{E}\mathcal{Y}(B)$, second moment measure, stationarity, reduced second moment measure (Definition 1.24), SOIRS (Definition 1.25), $\mathcal{K}_{inhom}$ which are related to both fibre process and its length measure.

Definition 1.34: *Inhomogeneous* K*-function of SOIRS fibre process* $\mathcal{Y}$ *is*

$$K(t) = \frac{1}{|B|} \mathbb{E} \left[\int_B \int \frac{\mathbf{1}_{(||y-x|| \leq t)}}{\lambda(y)\lambda(x)} \mathcal{Y}(dx)\mathcal{Y}(dy) \right] \tag{1.26}$$

for $B \in \mathcal{B}_0^d$, $|B| > 0$. *This expression does not depend on the choice of* B.

1.5 Markov Chain Monte Carlo

Consider that we need to simulate from a target probability distribution on a measurable space and we do not know how to do it directly. In this case we can use the Markov Chain Monte Carlo methods (MCMC) [32]. The idea is based on simulating of a Markov chain with a state space $\mathcal{X}$ and with stationary distribution equal to the target distribution. Metropolis-Hastings algorithm and Gibbs algorithm are basic MCMC algorithms.

Markov chain is a sequence $X = (X_0, X_1, \ldots)$ of a random elements in a complete separable metric space space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ such that conditional distribution of X_{n+1} given $(X_0, \ldots, X_n)$ is equal to the conditional distribution of X_{n+1} given X_n, $n \in \mathbb{N}$. The Markov kernel

$$P_n(x, A) = \mathsf{P}(X_{n+1} \in A | X_n = x), \; n \in \mathbb{N}, \; x \in \mathcal{X}, A \in \mathcal{B}(\mathcal{X})$$

is called a transition kernel of the Markov chain X. If $P_n = P$ does not depend on n, X is called homogeneous Markov chain. The n-th power of the kernel P is defined by a recursive formula

$$P^n(x, A) = \int_{\mathcal{X}} P^{n-1}(y, A) P(x, dy),$$

where we set $P^0(x, A) = \delta_x(A)$ (the Dirac measure). It is interpreted as the probability that the chain gets from state x to A in n steps.

Definition 1.35: *Let X be a homogeneous Markov chain on $\mathcal{X}$ with a transition kernel P. A probability distribution Π on $\mathcal{B}(\mathcal{X})$ is called a limiting distribution of X if*

$$\lim_{n\to\infty} P^n(x, A) = \Pi(A) \quad \textit{for } \Pi\textit{-almost all } \; x \in \mathcal{X} \; \textit{and } \forall A \in \mathcal{B}(\mathcal{X}).$$

Definition 1.36: *A probability distribution Π on $\mathcal{B}(\mathcal{X})$ is called a stationary distribution of a homogeneous Markov chain X with the transition kernel P if*

$$\Pi(A) = \int_{\mathcal{X}} P(x, A) \Pi(dx) \quad \textit{for all } A \in \mathcal{B}(\mathcal{X}).$$

Theorem 1.5: *If Π is a limiting distribution of a homogeneous Markov chain X then Π is also stationary distribution of X.*

Homogeneous Markov chain X with the transition kernel P is reversible with respect to a distribution ϕ on $\mathcal{B}(\mathcal{X})$ if

$$\int_A P(x, B)\phi(dx) = \int_B P(y, A)\phi(dy)$$

for $A, B \in \mathcal{B}(\mathcal{X})$. Markov chain X is irreducible with respect to a distribution ϕ if

$$\phi(A) > 0, \; A \in \mathcal{B}(\mathcal{X}) \Rightarrow \mathsf{P}(\min\{n \geq 1 : X_n \in A\} < \infty | X_0 = x) > 0$$

for all $x \in \mathcal{X}$.

Irreducibility and reversibility with respect to the target distribution imply that the chain has a limiting distribution which is according to Theorem 1.5 the desired target distribution.

1.5.1 Metropolis-Hastings algorithm

Let $Q(x, .)$ be a Markov kernel and $Q(x, dz) = q(x, z)dz,\ x, z \in \mathcal{X}$, q is called the proposal density. Let the target distribution have probability density f. Define acceptance probability by

$$\alpha(x, z) = \min\left\{1, \frac{f(z)q(z, x)}{f(x)q(x, z)}\right\}.$$

The Metropolis-Hastings algorithm of MCMC with T iterations is:

1. Choose $x^{(0)} \in \mathcal{X}$ and put $t = 0$,

   For $t < T$

2. generate z from the distribution $Q(x^{(t)}, .)$ and put $x^{(t+1)} = z$ with the probability $\alpha(x^{(t)}, z)$, otherwise put $x^{(t+1)} = x^{(t)}$.

Example 1.5: [Gaussian random walk] Let $\mathcal{X} = \mathbb{R}$ and $q(x, y) = q_0(y - x)$ where q_0 is the density of Gaussian distribution. This proposal is called a Gaussian random walk.

Theorem 1.6: *The target distribution is a stationary distribution of the Markov chain generated by the Metropolis-Hastings algorithm.*

1.5.2 Gibbs algorithm

Gibbs sampling uses the full condition distributions on a finite product $\mathcal{X}$ of complete separable metric spaces $\mathcal{X}_i,\ i = 1, \ldots, d$. Assume that the target density (w.r.t. a $\sigma-$finite Borel measure on $\mathcal{X}$) is $f(\theta)$ where $\theta = (\theta_1, \ldots, \theta_d) \in \mathcal{X}$ and consider that full condition distributions

$$f(\theta_j|\theta_1, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_d) = f(\theta_j|\theta_{-j}), \quad 1 \leq j \leq d,$$

are known and we are able to sample from them. Gibbs sampling can be described in the following way:

1. Choose $\theta^{(0)} \in \mathcal{X}$ and put $t = 0$,

   For $t < T$, for $1 \leq j \leq d$

2. generate $\theta_j^{(t+1)}$ from the distribution

$$f(\theta_j|\theta_1^{(t+1)}, \ldots, \theta_{j-1}^{(t+1)}, \theta_{j+1}^{(t)}, \ldots, \theta_d^{(t)}).$$

1.5.3 Simulation of point processes

Often we need to simulate a realization of the point process X. In this part we will show how to do that for some of the finite point processes X on $B \in \mathcal{B}^d$ (see [36]).

The easiest case is when X is homogeneous Poisson process on B with intensity λ. To simulate this process we need to simulate a random variable N from Poisson distribution with parameter $\lambda|B|$. Conditionally on $N = n$, according to Theorem 1.2 n independent uniformly distributed d-dimensional random vectors $X_1, X_2, \ldots, X_n$ on B are simulated. Then $\{X_1, X_2, \ldots, X_n\}$ is the realization of X.

Another case is if X is the inhomogeneous Poisson process with intensity function $\lambda(x) \leq \lambda_0$ on B. Then we can simulate the homogeneous Poisson process Y on B with intensity λ_0 and then we accept each point x_i of Y with probability $\lambda(x_i)/\lambda_0$.

For more complex processes we can use Markov Chain Monte Carlo methods. Suppose that X is a finite point process on B with target density p with respect to the unit Poisson process on B. Let $X^{(0)}$ be the initial realization of a point process (e.g. unit Poisson process). Then if we have $X^{(n-1)}$ we can simulate $X^{(n)}$ by adding or deleting a point or put $X^{(n)} = X^{(n-1)}$ (birth and death algorithm, see [36], Chapter 7).

Let $Q(X^{(n-1)})$ be the probability of proposing adding a point ψ to $X^{(n-1)}$, ψ has a birth density $b(x^{(n-1)}, \psi)$ on B. We accept the new realization with the new point ψ with probability

$$\alpha_b(x^{(n-1)}, \psi) = \min\left(1, h_b(x^{(n-1)}, \psi)\right),$$

where

$$h_b(x^{(n-1)}, \psi) = \frac{p(x^{(n-1)} \cup \psi)}{p(x^{(n-1)})} \frac{1 - Q(X^{(n-1)})}{Q(X^{(n-1)})} \frac{d(x^{(n-1)} \cup \psi, \psi)}{b(x^{(n-1)}, \psi)}, \qquad (1.27)$$

where $d(y, \psi)$ is a death probability of $\psi \in y,\ y \in \mathcal{N}$.

Then $1 - Q(X^{(n-1)})$ is the probability of proposing reducing a point ψ from $x^{(n-1)}$. If $x^{(n-1)} = \emptyset$ then put $x^{(n)} = \emptyset$. Otherwise sample a point $\psi \in x^{(n-1)}$ according to probabilities $d(x^{(n-1)}, \psi)$. We accept the new realization $x^{(n-1)} \setminus \{\psi\}$ with probability

$$\alpha_d(x^{(n-1)}, \psi) = \min\left(1, h_d(x^{(n-1)}, \psi)\right)$$

where

$$h_d(x^{(n-1)}, \psi) = \frac{p(x^{(n-1)} \setminus \psi)}{p(x^{(n-1)})} \frac{Q(X^{(n-1)})}{1 - Q(X^{(n-1)})} \frac{b(x^{(n-1)} \setminus \psi, \psi)}{d(x^{(n-1)}, \psi)}. \qquad (1.28)$$

We can simplify the equations (1.27) and (1.28) by setting $Q(.) = 1/2$, $b(.,.) = \frac{1}{|B|}$ and $d(x \cup \psi, .) = \frac{1}{x(B)+1}$. Then

$$h_b(x^{(n-1)}, \psi) = \frac{p(x^{(n-1)} \cup \psi)}{p(x^{(n-1)})} \frac{x^{(n-1)}(B) + 1}{|B|} \qquad (1.29)$$

and

$$h_d(x^{(n-1)}, \psi) = \frac{p(x^{(n-1)} \setminus \psi)}{p(x^{(n-1)})} \frac{|B|}{x^{(n-1)}(B)}. \qquad (1.30)$$

Also we can see that the probabilitiy of accepting the new realization does not depend on the normalizing constant of the density p so it is enough to know some function f such that $p \propto f$.

Under mild conditions birth and death algorithm yields the chain which is reversible and irreducible with respect to the target distribution, see [36], Chapter 7.

Chapter 2

Spatio-temporal Cox processes

2.1 Lévy based Cox processes

In this Chapter we will study the type of spatio-temporal point processes which are constructed from Lévy basis (see [5],[30]), which is a generalization of Lévy process. The background theory of Lévy processes can be found in [15] and [42].

Definition 2.1: *A random variable Y with a distribution F on $\mathbb{R}^d$ is said to be infinitely divisible if for any integer $n \geq 2$, there exist i.i.d. random variables $Y_1, \ldots, Y_n$ such that $Y_1 + \ldots + Y_n$ has the distribution F.*

Definition 2.2: *A stochastic process $\{X_t\}_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathsf{P})$ with values in $\mathbb{R}^d$ such that $X_0 = 0$ is called an additive process if*

1. *[Independent increments] for every increasing sequence of times $t_0, \ldots, t_n$ the random variable $X_{t_o}, (X_{t_1} - X_{t_0}), \ldots, (X_{t_n} - X_{t_{n-1}})$ are independent,*
2. *[Stochastic continuity] $\forall t \geq 0,\ \forall \epsilon > 0,\ lim_{h\to 0}\mathbb{P}\left(|X_{t+h} - X_t| > \epsilon\right) = 0,$*
3. *$\forall \omega \in \Omega$ the trajectory $X_t(\omega)$ is right continuous function for $\forall t \geq 0$ and have left limits for $\forall t > 0$.*

Definition 2.3: *An additive process $\{X_t\}_{t\geq 0}$ is called a Lévy process if the distribution of $X_{t+h} - X_t$ does not depend on t for any $h > 0,\ t \geq 0$.*

Theorem 2.1: *Let $\{X_t\}_{t\geq 0}$ be a Lévy process, resp. an additive process, then X_t has infinitely divisible distribution for $\forall t > 0$.*

Proof: See [42].

