

Spatio-temporal càdlàg functional marked point processes: Unifying spatio-temporal frameworks

Ottmar Cronie*

Stochastics research group, CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

Jorge Mateu ‡

Department of Mathematics, Universitat Jaume I, Campus Riu Sec, 12071 Castellón, Spain

Abstract

This paper defines the class of càdlàg functional marked point processes (CFMPPs). These are (spatio-temporal) point processes marked by random elements which take values in a càdlàg function space, i.e. the marks are given by càdlàg stochastic processes. We generalise notions of marked (spatio-temporal) point processes and indicate how this class, in a sensible way, connects the point process framework with the random fields framework. We also show how they can be used to construct a class of spatio-temporal Boolean models, how to construct different classes of these models by choosing specific mark functions, and how càdlàg functional marked Cox processes have a double connection to random fields. We also discuss finite CFMPPs, purely temporally well-defined CFMPPs and Markov CFMPPs. Furthermore, we define characteristics such as product densities, Palm distributions and conditional intensities, in order to develop statistical inference tools such as likelihood estimation schemes.

Key words: Boolean model, Càdlàg stochastic process, Conditional intensity, Discrete sampling, Geostatistics with random sampling locations, Intensity functional, LISTA function, Marked reduced Palm measure, Markov process, Maximum (pseudo)likelihood, Pair correlation functional, Papangelou conditional intensity, Product density, Random field, Spatio-temporal functional marked point process, Spatio-temporal geostatistical marking, Spatio-temporal intensity dependent marks, Wiener measure

*e-mail: ottmar@cw.nl, ottmar@alumni.chalmers.se (corresponding author)

†National Research Institute for Mathematics & Computer Science

‡e-mail: mateu@mat.uji.es

1 Introduction

Point processes [Cox and Isham, 1980, Daley and Vere-Jones, 2003, 2008, Karr, 1991, Van Lieshout, 2000, Møller and Waagepetersen, 2004, Stoyan et al., 1995], which may be treated as random collections of points falling in some measurable space, have found use in describing an increasing number of naturally arising phenomena, in a wide variety of applications, including epidemiology, ecology, forestry, mining, hydrology, astronomy, ecology, and meteorology [Cox and Isham, 1980, Daley and Vere-Jones, 2003, Karr, 1991, Møller and Waagepetersen, 2004, Ripley, 1981, Schoenberg and Tranbarger, 2008, Schoenberg, 2011, Tranbarger and Schoenberg, 2010].

Point processes evolved naturally from renewal theory and the statistical analysis of life tables, dating back to the 17th century, and in the earliest applications each point represented the occurrence time of an event, such as a death or an incidence of disease (see e.g. [Daley and Vere-Jones, 2003, Chapter 1] for a review). In the mid-20th century interest expanded to spatial point processes, where each point represents the location of some object or event, such as a tree or a sighting of a species [Cressie, 1993, Diggle, 2003, Ripley, 1981, Stoyan et al., 1995]. More recent volumes have a strong emphasis on spatial processes and address mathematical theory [Daley and Vere-Jones, 2008, Gelfand et al., 2010, Van Lieshout, 2000, Schneider and Weil, 2008], methodology of statistical inference [Van Lieshout, 2000, Møller and Waagepetersen, 2004], and data analysis in a range of applied fields [Diggle, 2003, Ripley, 1981, Baddeley et al., 2000, Illian et al., 2008], although the distinction between these three areas is far from absolute and there are substantial overlaps in coverage between the cited references.

The classical model for temporal or spatial point processes is the Poisson process, where the number of points in disjoint sets are independent Poisson distributed random variables. Alternative models for spatial point processes ([Cressie, 1993, Chapter 8] or [Møller and Waagepetersen, 2004]) grew quite intricate over the course of the 20th century, and among the names associated with these models are some of the key names in the history of statistics, including Jerzy Neyman and David Cox. Today, much attention is paid to spatio-temporal point processes, where each point represents the time and location of an event, such as the origin of an earthquake or wildfire, a lightning strike, or an incidence of a particular disease [Tranbarger and Schoenberg, 2010, Vere-Jones, 2009].

The intimate relationship between point processes and time series is worth noting. Indeed, many data sets that are traditionally viewed as realisations of point processes could in principle also be regarded as time series, and vice versa [Cox and Isham, 1980, Tranbarger and Schoenberg, 2010]. For instance, a sequence of earthquake origin times is typically viewed as a temporal point process, though one could also store such a sequence as a time series consisting of zeros and ones, with the ones representing earthquakes. The main difference is that for a point process, a point can occur at any time in a continuum, whereas for time series, the time intervals are discretised. In addition, if the points are sufficiently sparse, one can see that it may be far more practical to store and analyse the data as a point process, rather than dealing with a long list containing mostly zeros. By the mid 1990s, models for spatial-temporal point processes had become plentiful and often quite intricate.

A probabilistic view of spatio-temporal processes, in principle, can just regard time as one more coordinate and, hence, a special case of a higher-dimensional spatial approach. Of course, this is not appropriate for dynamic spatially referenced processes, as time has a different character than space. There has been a lot of recent work on spatio-temporal models, and a variety of ad hoc approaches have been suggested. Processes that are both spatially and temporally discrete are more naturally considered as binary-valued random fields. Processes that are temporally discrete with only a small number of distinct event-times can be considered initially as multivariate point processes, but with

the qualification that the temporal structure of the type-label may help the interpretation of any inter-relationships among the component patterns. Conversely, spatially discrete processes with only a small number of distinct event-locations can be considered as multivariate temporal point processes, but with a spatial interpretation to the component processes. The other more common end is considering processes that are temporally continuous and either spatially continuous or spatially discrete on a sufficiently large support to justify formulating explicitly spatio-temporal models for the data.

A marked point pattern is one in which each point of the process carries extra information called a mark, which may be a random variable, several random variables, a geometrical shape, or some other information. A multivariate or multi-type point pattern is the special case where the mark is a categorical variable. Marked point patterns with nonnegative real-valued marks are also of interest. A spatial pattern of geometrical objects, such as disks or polygons of different sizes and shapes, can be treated as a marked point process where the points are the centres of the objects, and the marks are parameters determining the size and shape of the objects [Ripley and Sutherland, 1990, Stoyan and Stoyan, 1994].

Marked point patterns raise new and interesting questions concerning the appropriate way to formulate models and pursue analyses for particular applications. In the analysis of a marked point pattern, an important choice is whether to analyse the marks and locations jointly or conditionally. Schematically, if we write X for the points and M for the marks, then we could specify a model for the marked point process $[X, M]$. Alternatively we may condition on the locations of the points, treating only the marks as random variables $[M|X]$. In some cases, we may condition on the marks, treating the locations as a random point process $[X|M]$. This is meaningful if the mark variable is a continuous real-valued quantity, such as time, age or distance. The concept of marking refers to methods of constructing marked point processes from unmarked ones. Two special cases, independent and geostatistical markings, are among the known simple examples of marking strategies and are often used in practice. However, these markings are not able to model density-dependence of marks, the case where the local point intensity affects the mark distribution.

One important situation is where the marks are provided by a (random) field – geostatistical/random field marking. A random field is a quantity $Z(u)$ observable at any spatial location u . A typical question is to determine whether X and Z are independent. If X and Z are independent, then we may condition on the locations and use geostatistical techniques to investigate properties of Z . However, in general, geostatistical techniques, such as the variogram, have a different interpretation when applied to marked point patterns. In this context, the analysis of dependence between marks and locations is of interest. [Schlather et al., 2004] defined the conditional mean and conditional variance of the mark attached to a typical random point, given that there exists another random point at a distance r away from it. These functions may serve as diagnostics for dependence between the points and the marks. Another way to generate non-Poisson marked point processes is to apply dependent thinning to a Poisson marked point process. Interesting examples occur when the thinning rule depends on both the location and the mark of each point.

Despite the relatively long history of point process theory, few approaches have been considered to analyse spatial point patterns where the features of interest are functions (i.e. curves) instead of qualitative or quantitative variables. For instance, an explicit example is given by the *growth-interaction process* [Comas, 2009, Comas et al., 2011, Cronie, 2012, Cronie and Särkkä, 2011, Cronie et al., 2013, Renshaw and Comas, 2009, Renshaw et al., 2009, Renshaw and Särkkä, 2001, Särkkä and Renshaw, 2006], which has been used to model the collective development of tree locations and diameters in forest stands. Moreover, [Illian et al., 2006] consider for each point a transformed Ripley’s K -function to characterise spatial point patterns of ecological plant communities, whilst

[Mateu et al., 2008] build new marked point processes formed by spatial locations and curves defined in terms of LISA functions, which define local characteristics of the point pattern. They use this approach to classify and discriminate between points belonging to a clutter and those belonging to a feature. The study of such configurations permits to analyse the effects of the spatial structure on individual functions. For instance, the analysis of point patterns where the associated curves depend on time may permit the study of spatio-temporal interdependencies of such dynamic processes.

Functional data analysis describes and models data based on curves [Ramsay and Silverman, 2002, 2005]. This theory considers each curve as an observation rather than a set of numbers [Ramsay and Silverman, 2002]. Therefore, functional data analysis together with point process theory provides the theoretical framework to analyse point patterns with associated curves. The use of functional data analysis has already been considered to analyse geostatistical data involving functions instead of single observations. For instance, [Delicado et al., 2010, Giraldo et al., 2010, 2011] develop new geostatistical tools to predict unobserved curves representing daily temperature throughout a year, and analyse a data set consisting of daily meteorological measurements recorded at several weather stations of Canada. However, the use of functional tools in point pattern analysis is limited to just a few references and none of them provides new second order characteristics.

It is clear that there is a wealth of approaches in the theory of spatial point processes. However, the large number of derived spatial point process approaches and methods reduces significantly when handling a spatio-temporal structure in combination with such associated marks. Our aim here is to propose a new class of (spatio-temporal) functional marked point processes, where the marks are random elements which take values in a càdlàg function space. The reason for this choice of function class is its generality and flexibility, and thus its ability to accommodate a variety of different models and structures. With this new setup, we generalise most of the usual notions of (spatio-temporal) marked point processes, hence providing a unifying framework. In addition, we indicate how this framework in a natural way unifies the frameworks of marked point processes and random fields, and we indicate a geometrical interpretation which connects this framework with (spatio-temporal) Boolean models. We develop characteristics such as product densities, Palm distributions and (Papangelou) conditional intensities, as these play a significant role in both theoretical as well as practical aspects of point process analysis. We also we discuss different explicit marking structures and give a thorough description of the statistical framework when the marks are sampled discretely.

The paper is structured as follows. Section 2 presents the new class of càdlàg functional marked point processes, both in its spatial and spatio-temporal versions. Here also some geometric interpretations are discussed. Some motivating examples and connections with other spatio-temporal frameworks are given in Section 3. Section 4 develops certain point process characteristics, such as product densities and Papangelou conditional intensities, which are needed for the development of the statistical theory underlying these processes. Section 5 discusses certain specific marking structures, which may be considered within this framework. The point process characteristics are particularised to Poisson, Cox, temporally well-defined, finite and Markov càdlàg functional marked point processes in Section 6. Finally, in Section 7, the scenario where the functional marks are sampled discretely is covered.

2 Càdlàg functional marked point processes

We here describe the construction of two types of point processes, where the second type is a spatio-temporal version of the first type. Heuristically, the first type may be described as a collection $\Psi = \{(X_i, (L_i, M_i))\}_{i=1}^N$ of Euclidean spatial locations X_i with associated function-valued marks M_i and auxiliary marks L_i , i.e. random parameters/variables with the purpose of controlling M_i . For the second type we further add a random temporal event T_i to each point of Ψ so that the point process Ψ may be described as the collection $\Psi = \{((X_i, T_i), (L_i, M_i))\}_{i=1}^N$. Note that the extra parentheses here are meant to emphasise which part is the space-time location and which part is the mark.

We start by defining the two product spaces on which these two types of point processes are defined. We then continue to define the two types of point processes Ψ as random measures on these two spaces.

2.1 Notation

Let the underlying probability space be denoted by $(\Omega, \mathcal{F}, \mathbb{P})$. Due to the inherent temporally evolving nature of the functional marks and/or the spatio-temporal point process part, at times we will further consider some filtration $\mathcal{F}_{\mathbb{T}}$ and thus obtain a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{\mathbb{T}}, \mathbb{P})$. We let $\mathbb{Z}_+ = \{1, 2, \dots\}$ and $\mathbb{N} = \{0\} \cup \mathbb{Z}_+$, and let \mathcal{P}_N denote the power set of $\{1, \dots, N\}$.

For any x, y in d -dimensional Euclidean space \mathbb{R}^d , $d \geq 1$, we denote the Euclidean norm by $\|x\| = (\sum_{i=1}^d x_i^2)^{1/2}$ (or sometimes $|x|$) and the Euclidean metric by $d_{\mathbb{R}^d}(x, y) = \|x - y\|$. Given some topological space \mathcal{X} , we will call \mathcal{X} a *csm* space if it is a complete separable metric space, and as usual we will denote the Borel sets of \mathcal{X} by $\mathcal{B}(\mathcal{X})$. Given Borel σ -algebras $\mathcal{B}(\mathcal{X}_i)$, $i = 1, \dots, n$, we denote the product σ -algebra by $\bigotimes_{i=1}^n \mathcal{B}(\mathcal{X}_i)$ and by $\mathcal{B}(\mathcal{X})^n$ if $\mathcal{X}_i = \mathcal{X}$, $i = 1, \dots, n$. For measures $\nu_i(\cdot)$ defined on $\mathcal{B}(\mathcal{X}_i)$, $i = 1, \dots, n$, we write $\bigotimes_{i=1}^n \nu_i(\cdot)$ for the product measure and we write ν^n if the measure spaces are identical. We will denote Lebesgue measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ by ℓ and use both $\int_G f(x)\ell(dx)$ and $\int_G f(x)dx$ interchangeably to denote the integral of some measurable function $f: \mathbb{R}^d \rightarrow \mathbb{R}$, with respect to ℓ and $G \subseteq \mathbb{R}^d$. When we need to emphasise the dimension of the space on which we apply ℓ , we write e.g. ℓ_d to denote Lebesgue measure on \mathbb{R}^d .

For any set A , we let $\mathbf{1}_A(a) = \mathbf{1}\{a \in A\}$ denote the indicator function of A and $|A|$ will denote the related cardinality (it will be clear from context whether we consider the norm or the cardinality). Given some measurable space \mathbb{Y} , we let $\delta_y(\cdot)$ denote the Dirac measure of the measurable singleton $\{y\} \subseteq \mathbb{Y}$ and sometimes this notation will also be used for Dirac deltas. As usual, a.s. will be short for *almost surely* and a.e. will be used for *almost everywhere*.

Throughout, by a *kernel* we understand a family $\mu = \{\mu(x, A) : x \in \mathcal{X}, A \in \mathcal{F}\}$ such that, for a fixed $x \in \mathcal{X}$, $\mu(x, \cdot)$ is a measure on some σ -algebra \mathcal{F} and $\mu(\cdot, A)$ is a measurable function for a fixed $A \in \mathcal{F}$. When μ is a kernel such that each $\mu(x, \cdot)$ is a probability measure on \mathcal{F} , we call μ a family of *regular (conditional) probabilities*.

2.2 The state spaces

The spaces \mathbb{X} , \mathbb{T} , \mathbb{A} and \mathbb{F} below will be used as underlying spaces in the construction of (spatio-temporal) càdlàg functional marked point processes. For instance, as we shall see, a spatio-temporal càdlàg functional marked point process will be defined as a marked point process with ground space $\mathbb{X} \times \mathbb{T}$ and mark space $\mathbb{A} \times \mathbb{F}$.

2.2.1 The spatial ground space

Turning now to the purely spatial domain, throughout we will assume that it is given by some subset $\mathbb{X} \subseteq \mathbb{R}^d$, $d \geq 1$, with Borel sets $\mathcal{B}(\mathbb{X})$. Hereby, when we construct our point processes, each point of the point process will have some spatial location $x \in \mathbb{X}$. We note that the most common assumption here is that $\mathbb{X} = \mathbb{R}^d$. However, if $\mathbb{X} \neq \mathbb{R}^d$ or $\mathbb{X} \neq \pm[0, \infty)^d$ we will require that \mathbb{X} is compact in order to make it csm (note that this includes the case of identifying the sides of a (hyper)rectangle in order to construct a torus). When this is the case Ψ becomes a finite point process.

2.2.2 The temporal ground space

In the case of spatio-temporal point processes we also consider the temporal interval domain $\mathbb{T} \subseteq \mathbb{R}$, which contains a point's (main) temporal occurrence/event time $t \in \mathbb{T}$ (in some applications t symbolises e.g. a birth/arrival time). Note that $\mathbb{T} \in \mathcal{B}(\mathbb{T}) \subseteq \mathcal{B}(\mathbb{R})$ and usually $\mathbb{T} = [0, T^*] \subseteq \{0\} \cup \mathbb{R}_+ = [0, \infty)$.

2.2.3 The auxiliary mark space

Being marked point process models, at times we need to connect some auxiliary variable to each point of the process. Such auxiliary information may possibly represent one of the following things.

1. A classification of type: Let $\mathbb{A} = \mathbb{A}_d = \{1, \dots, k_{\mathbb{A}}\}$, $k_{\mathbb{A}} \in \mathbb{Z}_+$, i.e. each auxiliary mark will be of discrete type. Note here that the resulting (spatio-temporal) point process models will be of *multivariate* type [Daley and Vere-Jones, 2003, Van Lieshout, 2000]. The metric chosen is $d_{\mathbb{A}}(l_1, l_2) = |l_1 - l_2|$, $l_1, l_2 \in \mathbb{A}$ and the Borel sets are given by $\mathcal{P}_{k_{\mathbb{A}}}$.
2. Continuous auxiliary information: Let $\mathbb{A} = \mathbb{A}_c \subseteq \mathbb{R}^{m_{\mathbb{A}}}$ for some $m_{\mathbb{A}} \in \mathbb{Z}_+$ (usually $\mathbb{A} = [0, \infty)$). This corresponds to e.g. some additional temporal information, such as a *lifetime*, which possibly controls the behaviour of the functional mark. Here the metric $d_{\mathbb{A}}(\cdot, \cdot)$ on \mathbb{A} will be given by the Euclidean metric $\|\cdot\|$.
3. The combination of the above: Let $\mathbb{A} = \mathbb{A}_d \times \mathbb{A}_c$, with the metric $d_{\mathbb{A}}(l_1, l_2) = \|l_{12} - l_{22}\| + |l_{11} - l_{21}|$, $(l_1, l_2) = ((l_{11}, l_{12}), (l_{21}, l_{22})) \in \mathbb{A}^2$. Note that case 2 may be considered superfluous since we here simply may let $k_{\mathbb{A}} = 1$, whereby each auxiliary mark will take values in $\{1\} \times \mathbb{A}_c$.

Note that under each of the proposed metrics, the corresponding space becomes a csm space and we denote the Borel sets by $\mathcal{B}(\mathbb{A})$.

2.2.4 The functional mark space

In a functional marked point process, a (functional) mark may represent an array of things, ranging from e.g. some feature's growth over time to some function describing spatial dependence. In order to accommodate a large range of models and applications, we choose to allow for the functional marks to take values in a Skorohod space (see e.g. [Billingsley, 1999, Ethier and Kurtz, 1986, Jacod and Shiryaev, 1987, Silvestrov, 2004]).

More specifically, consider some $\mathcal{T} \subseteq [0, \infty)$, with $T^* = \sup \mathcal{T}$ (with $T^* = \infty$ if $\mathcal{T} = [0, \infty)$), and consider the function space

$$\mathbb{F} = D_{\mathcal{T}}(\mathbb{R}) = \{f : \mathcal{T} \rightarrow \mathbb{R} \mid f \text{ càdlàg}\},$$

which is the set of càdlàg (right continuous with existing left limits) functions $f : \mathcal{T} \rightarrow (\mathbb{R}, d_{\mathbb{R}}(\cdot, \cdot))$ (see e.g. [Billingsley, 1999]). Consider now the collection Λ of all strictly increasing, surjective and Lipschitz continuous functions $\lambda : \mathcal{T} \rightarrow \mathcal{T}$, $\lambda(0) = 0$, $\lim_{t \rightarrow \infty} \lambda(t) = T^*$, such that

$$\gamma(\lambda) = \sup_{s, t \in \mathcal{T} : t < s} \left| \log \frac{\lambda(s) - \lambda(t)}{s - t} \right| < \infty.$$

Since $(\mathbb{R}, d_{\mathbb{R}}(\cdot, \cdot))$ is a csm space, by endowing \mathbb{F} with the metric

$$d_{\mathbb{F}}(f, g) = \inf_{\lambda \in \Lambda} \left\{ \gamma(\lambda) \vee \int_{\mathcal{T}} e^{-u} \sup_{t \in \mathcal{T}} \{d_{\mathbb{R}}(f(t \wedge u), g(\lambda(t) \wedge u)) \wedge 1\} du \right\},$$

we turn it into a csm space [Ethier and Kurtz, 1986]. The Borel sets generated by the corresponding topology will be denoted by $\mathcal{B}(\mathbb{F})$ and it follows that $\mathcal{B}(\mathbb{F}^n) = \mathcal{B}(\mathbb{F})^n$ [Jacod and Shiryaev, 1987]. Consider now the following definition, given in accordance with [Silvestrov, 2004, 1.6.1].