Theorem 2.2: *Let X be a random variable with infinitely divisible distribution on $\mathbb{R}^d$. Then*

$$\log(\mathbb{E}(\exp(i\zeta X))) = -\frac{1}{2}\langle A\zeta, \zeta\rangle + i\langle a, \zeta\rangle + \int_{\mathbb{R}^d} \left(\exp(i\langle \zeta, x\rangle) - 1 - i\langle \zeta, x\rangle \mathbf{1}_{[|x|\leq 1]}\right) \nu(dx)$$

where $\zeta \in \mathbb{R}^d$, A is a symmetric positive definite $d \times d$ matrix, $a \in \mathbb{R}^d$ and ν is a positive measure on $\mathbb{R}^d$ verifying:

$$\nu(\{0\}) = 0 \text{ and } \int_{\mathbb{R}^d} \left(|x|^2 \wedge 1\right) \nu(dx) < \infty.$$

Proof: See [42].

Definition 2.4: *The triplet (a, A, ν) from Theorem 2.2 is called the generating triplet (of the distribution of X) and the measure ν is called the Lévy measure. The generating triplet of the Lévy process $\{X_t\}_{t\geq 0}$ is the generating triplet of X_1.*

Consider $\mathbb{R}^d$ with the Borel σ-algebra $\mathcal{B}^d$. Let $Z = \{Z(A);\ A \in \mathcal{B}^d\}$ be an independently scattered random measure. That means that for every sequence $\{A_1, \ \ldots, \ A_n; n \in \mathbb{N}_0\}$ of disjoint sets in $\mathcal{B}^d$, the random variables $Z(A_n)$ are mutually independent and $Z(\cup_n A_n) = \sum_n Z(A_n)$ almost surely. Moreover assume that $Z(A)$ is infinitely divisible for all $A \in \mathcal{B}^d$, in this case Z is called a Lévy basis. Then the cumulant transform of the Lévy basis Z $\mathcal{C}\{\zeta \ddagger Z(A)\} = \log \mathbb{E}(e^{i\zeta Z(A)})$ can be written as

$$\mathcal{C}\{\zeta \ddagger Z(A)\} = i\zeta a(A) - \frac{1}{2}\zeta^2 b(A) + \int_{\mathbb{R}} \{e^{i\zeta x} - 1 - i\zeta x \mathbf{1}_{[|x|\leq 1]}\}\nu(dx, A), \qquad (2.1)$$

where a is a signed measure, b is a positive measure and $\nu(dx, A)$ is a Lévy measure on $\mathbb{R}$ for fixed $A \in \mathcal{B}$ and a measure on $\mathcal{B}^d$ in the second variable. The triplet (a, b, ν) is called the generating triplet of the Lévy basis.

Example 2.1: [Poisson Lévy basis] The example of Lévy basis is Poisson basis for which Z is a Poisson process with intensity measure λ. Because $Z(A)$ has a Poisson distribution we have

$$\mathcal{C}\{\zeta \ddagger Z(A)\} = i\zeta\lambda(A) + \int_{\mathbb{R}} \{e^{i\zeta x} - 1 - i\zeta \mathbf{1}_{[|x|\leq 1]}\}\delta_1(dx)\lambda(A) = \lambda(A)\left(e^{i\zeta} - 1\right),$$

where δ_1 is the Dirac measure at 1. The generating triplet of this basis is $(\lambda, 0, \delta_1\lambda)$.

Example 2.2: [Gaussian Lévy basis] The other example of a Lévy basis is Gaussian Lévy basis which has the generating triplet $(a, b, 0)$. Then for $A \in \mathcal{B}^d$

$$\mathcal{C}\{\zeta \ddagger Z(A)\} = i\zeta a(A) - \frac{1}{2}\zeta^2 b(A).$$

and $Z(A)$ has a Gaussian distribution $N(a(A), b(A))$.

The following background comes from [41]. Let $(a, 0, \nu)$ be a generating triplet of a Lévy basis.. Zero in the second place of the triplet implies that the cumulant transform (2.1) is

$$\mathcal{C}\{\zeta \ddagger Z(A)\} = i\zeta a(A) + \int_{\mathbb{R}} \{e^{i\zeta x} - 1 - i\zeta x \mathbf{1}_{[|x|\le 1]}\}\nu(dx, A), \quad \zeta \in \mathbb{R}.$$

It is important that ν can be factorized with no essential loss of generality as

$$\nu(dx, d\xi) = \mu(dx, \xi)U(d\xi), \tag{2.2}$$

where $\mu(dx, \xi)$ is a Lévy measure on $\mathbb{R}$ for fixed $\xi \in \mathbb{R}^d$ and $U(d\xi)$ is a measure on $\mathcal{B}^d$.

Then assuming that the density a' exists, $a(d\eta) = a'(\eta)U(d\eta)$, $\eta \in \mathbb{R}^d$, we can write

$$\mathcal{C}\{\zeta \ddagger Z'(\eta)\} = i\zeta a'(\eta) + \int_{\mathbb{R}} \{e^{i\zeta x} - 1 - i\zeta x \mathbf{1}_{[|x|\le 1]}\}\mu(dx, \eta), \tag{2.3}$$

$\zeta \in \mathbb{R}$, for an additive process $Z'(\eta)$. For a fine discussion about the correspondence of Z and Z' see [39]. We will in the end consider the situation when Z' is a compound Poisson process which is based on a Poisson process of jumps in $\mathbb{R}^d$, possibly inhomogeneous, and a fixed distribution of jump sizes (jump size being independent of location). In this case $\mu(A, \xi)$ is a finite measure for each $\xi \in \mathbb{R}^d$.

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.

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

$$\mathcal{C}\{\zeta \ddagger \int_{\mathbb{R}^d} f dZ\} = \int_{\mathbb{R}^d} \mathcal{C}\{\zeta f(\xi) \ddagger Z'(\xi)\}U(d\xi). \tag{2.4}$$

Proof: See [41]. □

We will apply Lévy basis to the theory of simple point processes in $\mathbb{R}^d$ and specially in space and time [18]. 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

$$\Lambda(\xi) = \int_{\mathbb{R}^d} g(\xi, \eta)Z(d\eta), \ \xi \in \mathbb{R}^d, \tag{2.5}$$

where g is a measurable function on $\mathbb{R}^{2d}$. For the compound Poisson process a sufficient condition for local integrability follows from the Campbell theorem [36]: the mean jump size distribution has to be finite and $h(\xi) = \int g(\xi, \eta)U(d\eta)$ should be an integrable function of ξ on each bounded set. See also [30] for the discussion of this issue.

For a Cox process in $\mathbb{R}^d$ (see Definition 1.10) let Λ_m be the driving measure and Λ its driving intensity function. The generating functional of a point process is defined as

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

for measurable functions $u : \mathbb{R}^d \mapsto [0,1]$ with bounded support, where $x_i,\ i = 1,\dots N$ are events of the point process observed within the support of u. For a Cox process X the generating functional has form

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

Theorem 2.3: *Consider a Lévy basis Z on $\mathbb{R}^d$ with triplet $(a,0,\nu)$ and a nonnegative locally integrable random field Λ (2.5). Then the generating functional of a Cox point process X driven by Λ is*

$$G(u) = \exp\left[-\int_{\mathbb{R}^d} f(\xi)a'(\xi)U(d\xi)+\right. \tag{2.6}$$

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

where

$$f(\xi) = \int_{\mathbb{R}^d}(1-u(\sigma))g(\sigma,\xi)d\sigma. \tag{2.7}$$

Proof: For $h(\xi) = i\int_{\mathbb{R}^d}(1-u(\sigma))\,g(\sigma,\xi)d\sigma$

$$\mathcal{C}\{1\ddagger\int_{\mathbb{R}^d} hdZ\} = \log\mathbb{E}\left[\exp\left\{-\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(1-u(\sigma))\,g(\sigma,\xi)d\sigma dZ\right\}\right] = \log G(u).$$

From the Lemma 2.1 we have

$$\mathcal{C}\{1\ddagger\int_{\mathbb{R}^d} hdZ\} = \int_{\mathbb{R}^d}\mathcal{C}\{h(\xi)\ddagger Z'(\xi)\}U(d\xi)$$

$$\int_{\mathbb{R}^d}\left\{ih(\xi)a'(\xi)+\int_{\mathbb{R}}\{e^{ih(\xi)x}-1-ih(\xi)x\mathbf{1}_{[|x|\le 1]}\}\mu(dx,\xi)\right\}U(d\xi)$$

□

Corollary 2.1: Specially for $a'(\xi) = \int_{-1}^{1} r\mu(dr,\xi)$ (zero drift) it holds

$$G(u) = \exp\left[\int_{\mathbb{R}^d}\int_{\mathbb{R}}(e^{-rf(\xi)}-1)\mu(dr,\xi)U(d\xi)\right]. \tag{2.8}$$

The distribution of a point process is determined by void probabilities (Theorem 1.1). The void probabilities

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

have under the assumptions of Theorem 2.3 form (2.6) with $u = 1-\mathbf{1}_D$, i.e.

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

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

$$M(D) = \mathbb{E}X(D) = -\frac{\partial}{\partial z} G(1 - z\mathbf{1}_D)\,|_{z=0}, \quad D \in \mathcal{B}^d, \tag{2.9}$$

and the factorial second moment measure

$$\alpha^{(2)}(C) = \mathbb{E} \sum_{\xi,\eta \in X}^{\neq} \mathbf{1}_{[(\xi,\eta)\in C]}, \quad C \subset \mathcal{B}^{2d} \tag{2.10}$$

as

$$\alpha^{(2)}(D_1, D_2) = \frac{\partial}{\partial z_1}\frac{\partial}{\partial z_2} G(1 - z_1\mathbf{1}_{D_1} - z_2\mathbf{1}_{D_2})\,|_{z_1=z_2=0}, \tag{2.11}$$

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

By [5] positive Lévy bases have Lévy-Itó representation

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

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

$$\Lambda(\xi) = \int_{\mathbb{R}^d} g(\xi, \sigma) \left(\bar{a}(d\sigma) + \int_{\mathbb{R}_+} r\Phi(dr, d\sigma) \right) \tag{2.12}$$

and a connection with the class of shot-noise Cox processes (SNCP), see Definition 1.18.

The class of non-Gaussian Ornstein-Uhlenbeck processes was extended by means of superpositions in [4] 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 (2.5) 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 (2.11) we obtain for $u = 1 - z_1\mathbf{1}_A - z_2\mathbf{1}_B$

$$\alpha^{(2)}_{\Lambda}(A,\, B) = \alpha^{(2)}_{\Lambda_1}(A,\, B) + \alpha^{(2)}_{\Lambda_2}(A,\, B) + \tag{2.13}$$

$$+ \left[\frac{\partial}{\partial z_1} G_1(u) \frac{\partial}{\partial z_2} G_2(u) + \frac{\partial}{\partial z_1} G_2(u) \frac{\partial}{\partial z_2} G_1(u) \right]_{z_1=z_2=0}.$$

In the following we will mainly study a special case of the model (2.5) suggested for the purpose of spatio-temporal modelling by [5]. They define an Ornstein-Uhlenbeck (OU) type process $\Lambda(t,\sigma)$, $t \in \mathbb{R}$ (time), $\sigma \in \mathbb{R}^d$ (space) by

$$\Lambda(t,\sigma) = \int_{-\infty}^{t} e^{\gamma(s-t)} Z(B_{s-t}(\sigma) \times ds), \; \sigma \in \mathbb{R}^d, \; t \in \mathbb{R}, \tag{2.14}$$

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

$$B_s(\sigma) = \{\rho \in \mathbb{R}^d; \; \chi(\rho,\sigma) \leq -us\}$$

for a metric χ on $\mathbb{R}^d$, $u > 0$ is a parameter. The form of $B_s(\sigma)$ determines the ambit set $A_t(\sigma)$ [5] for which

$$\Lambda(t,\sigma) = \int_{A_t(\sigma)} e^{\gamma(s-t)} Z(d(s,\xi)).$$

Definition 2.5: *A spatio-temporal Cox process driven by nonnegative locally integrable Ornstein-Uhlenbeck type process is denoted OUCP.*

Corollary 2.2: On $\mathbb{R}^d \times \mathbb{R}$ consider a Cox process X with driving intensity (2.14). Then the generating functional has form (2.6) with

$$f(s,\rho) = \int_{s}^{\infty} \int_{B_{s-t}(\rho)} (1 - u(t,\sigma)) e^{\gamma(s-t)} d\sigma dt. \tag{2.15}$$

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 (2.6) with

$$f(s,\rho) = \int_{s}^{\infty} |B_{s-t}(\rho) \cap D_t| e^{\gamma(s-t)} dt.$$

Proof: (2.14) is of type (2.5) with

$$g(\xi,\eta) = g((t,\sigma),(s,\rho)) = \mathbf{1}_{[-\infty,t]}(s) \mathbf{1}_{B_{s-t}(\sigma)}(\rho) e^{\gamma(s-t)} \tag{2.16}$$

and so

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

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

Corollary 2.3: Let

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

Z_j be independent identically distributed. Under the conditions (2.8) and $\nu(dx, d\xi) = \mu(dx)d\xi$ 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 $m_1 = \int_{\mathbb{R}} x\mu(dx)$ and for $C = C_1 \times C_2$, $C_1 \subset \mathbb{R}$

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

Proof: Use (2.13), (2.8) and (2.15). □

2.2 Cox processes on a curve

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

$$Y = \{(t, y_t), t \in [0, T]\} \in \mathbb{R}^{d+1}. \tag{2.17}$$

Let Λ be a locally integrable random function with realizations

$$\Lambda : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}_+. \tag{2.18}$$

Definition 2.6: *Denote X_Y a Cox point process on $[0, T] \times \mathbb{R}^d$ with driving intensity measure*

$$\Lambda_Y([t_1, t_2] \times B) = \int_{t_1}^{t_2} \mathbf{1}_B(y_t)\Lambda(t, y_t)dt, \quad 0 \le t_1 \le t_2 \le T, \quad B \in \mathcal{B}^d. \tag{2.19}$$

The process X_Y is called a Cox process on the curve.