Definition 1. *A stochastic process $X(t) = (X_1(t), \dots, X_n(t))$, $n \geq 1$, $t \in \mathcal{T}$, is called an n -dimensional càdlàg stochastic process if each of its sample paths $X(\omega) = \{X(t; \omega)\}_{t \in \mathcal{T}}$, $\omega \in \Omega$, is an element of \mathbb{F}^n .*

In light of this definition, we note that functions in \mathbb{F} include e.g. sample paths of Markov processes, Lévy processes and semi-martingales, as well as empirical distribution functions. We further note that the space $C_{\mathcal{T}}(\mathbb{R}) = \{f : \mathcal{T} \rightarrow \mathbb{R} : f \text{ continuous}\}$ is a subspace of \mathbb{F} and for these functions $d_{\mathbb{F}}$ reduces to the uniform metric $d_{\infty}(f, g) = \sup_{t \in \mathcal{T}} |f(t) - g(t)|$. In addition, the Borel σ -algebra $\mathcal{B}(C_{\mathcal{T}}(\mathbb{R}))$ generated by $d_{\infty}(\cdot, \cdot)$ on $C_{\mathcal{T}}(\mathbb{R})$ satisfies $\mathcal{B}(C_{\mathcal{T}}(\mathbb{R})) = \{E \cap C_{\mathcal{T}}(\mathbb{R}) : E \in \mathcal{B}(\mathbb{F})\} \subseteq \mathcal{B}(\mathbb{F})$ [Jacod and Shiryaev, 1987, Chapter VI]. For details on filtrations with respect to càdlàg stochastic processes, see [Jacod and Shiryaev, 1987, Chapter VI]. Hence, we can accommodate e.g. diffusion processes or some other class of processes with continuous sample paths (note also that each space $C_{\mathcal{T}}^k(\mathbb{R})$, $k \in \mathbb{N}$, of k times continuously differentiable functions is a subspace of $C_{\mathcal{T}}(\mathbb{R})$).

2.3 The spatial and spatio-temporal state spaces

Since both \mathbb{A} and \mathbb{F} are csm, by endowing $\mathbb{M} = \mathbb{A} \times \mathbb{F}$ with the supremum metric

$$d_{\mathbb{M}}((l_1, f_1), (l_2, f_2)) = \max\{d_{\mathbb{A}}(l_1, l_2), d_{\mathbb{F}}(f_1, f_2)\}, \quad (l_1, f_1), (l_2, f_2) \in \mathbb{M},$$

(or any other equivalent metric) \mathbb{M} itself becomes csm [Daley and Vere-Jones, 2003, p. 377] and its Borel sets are given by $\mathcal{B}(\mathbb{M}) = \mathcal{B}(\mathbb{A} \times \mathbb{F}) = \mathcal{B}(\mathbb{A}) \otimes \mathcal{B}(\mathbb{F})$ (see e.g. [Bogachev, 2007, Lemma 6.4.2.]).

2.3.1 The spatio-temporal state space

Let $\mathbb{G} = \mathbb{X} \times \mathbb{T}$ and endow it with the supremum norm $\|(x, t)\|_{\infty} = \max\{\|x\|, |t|\}$ and the supremum metric

$$d_{\mathbb{G}}((x_1, t_1), (x_2, t_2)) = \|(x_1, t_1) - (x_2, t_2)\|_{\infty} = \max\{d_{\mathbb{R}^d}(x_1, x_2), d_{\mathbb{R}}(t_1, t_2)\},$$

where $(x_1, t_1), (x_2, t_2) \in \mathbb{G}$, so that \mathbb{G} becomes a csm space and $\mathcal{B}(\mathbb{G}) = \mathcal{B}(\mathbb{X} \times \mathbb{T}) = \mathcal{B}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{T})$. We note that there are other possible equivalent metrics, which measure space and time differently (this is needed since it is the defining property of spatio-temporal point processes). However, for our purposes, this is the preferable choice.

Remark 1. Note that if the scales of time and space need to be altered, we may rescale e.g. time by letting $\|(x, t)\|_\infty = \max\{\|x\|, \beta|t|\}$, $\beta > 0$. The current construction amounts to $\beta = 1$.

The resulting underlying measurable spatio-temporal space which we will consider is given by

$$(\mathbb{Y}, \mathcal{B}(\mathbb{Y})) = (\mathbb{G} \times \mathbb{M}, \mathcal{B}(\mathbb{G} \times \mathbb{M})) = ((\mathbb{X} \times \mathbb{T}) \times (\mathbb{A} \times \mathbb{F}), \mathcal{B}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{T}) \otimes \mathcal{B}(\mathbb{A}) \otimes \mathcal{B}(\mathbb{F}))$$

and we note that \mathbb{Y} is a Polish space, as a product of Polish spaces. In fact, by endowing $\mathbb{Y} = \mathbb{G} \times \mathbb{M}$ with the supremum metric

$$d((x_1, t_1, l_1, f_1), (x_2, t_2, l_2, f_2)) = \max\{d_{\mathbb{G}}((x_1, t_1), (x_2, t_2)), d_{\mathbb{M}}((l_1, f_1), (l_2, f_2))\},$$

\mathbb{Y} itself becomes a csm space [Van Lieshout, 2000, p. 8].

Concerning \mathbb{G} , for any $(x, t) \in \mathbb{G}$ and $u, v \geq 0$, consider the cylinder set

$$\begin{aligned} (x, t) + C_u^v &= (x, t) + \{(y, s) \in \mathbb{G} : \|y\| \leq u, |s| \leq v\} \\ &= \{(y, s) \in \mathbb{G} : d_{\mathbb{R}^d}(x, y) \leq u, d_{\mathbb{R}}(t, s) \leq v\}. \end{aligned} \quad (1)$$

We see that in this metric space closed balls satisfy $B[(x, t), u] = (x, t) + C_u^u$.

Remark 2. We note that in many, if not most, cases it is desirable to set $\mathcal{T} = \mathbb{T}$ so that \mathbb{T} describes the total part of time which we are considering for the constructed point process on \mathbb{Y} .

2.3.2 The spatial state space

The same reasoning gives us the (explicitly) non-temporal space

$$(\mathbb{Y}, \mathcal{B}(\mathbb{Y})) = (\mathbb{G} \times \mathbb{M}, \mathcal{B}(\mathbb{G}) \otimes \mathcal{B}(\mathbb{M})) = (\mathbb{X} \times (\mathbb{A} \times \mathbb{F}), \mathcal{B}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{A}) \otimes \mathcal{B}(\mathbb{F}))$$

with \mathbb{G} having underlying norm $\|\cdot\|$ and metric $d_{\mathbb{G}}(x_1, x_2) = d_{\mathbb{R}^d}(x_1, x_2)$. We see here that the only temporal information present is found implicitly in each $f = \{f(t) : t \in \mathcal{T}\} \in \mathbb{F}$, provided that $t \in \mathcal{T}$ describes time.

2.4 Reference measures and reference càdlàg stochastic processes

When constructing marked point processes, for various reasons, including the derivation of explicit structures for different summary statistics, one has to choose a sensible reference measure $\nu_{\mathbb{M}}$ for the mark space $(\mathbb{M}, \mathcal{B}(\mathbb{M}))$. For similar reasons one also usually considers some reference measure $\nu_{\mathbb{G}}$ on the ground space $(\mathbb{G}, \mathcal{B}(\mathbb{G}))$. We here let the reference measure on $(\mathbb{Y}, \mathcal{B}(\mathbb{Y}))$ be given by

$$\nu(\cdot) = [\nu_{\mathbb{G}} \otimes \nu_{\mathbb{M}}](\cdot) = [\ell \otimes [\nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}}]](\cdot), \quad (2)$$

where each component measure in ν governs the probabilistic structures of Ψ on \mathbb{G} , \mathbb{A} and \mathbb{F} , respectively.

Regarding the measure on \mathbb{G} , we let it be given by Lebesgue measure ℓ , where $\ell = \ell_d$ if $\mathbb{G} = \mathbb{X}$ and $\ell = \ell_{d+1} = \ell_d \otimes \ell_1$ if $\mathbb{G} = \mathbb{X} \times \mathbb{T}$. This is the usual choice when constructing point processes on \mathbb{R}^d or spatio-temporal point processes on $\mathbb{R}^d \times \mathbb{R}$ (recall that the metrics are different).

Recall the different auxiliary mark spaces given in Section 2.2.3. Irrespective of whether $\mathbb{A} = \mathbb{A}_d$, $\mathbb{A} = \mathbb{A}_c$ or $\mathbb{A} = \mathbb{A}_d \times \mathbb{A}_c$, we let the auxiliary mark reference measure $\nu_{\mathbb{A}}$ be given by some locally

finite Borel measure on $\mathcal{B}(\mathbb{A})$, i.e. $\nu_{\mathbb{A}}(D) < \infty$ for bounded $D \in \mathcal{B}(\mathbb{A})$. In Section 5.3 we discuss in detail some possible choices for $\nu_{\mathbb{A}}$.

Turning to the functional mark space $(\mathbb{F}, \mathcal{B}(\mathbb{F}))$, consider some suitable reference càdlàg stochastic process

$$\begin{aligned} X_{\mathbb{F}} : (\Omega, \mathcal{F}, \mathbb{P}) &\rightarrow (\mathbb{F}, \mathcal{B}(\mathbb{F})), \\ \Omega \ni \omega &\mapsto X_{\mathbb{F}}(\omega) = \{X_{\mathbb{F}}(t; \omega)\}_{t \in \mathcal{T}} \in \mathbb{F}, \end{aligned} \quad (3)$$

where each $X_{\mathbb{F}}(\omega)$ is commonly referred to as a sample path/realisation of $X_{\mathbb{F}}$, and consider the induced probability measure

$$\nu_{\mathbb{F}}(E) = \mathbb{P}(\{\omega \in \Omega : X_{\mathbb{F}}(\omega) \in E\}), \quad E \in \mathcal{B}(\mathbb{F}),$$

which will be the canonical reference measure under consideration. Note that the joint distribution on $(\mathbb{F}^n, \mathcal{B}(\mathbb{F}^n))$ of n independent copies of $X_{\mathbb{F}}$ is given by $\nu_{\mathbb{F}}^n$, the n -fold product measure of $\nu_{\mathbb{F}}$ with itself. Also, we may conversely first choose the measure $\nu_{\mathbb{F}}$ and then consider the corresponding process $X_{\mathbb{F}}$.

For reasons which will become clear, $\nu_{\mathbb{F}}$ or $X_{\mathbb{F}}$ should be chosen so that suitable absolute continuity/change-of-measure results can be applied. More specifically, the distribution P_X on $(\mathbb{F}^n, \mathcal{B}(\mathbb{F}^n))$, $n \geq 1$, of some stochastic process $X = \{X(t)\}_{t \in \mathcal{T}} \in \mathbb{F}^n$ of interest should have some (functional) Radon-Nikodym derivative f_X with respect to $\nu_{\mathbb{F}}^n$, i.e. $P_X(E) = \int_E f_X(f) \nu_{\mathbb{F}}^n(df) = \mathbb{E}_{\nu_{\mathbb{F}}^n}[\mathbf{1}_E f_X]$, $E \in \mathcal{B}(\mathbb{F}^n)$ (see [Skorohod, 1967] for a discussion on such densities). In Section 5.2 we discuss such choices further and we look closer at Wiener measure as reference measure, i.e. the measure induced by a Brownian motion $X_{\mathbb{F}} = W = \{W(t)\}_{t \in \mathcal{T}}$.

2.5 Point processes

Having defined the state spaces for the two types of point processes defined here, we now turn to their actual definitions.

Let $(\mathbb{Y}, \mathcal{B}(\mathbb{Y}))$ be given by any of the two state spaces defined above. Furthermore, let $\mathcal{N}_{\mathbb{Y}}$ be the collection of all locally finite counting measures $\varphi = \sum_{y \in \mathbb{Y}} \delta_y$ on $\mathcal{B}(\mathbb{Y})$, i.e. $\varphi(A) < \infty$ for bounded $A \in \mathcal{B}(\mathbb{Y})$ and denote the corresponding counting measure σ -algebra by $\Sigma_{\mathcal{N}_{\mathbb{Y}}}$ (see [Daley and Vere-Jones, 2008, Chapter 9]). Note that in what follows we will not distinguish in the notation between a measure $\varphi \in \mathcal{N}_{\mathbb{Y}}$ and its support $\varphi \subseteq \mathbb{Y}$ whereby $\varphi(\{y\}) > 0$ and $y \in \varphi$ (or $|\varphi \cap \{y\}| \neq 0$) will mean the same thing for any $y \in \mathbb{Y}$.

Definition 2. If $\Psi : \Omega \rightarrow \mathcal{N}_{\mathbb{Y}}$, $\omega \mapsto \Psi(\cdot; \omega)$, is a measurable mapping from the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ into the space $(\mathcal{N}_{\mathbb{Y}}, \Sigma_{\mathcal{N}_{\mathbb{Y}}})$, we call Ψ a point process on \mathbb{Y} .

Denote further by $\mathcal{N}_{\mathbb{Y}}^*$ the sub-collection of $\varphi \in \mathcal{N}_{\mathbb{Y}}$ such that the ground measure $\varphi_G(\cdot) = \varphi(\cdot \times \mathbb{M})$ is a locally finite simple counting measure on $\mathcal{B}(\mathbb{G})$ (simple means that $\varphi_G(\{g\}) \in \{0, 1\}$ for any $g \in \mathbb{G}$). We note that the simplicity of the ground measure further implies that $\varphi(\{(g, m)\}) \leq \varphi_G(\{g\}) \in \{0, 1\}$ for any $(g, m) \in \mathbb{G} \times \mathbb{M}$.

Throughout, irrespective of the choice of \mathbb{G} , for any $\varphi = \sum_{(g, l, f) \in \varphi} \delta_{(g, l, f)} \in \mathcal{N}_{\mathbb{Y}}$ (where $g \in \mathbb{G}$ and $(l, f) \in \mathbb{M}$) we will write $\varphi + z = \sum_{(g, l, f) \in \varphi} \delta_{(g+z, l, f)}$ to denote a shift of φ in the ground space by the vector $z \in \mathbb{G}$. This notation will, in particular, be used in the definition of stationarity.

Recalling Definition 1, we see that any collection of elements $\{(g_1, l_1, f_1), \dots, (g_n, l_n, f_n)\} \subseteq \Psi$ consists of the combination of a) a collection of spatial(-temporal) points $g_1, \dots, g_n \in \mathbb{G}$, b) a collection l_1, \dots, l_n of random variables taking values in \mathbb{A} , and c) an n -dimensional càdlàg stochastic process $(f_1(t), \dots, f_n(t))$, $t \in \mathcal{T}$, all tied together.

2.6 Càdlàg functional marked point processes

Following the terminology and structure given in [Comas et al., 2011], we now have the following definition.

Definition 3. Let $\mathbb{Y} = \mathbb{G} \times \mathbb{M} = \mathbb{X} \times (\mathbb{A} \times \mathbb{F})$ and let $\Psi : \Omega \rightarrow \mathcal{N}_{\mathbb{Y}}$ be a point process on $\mathbb{Y} = \mathbb{X} \times (\mathbb{A} \times \mathbb{F})$. If $\Psi \in \mathcal{N}_{\mathbb{Y}}^*$ a.s., we call

$$\Psi = \sum_{y \in \Psi} \delta_y = \sum_{(x,l,f) \in \Psi} \delta_{(x,l,f)}$$

a (simple) càdlàg functional marked point process (CFMPP) on \mathbb{Y} .

- If either $\mathbb{A} = \mathbb{A}_d$ or $\mathbb{A} = \mathbb{A}_d \times \mathbb{A}_c$, with $k_{\mathbb{A}} \geq 2$ different type classifications, we call Ψ a multivariate CFMPP.
- If further Ψ a.s. takes its values in $\mathcal{N}^f = \{\varphi \in \mathcal{N}_{\mathbb{Y}}^* : \varphi(\mathbb{Y}) < \infty\} \subseteq \mathcal{N}_{\mathbb{Y}}$, we call it a finite CFMPP.

We note that through a unique measurable enumeration (see [Daley and Vere-Jones, 2008, Chapter 9.1]), we may write

$$\Psi = \sum_{i=1}^N \delta_{(X_i, L_i, M_i)}, \quad 0 \leq N \equiv \Psi(\mathbb{Y}) \leq \infty,$$

for some sequence $\{(X_i, L_i, M_i)\}_{i=1}^N$ of random vectors, which geometrically corresponds to the support of Ψ . Here $X_i \in \mathbb{R}^d$ represents the spatial location of the i th point, $L_i \in \mathbb{A}$ its auxiliary mark and $M_i \in \mathbb{F}$ its functional mark. Note that when Ψ is multivariate and $\mathbb{A} = \mathbb{A}_d \times \mathbb{A}_c$, to emphasise this aspect we often write $L_i = (L_{i1}, L_{i2})$. It should further be noted that by construction the *ground process* (unmarked process)

$$\Psi_{\mathbb{X}}(\cdot) = \Psi_G(\cdot) = \sum_{x \in \Psi_G} \delta_x(\cdot) = \sum_{y \in \Psi} \delta_y(\cdot \times \mathbb{A} \times \mathbb{F}),$$

with support $\Psi_G = \{X_i\}_{i=1}^N \subseteq \mathbb{X}$, is a well-defined simple point process on \mathbb{X} with $\Psi_G(B) = |\Psi_G \cap B| = \Psi(B \times \mathbb{A} \times \mathbb{F}) < \infty$ a.s. for bounded $B \in \mathcal{B}(\mathbb{X})$. Note that the dual notation $\Psi_G = \Psi_{\mathbb{X}}$ is introduced for later convenience.

Remark 3. Implicitly in the definition of a CFMPP we assume that Ψ is simple (since $\Psi \in \mathcal{N}_{\mathbb{Y}}^*$ a.s.). Furthermore, if $\Psi_{\mathbb{F}}(\cdot) := \Psi(\mathbb{X} \times \cdot)$ is locally finite, then $\Psi_{\mathbb{F}}$ becomes a well-defined point process on \mathbb{F} and we refer to $\Psi_{\mathbb{F}}$ as the associated mark space point process. However, we will not necessarily make that assumption here.

As already noted, by construction the collection of marks $\Psi_{\mathbb{F}} = \{M_i\}_{i=1}^N$, $M_i = \{M_i(t)\}_{t \in \mathcal{T}}$, consists of random elements in \mathbb{F} (functional random variables), which simply are càdlàg stochastic processes with sample paths/realisations $M_i(\omega) = \{M_i(t; \omega)\}_{t \in \mathcal{T}} \in \mathbb{F}$, $\omega \in \Omega$. As we will see, by letting M_i be given by a point mass δ_f on $(\mathbb{F}, \mathcal{B}(\mathbb{F}))$ we also have the possibility to consider marks which are given by deterministic functions f .

2.7 Spatio-temporal càdlàg functional marked point processes

We now turn to the case where we include the explicit temporal space \mathbb{T} and, consequently, deal with spatio-temporal CFMPPs. Recall that $\mathbb{Y} = \mathbb{G} \times \mathbb{M} = (\mathbb{X} \times \mathbb{T}) \times (\mathbb{A} \times \mathbb{F})$.

Definition 4. Let $\Psi : \Omega \rightarrow \mathcal{N}_{\mathbb{Y}}$ be a point process on $\mathbb{Y} = \mathbb{G} \times \mathbb{M} = (\mathbb{X} \times \mathbb{T}) \times (\mathbb{A} \times \mathbb{F})$. If $\Psi \in \mathcal{N}_{\mathbb{Y}}^*$ a.s., we call

$$\Psi = \sum_{y \in \Psi} \delta_y = \sum_{(x,t,l,f) \in \Psi} \delta_{(x,t,l,f)}$$

a (simple) spatio-temporal càdlàg functional marked point process (STCFMPP) on \mathbb{Y} .

- If either $\mathbb{A} = \mathbb{A}_d$ or $\mathbb{A} = \mathbb{A}_d \times \mathbb{A}_c$, with $k_{\mathbb{A}} \geq 2$ different type classifications, we call Ψ a multivariate STCFMPP.
- When $\Psi \in \mathcal{N}^f = \{\varphi \in \mathcal{N}_{\mathbb{Y}}^* : \varphi(\mathbb{Y}) < \infty\} \subseteq \mathcal{N}_{\mathbb{Y}}$ a.s., we call Ψ a finite STCFMPP.

A few things should be mentioned at this point. To begin with we note that we may write

$$\Psi = \sum_{i=1}^N \delta_{(X_i, T_i, L_i, M_i)}, \quad 0 \leq N \equiv \Psi(\mathbb{Y}) \leq \infty,$$

where all $X_i \in \mathbb{R}^d$ represent the spatial locations, $T_i \in \mathbb{T}$ the occurrence times, $L_i \in \mathbb{A}$ the related auxiliary marks and $M_i = \{M_i(t)\}_{t \in \mathcal{T}} \in \mathbb{F}$ the functional marks. In connection hereto, an interesting feature which sets this scenario apart from the non-spatio-temporal CFMPP case is that we here have a natural enumeration/order of the points, which is obtained by assigning the indices $1, \dots, N$ to the points according to their ascending occurrence times $T_1 < \dots < T_N$. Hereby the support may be written as $\Psi = \{(X_i, T_i), (L_i, M_i)\}_{i=1}^N = \{(X_i, T_i, L_i, M_i)\}_{i=1}^N$. Also here, when Ψ is multivariate and $\mathbb{A} = \mathbb{A}_d \times \mathbb{A}_c$, we sometimes write $L_i = (L_{i1}, L_{i2})$.

We note further that, by construction, the ground process is a well-defined simple point process on $\mathbb{X} \times \mathbb{T}$, i.e.

$$\Psi_{\mathbb{X} \times \mathbb{T}}(B \times C) = \Psi_G(B \times C) = \sum_{(x,t) \in \Psi_G} \delta_{(x,t)}(B \times C) = \Psi(B \times C \times (\mathbb{A} \times \mathbb{F})) < \infty,$$

for bounded $B \times C \in \mathcal{B}(\mathbb{X} \times \mathbb{T})$, with support $\Psi_G = \{(X_i, T_i)\}_{i=1}^N$. However, at times it may be useful to require that also $\Psi_{\mathbb{X}} = \{X_i\}_{i=1}^N$ and/or $\Psi_{\mathbb{T}} = \{T_i\}_{i=1}^N$ constitute well-defined point processes.

Definition 5. Let Ψ be a STCFMPP.