Events of X_Y lie on Y since outside Y we have Λ_Y zero. The measure Λ_Y does not have the density w.r.t. Lebesgue measure on $[0, T] \times \mathbb{R}^d$.

For Λ given in (2.18) we will consider also a driving intensity function $\Lambda(t, y_t),\ t \in [0, T]$, of a temporal Cox process. Finally put

$$\tilde{\Lambda}(y_t) = \frac{1}{v(t)}\Lambda(t, y_t),$$

where

$$v(t) = \sqrt{1 + \sum_{i=1}^{d} \left(\frac{dy_t^{(i)}}{dt}\right)^2}. \tag{2.20}$$

Proposition 2.1: *Let*

$$Y_t = \{(s, y_s), s \in [0, t]\},\ t \in (0, T].$$

The curvilinear integral $\int_{Y_t} \tilde{\Lambda}(y)dc$ is the driving intensity measure of the temporal Cox process on $[0, t]$.

Proof: From the definition of the curvilinear integral, we have $\forall\ 0 \le t \le T$

$$\int_{Y_t} \tilde{\Lambda}(y)dc = \int_0^t \tilde{\Lambda}(y_s)v(s)ds = \int_0^t \Lambda(s, y_s)ds$$

□

$\tilde{\Lambda}$ can be interpreted as a driving intensity function on the curve Y. Denote the intensity measure $M(.) = \mathbb{E}X_Y(.) = \mathbb{E}\Lambda_Y(.)$. According to [34] define the Fano factor [26] for X_Y as

$$F_a = 1 + \frac{var\left(\int_0^T \mathbf{1}_B(y_t)\Lambda(t, y_t)dt\right)}{\int_0^T \mathbf{1}_B(y_t)\mathbb{E}\Lambda(t, y_t)dt} \tag{2.21}$$

which is the event number variance to mean ratio, $B \in \mathcal{B}^d$, since

$$var(X_Y([0,T] \times B)) = \mathbb{E}\left[var(X_Y([0,T] \times B))|\Lambda\right] + var\mathbb{E}\left[(X_Y([0,T] \times B))|\Lambda\right]$$

$$= M([0,T] \times B) + var\left(\int_0^T \mathbf{1}_B(y_t)\Lambda(t, y_t)dt\right).$$

F_a is equal to 1 for a Poisson process and it is a measure of overdispersion.

Further Λ is of form $\Lambda = \int gdZ$, (cf. (2.5)), with deterministic function g and a Lévy basis Z with zero drift condition (2.8). Let

$$\nu(dx, d\xi) = \mu(dx)\rho(\xi)d\xi, \tag{2.22}$$

where μ is a finite measure. This corresponds to the compound Poisson process Z' where normalized μ is the jump size distribution and ρ the spatio-temporal intensity (density of the measure U). Denote m_j the $j-$th moment of μ i.e.

$$m_j = \int x^j \mu(dx),\ j = 1, 2, \ldots.$$

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 2.4: *Denote*

$$f_C(\xi) = \int_{C_1} \mathbf{1}_{C_2}(y_t)g((t, y_t), \xi)dt, \quad \xi \in \mathbb{R}^{d+1} \tag{2.23}$$

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

$$M(C) = m_1 \int f_C(\xi)\rho(\xi)d\xi, \tag{2.24}$$

and the factorial second moment measure of X_Y

$$\alpha^{(2)}(C,\ D) = M(C)M(D) + m_2 \int f_C(\xi)f_D(\xi)\rho(\xi)d\xi. \tag{2.25}$$

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

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

Using (2.5) and Fubini Theorem we obtain

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

and from Lemma 2.1 we have

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

By differentiating the result for intensity follows. Analogously we obtain

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

and by differentiating the factorial second moment measure. □

Since the measure μ is finite we get from (2.5) a representation

$$\Lambda(\xi) = \sum_j w_j g(\xi, \eta_j) \tag{2.26}$$

where η_j are events of a Poisson process with intensity function ρ and w_j are jump sizes. In fact formula (2.24) follows then from the Campbell Theorem

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

We can extend the Definition 2.5 of OUCP to a Cox process on a curve Y by using an Ornstein-Uhlenbeck type process Λ in (2.18). Specially we have

Corollary 2.4: Consider the random function Λ from (2.14) and an OUCP X_Y. For the intensity measure $M(.)$ of a product set $C = C_1 \times C_2$ and for the factorial second moment measure formulas (2.24), (2.25) of Theorem 2.4 hold, respectively, with

$$f_C(s, \sigma) = \int_{C_1\cap[s,\infty)} e^{\gamma(s-t)}\mathbf{1}_{C_2\cap B_{s-t}(\sigma)}(y_t)dt. \tag{2.27}$$

Proof: Put (2.16) into (2.23). □

We will consider two models for the intensity function ρ in the special case of dimension $d = 3$. For the model of a piecewise constant ρ in $\mathbb{R}^3$

$$\rho(\xi) = \sum_{ijk} \rho_{ijk}\mathbf{1}_{A_{ijk}}(\xi), \tag{2.28}$$

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

Corollary 2.5: Under the assumptions of Corollary 2.4 and with the model (2.28) it holds

$$M(C) = m_1 \int_{C_1} \mathbf{1}_{C_2}(y_t) \sum_{ijk} \rho_{ijk} \int_{(-\infty,t]\cap A_i} e^{\gamma(s-t)} |B_{s-t}(y_t) \cap A_{jk}| ds dt \tag{2.29}$$

and

$$\alpha^{(2)}(C, D) = m_2 \int_{C_1} \mathbf{1}_{C_2}(y_t) \int_{D_1} \mathbf{1}_{D_2}(y_u) \sum_{ijk} \rho_{ijk} \times \tag{2.30}$$

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

Proof: Formula (2.29) follows putting (2.28) and (2.27) in (2.24) and similarly using Fubini Theorem we obtain (2.30). □

Another way how to model intensity function ρ is using the Zernicke polynomials. Consider that intensity function ρ is inhomogeneous in space $\mathbb{R}^2$ and homogeneous in time and ρ is given by

$$\rho(t, x, y) = \exp\left\{ \sum_{j=0}^{n} \sum_{i=-j}^{j} \psi^{j,i} R_j^i(x, y)) \right\}, \tag{2.31}$$

where R_j^i is the $i-$th component of $j-$th order Zernicke polynomials and $\psi^{j,i}$ are parameters of the model. When (r, φ) are the polar coordinates of (x, y) in the plane, we have

$$\begin{aligned}
R_0^0(x,y) &= 1\\
R_1^{-1}(x,y) &= r\sin(\varphi)\\
R_1^0(x,y) &= 0\\
R_1^1(x,y) &= r\cos(\varphi)\\
R_2^{-2}(x,y) &= 2r^2\sin(\varphi)\\
R_2^{-1}(x,y) &= 0\\
R_2^0(x,y) &= 2r^2-1\\
R_2^1(x,y) &= 0\\
R_2^2(x,y) &= 2r^2\cos(\varphi).
\end{aligned}$$

Corollary 2.6: In this case

$$M(C) = m_1 \int_{C_1} \mathbf{1}_{C_2}(y_t) \int_{(-\infty,t)} e^{\gamma(s-t)} \int_{B_{s-t}(y_t)} \exp\left\{ \sum_{j=0}^{n} \sum_{i=-j}^{j} \psi^{j,i} R_j^i(x, y)) \right\} dx ds dt \tag{2.32}$$

and

$$\alpha^{(2)}(C, D) = M(C)M(D) + m_2 \int_{C_1} \mathbf{1}_{C_2}(y_t) \int_{D_1} \mathbf{1}_{D_2}(y_u) \tag{2.33}$$

$$\int_{(-\infty,\min(u,t))} e^{\gamma(2s-t-u)} \int_{B_{s-t}(y_t)\cap B_{s-u}(y_u)} \exp\left\{ \sum_{j=0}^{n} \sum_{i=-j}^{j} \psi^{j,i} R_j^i(x, y)) \right\} dx ds du dt.$$

2.3 Bayesian Markov Chain Monte Carlo

One of important questions in the analysis of Cox point processes is the inference on the driving intensity and its characteristics. A rigorous approach to this problem is the filtering, see [28], [35]. Filtering and transition together yield prediction, cf. [11] 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 (2.14).

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 $\mathbb{E}[\Lambda|X]$.

Since $\mathbb{E}[\Lambda|X]$ is not explicitly available the Bayes formula for probability densities enters

$$f(\lambda|x) \propto f(x|\lambda)f(\lambda). \tag{2.34}$$

From the definition of the Cox process $f(x|\lambda)$ is a known density of type (1.16) of an inhomogeneous Poisson process with intensity function λ, $f(\lambda)$ is a prior density. The aim is to simulate a sample 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, see Section 1.5.

The benefit of this approach for applications would become mostly evident in case of spatio-temporal data rather than for data on a curve. When the curve is known it yields further information which enables other approaches that may probably be more efficient, especially for long time measurements. See Section 4.1 for an approach based on sequential methods [24]. Here we will develop filtering for the Cox process on a curve using the model from Section 2.2.

Let $W = [0, T] \times \mathcal{A}$, $\mathcal{A} \in \mathcal{B}^2$ be a bounded window, $Y \subset W$ a known curve. In W we observe the data $x = \{\tau_j\}$, a realization of a Cox process X_Y driven by Λ_Y (see Definition 2.6, where Λ is unknown random function). Each τ_j reflects time and location of an event on Y. We have now in Bayesian setting

$$f(\psi, b|x) \propto f(x|\psi, b)f(\psi|b)f(b), \tag{2.35}$$

where $\psi = \{t_j, z_j, w_j\}$ represents the auxiliary compound Poisson process Z', t_j are times, z_j are locations, w_j are jumps of points of Z'. Further b is a vector of unknown parameters of models for the intensity function ρ in (2.28) or (2.31), the jump size distribution and for function g in Λ (2.26).

Since the Cox process is conditionally Poisson we used the likelihood in the form

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

which corresponds to a density (w.r.t. a unit Poisson process on the time axis) of the temporal Poisson process on Y given $\Lambda = \lambda$. The second term on the right side of (2.35) can be approximate by

$$f(\psi|b) \approx e^{|W_0|} \exp\left(-\int_{W_0} \rho(v)dv\right) \prod_{(t_j, z_j, w_j) \in \psi} \rho(t_j, z_j)h(w_j),$$

where h is the probability density of jump size. Theoretically unbounded domain of ρ is substituted by some W_0 bounded, $W \subset W_0$, containing also events at negative times t_j. See Fig. 2.1 for a scheme of this model. 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 (2.35). For ψ the birth-death algorithm (1.27)- (1.30) is available. Parameters, which are real numbers, are updated by a Gaussian random walk (1.5). Geometric ergodicity of the chain follows under mild conditions, cf. [35].