- If $\Psi_{\mathbb{X}}(\cdot) = \Psi_G(\cdot \times \mathbb{T}) = \Psi(\cdot \times \mathbb{T} \times \mathbb{A} \times \mathbb{F})$ is simple and locally finite, i.e. the spatial part $\Psi_{\mathbb{X}}$ of the ground process also constitutes a well-defined simple point process on \mathbb{X} , we say that Ψ is spatially grounded.
- Similarly, if $\Psi_{\mathbb{T}}(\cdot) = \Psi_G(\mathbb{X} \times \cdot) = \Psi(\mathbb{X} \times \cdot \times \mathbb{A} \times \mathbb{F})$ is simple and locally finite, so that $\Psi_{\mathbb{T}}$ constitutes a well-defined point process on \mathbb{T} , we say that Ψ is temporally grounded.

To additionally ground Ψ spatially and/or temporally can be of importance for different reasons. For instance, as we shall see, we may speak of three different types of stationarity of Ψ and when Ψ is temporally grounded we may e.g. define conditional intensities, as considered by e.g. [Ogata, 1998, Schoenberg, 2004, Vere-Jones, 2009].

2.8 The point process distribution

In what follows, we only distinguish in the notation between CFMPPs and STCFMPPs when necessary. Let Ψ be a (ST)CFMPP and let the induced probability measure of Ψ on $\Sigma_{\mathcal{N}_{\mathbb{Y}}}$ be denoted by P , i.e.

$$\mathbb{P}(\{\omega \in \Omega : \Psi(\omega) \in R\}) = P(R) = \int_R P(d\varphi), \quad R \in \Sigma_{\mathcal{N}_{\mathbb{Y}}}.$$

Note that since Ψ is a simple point process on the csm space \mathbb{Y} , P is completely and uniquely determined by its finite dimensional distributions, i.e. the collection of joint distributions of $(\Psi(A_1), \dots, \Psi(A_n))$ for all collections of bounded $A_i \in \mathcal{B}(\mathbb{Y})$, $i = 1, \dots, n$, $n \in \mathbb{Z}_+$, as well as by its void probabilities $v(A) = \mathbb{P}(\Psi(A) = 0)$, $A \in \mathcal{B}(\mathbb{Y})$ (see e.g. [Van Lieshout, 2000, Chapter 1]).

2.9 Stationarity and isotropy

We next give the definition of stationarity which, irrespective of the choice of \mathbb{G} , is the usual definition for marked point processes, i.e. translational invariance of the ground process.

Definition 6. *Let Ψ be a (ST)CFMPP.*

- *Then Ψ is stationary if $\Psi + z \stackrel{d}{=} \Psi$ for any $z \in \mathbb{G}$.*
- *In the case of a STCFMPP, for $z = (a, b) \in \mathbb{X} \times \mathbb{T}$, if stationarity only holds when $b = 0$ we say that Ψ is spatially stationary and if it only holds when $a = 0$ we say that Ψ is temporally stationary.*
- *Ψ is isotropic if it is stationary and, in addition, Ψ_G is rotation invariant with respect to rotations about the origin $0 \in \mathbb{G}$.*

We see e.g. that when Ψ is temporally grounded, temporal stationarity implies that $\Psi_{\mathbb{T}}$ is a stationary point process on \mathbb{T} . Note further that one often refers to Ψ as *homogeneous* if it is both stationary and isotropic, and as *inhomogeneous* if it is not stationary.

2.10 Supports

Being a stochastic process, which may have zeroes on \mathcal{T} , conditionally on Ψ_G and the auxiliary marks $L_i \in \mathbb{A}$, we may consider different types of supports for each functional mark $M_i = \{M_i(t)\}_{t \in \mathcal{T}}$. We distinguish between the *deterministic support* $\text{supp}(M_i) = \{t \in \mathcal{T} : M_i(t; \omega) \neq 0 \text{ for any } \omega \in \Omega\}$ and the *stochastic support* $\text{supp}^*(M_i) = \{t \in \mathcal{T} : M_i(t) \neq 0\}$. We note that $\text{supp}^*(M_i) \subseteq \text{supp}(M_i)$ is a random subset of \mathcal{T} and, moreover,

$$\text{supp}(M_i) = \bigcup_{\omega \in \Omega} \{t \in \mathcal{T} : M_i(t; \omega) \neq 0\} = \bigcup_{\omega \in \Omega} \text{supp}^*(M_i; \omega).$$

For a STCFMPP Ψ , when $L_i \geq 0$ represents a (random) time and we condition on the T_i 's and the L_i 's, a natural construction would be to let $D_i = (T_i + L_i) \wedge T^*$ and $\text{supp}(M_i) = [T_i, D_i)$ so that $M_i(t) = 0$ for all $t \notin [T_i, D_i)$. The interpretation here would be that T_i symbolises the *birth time* of the i th point, L_i its *lifetime* and D_i its *death time*. Furthermore, we note that if e.g. $M_i(t) = \mathbf{1}_{[T_i, D_i)}(t)W_i((t - T_i) \wedge 0)$ for a Brownian motion W_i , then $\ell_1((\text{supp}(M_i) \setminus \text{supp}^*(M_i))^c) = \ell_1(\{t \in [0, L_i) : W_i(t) = 0\}) = 0$ a.s. (see e.g. [Klebaner, 2005]). In the case of a CFMPP one could similarly let $\text{supp}(M_i) = [0, L_i)$.

2.11 Geometric representation and spatio-temporal Boolean models

Let Ψ be a STCFMPP where $\mathbb{X} \subseteq \mathbb{R}^2$ and where each M_i a.s. takes values in the sub-space of continuous functions on \mathcal{T} . One possibility for interpretation is obtained by letting the disk (ball) $B_{\mathbb{X}}[X_i, M_i(t)] = \{x \in \mathbb{R}^2 : d_{\mathbb{R}^2}(X_i, x) \leq M_i(t)\}$ with centre X_i and radius $M_i(t)$ illustrate the space occupied by the i th point of Ψ at time $t \in \mathbb{T}$ (with the convention that $B_{\mathbb{X}}[X_i, r] = \emptyset$ if $r \leq 0$). Hereby, at time t we may illustrate Φ_M as the Boolean model (see e.g. [Stoyan et al., 1995])

$$\Xi(t) = \bigcup_{i=1}^N B_{\mathbb{X}}[X_i, M_i(t)] = \bigcup_{(X_i, T_i, L_i, M_i) \in \Psi: t \in \text{supp}^*(M_i)} B_{\mathbb{X}}[X_i, M_i(t)].$$

Consequently, Ψ may be represented by the collection

$$\Xi = \int_{\mathbb{T}} \Xi(dt) = \bigcup_{i=1}^N \Xi_i = \bigcup_{i=1}^N \{(x, y, z) \in \mathbb{R}^3 : z \in \text{supp}^*(M_i), d_{\mathbb{R}^2}(X_i, (x, y)) \leq M_i(z)\}$$

and we see that whenever $\text{supp}(M_i)$ is bounded, each *deformed cone* Ξ_i a.s. is a compact subset of \mathbb{R}^3 . Figure 1 illustrates a realisation of such a random set Ξ . Hence, the cross section of Ξ at $z = t$ gives us $\Xi(t)$ and in the context of e.g. forest stand modelling, we find that $\Xi(t)$ gives us the geometric representation of the cross section of the forest stand at time t , at some given height (usually *breast height*). Note that when in addition $\ell(\mathbb{X}) < \infty$, depending on the form of the functional marks, we may derive geometric properties such as the expected coverage proportion $\frac{\pi}{\ell(\mathbb{X})} \sum_{n=0}^{\infty} \sum_{i=1}^n \mathbb{E}[M_i(t)^2] \mathbb{P}(N = n)$ of \mathbb{X} at time t (provided that the disks do not overlap).

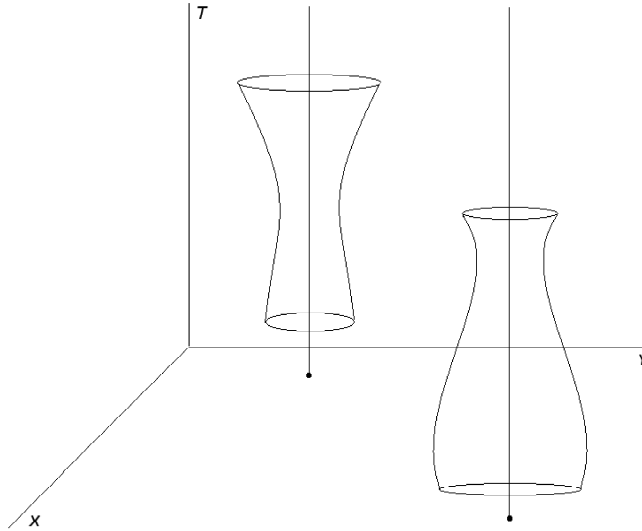


Figure 1: A realisation of a random set Ξ .

3 Examples of (ST)CFMPPs

The class of (ST)CFMPPs provides a framework to give structure to a series of existing models and it allows for the construction of new important models and modelling frameworks, which have uses in different applications. Below we present a few such instances.

3.1 Marked (spatio-temporal) point processes

Not surprisingly, (ST)CFMPPs generalise ordinary marked (spatio-temporal) point processes. To see this, by considering the class $\mathbb{F}_c = \{f \in \mathbb{F} : f \text{ is constant}\}$, we find that when Ψ a.s. is restricted to the space $\mathbb{G} \times (\mathbb{A} \times \mathbb{F}_c)$, its functional marks become constant functions $M_i(t) = \xi_i$ with (random) values $\xi_i \in \mathbb{R}$, $i = 1, \dots, N$. Hereby, in the case of a CFMPP Ψ , $\bar{\Psi} = \{(X_i, L_i, M_i(0)) : X_i \in \Psi_G\} = \{(X_i, L_i, \xi_i) : X_i \in \Psi_G\}$ gives us a classical definition of a marked point process (provided $0 \in \text{supp}(M_i)$). Similarly, in the case of a STCFMPP Ψ , $\bar{\Psi} = \{(X_i, T_i, L_i, M_i(0)) : X_i \in \Psi_G\} = \{(X_i, T_i, L_i, \xi_i) : X_i \in \Psi_G\}$ gives the classical definition of a marked spatio-temporal point process.

Here, when Ψ is multivariate with, say, $\mathbb{A} = \mathbb{A}_c$, L_i would be a discrete variable which describes a point's type and ξ_i would be the size of the quantitative mark. As a consequence $\bar{\Psi}$ would become multivariate.

A slightly more direct way of creating, say, a marked spatio-temporal point process $\bar{\Psi}$ through a STCFMPP is to let

$$\bar{\Psi}(B \times C \times D) = \Psi(B \times C \times D \times \mathbb{F}) = \sum_{(x,t,l,f) \in \Psi} \delta_{(x,t,l,f)}(B \times C \times D \times \mathbb{F}),$$

$B \times C \times D \in \mathcal{B}(\mathbb{X}) \times \mathcal{B}(\mathbb{T}) \times \mathcal{B}(\mathbb{A})$ so that the i th mark is given by $L_i \in \mathbb{A}$. This is naturally possible only if $\bar{\Psi}$ constitutes a well-defined point process in its own right.

3.2 Spatio-temporal geostatistical marking and geostatistics with uncertainty in the sampling locations

For classic marked point processes $\bar{\Psi} = \{(X_i, M_i)\}_{i=1}^N$ one often speaks of *geostatistical marking* [Illian et al., 2008]. This is the case where, conditionally on X_i , the marks $M_i = Z_{X_i}$, $i = 1, \dots, N$, are provided by some random field $Z = \{Z_x\}_{x \in \mathbb{X}}$. This may be regarded as *sampling the random field Z at random locations*, provided by $\{X_i\}_{i=1}^N$. Within the CFMPP-context this idea may further be extended to the case of marks coming from a spatio-temporal random field $Z = \{Z_x(t)\}_{(x,t) \in \mathbb{X} \times \mathcal{T}}$.

Definition 7. Consider a spatio-temporal random field $Z = \{Z_x(t)\}_{(x,t) \in \mathbb{X} \times \mathcal{T}}$. If, conditionally on Ψ_G and $\{L_i\}_{i=1}^N$, the marks of a (ST)CFMPP Ψ are given by $M_i = \{Z_{X_i}(t)\}_{t \in \mathcal{T}}$, $i = 1, \dots, N$, we say that Ψ has a spatio-temporal geostatistical marking.

We may also refer to this type of marking as *sampling a spatio-temporal random field at random spatial locations*.

Given (dependent) random fields $Z_j = \{Z_j(x, t)\}_{(x,t) \in \mathbb{X} \times \mathcal{T}}$, $j = 1, \dots, k$, when Ψ is multivariate, natural constructions include

- $M_i(t) = \sum_{j=1}^{k_{\mathbb{A}}} \mathbf{1}\{L_i = j\} Z_j(X_i, t)$, when $\mathbb{A} = \mathbb{A}_d$,
- $M_i(t) = \mathbf{1}_{[T_i, T_i + L_{i2})}(t) \sum_{j=1}^{k_{\mathbb{A}}} \mathbf{1}\{L_{i1} = j\} Z_j(X_i, t - T_i)$, when $\mathbb{A} = \mathbb{A}_d \times \mathbb{A}_c = \{1, \dots, k_{\mathbb{A}}\} \times [0, \infty)$.

3.2.1 Geostatistical functional data

When observations have been made of a spatio-temporal random field, at a set of fixed known locations $x_i \in \mathbb{X}$, $i = 1, \dots, n$, one often speaks of *geostatistical functional data*. The class of related data types comprise a broad family of spatially dependent functional data. For a good account of these types of data, the reader is referred to [Delicado et al., 2010, Giraldo et al., 2010, 2011].

Here, given some spatial functional process $\{Z_x : x \in \mathbb{X} \subseteq \mathbb{R}^d\}$, we assume to observe a set of functions, or rather spatially located curves, $(Z_{x_1}(t), \dots, Z_{x_n}(t))$ at locations $x_i \in \mathbb{X}$, $i = 1, \dots, n$, for $t \in \mathcal{T} = [a, b]$, which define the set of functional observations. Each function is assumed to belong to a Hilbert space

$$L_2(\mathcal{T}) = \{f : \mathcal{T} \rightarrow \mathbb{R} : \int_{\mathcal{T}} f(t)^2 dt < \infty\}$$

with the inner product $\langle f, g \rangle = \int_{\mathcal{T}} f(t)g(t)dt$. Moreover, for a fixed site x_i , the observed data is assumed to follow the model

$$Z_{x_i}(t) = \mu_{x_i}(t) + \epsilon_{x_i}(t), \quad i = 1, \dots, n,$$

where $\epsilon_{x_i}(t)$ are zero-mean residual processes and each $\mu_{x_i}(\cdot)$ is a mean function which summarises the main structure of Z_{x_i} . For each t , we assume that the process is a second-order stationary functional random process. That formally means that the expected value $\mathbb{E}[Z_x(t)] = \mu(t)$, $t \in \mathcal{T}$, $x \in \mathbb{X}$, and the variance $\text{Var}(Z_x(t)) = \sigma^2(t)$, $t \in \mathcal{T}$, $x \in \mathbb{X}$, do not depend on the spatial location. In addition, we have that

- $\text{Cov}(Z_{x_i}(t), Z_{x_j}(t)) = \mathbb{C}(h, t)$, where $h = \|x_i - x_j\|$, for all $t \in \mathcal{T}$ and all $x_i, x_j \in \mathbb{X}$.
- $\frac{1}{2} \text{Var}(Z_{x_i}(t) - Z_{x_j}(t)) = \gamma(h, t) = \gamma_{x_i x_j}(t)$, where $h = \|x_i - x_j\|$, for all $t \in \mathcal{T}$ and all $x_i, x_j \in \mathbb{X}$.

Note that under the second-order stationarity assumption one may write $\frac{1}{2} \text{Var}(Z_x(t) - Z_{x+h}(t))$, $h = \|x_i - x_j\|$, for $\frac{1}{2} \text{Var}(Z_{x_i}(t) - Z_{x_j}(t))$. However, for clarity we do prefer the more general formulation. Since we are assuming that the mean function is constant over \mathbb{X} , the function $\gamma(h, t)$, called the *variogram* of $Z_x(t)$, can be expressed by

$$\gamma(h, t) = \gamma_{x_i x_j}(t) = \frac{1}{2} \text{Var}(Z_{x_i}(t) - Z_{x_j}(t)) = \frac{1}{2} \mathbb{E} [Z_{x_i}(t) - Z_{x_j}(t)]^2.$$

By integrating this expression over \mathcal{T} , using Fubini's theorem and following [Giraldo et al., 2010], a measure of spatial variability is given by

$$\gamma(h) = \frac{1}{2} \mathbb{E} \left[\int_{\mathcal{T}} (Z_{x_i}(t) - Z_{x_j}(t))^2 dt \right] \quad (4)$$

for $x_i, x_j \in \mathbb{X}$ with $h = \|x_i - x_j\|$. This is the so-called *trace-variogram* and it is used to describe the spatial variability among functional data across an entire spatial domain. In this case, all possible location pairs are considered.

Consider now the scenario where one would perform some geostatistical analysis, such as spatio-temporal prediction [Giraldo et al., 2010], in a spatio-temporal random field when, in addition, there is uncertainty in the monitoring locations x_i , $i = 1, \dots, n$. Note that one then instead samples the random field/spatial functional process Z at locations $X_i = x_i + \varepsilon_i$, $i = 1, \dots, n$, where each ε_i follows some suitable spatial distribution. Here the CFMPP framework is the correct one since $\{X_i\}_{i=1}^n$ constitutes a spatial point process. Consequently, the above geostatistical framework could be extended to incorporate such randomness in the sampling locations. In the deterministic case, i.e. when $\varepsilon_i \equiv 0$, [Giraldo et al., 2010] proposed the estimator $\widehat{Z_{x_0}}(t) = \sum_{i=1}^n \lambda_i(t) Z_{X_i}(t) = \sum_{i=1}^n \lambda_i(t) M_i(t)$, $\lambda_i : \mathcal{T} \rightarrow \mathbb{R}$, $i = 1, \dots, n$, for the marginal random process $\{Z_{x_0}(t)\}_{t \in \mathcal{T}}$, $x_0 \in \mathbb{X}$.

Assuming that the X_i 's are in fact random, following [Giraldo et al., 2010], the associated prediction problem may be expressed as

$$\begin{aligned} \min_{\lambda_1, \dots, \lambda_n \in L_2(\mathcal{T})} \mathbb{E} \left[\int_{\mathcal{T}} (\widehat{Z_{x_0}}(t) - Z_{x_0}(t))^2 dt \right] &= \min_{\lambda_1, \dots, \lambda_n \in L_2(\mathcal{T})} \mathbb{E} \left[\int_{\mathcal{T}} \left(\sum_{i=1}^n \lambda_i(t) M_i(t) - Z_{x_0}(t) \right)^2 dt \right] \\ &= \min_{\lambda_1, \dots, \lambda_n \in L_2(\mathcal{T})} \int_{\mathbb{X}^n} \int_{\mathcal{T}} \mathbb{E} \left[\left(\sum_{i=1}^n \lambda_i(t) Z_{x_i}(t) - Z_{x_0}(t) \right)^2 \middle| X_1 = x_1, \dots, X_n = x_n \right] dt \\ &\quad \times \frac{j_n^G(x_1, \dots, x_n)}{n!} d(x_1, \dots, x_n) \end{aligned}$$

by Fubini's theorem, where $j_n^G(\cdot)$ is the n th Janossy density of the ground process $\Psi_G = \{X_i\}_{i=1}^n$ (see Section 6.4). Hence, one obtains a geostatistical analysis based on CFMPPs.

3.3 LISA and LISTA functions

In the context of spatial point processes, [Collins and Cressie, 2001] developed exploratory data analytic tools, in terms of Local Indicators of Spatial Association (LISA) functions based on the product density, to examine individual points in the point pattern in terms of how they relate to their neighbouring points. For each point X_i of the point process/pattern we can attach to it a LISA function $M_i(h)$, $h = \|X_i - x\| \geq 0$, $x \in \mathbb{X}$, which determines the local spatial structure associated to each event of the pattern. These functions can be regarded as functional marks [Mateu et al., 2007]. To perform statistical inference, which is needed for example in testing for local clustering, [Collins and Cressie, 2001] developed closed form expressions of the auto-covariance and cross-covariance between any two such functions. These covariance structures are complicated to work with as they live in high-dimensional spaces.

If the ground point pattern evolves in time, i.e. if we have a spatio-temporal point pattern, then we can extend the ideas of LISA functions to incorporate time in their structure. In this case, local versions of spatio-temporal product densities provide the concept of LISTA surfaces [Rodríguez-Cortés et al., 2014]. Attached to each spatio-temporal location (X_i, T_i) we now have surfaces $M_i(x, t)$, $(x, t) \in \mathbb{X} \times \mathbb{T}$ (i.e. with dimensions space and time). When we assume that $M_i(x, t) = M_i(h)$, $h = d_{\mathbb{G}}((X_i, T_i), (x, t))$, these surfaces can again be regarded as functional marks. The LISTA surfaces provide information on the local spatio-temporal structure of the point pattern.

3.4 The (stochastic) growth-interaction process

One of the models which has given rise to a substantial part of the ideas underlying the construction of STCFMPPs is the *growth-interaction process*. It has been extensively studied in a series of papers (see e.g. [Comas, 2009, Comas et al., 2011, Cronie, 2012, Cronie and Särkkä, 2011, Cronie et al., 2013, Renshaw and Comas, 2009, Renshaw et al., 2009, Renshaw and Särkkä, 2001, Särkkä and Renshaw, 2006]), mainly within the forestry context. However, its representation as a functional marked point process has only been noted in [Comas et al., 2011, Cronie, 2012].

It is a STCFMPP for which the ground process Ψ_G is generated by a spatial birth-death process, which has Poisson arrivals T_i , with intensity $\alpha > 0$, and uniformly distributed spatial locations X_i . Furthermore, the auxiliary marks are the associated holding times L_i , which are independently

$Exp(\mu)$ -distributed, $\mu > 0$, and, conditionally on the previous components, the functional marks are given by a system of ordinary differential equations,

$$\frac{dM_i(t)}{dt} = g(M_i(t); \theta) - \sum_{(X_j, T_j, L_j, M_j(t)) \in \Psi, j \neq i} h((X_i, T_i, L_i, M_i(t)), (X_j, T_j, L_j, M_j(t)); \theta),$$

$i = 1, \dots, N$, where $t \in \text{supp}(M_i) = [T_i, D_i)$, $D_i = (T_i + L_i) \wedge T^*$. Here $g(\cdot)$ represents the individual growth of the i th *individual*, in absence of spatial interaction with other individuals, and $h((X_i, T_i, L_i, M_i(t)), (X_j, T_j, L_j, M_j(t)); \theta)$ the amount of spatial interaction to which individual i is subjected by individual j during $[t, t + dt]$.