Consider a compound Poisson process Z' and the model (2.14) (with $d = 2,\ \gamma = 1$). From (2.26) and (2.16) we have a representation

$$\Lambda(t, v) = \sum_{t_j \le t} w_j e^{t_j - t}\, \mathbf{1}_{B_{t_j - t}(v)}(z_j). \tag{2.36}$$

Let the jumps have an exponential distribution with density $h(w) = \frac{1}{\alpha} \exp(-\frac{w}{\alpha}),\ w \ge 0$, where $\alpha > 0$ is a parameter. Further let

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

Alternatively we can use $B_s(x_1, x_2) = b((x_1, x_2), us)$ a ball. 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 (2.28) the vector of parameters is

$$b = (\alpha, u, \{\rho_{ijk},\ i, j, k = 1, \ldots, n\}).$$

The prior distributions are chosen one-dimensional exponential with fixed hyperparameters $l_\alpha, l_u, l_{ijk} >> 0$ for random α, u, ρ_{ijk}, respectively.

Under these assumptions, denoting N_ψ the number of events of ψ in W_0, we can rewrite (2.35) as

$$f(\psi, b|x) \propto \exp\left(-\int_0^T \sum_{t_j \le t} w_j e^{t_j - t} \mathbf{1}_{B_{t_j - t}(y_t)}(z_j) dt\right) \times \tag{2.37}$$

$$\times \prod_{\tau_i \in x} \lambda(\tau_i) \prod_{ijk} \exp\left(-\rho_{ijk} |A_{ijk} \cap W_0|\right) \alpha^{-N_\psi} \exp\left(-\sum_i \frac{w_i}{\alpha}\right) \times$$

$$\times \left(\prod_{(t,z) \in \psi} \sum_{jlk} \rho_{jlk} \mathbf{1}_{A_{jlk}}(t, z)\right) l_\alpha^{-1} e^{-\frac{\alpha}{l_\alpha}} l_u^{-1} e^{-\frac{u}{l_u}} \prod_{jk} l_{ijk}^{-1} e^{-\frac{\rho_{ijk}}{l_{ijk}}}.$$

The full-conditional distributions for the Gibbs sampler are then

$$f(\psi|b, x) \propto \exp\left(-\int_0^T \sum_{t_j \le t} w_j e^{t_j - t} \mathbf{1}_{B_{t_j - t}(y_t)}(z_j) dt\right) \prod_{\tau_i \in x} \lambda(\tau_i) \times$$

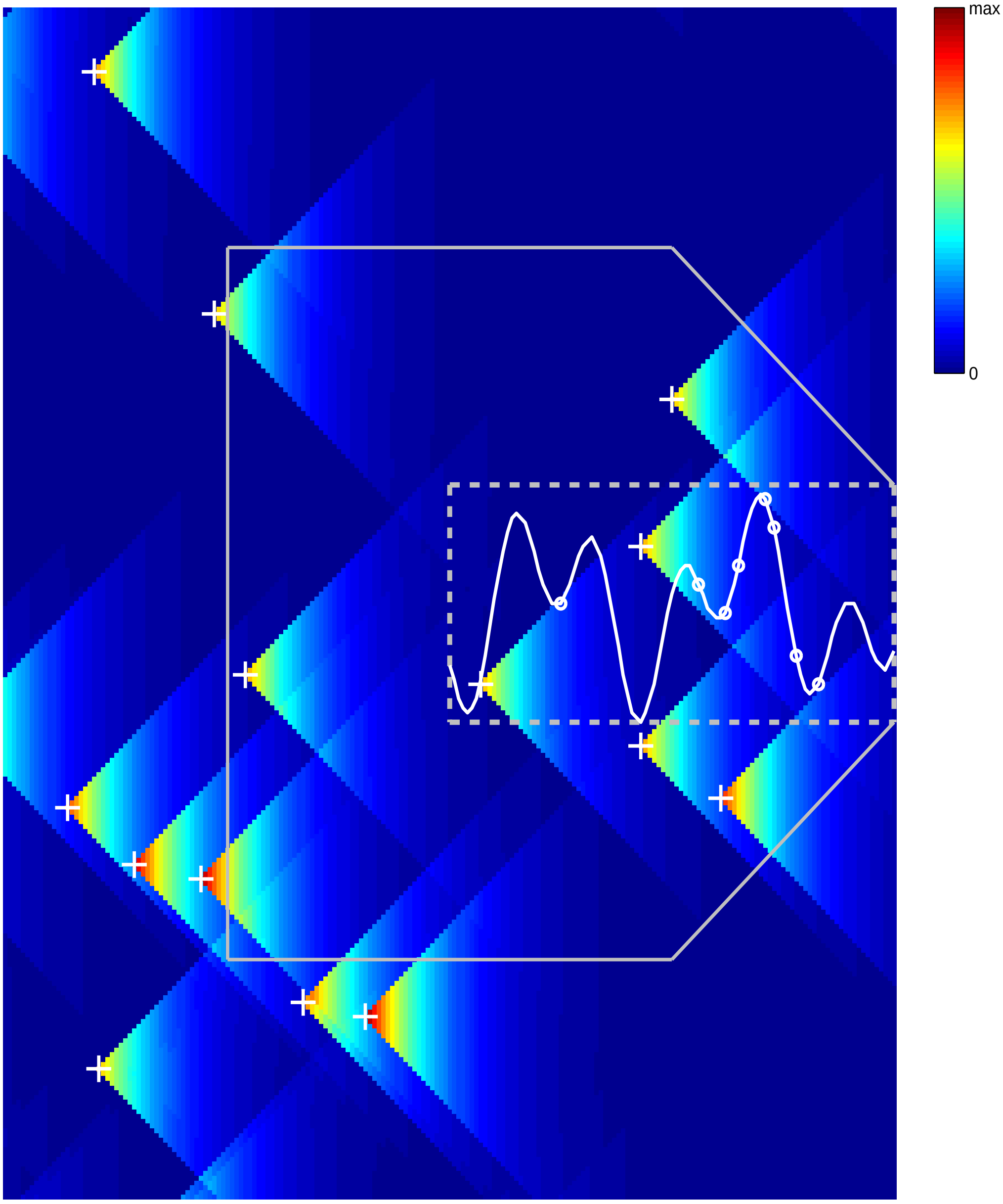

Figure 2.1: A simplified representation of the model (2.36). Here the horizontal axis presents time and the vertical axis space. In the window W (the rectangle delimited by the dashed white line) the track with spikes (circles) is drawn. The crosses denote events of the auxiliary point process ψ, which lie theoretically within the whole space and time. For numerical evaluation they are limited to W_0 (region delimited by the white full line). The numerical contribution of each event of ψ to the intensity (2.36) is expressed by spectral colours.

$$\times \alpha^{-N_\psi} \exp\left(-\sum_i \frac{w_i}{\alpha}\right)\left(\prod_{(t,z)\in\psi}\sum_{jlk}\rho_{jlk}\mathbf{1}_{A_{jlk}}(t,z)\right),$$

$$f(\rho_{ijk}|\psi,u,\alpha,x) \propto \exp\left(-\rho_{ijk}|A_{ijk}\cap W_0|\right)\left(\prod_{(t,z)\in\psi}\sum_{jlm}\rho_{jlm}\mathbf{1}_{A_{jlm}}(t,z)\right)\times$$

$$\times e^{-\frac{\rho_{ijk}}{l_{ijk}}}, \quad i,j,k = 1,\dots,n,$$

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

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

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.

2.4 Model Selection

Using ergodicity properties of the MCMC we can try to estimate statistical characteristics of Λ. Denote $\Lambda^{(l)}(t,v)$ from (2.36) evaluated in the $l-$th iteration 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 given all data X and the curve Y is estimated by the average value

$$\hat{\Lambda}(t,v) = \frac{1}{k}\sum_{l=K+1}^{J}\Lambda^{(l)}(t,v), \tag{2.38}$$

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 $V(x,Y)$ [36] computed from the data and compare it with $V(X,Y)$ where X is a Cox process with the estimated driving intensity.

We can 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 $C_j \subset W$ hitting Y, $j = 1, \ldots, k$. A measure of discrepancy of the model is e.g.

$$C = \sum_{j=1}^{k} (M(C_j) - N(C_j))^2, \tag{2.39}$$

We approximate the mean value $M(.) = \mathbb{E}\Lambda_Y(.)$ as

$$\hat{\Lambda}_Y([t,s] \times B) \approx \triangle \sum_{p=1}^{m} \mathbf{1}_B(y_{t_p}) \hat{\Lambda}(t_p, y_{t_p}), \tag{2.40}$$

where $t_p = t + p\triangle$, $\triangle = (s-t)/m$, where $\hat{\Lambda}(t_p, y_{t_p})$ is evaluated from (2.38). Thus we obtain an estimate of $M([t,s] \times B)$ based on the auxiliary process iterations since Λ comes from (2.36).

For the second-order analysis we can evaluate the factorial second moment measure $\alpha^{(2)}$ for pairs of subsets of the window and compare it with the estimator

$$\hat{\alpha}^{(2)}(C,D) = \sum_{\xi,\eta \in X}^{\neq} \mathbf{1}_{[\xi \in C, \eta \in D]}, \quad C,\ D \subset \mathbb{R}^2 \tag{2.41}$$

unbiased from (2.10). The statistics

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

is another measure of discrepancy.

Fix M large integer, let $\triangle = \frac{T}{M}$. Denote

$$\mathcal{E}(B) = \frac{\triangle}{M} \sum_{q=0}^{M-1} \hat{\Lambda}(q\triangle,\ y_{q\triangle}) \mathbf{1}_{[y_{q\triangle} \in B]},\ B \subset \mathcal{A}$$

and

$$\mathcal{V}(B) = \frac{1}{k} \sum_{l=K+1}^{J} \left(\frac{\triangle}{M} \sum_{q=0}^{M-1} \Lambda(q\triangle,\ y_{q\triangle}) \mathbf{1}_{[y_{q\triangle} \in B]} \right)^2 - \mathcal{E}(B)^2.$$

Then we can estimate the Fano factor as

$$\hat{F}(B) = 1 + \frac{\mathcal{V}(B)}{\mathcal{E}(B)}. \tag{2.42}$$

2.5 Residual analysis

Another way to quantify the fit of the model and data is the residual analysis. For temporal and spatio-temporal point processes it is well developed, see [35], 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 [2]. The authors of that paper 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

$$I(B) = X(B) - \mathbb{E}[\Lambda_m(B) \mid X], \quad B \in \mathcal{B}, \tag{2.43}$$

it holds

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

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

$$R_{\hat{\theta}}(B) = X(B) - \mathbb{E}_{\hat{\theta}}[\Lambda_m(B) \mid X] \tag{2.44}$$

oscillates around zero. A possibility to perform a statistical test depends on the way how exactly the conditioning in (2.43),(2.44) 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

$$\lambda_t^* = \lim_{\triangle t \downarrow 0} \frac{1}{\triangle t} \mathbb{E}[N_{t+\triangle t} - N_t \mid N_s,\ s < t] =$$

$$= \lim_{\triangle t \downarrow 0} \frac{1}{\triangle t} \mathbb{E}[\mathbb{E}[N_{t+\triangle t} - N_t \mid N_s,\ s < t;\ \Lambda_p,\ t \leq p < t + \triangle t]] =$$

$$= \lim_{\triangle t \downarrow 0} \frac{1}{\triangle t} \mathbb{E}\left[\int_t^{t+\triangle t} \Lambda(s) ds \mid N_s,\ s < t\right] = \mathbb{E}[\Lambda(t) \mid N_s,\ s < t]$$

the innovation process $N_t - \int_0^t \lambda_s^* ds$ is a martingale [28]. In the spatio-temporal case denote $N_s(C) = card\{x \in X;\, x \in [0,s] \times C\},\ C \in \mathcal{B}^d$. Analogously the conditional intensity λ^* (see Subsection 1.2.1 for a definition)

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

of a Cox process corresponds to

$$\mathbb{E}[\Lambda(t,\xi) \mid N_s(C),\ s < t,\ C \in \mathcal{B}^d] \tag{2.45}$$

and

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

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

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

where H is a predictable process, are investigated.