As can be found in the above mentioned references, the usual application of this model is the modelling of the collective development of trees in a forest stand; X_i is the location of the i th tree, T_i is its birth time, D_i its death time, and $M_i(t)$ its radius (at breast height) at time t .

As one may argue that this approach does not sufficiently incorporate individual growth features in the radial growth, [Cronie, 2012] suggested that a scaled white noise processes should be added to each functional mark equation, i.e.

$$dM_i^*(t) = dM_i(t) + \sigma(M_i(t); \theta) dW_i(t),$$

where $W_1(t), \dots, W_N(t)$, are independent standard Brownian motions and $\sigma(\cdot)$ is some suitable diffusion coefficient. Here the noise would represent measurement errors and give rise to individual growth deviations. The resulting stochastic differential equation marked point process, the *stochastic growth-interaction process*, was then studied in the simplified case where the spatial interaction is negligible, i.e. $h(\cdot) \equiv 0$.

3.5 Applications

Besides the applications mentioned previously, we here give a list of further possible applications of (ST)CFMPPs, providing a wide scope of the current framework.

1. *Modelling nerve fibres*: X_i gives the location of the root of the nerve. A mark M_i (here continuous) provides the shape of the actual nerve fibre and the related auxiliary variable is given by $L_i = (L_{i1}, L_{i2}) \in \mathbb{A}_c = [0, \infty) \times [0, 2\pi)$, where $\text{supp}(M_i) = [0, L_{i1})$ and L_{i2} represents a random rotation angle of M_i , which gives the direction of the fibre.
2. *Spread of pollutant*: X_i is the pollution location, $M_i(h)$ gives us the ground concentration of the contaminant at distance $h = \|X_i - x\|$, $x \in \mathbb{X}$, from X_i .
3. *Modelling tumours*: \mathbb{X} represents (a region in) the human body, X_i is the location of the i th tumour and $M_i(t)$ its approximate radius at time t .
4. *Disease incidences in epidemics*: Each $M_i(t)$ is a stochastic process with piecewise constant sample paths (e.g. a Poisson process), which counts the number of incidences having occurred by time t at epidemic centre X_i .
5. *Population growth*: X_i is the location of a village/town/city, T_i the time point at which it was founded and $M_i(t)$ its total population at time t .

6. *Mobile communication:* Letting $\mathbb{T} = \mathcal{T}$, consider a STCFMPP Ψ where each X_i represents the location of a cellphone caller who makes a call at time T_i , which lasts for $L_i \in \mathbb{A}_c = [0, \infty)$ time units, i.e. the call ends at time $D_i = (T_i + L_i) \wedge T^*$. Then the function $M_i(t) = \mathbf{1}_{[B_i, D_i)}(t)$ represents the phone call in question and the total load of a server/antenna located at $s \in \mathbb{X}$, which has spatial reach within the region $B \subseteq \mathbb{X}$, $s \in B$, is $N_s(t) = \sum_{i=1}^N \mathbf{1}_B(X_i) M_i(t)$. Assuming that the server has capacity $c_s(t)$ at time t , it breaks down if $\sup_{t \in \mathcal{T}} c_s(t) - N_s(t) \leq 0$. Note the connection with [Baum and Kalashnikov, 2001].

An extension here could be to let $M_i(t) = \xi_i \mathbf{1}_{[T_i, D_i)}(t)$ for some random quantity $\xi_i = \xi_i(X_i, T_i, L_i)$, which represents the specific load that call i puts on the network.

4 Point process characteristics of (ST)CFMPPs

For a wide range of summary statistics, the core elements are the product densities and the intensity function(al). We here derive these for (ST)CFMPPs under a few usual assumptions. In addition, we define two further, highly important, building blocks for different statistics for point processes; the Palm measures and the Papangelou conditional intensities.

Recall that for both types of processes the mark space is given by $\mathbb{M} = \mathbb{A} \times \mathbb{F}$ and to provide a general notation, which may be used to describe both CFMPPs and STCFMPPs, we write $\Psi = \sum_{(g,m) \in \Psi} \delta_{(g,m)} = \sum_{(g,l,f) \in \Psi} \delta_{(g,l,f)}$, where $x = g \in \mathbb{G} = \mathbb{X}$ in the CFMPP case and $(x, t) = g \in \mathbb{G} = \mathbb{X} \times \mathbb{T}$ in the STCFMPP case.

Throughout, for different measures constructed, we will use the following measure extension approach. When some set function $\mu(A)$ is defined for the bounded Borel sets A in some Borel space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, by assuming that $\mu(\cdot)$ is locally finite, $\mu(\cdot)$ becomes a finite measure on the ring of bounded Borel sets. Hereby one may extend μ to a measure on the whole σ -algebra $\mathcal{B}(\mathcal{X})$ (see e.g. [Halmos, 1974, Theorem A, p. 54]).

4.1 Product densities and intensity functionals

We first consider the (factorial) moment measures and the product densities of a (ST)CFMPP Ψ . The construction of the product densities paves the way for the construction of certain likelihood functions and summary statistics.

We start by defining the *factorial moment measures*.

Definition 8. For any $n \geq 1$ and bounded $A_1, \dots, A_n = (G_1 \times H_1), \dots, (G_n \times H_n) \in \mathcal{B}(\mathbb{Y})$, define

$$\alpha^{(n)}(A_1 \times \dots \times A_n) = \mathbb{E} \left[\sum_{(g_1, l_1, f_1), \dots, (g_n, l_n, f_n) \in \Psi}^{\neq} \prod_{i=1}^n \mathbf{1}\{(g_i, l_i, f_i) \in A_i\} \right],$$

where \sum^{\neq} denotes a sum over distinct elements. Note that $\alpha^{(n)}$ may be extended to a measure on the n -fold product σ -algebra $\mathcal{B}(\mathbb{Y})^n = \bigotimes_{i=1}^n \mathcal{B}(\mathbb{Y})$, the n th factorial moment measure.

Note that in the STCFMPP setting, $G_i = B_i \times C_i \in \mathcal{B}(\mathbb{X} \times \mathbb{T})$ and $H_i = D_i \times E_i \in \mathcal{B}(\mathbb{A} \times \mathbb{F})$, $i = 1, \dots, n$.

Recall next the reference measure ν in (2) and assume that $\alpha^{(n)} \ll \nu^n$, i.e. that $\alpha^{(n)}$ is absolute continuous with respect to the n -fold product measure of the reference measure ν with itself. This leads to the definition of (functional) product densities.

Definition 9. The permutation invariant Radon-Nikodym derivatives (measurable functionals) $\rho^{(n)}$, which are defined by the integral formula

$$\begin{aligned} \mathbb{E} \left[\sum_{(g_1, l_1, f_1), \dots, (g_n, l_n, f_n) \in \Psi}^{\neq} h((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) \right] &= \\ &= \int_{\mathbb{Y}^n} h((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) \alpha^{(n)}(d(g_1, l_1, f_1) \times \dots \times d(g_n, l_n, f_n)) \\ &= \int_{\mathbb{Y}} \dots \int_{\mathbb{Y}} h((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) \rho^{(n)}((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) \prod_{i=1}^n \nu(d(g_i, l_i, f_i)) \end{aligned} \quad (5)$$

for any measurable functional $h : \mathbb{Y} \rightarrow [0, \infty)$, are referred to as the n th product densities.

Note here that (5) is referred to as the *Campbell theorem*. Furthermore, $\rho^{(n)}$ is partly a (functional) density on \mathbb{F} , as discussed in Section 2.4. We note that, heuristically, the interpretation of $\rho^{(n)}((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) \prod_{i=1}^n \nu(d(g_i, l_i, f_i))$ is the probability of finding ground process points in the infinitesimal regions $dg_1, \dots, dg_n \subseteq \mathbb{G}$, with associated marks in the infinitesimal regions $d(l_1, f_1), \dots, d(l_n, f_n) \subseteq \mathbb{A} \times \mathbb{F}$.

Similarly to $\alpha^{(n)}$ we may also define the n th moment measures

$$\mu^{(n)}(A_1 \times \dots \times A_n) = \mathbb{E}[\Psi(A_1) \dots \Psi(A_n)], \quad A_1, \dots, A_n \in \mathcal{B}(\mathbb{Y}), n \geq 1.$$

Furthermore, since the *intensity measure* μ of a simple point process coincides with its first moment measure and its first factorial moment measure, i.e.

$$\mu(A) = \mathbb{E}[\Psi(A)] = \alpha^{(1)}(A) = \int_A \rho^{(1)}(g, l, f) \nu(d(g, l, f)), \quad A \in \mathcal{B}(\mathbb{Y}), \quad (6)$$

we obtain a definition of intensity functionals for (ST)CFMPPs.

Definition 10. The intensity functional of a (ST)CFMPP Ψ is given by

$$\lambda(g, l, f) = \rho^{(1)}(g, l, f).$$

Note that we use the term *functional* since the mapping $\lambda : \mathbb{G} \times \mathbb{A} \times \mathbb{F} \rightarrow [0, \infty)$ takes a càdlàg function $f \in \mathbb{F}$ as one of its arguments.

Turning to the ground process Ψ_G , through $\alpha^{(n)}$ we may define the n th ground factorial moment measure $\alpha_G^{(n)}$ and its Radon-Nikodym derivative $\rho_G^{(n)}$ with respect to ℓ^n , the n th ground product density.

Definition 11. When the measure

$$\alpha_G^{(n)}(G_1 \times \dots \times G_n) = \alpha^{(n)}((G_1 \times \mathbb{M}) \times \dots \times (G_n \times \mathbb{M})),$$

is assumed to be locally finite, it becomes the n th factorial moment measure of the ground process Ψ_G . When $\alpha_G^{(n)} \ll \ell^n$, we refer to the corresponding Radon-Nikodym derivative $\rho_G^{(n)}$ as the n th ground product density.

The ground intensity measure is given by

$$\mu_G(G) = \mathbb{E}[\Psi_G(G)] = \alpha_G^{(1)}(G) = \int_G \lambda_G(g) dg, \quad G \in \mathcal{B}(\mathbb{G}), \quad (7)$$

where $\lambda_G(g) = \rho_G^{(1)}(g)$ is referred to as the ground intensity function.

We note that here $\alpha_G^{(n)}(d(g_1, \dots, g_n)) = \rho_G^{(n)}(g_1, \dots, g_n) dg_1 \dots dg_n$ is interpreted as the infinitesimal probability of finding points of Ψ_G in $dg_1, \dots, dg_n \subseteq \mathbb{G}$.

4.1.1 Product densities in terms of conditional mark distribution densities

The following observation regarding regular probabilities on the product spaces $\mathcal{B}((\mathbb{A} \times \mathbb{F})^n)$, $n \geq 1$, which we state in the form of a lemma, will be exploited frequently.