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

$$\lambda_s^* = \mathbb{E}[\Lambda(s, y_s)|N_u, u < s], \tag{2.46}$$

$N_t - \int_0^t \lambda_s^* ds$ 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

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

Theorem 2.5: *For a nonnegative predictable process $\{H(t),\ t \in [0,T]\}$ the scaled innovation has variance*

$$varV_C = \mathbb{E}\left[\int_0^T \mathbf{1}_C(y_t)H^2(t)\lambda_t^* dt\right].$$

Proof: 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

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

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

$$varV_C = \mathbb{E}\left(\left[\int G(t)N(dt)\right]^2\right) +$$

$$+\mathbb{E}\left(\left[\int G(t)\lambda_t^* dt\right]^2\right) - 2\mathbb{E}\left[\int G(t)N(dt)\int G(t)\lambda_t^* dt\right].$$

Using Fubini and Theorem 1 from [25] we have

$$varV_C = \mathbb{E}\int G^2(t)\lambda_t^* dt +$$

$$-2\mathbb{E}\left[\int G(s)\lambda_s^* ds \int G(t)[N(dt) - \lambda_t^* dt]\right]$$

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

$$V_p = \int_D \mathbf{1}_C(y_t)[(\lambda_t^*)^{-\frac{1}{2}} N(dt) - (\lambda_t^*)^{\frac{1}{2}} dt] \tag{2.47}$$

with

$$varV_p = |\{t \in D;\, y_t \in C\}|$$

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 (2.47), time t and a measurable set $C \subset \mathcal{A}$ is then

$$R_{\hat{\theta}}(t, C) = \sum_{\substack{(s_l,\eta_l) \\ s_l \le t,\, \eta_l \in C}} \hat{\Lambda}(s_l)^{-\frac{1}{2}} - \int_0^t \mathbf{1}_C(y_s)[\hat{\Lambda}(s)]^{\frac{1}{2}} ds, \tag{2.48}$$

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 (2.48) which desires either more chains (and this is 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 $0 < t_1 < \cdots < t_n = T$ with bounds $2\sigma_i$ at t_i,

$$2\sigma_i = 2[|\{t \le t_i;\, y_t \in C\}|]^{\frac{1}{2}}. \tag{2.49}$$

Chapter 3

Numerical results

A spatio-temporal point process on a curve is a mathematical model for an experiment in neurophysiology. Consider a bounded arena $\mathcal{A} \in \mathbb{R}^2$ and an experimental animal (typically a rat) moving in time in the arena. Action potentials (spikes) of a neuronal cell in the brain are discrete events during the movement, so that the time and the location of each spike can be monitored. The curve represents a track of the rat.

In Section 3.1 we present simulations corresponding to the stochastic model, in Section 3.2 real data from an experiment are evaluated.

3.1 Simulation

It is desirable to simulate models from Chapter 2 in order to test various methods (filtering, model selection, residual analysis) before evaluating real data.

A simple simulation on a grid demonstrates some properties of the model in Subsection 3.1.1. In Subsection 3.1.2 a numerical study of the model from Chapter 2 is presented with evaluation of simulated data based on MCMC.

3.1.1 A model on a grid

Consider that the random function Λ generating driving intensities is homogeneous in time and inhomogeneous in space, i.e.

$$\mathbb{E}(\Lambda(x,t)) = \mu_x$$

and

$$var(\Lambda(x,t)) = \sigma_x^2,\ x \in \mathcal{A},$$

do not depend on t. Assume that $\mathcal{A} \subset \mathbb{R}^2$ bounded is divided in l boxes where μ_x is piecewise constant and denote μ_i intensity mean in $i-$th box. Consider experimental data in a form $(t_{ij}, n_{ij}),\ i = 1, \ldots, l;\ j = 1, \ldots, k_i$, ([27]) where t_{ij} is the duration of $j-$th stay of the animal in $i-$th box and n_{ij} is the number of events during this stay. A natural estimator of the expected intensity μ_i is then

$$\hat{\mu}_i = \frac{\sum_j n_{ij}}{\sum_j t_{ij}}. \tag{3.1}$$

Consider Λ in a bounded window $W = \mathcal{A} \times [0, T]$

$$\Lambda(t,u) = \int_{-\infty}^{t} e^{s-t}\, Z(B_{s-t}(u) \times ds). \tag{3.2}$$

To simulate the model we need to consider a larger region W_0 (see Fig. 3.1). The shape of this region is derived from the form of sets B_s in (3.2). where Z is the Poisson basis. We put

$$B_s(x_1, x_2) = [x_1 + s,\ x_1 - s] \times [x_2 + s,\ x_2 - s],\ s \leq 0. \tag{3.3}$$

Lemma 3.1: *For B_s given by (3.3) and the homogeneous Poisson basis Z with intensity constant a we have*

$$\mathbb{E}\Lambda(t,u) = \int_{-\infty}^{t} e^{s-t}(2(s-t))^2 a ds = \int_{-n}^{0} e^{s}(2s)^2 a ds + \int_{-\infty}^{-n} e^{s}(2s)^2 a ds$$

and

$$\int_{-\infty}^{-n} e^{s}(2s)^2 a ds = e^{-n} a(4n^2 + 8n + 8). \tag{3.4}$$

Because of the Lemma 3.1 we can substitute $-\infty$ in the integral in (3.2) by a finite time, which leads to a good approximation.

Let us start with a discretization step $\triangle > 0$ in space and $\triangle' > 0$ in time. Here W is a parallelepiped divided into cubes, each (i,j,k)-th cube is represented by its central point. We get an approximation of (3.2) in the form

$$\widetilde{\Lambda}(k,i,j) = \sum_{r=k}^{k+n} e^{(k-r)\triangle'} \sum_{l=i-r+k}^{i+r-k} \sum_{q=j-r+k}^{j+r-k} \widetilde{Z}(r,l,q), \tag{3.5}$$

where $\widetilde{Z}(r,l,q)$ are independent Poisson distributed random variables with mean equal to a constant multiple of the volume of a cube. To simulate an inhomogeneous Λ we may vary the parameter of the Poisson distribution in cubes.

A simple model of the curve Y is considered on a discrete square arena $\mathcal{A} = \{1, \ldots, m\}^2$ putting Y_t a symmetric random walk on $\mathcal{A}$ with reflecting walls. The animal starts from the position $(\frac{m+1}{2}, \frac{m+1}{2})$, m odd, with constant speed in each time interval moves a step $\triangle$ in a random direction on the square grid with equal probability $\frac{1}{4}$. Y is here random and independent of Λ.

Finally we simulate the spatio-temporal Cox point process X. Given Λ, the number of points during time interval $[q, q + \triangle']$ is Poisson distributed with mean $\triangle'\widetilde{\Lambda}(q, y_q)$.

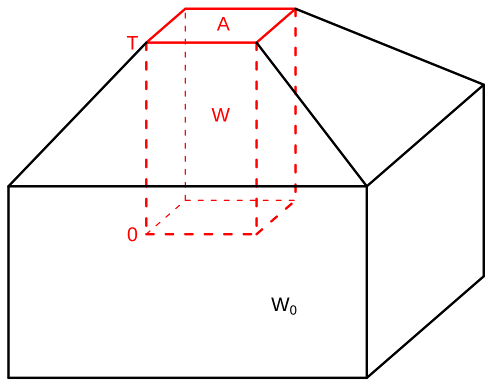

Figure 3.1: The enlarged window W_0, points outside this region have no or very small contribution to the driving intensity measure.

Simulation:

1. Simulate a symmetric random walk $Y_i,\ i = 1\ldots, M$,
2. for $r = -n, \ldots, M$
   for $q, l = -i_r, \ldots, i_r$ sample a random variable $\widetilde{Z}(r, l, q)$ with Poisson distribution with parameter $\rho_{lqr}\triangle^2\triangle'$,
3. for $i = 1, \ldots, M$ compute $\widetilde{\Lambda}(i, y_i)$ from equation (3.5),
4. for $i = 1, \ldots, M$ simulate random variable p_i Poisson distributed with parameter $\widetilde{\Lambda}(i, y_i)\triangle'$ and for $j = 1, \ldots p_i$ simulate random variable $x_{\sum_{k=1}^{i-1} p_k + j}$ with uniform distribution on $[(i-1)\triangle', i\triangle')$.

For the evaluation of Fano factor in (2.21) we get an approximation

$$\int_0^T \mathbf{1}_B(y_t)\Lambda(t, y_t)dt \approx \triangle' \sum_{i=1}^{M} \mathbf{1}_B(y_i)\widetilde{\Lambda}(i, y_i), \tag{3.6}$$

where y_i is the location at time i. Empirical mean and variance is substituted in (2.21), obtained from N realisations of Λ, respectively.

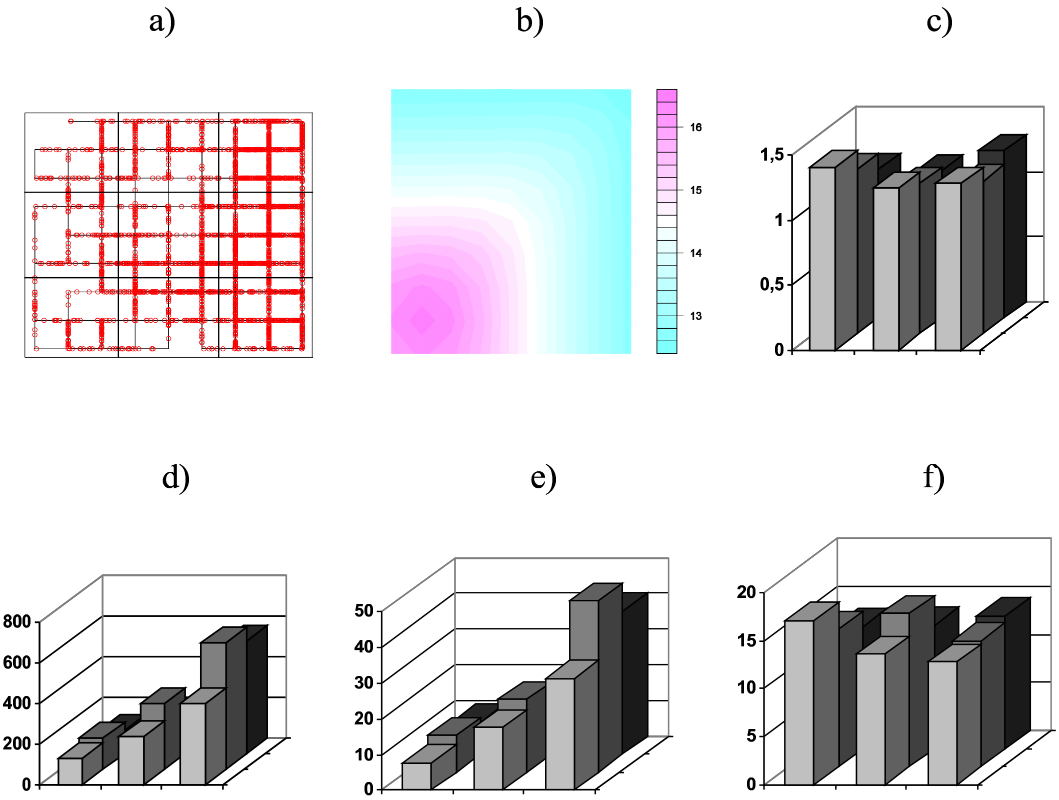

Figure 3.2: 3×3 grid of boxes, a) simulated random walk with spikes, b) a simulated inhomogeneous intensity, c) Fano factors, d) histogram of the numbers of spikes, e) histogram of time units spent in a box, f) estimated expected intensity.

We present numerical results of the simulation for the grid size $m = 9$. The grid $\mathcal{A}$ is subdivided onto boxes $A_{ij},\ i, j = 1, \ldots, 3$, see the thick lines in Fig. 3.2a. The parameters are $M = 1000$ (number of time steps), $\triangle = 0.1,\ \triangle' = 0.1,\ n = 65$ in (3.5). Then the approximation of the term (3.4) is equal to $0.38a$ which is relatively small compared to $\mathbb{E}\Lambda = 8a$ and $N = 20$ (number of realizations of Λ).

The data and results (estimation of Fano factor and intensity) are in Fig. 3.2. The histograms in d) and e) are naturally of a similar shape resulting from a). The Fano factors estimated in c) are clearly above 1. Using (3.1) the estimated expected intensity in f) corresponds to the theoretical intensity in b). Unfortunately for further characteristics estimation, e.g. for the variance of the driving intensity, there are no simple estimators available. Therefore we proceed using the filtering techniques.

3.1.2 Cox process on a curve

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

$$Y = \{t, t\cos(\beta t), t\sin(\beta t)\},\ 0 < t < T, \tag{3.7}$$

$\beta > 0$ a parameter. Here the parametrization leads to the function $v(t)$ in (2.20) of form

$$v(t) = \sqrt{2 + \beta^2 t^2}.$$

We simulated realization of a Cox point process on the curve Y (see Section 2.2) in the following way. The intensity of the auxiliary point process ψ generating the Poisson basis Z was homogeneous with parameter $\rho = 0.125$, further parameters were put $\beta = 0.375$, $\gamma = 1$ and $u = 1$. First we simulated auxiliary Poisson point process in space and time with intensity ρ and then we computed the function Λ

$$\Lambda(t, v) = \sum_{\substack{(s_j, z_j) \in \psi \\ s_j \leq t}} e^{s_j - t} \, \mathbf{1}_{B_{s_j - t}(v)}(z_j)$$

for (t_i, y_{t_i}) $i = 1, \ldots, 1000$, $t_i = 0.08i$. Finally we simulated the process on a curve analogously to algorithm from previous subsection. In Fig. 3.3 there are points of a realization projected onto the plane together with Y.

We tested the method from Section 2.3 based on Bayesian MCMC. The number of iterations was equal to 4×10^3 and we obtained estimators of $\hat{\Lambda}$ in (2.38), see Fig. 3.3.

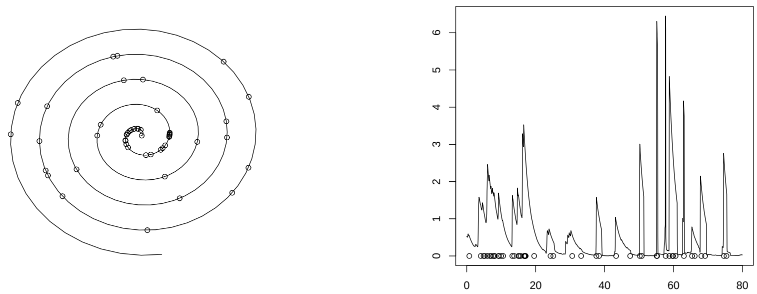

Figure 3.3: Left a) Simulated realization with 50 events (circles) of a point process on the curve (3.7) projected on the plane. Right b) The evolution of the input events (circles) in time (horizontal axis) and the filtered graph of $\hat{\Lambda}(t, y_t)$.

The posterior predictive distributions in space and time were evaluated using formula (2.40). In Fig. 3.4 the empirical counts $N(C)$ and theoretical values $M(C)$ are compared in planar subregions and finally in time in Fig. 3.5a left. We observe a reasonably good fit in both space and time for given data.

3.2 Analyses of neurophysiological data

In this section experimental data [33] of times of occurences of action potentials of a hippocampal neuron together with the track of a rat are investigated.

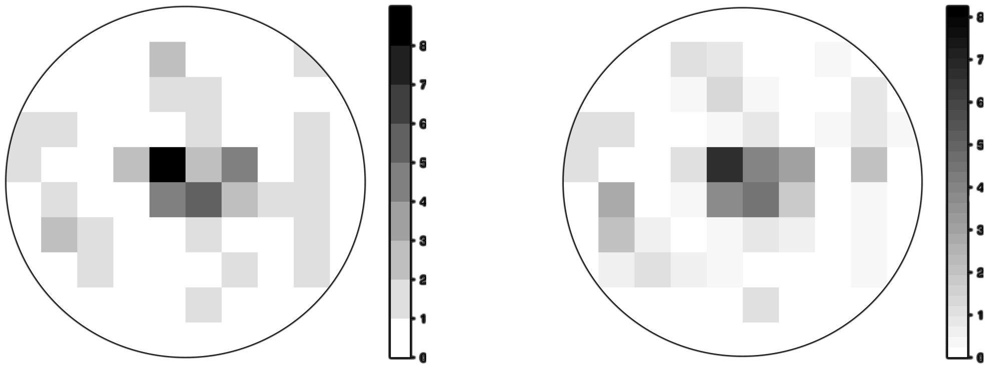

Figure 3.4: Left a) - the counts $N(D)$ (increasing with grey level), Right b) - evaluation of $M(D)$ from (2.40), here $D = [0, T] \times A_{ij}$, $\{A_{ij}\}$ is 10×10 planar grid restricted by circular $\mathcal{A}$.

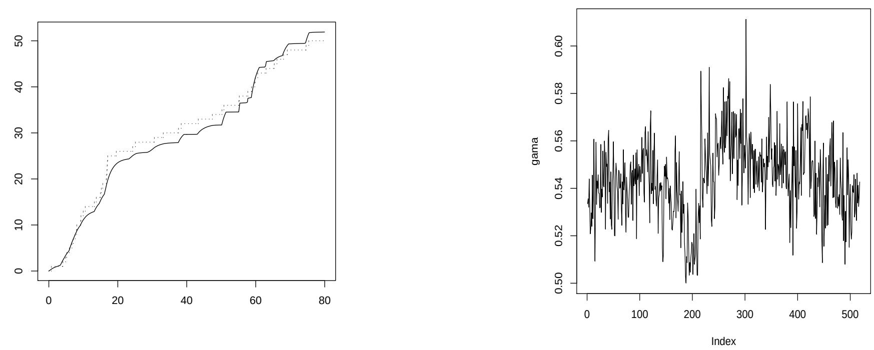

Figure 3.5: Left a) - $N(C)$ (grey dotted curve) and $M(C)$ (black curve) from (2.40) for $C = \mathcal{A} \times [0, t]$ with increasing time t (horizontal axis). Right b) (parameter estimation) 5×10^5 iterations (with step 1000) of parameter γ .

The shape of action potentials is considered to be irrelevant. Therefore the pulses may be seen as a realization of a spatio-temporal point process. The spikes were recorded with 0.1μsec precision from a rat searching for food and at the same time avoiding a northern part of a 75 cm wide circular arena $\mathcal{A}$. Each $\frac{1}{60}s$ the location of the rat was monitored. We chose first the recording of length 35 sec., there were 51 spikes observed, the average firing rate at this segment was 1.46 Hz. In Fig. 3.6a there is a planar plot of the measurement in space. The neuron fires mostly when the rat visits the east part of the arena. The temporal behaviour of the recorded neuron is such that short periods of high activity are separated by longer periods of small activity (Fig. 3.6b), there is an apparent clustering and suggests a Cox process model.

For this data we used the measurement of length 37 sec. with 101 spikes, see Fig. 3.6. The model of ρ based on Zernicke polynomials (see (2.31)) was used. The graph of filtered function $\hat{\Lambda}(t, y_t)$ is presented in Fig. 3.7a. The posterior predictive distribution is evaluated in Fig. 3.7c and compared with data counts in Fig. 3.7b.

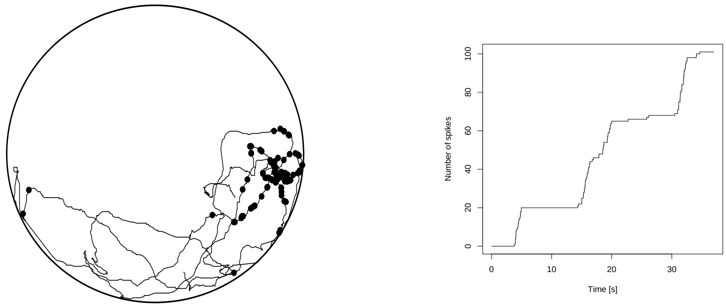

Figure 3.6: Positional firing of the hippocampal neuron. Left a) - rat's track in the arena is displayed by the line. Places at which the neuron fired are indicated by dots. Right b) - the temporal evolution of spikes with graphical presentation as a point process realization. At each spike time the graph is increased by 1 so that at each time it corresponds to the total number of spikes.([33])

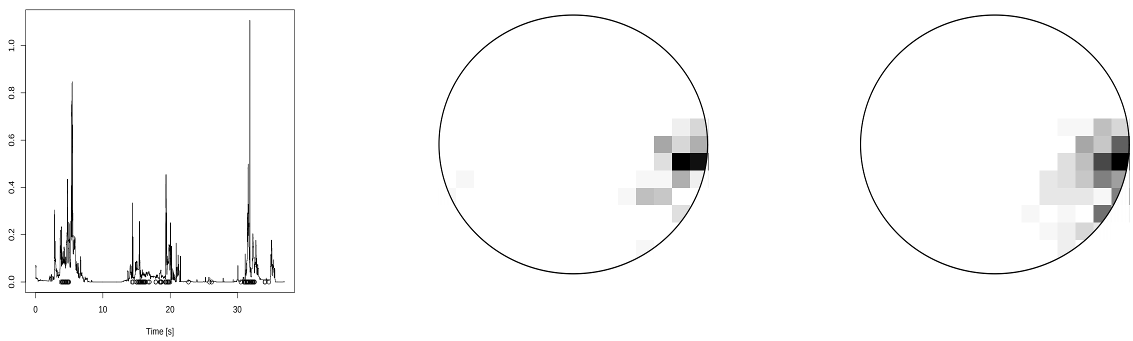

Figure 3.7: Left a) - Circles on the horizontal (temporal) axis denote spikes from Fig. 3.6, the results of filtering are presented by a graph of $\hat{\Lambda}(t, y_t)$, cf. (2.38). The intensity function was based on Zernicke polynomials, Middle b) - the counts $N(D)$, (increasing with grey level), Right c) - evaluation of $M(D)$ from (2.40), here $D = [0, T] \times A_{ij}$, $\{A_{ij}\}$ is 15×15 planar grid restricted by circular $\mathcal{A}$.

Summarizing our computational efforts a variability of results shall be admitted, especially for the real data. Therefore we study also other approaches in the next chapter.

Chapter 4

Discussion: other approaches

In this chapter we will discuss other approaches how to deal with the neurophysiological data. The first approach is known from literature (in Section 4.1) and it is based on recursive equations ([24]). In the second section we apply recently introduced theory of random marked sets ([3]) and show its connection with weighted random measures from Section 1.3. A test for the so-called random-field model is developed.

4.1 Recursive filtering

Suppose a parametric shape of the conditional intensity λ^* of a temporal point process with parameter vector $\psi \in \mathbb{R}^d$. An approach which enables to evaluate large data set of events $\{\tau_i,\ i = 1, \dots, n\}$ is such that the parameters of the model vary in time.

Consider $\triangle > 0$ small such that there is at most one event in each interval of length $\triangle$. Let $\triangle N_k$ be an indicator of an event (spike) in the interval $((k-1)\triangle, k\triangle],\ k \geq 1$. Let $N_{1:k} = [\triangle N_1, \dots, \triangle N_k]$ and $\psi_{1:k} = [\psi_1, \dots, \psi_k]$ be the values of ψ in each subinterval of length $\triangle$. The conditional intensity $\lambda^*(k\triangle|\psi_k, N_{1:k-1})$ is denoted λ_k^*. The state equation is

$$\psi_k = F\psi_{k-1} + \eta_k \tag{4.1}$$

with fixed system evolution matrix F and zero mean Gaussian noise η_k with covariance matrix Q_k. In the Bayesian approach parameter ψ is random.

The recursive system for computing the posterior density $p(\psi_k|N_{1:k})$ is

$$p(\psi_k|N_{1:k}) = \frac{p(\psi_k|N_{1:k-1})p(\triangle N_k|N_{1:k-1}, \psi_k)}{p(\triangle N_k|N_{1:k-1})} \tag{4.2}$$

$$p(\psi_k|N_{1:k-1}) = \int p(\psi_k|\psi_{k-1})p(\psi_{k-1}|N_{1:k-1})d\psi_{k-1}. \tag{4.3}$$

Here the transition probability density $p(\psi_k|\psi_{k-1})$ is defined by the state equation (4.1) and

$$p(\triangle N_k|N_{1:k-1}, \psi_k) = \lambda_k^{\triangle N_k} \exp(-\lambda_k \triangle)$$

is an approximation to the point process interval likelihood, valid for small $\triangle$. Eden et al. in [24] developed stochastic state point process filter, an algoritm using Gaussian approximation to both equations (4.2) and (4.3). Denoting $\psi_{k|k}, W_{k|k}$ the mean vector and covariance matrix of the posterior distribution (4.2), $\psi_{k|k-1}, W_{k|k-1}$ the mean vector and covariance matrix of the posterior distribution (4.3), the algorithm is given by the following equations (F' is the transpose of F):

$$\psi_{k|k-1} = F\psi_{k-1|k-1}, \; W_{k|k-1} = FW_{k-1|k-1}F' + Q_k$$

$$W_{k|k}^{-1} = W_{k|k-1}^{-1} + \left[\left(\frac{\partial \log \lambda_k^*}{\partial \psi_k} \right)' \lambda_k^* \triangle \frac{\partial \log \lambda_k^*}{\partial \psi_k} - (\triangle N_k - \lambda_k^* \triangle) \frac{\partial^2 \log \lambda_k^*}{\partial \psi_k \partial \psi_k'} \right]_{|\psi_{k|k-1}}$$

$$\psi_{k|k} = \psi_{k|k-1} + W_{k|k} \left[\left(\frac{\partial \log \lambda_k^*}{\partial \psi_k} \right)' (\triangle N_k - \lambda_k^* \triangle) \right]_{|\psi_{k|k-1}}$$

for $k = 1, 2, \ldots$

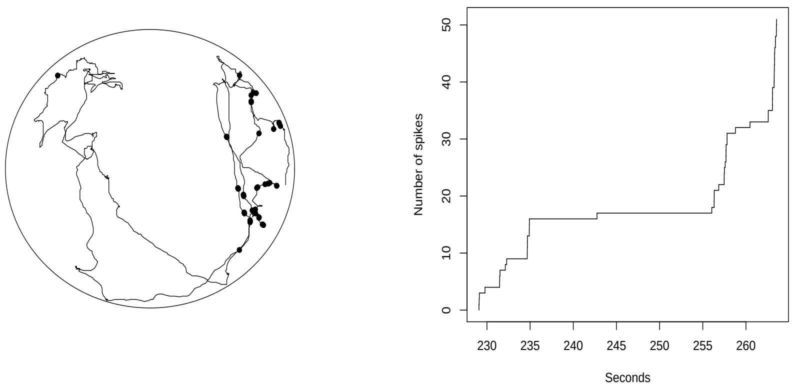

Figure 4.1: Positional firing of the hippocampal neuron. Left a) - rat's track in the arena is displayed by the line. Places at which the neuron fired are indicated by dots. Right b) - the temporal evolution of spikes with graphical presentation as a point process realization. At each spike time the graph is increased by 1 so that at each time it corresponds to the total number of spikes.([33])

This algorithm was evaluated with the data of length 35 sec. with 51 spikes (Fig. 4.1). For the model of conditional intensity λ^* we used Zernicke polynomials (see (2.31)).

$$\lambda^*(x, z) = \exp \left\{ \sum_{j=0}^{n} \sum_{i=-j}^{j} \psi^{j,i} R_j^i(x, z)) \right\},$$

where R_j^i is the $i-$th component of $j-$th order Zernicke polynomials, (x, z) planar coordinates, $\psi^{j,i}$ parameters.

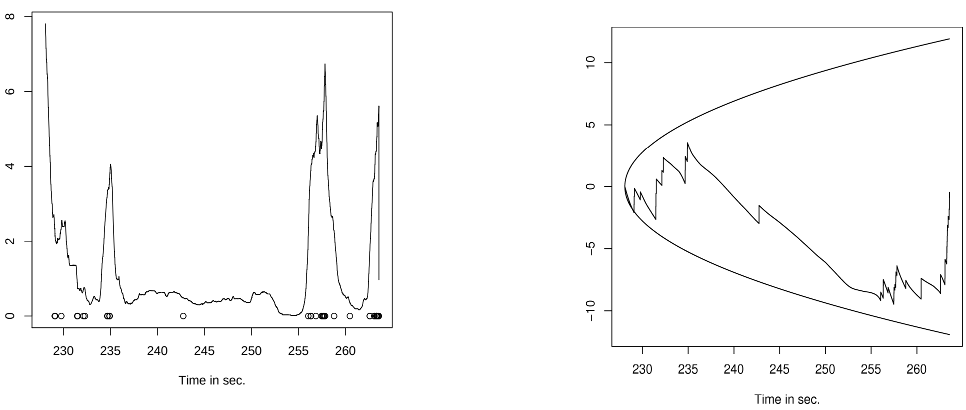

Figure 4.2: Left a) - Conditional intensity estimated by sequential analysis - temporal graph, Right b) - Residual analysis based on sequential Monte Carlo - graph of Pearson residuals (2.48) with bounds (2.49). We observe a good fit on a short interval.

So we can compute λ_k^* as

$$\lambda_k^* = \exp\left\{\sum_{j=0}^{n}\sum_{i=-j}^{j}\psi_k^{j,i}R_j^i(y(k\triangle))\right\}, \tag{4.4}$$

$y(k\triangle)$ is the location on the track and $\psi_k^{j,i}$ the value of parameters at time $k\triangle$. The resulting graph is in Fig. 4.2a, residual analysis in Fig. 4.2b.

We can also compute the function λ_k^* in the whole arena $\mathcal{A}$ according to (4.4) once we have parameters $\psi_k^{j,i}$ estimated. Thus planar maps are produced at each time $k\triangle$. To see the temporal evolution we divide the time interval of observations into subintervals and average estimated coefficients over each subinterval. The graph of resulting planar function is in Fig. 4.3 for two subsequent subintervals of length 18.5 sec.

The Gaussian approximation of posteriors can be further relaxed using sequential Monte Carlo [22]. Particle filter was applied in [25] to evaluate data from the experiment.

4.2 Random marked sets

At this section we model the Cox process on a curve as a random marked set. First the background of recently developed theory of random marked sets [3] is presented.

Denote by $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ the extended real line. Let

$$\phi_{usc} = \left\{(X, f) : X \subset \mathbb{R}^d \text{ is closed, } f : X \to \bar{\mathbb{R}} \text{ is upper semi-continuous}\right\}.$$

ϕ_{usc} is isomorphic to the system $\mathcal{U}_{cl}$ of all closed sets $A \subseteq \mathbb{R}^d \times \bar{\mathbb{R}}$ which satisfy

$$\forall x \in \mathbb{R}^d \forall t \in \bar{\mathbb{R}} : (x,t) \in A \Rightarrow \{x\} \times [-\infty, t] \subseteq A$$

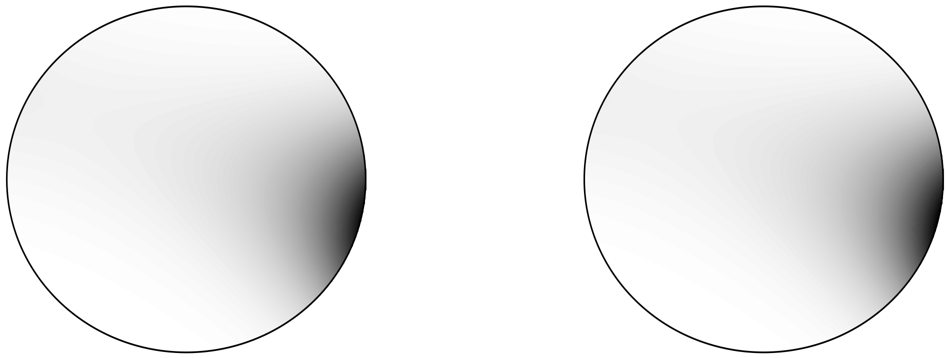

Figure 4.3: Using data from Fig. 3.6 divided in two subsequent intervals of equal length 18.5 sec, conditional intensity function in space was computed from average of Zernicke coefficients.

by the bijection $\tau : \phi_{usc} \to \mathcal{U}_{cl}$

$$(X, f) \mapsto \left\{(x, t) \in X \times \bar{\mathbb{R}} : t \leq f(x)\right\}, \quad (X, f) \in \phi_{usc}.$$

Definition 4.1: *Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space and let $(\Theta, Z) : \Omega \to \phi_{usc}$ be a mapping with*

$$\{\omega \in \Omega : \tau(\Theta, Z) \cap B \neq \emptyset\} \in \mathcal{F}$$

for every compact set B in $\mathbb{R}^d \times \bar{\mathbb{R}}$. Then (Θ, Z) is called a random marked closed set (RMCS).

Random marked closed set is a generalization of the marked point process. In this case the dimension of Θ is zero. Another example of RMCS (Θ, Z) is the model where Z is a Gaussian random field in $\mathbb{R}^d$, $t \in \mathbb{R}$ and

$$\Theta_t = \left\{x \in \mathbb{R}^d : Z(x) \geq t\right\}. \tag{4.5}$$

Here the dimension of Θ_t is d. We are interested in RMCSs with integer Hausdorff dimension $k,\ 0 < k < d$.

Let $X \subset \mathbb{R}^d$ be $\mathcal{H}^k-$rectifiable (Definition 1.29). The relation between RMCS and weighted random measure (Definition 1.26) can be demonstrated by the following diagram, where Ξ is the system of weighted measures in $\mathbb{R}^d$ with weight space $\mathbb{R}$.

$$\begin{array}{ccc} \Phi_{usc} = \{(X, f)\} & & \\ \uparrow & \searrow & \\ (\Omega, \mathcal{F}) & \to & \Xi = \{(\psi, w)\} \end{array}$$

The diagram comumtes for $\psi(.) = \mathcal{H}^k(X \cap .)$ and $w(x,\psi) = f(x)$ on X. Specially in the case $k = 1$ ($k = d-1$) RMCS given by a random fibre (surface) system (see Section 1.4) and a random field Z on $\mathbb{R}^d$ ca be described by a weighted random length (surface area) measure, respectively, with weight given by Z restricted to Θ.

4.2.1 Random-field model

Definition 4.2: *Let $\tilde{Z}$ be an upper semicontinous random field on $\mathbb{R}^d$. (Θ, Z) a RMCS such that $\forall x \in \Theta \;\; Z(x) = \tilde{Z}(x)$. If $\tilde{Z}$ and Θ are stochastically independent, that (Θ, Z) is called a random-field model (with dash).*

In the random-field case from statistical point of view we can study Θ and Z independently. Obviously in the example when Θ_t is a level set (4.5) of a Gaussian random field $\tilde{Z}$ we have that (Θ_t, Z) is not a random-field model.

In Chapter 2 the Cox process on a curve was investigated as a spatio-temporal process. In the case of a random curve, $(\mathcal{Y}, \Lambda)$ is a random marked closed set, where $\mathcal{Y}$ is a special case of a random fibre process (see Section 1.4) in space and time and Λ is a spatio-temporal random field. It is not clear wheather this is a random-field model for given neurophysiological experiment.

We will therefore concentrate on the problem of testing the null-hypothesis H_0 of a random-field model. This was investigated first by Schlather et al. [43] for marked point processes. Later in the book by Illian et al. [31], for stationary marked point processes in $\mathbb{R}^2$, the test is based on a mark-weighted $L-$function

$$L_m = \sqrt{\frac{K_m(r)}{r}}$$

where K_m is the mark-weighted $K-$function.

The testing procedure will be generalized in two ways: firstly to avoid stationarity assumption we will use the Definition 1.25 of second-order intensity reweighted stationary (SOIRS) random measure. Secondly, a general class of RMCSs with $\mathcal{H}^k-$sets Θ in Definition 4.2 can be tested.

Specially we will formulate the test for SOIRS fibre process $\mathcal{Y}$ in the plane considering the random length measure, cf. Section 1.4, Λ is a planar random field. The fibre process here is specially a random curve, therefore we use the temporal parametrization and have thus an unbiased estimator of K-function according to (1.26):

$$\hat{K}(r) = \frac{\triangle_x \triangle_y}{|\mathcal{A}|} \sum_{x \in V_{\mathcal{Y}}^x} \sum_{\substack{y \in V_{\mathcal{Y}}^y \\ y \neq x}} \frac{\mathbf{1}_{(||x-y||<r)}}{L_A(x) L_A(y)}, \quad r > 0. \tag{4.6}$$

Here the curve is approximated by a piecewise linear curve (union of segments) with set of vertices $V_{\mathcal{Y}}^y$ and equal segment length $\triangle_y$. For test points $x \in V_{\mathcal{Y}}^x$ larger intervals (of length $\triangle_x$) can be chosen. $L_A(x)$ is the intensity function of the inhomogeneous length intensity

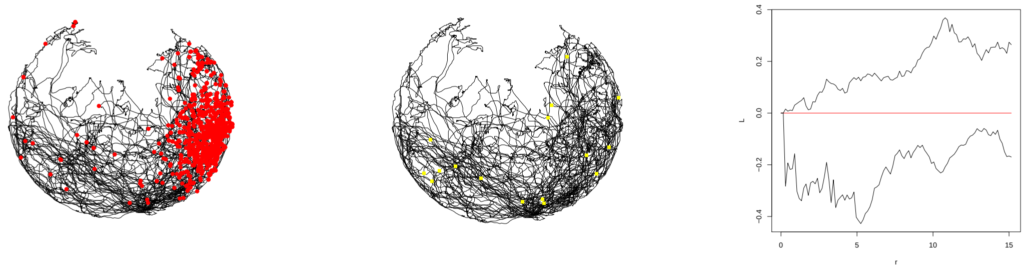

Figure 4.4: The random field model test: data and results. Left a)- The track with spikes (red dots), Middle b) - The test points $x_i \in V_{\mathcal{Y}}^x$ (yellow dots) together with the track, Right c) - The envelopes (4.8) do not hit horizontal axis, we cannot reject the hypotheses of a random-field model.

measure of $\mathcal{Y}$ in $\mathcal{A}$. For data in a bounded set $\mathcal{A}$ standard edge-effect corrections [31] of (4.6) can be used.

Our test is analogous to the case of marked point processes (cf. [31]) and it can be suggested in the following algorithm:

a) evaluate a kernel estimator of $L_A(x)$ from the observed curve,

b) divide the curve in time onto m pieces of equal length $\triangle_x$, $V_{\mathcal{Y}}^x$ is the set of midpoints $x_i \in \mathcal{Y}$ of each piece,

c) evaluate the Λ-weighted K-function estimator ($\triangle_y$ very small)

$$\hat{K}_\Lambda(r) = \frac{\triangle_x \triangle_y}{|\mathcal{A}|} \sum_{x \in V_{\mathcal{Y}}^x} \triangle\Lambda(x) \sum_{\substack{y \in V_{\mathcal{Y}}^y \\ y \neq x}} \frac{\mathbf{1}(||x-y|| < r)}{L_A(x) L_A(y)}, \tag{4.7}$$

where $\triangle\Lambda(x) = \frac{\Lambda(x)}{\mathbb{E}\Lambda(x)}$.

d) Random reallocation: put subsequently n permutations of $\triangle\Lambda(x)$ values in (4.7) and evaluate the estimator for each,

e) draw envelopes

$$\hat{L}_{max}(r) - \hat{L}_\Lambda(r),\ \hat{L}_{min}(r) - \hat{L}_\Lambda(r); \quad \hat{L}_\Lambda(r) = \sqrt{\hat{K}_\Lambda(r)/\pi}. \tag{4.8}$$

If any of the envelopes hits the horizontal axis at a point, the null hypothesis of a random-field model is rejected.

We do not consider the temporal evolution, simply a planar random-field model $(\mathcal{Y}, \Lambda)$ is tested using data from Fig. 4.4a, where $\mathcal{Y}$ is the track (a random curve) and Λ is a planar random field which induces the driving intensity function of a Cox process of spikes along the track.

The above algorithm was used to evaluate the data from Section 3.2. The problem of estimating the weights $\triangle\Lambda(x_i) = \frac{\Lambda(x_i)}{\mathbb{E}\Lambda(x_i)}$ from a single realization was solved in the following way. Both the numerator and denominator is estimated as a ratio of number of events to length of track within a small neighborhood of x_i (or generalized to kernel estimator). While for the denominator the whole track $\mathcal{Y}$ is considered, for the numerator we consider to each x_i only the corresponding piece of track as defined in b).

The results of real data evaluation are in Fig. 4.4b, the null hypothesis of the random-field model is not rejected.

Bibliography

[1] Baddeley A, Moller J, Waagepetersen RP (2000), Non- and semi-parametric estimation of interaction in inhomogeneous point patterns, Statistica Neerlandica 54, 3, 329 – 350.

[2] Baddeley A, Turner R, Moller J, Hazelton M (2005), Residual analysis for spatial point processes (with discussion), J. Royal Stat. Soc. B 67, 617—666.

[3] Ballani F, Kabluchko Z, Schlather M (2009), Random marked sets, http://arxiv.org/abs/0903.2388 .

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

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

[6] Beneš V, Rataj J (2004), Stochastic Geometry: Selected Topics, Kluwer Acad. Publ., Boston.

[7] Beneš V, Bodlák K, Moller J, Waagepetersen R (2005), A case study on point process modelling in disease mapping Image Analysis and Stereology 24, 159–168.

[8] Beneš V, Frcalová B, Klement D, Lánský P (2008), Overdispersion in the Place Cell Discharge - Stochastic Modelling and Inference, Collective Dynamics: Topics on Competition and Cooperation in the Biosciences: A Selection of Papers in the Proceedings of the BIOCOMP2007 International Conference. AIP Conference Proceedings, Volume 1028, 186–197.

[9] Beneš V, Frcalová B (2008), Modelling and simulation of a neurophysiological experiment by spatio-temporal point processes. Image Analysis and Stereology 27, 45–72.

[10] Beneš V, Frcalová B, Klement D (2010), Spatio-temporal point process filtering methods with an application, Environmetrics 2010, 21, 240–252.

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

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

[13] Brown EN, Frank LM, Tang D, Quirk MC, Wilson MA (1998), A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. J Neuroscience 18, 7411–25.

[14] Brown EN, Nguyen DP, Frank LM, Wilson MA, Solo V (2001), An analysis of neural receptive field plasticity by point process adaptive filtering. Neurobiology 98, 21, 12261–66.

[15] Cont R, Tankov P (2004), Financial Modelling With Jump Processes, Chapman & Hall/CRC, London.

[16] Cox DR (1955), Some statistical models related with series of events, Journal of the Royal Statistical Society Series B 17, 129—164.

[17] Vere-Jones D, Musmeci F (1992), A Space-Time Clustering Model For Historical Earthquakes. Ann. Inst. Stat Math. 44, 1992, 1 – 11.

[18] Daley DJ, Vere-Jones D (1988), An Introduction to the Theory of Point Processes, 1st Ed., Springer, New York.

[19] Daley DJ, Vere-Jones D (2008), An Introduction to the Theory of Point Processes, Vol. II, 2nd Ed., Springer, New York.

[20] Diggle PJ (1983). Statistical Analysis of Spatial Point Patterns, Academic press, London.

[21] Diggle PJ (2005), Spatio-temporal point processes, partial likelihood, foot-and-mouth. Statistical Methods in Medical Research 15, 325—336.

[22] Doucet A, de Freitas N, Gordon N (2001), Sequential Monte Carlo Methods in Practice, Springer, New York.

[23] Duan JA, Gelfand AE, Sirmans CF (2009), Modeling space-time data using stochastic differential equations, Bayesian analysis 4, 4, 733 – 758.

[24] Eden UT, Frank LM, Barbieri R, Solo V, Brown EN (2004), Dynamic analysis of neural encoding by point process adaptive filtering. Neural Comp 16, 971–998.

[25] Ergun A, Barbieri R, Eden UT, Wilson MA, Brown EN (2007), Construction of point process adaptive filter algorithms for neural system using sequential Monte Carlo methods, IEEE Transactions on Biomedical Engineering 54, 3, 419–28.

[26] Fano U (1947), Ionization yields radiations. II. The fluctuations of the number of ions. Phys Rev 72, 26–29.

[27] Fenton AA, Muller RU (1998), Place cell discharge is extremely variable during individual passes of the rat through the firing field. Proc Nat Acad Sci USA 95, 3182–3187.

[28] Fishman PM, Snyder D (1976), The statistical analysis of space-time point processes, IEEE Trans. Inf. Theory 22, 257–274.

[29] Frcalová B, Beneš V (2009), Spatio-temporal modelling of a Cox point process sampled by a curve, filtering and inference. Kybernetika 45, 6, 912–930.

[30] Hellmund G, Prokešová M, Vedel Jensen EB (2008), Lévy-based Cox point processes, Adv. Appl. Probab. 40, 3, 603–629.

[31] Illian J, Penttinen A, Stoyan H, Stoyan D (2008), Statistical Analysis and Modelling of Spatial Point Patterns, Wiley, Chichester.

[32] Kendall WS, Liang F, Wang JS (2005), Markov Chain Monte Carlo: Innovations and Applications, World Scientic, Singapore.

[33] Klement D (2006), Stochastic Models in Neurophysiology, Thesis, Charles University, Faculty of Math and Physics, Prague.

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

[35] Lechnerová R, Helisová K, Beneš V (2008), Cox point processes driven by Ornstein-Uhlenbeck type processes, Method. Comp. Appl. Probab. 10, 315–335.

[36] Moller J, Waagepetersen RP (2003), Statistical Inference and Simulation for Spatial Point Processes, Chapman & Hall/CRC, Boca Raton.

[37] Moller J, Diaz-Avalos C (2010), Structured spatio-temporal shot-noise Cox point process models, with a view to modelling forest fires, Scandinavian Journals of Statistics, Theory and Applications, 37, 1, 2–25.

[38] Ogata Y (1998), Space-time point process models for earthquake occurences, Annals Inst. Stat. Math. 50, 379–402.

[39] Pedersen J (2003), The Lévy-Ito decomposition of independently scattered random measure, Res. Report 2, MaPhySto, University of Aarhus.

[40] Peng RD, Schoenberg FP, Woods J (2005), A space-time conditional intensity model for evaluating a wildfire hazard index, JASA 100 (469), 26–35.

[41] Rajput BS, Rosinski J (1989), Spectral representations of infinitely divisible processes, Probab. Th. Rel. Fields 82, 451–487.

[42] Sato KI (1999), Lévy processes and Infinitely Divisible Distributions, Cambridge University Press, Cambridge

[43] Schlather M, Ribeiro PJ, Diggle PJ (2004), Detecting dependence between marks and locations of marked point processes, J. R. Statist. Soc. B 66, 79—93.

[44] Schoenberg FP, Brillinger DR, Guttorp PM (2002). Point processes, spatial-temporal. In: El-Shaarawi A, Piegorsch W, eds., Encyclopedia of Environmetrics, Vol. 3, Wiley, New York, 1573–1577.

[45] Stoyan D, Kendall WS, Mecke J (1995), Stochastic Geometry and Its Applications, 2nd Ed., Wiley, Chichester.

[46] Jensen EV, Jónsdóttir KY, Schmiegel J, Barndorff-Nielsen O (2007), Spatio-temporal modelling – with a view to biological growth, In:Finkenstädt B, Held L, Isham V, eds., Statistical Methods for Spatio-Temporal Systems, Chapman & Hall/CRC, Boca Raton, 47–75.

[47] Zähle M (1982), Random process of Hausdorff rectifiable closed sets, Math. Nachrichten, 108, 49–72

[48] Zhuang J (2006), Second-order residual analysis of spatiotemporal point processes and applications in model evaluation, Royal Statist. Soc. B 68, 4, 635–653.