Lemma 1. *For any $n \geq 1$, assuming that there exists some family*

$$P^{\mathbb{M},n} = \{P_{g_1,\dots,g_n}^{\mathbb{M}}(H_1 \times \dots \times H_n) : (g_1, \dots, g_n) \in \mathbb{G}^n, H_1 \times \dots \times H_n \in \mathcal{B}(\mathbb{M}^n)\}$$

of regular probabilities, it follows that for any $(g_1, \dots, g_n) \in \mathbb{G}^n$ and any $(D_1 \times E_1), \dots, (D_n \times E_n) \in \mathcal{B}(\mathbb{A} \times \mathbb{F})$,

$$\begin{aligned} P_{g_1,\dots,g_n}^{\mathbb{M}}((D_1 \times E_1) \times \dots \times (D_n \times E_n)) &= \\ &= \int_{D_1 \times \dots \times D_n} P_{(g_1,l_1),\dots,(g_n,l_n)}^{\mathbb{F}}(E_1 \times \dots \times E_n) P_{g_1,\dots,g_n}^{\mathbb{A}}(d(l_1, \dots, l_n)), \end{aligned}$$

where $P_{g_1,\dots,g_n}^{\mathbb{A}}(\cdot \times \dots \times \cdot) = P_{g_1,\dots,g_n}^{\mathbb{M}}((\cdot \times \mathbb{F}) \times \dots \times (\cdot \times \mathbb{F}))$ and $P_{(g_1,l_1),\dots,(g_n,l_n)}^{\mathbb{F}}(\cdot \times \dots \times \cdot)$, $l_1, \dots, l_n \in \mathbb{A}$, are families of regular probabilities on $(\mathbb{A}^n, \mathcal{B}(\mathbb{A}^n))$ and $(\mathbb{F}^n, \mathcal{B}(\mathbb{F}^n))$, respectively.

Assume further that each $P_{g_1,\dots,g_n}^{\mathbb{M}}(\cdot)$ has a density $f_{g_1,\dots,g_n}^{\mathbb{M}}(\cdot)$ on $(\mathbb{A} \times \mathbb{F})^n$ with respect to $(\nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}})^n = \nu_{\mathbb{A}}^n \otimes \nu_{\mathbb{F}}^n$. Then it follows that $P_{g_1,\dots,g_n}^{\mathbb{A}}(\cdot) \ll \nu_{\mathbb{A}}^n$ and $P_{(g_1,l_1),\dots,(g_n,l_n)}^{\mathbb{F}}(\cdot) \ll \nu_{\mathbb{F}}^n$ for each of the regular probabilities on $\mathcal{B}(\mathbb{A}^n)$ and $\mathcal{B}(\mathbb{F}^n)$, whereby there exist associated densities

$$f_{g_1,\dots,g_n}^{\mathbb{A}} = \{f_{g_1,\dots,g_n}^{\mathbb{A}}(\cdot) : g_1, \dots, g_n \in \mathbb{G}\}, \quad (8)$$

$$f_{(g_1,l_1),\dots,(g_n,l_n)}^{\mathbb{F}} = \{f_{(g_1,l_1),\dots,(g_n,l_n)}^{\mathbb{F}}(\cdot) : ((g_1, l_1), \dots, (g_n, l_n)) \in (\mathbb{G} \times \mathbb{A})^n\}, \quad (9)$$

on \mathbb{A}^n and \mathbb{F}^n , such that

$$f_{g_1,\dots,g_n}^{\mathbb{M}}((l_1, f_1), \dots, (l_n, f_n)) = f_{(g_1,l_1),\dots,(g_n,l_n)}^{\mathbb{F}}(f_1, \dots, f_n) f_{g_1,\dots,g_n}^{\mathbb{A}}(l_1, \dots, l_n)$$

for all $(g_1, \dots, g_n) \in \mathbb{G}^n$ and almost all $(l_1, f_1), \dots, (l_n, f_n) \in \mathbb{A} \times \mathbb{F}$.

Proof. The existence of all $P_{g_1,\dots,g_n}^{\mathbb{A}}(\cdot)$ and $P_{(g_1,l_1),\dots,(g_n,l_n)}^{\mathbb{F}}(\cdot)$, $(g_1, l_1), \dots, (g_n, l_n) \in \mathbb{G} \times \mathbb{A}$, is a direct consequence of [Daley and Vere-Jones, 2003, Proposition A1.5.III], since \mathbb{A}^n and \mathbb{F}^n are csm spaces for any $n \geq 1$.

Assuming that $P_{g_1,\dots,g_n}^{\mathbb{M}}(\cdot) \ll \nu_{\mathbb{A}}^n \otimes \nu_{\mathbb{F}}^n$ results in

$$\begin{aligned} P_{g_1,\dots,g_n}^{\mathbb{M}}((D_1 \times E_1) \times \dots \times (D_n \times E_n)) &= \\ &= \int_{(D_1 \times E_1) \times \dots \times (D_n \times E_n)} f_{g_1,\dots,g_n}^{\mathbb{M}}((l_1, f_1), \dots, (l_n, f_n)) \prod_{i=1}^n \nu_{\mathbb{A}}(dl_i) \nu_{\mathbb{F}}(df_i). \end{aligned}$$

On the other hand, from the absolute continuity it clearly follows that $P_{g_1,\dots,g_n}^{\mathbb{A}}(\cdot) \ll \nu_{\mathbb{A}}^n$ and $P_{(g_1,l_1),\dots,(g_n,l_n)}^{\mathbb{F}}(\cdot) \ll \nu_{\mathbb{F}}^n$, whereby

$$\begin{aligned} P_{g_1,\dots,g_n}^{\mathbb{M}}((D_1 \times E_1) \times \dots \times (D_n \times E_n)) &= \\ &= \int_{(D_1 \times E_1) \times \dots \times (D_n \times E_n)} f_{(g_1,l_1),\dots,(g_n,l_n)}^{\mathbb{F}}(f_1, \dots, f_n) f_{g_1,\dots,g_n}^{\mathbb{A}}(l_1, \dots, l_n) \prod_{i=1}^n \nu_{\mathbb{A}}(dl_i) \nu_{\mathbb{F}}(df_i). \end{aligned}$$

The two integrands are equal a.e..

□

Next, through the above lemma, one finds that when the product densities exist, they may be expressed through conditional probability densities on the different mark spaces.

Proposition 1. *For any $n \geq 1$, given that the n th ground factorial moment measure $\alpha_G^{(n)}$ exists, there exist regular probabilities*

$$\begin{aligned} P^{\mathbb{M},n} &= \{P_{g_1,\dots,g_n}^{\mathbb{M}}(H) : (g_1, \dots, g_n) \in \mathbb{G}^n, H \in \mathcal{B}(\mathbb{M}^n)\}, \\ P^{\mathbb{A},n} &= \{P_{g_1,\dots,g_n}^{\mathbb{A}}(D) : g_1, \dots, g_n \in \mathbb{G}, D \in \mathcal{B}(\mathbb{A}^n)\} \ll \nu_{\mathbb{A}}^n, \end{aligned} \quad (10)$$

$$P^{\mathbb{F},n} = \{P_{(g_1,l_1),\dots,(g_n,l_n)}^{\mathbb{F}}(E) : (g_1, l_1), \dots, (g_n, l_n) \in \mathbb{G} \times \mathbb{A}, E \in \mathcal{B}(\mathbb{F}^n)\} \ll \nu_{\mathbb{F}}^n, \quad (11)$$

as indicated in Lemma 1. When $\rho^{(n)}$ exists, we can find conditional densities (8) and (9) such that

$$\begin{aligned} \rho^{(n)}((x_1, l_1, f_1), \dots, (x_n, l_n, f_n)) &= \\ &= f_{(x_1,l_1),\dots,(x_n,l_n)}^{\mathbb{F}}(f_1, \dots, f_n) f_{x_1,\dots,x_n}^{\mathbb{A}}(l_1, \dots, l_n) \rho_G^{(n)}(x_1, \dots, x_n) \end{aligned} \quad (12)$$

when Ψ is a CFMPP and

$$\begin{aligned} \rho^{(n)}((x_1, t_1, l_1, f_1), \dots, (x_n, t_n, l_n, f_n)) &= \\ &= f_{(x_1,t_1,l_1),\dots,(x_n,t_n,l_n)}^{\mathbb{F}}(f_1, \dots, f_n) f_{(x_1,t_1),\dots,(x_n,t_n)}^{\mathbb{A}}(l_1, \dots, l_n) \rho_G^{(n)}((x_1, t_1), \dots, (x_n, t_n)) \end{aligned} \quad (13)$$

when Ψ is a STCFMPP. Hereby the intensity functional $\lambda(g, l, f)$ turns into

$$\begin{aligned} \lambda(x, l, f) &= f_{(x,l)}^{\mathbb{F}}(f) f_x^{\mathbb{A}}(l) \lambda_G(x) && \text{if } \Psi \text{ is a CFMPP,} \\ \lambda(x, t, l, f) &= f_{(x,t,l)}^{\mathbb{F}}(f) f_{(x,t)}^{\mathbb{A}}(l) \lambda_G(x, t) && \text{if } \Psi \text{ is a STCFMPP.} \end{aligned}$$

Proof. For fixed $H_1, \dots, H_n \in \mathcal{B}(\mathbb{M})$ we have that $\alpha^{(n)}((\cdot \times H_1) \times \dots \times (\cdot \times H_n)) \ll \alpha_G^{(n)}$ and (since the underlying spaces are Borelian) a regular family $P^{\mathbb{M},n}$ of conditional regular probabilities on $\mathcal{B}(\mathbb{M})^n = \mathcal{B}(\mathbb{A} \times \mathbb{F})^n$ exists, as indicated in Lemma 1 (see e.g. [Stoyan et al., 1995]). An application of Lemma 1 further gives us $P^{\mathbb{A},n}$ and $P^{\mathbb{F},n}$.

By the existence of $\rho^{(n)}$, the underlying absolute continuity with respect to $\nu^n = (\ell \otimes \nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}})^n$ implies that $\alpha_G^{(n)} \ll \ell^n$, whence

$$\alpha_G^{(n)}(G_1 \times \dots \times G_n) = \int_{G_1} \dots \int_{G_n} \rho_G^{(n)}(g_1, \dots, g_n) dg_1 \dots dg_n$$

for any $G_1, \dots, G_n \in \mathcal{B}(\mathbb{G})$. In addition, it follows that each member of $P^{\mathbb{M},n}$ is absolute continuous with respect to $(\nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}})^n$, whereby an application of Lemma 1 gives us the specific structure of the product densities. □

The family $P^{\mathbb{M},n}$ is often referred to as the n -point mark distributions. We see here that each $P_{(g_1,l_1),\dots,(g_n,l_n)}^{\mathbb{F}}(\cdot)$ in expression (11) is the distribution of some (mark) càdlàg stochastic process $\{(M_1(t), \dots, M_n(t))\}_{t \in \mathcal{T}}$ on $(\mathbb{F}^n, \mathcal{B}(\mathbb{F}^n))$, which is absolutely continuous with respect to the reference measure $\nu_{\mathbb{F}}^n$ (i.e. the distribution of an n -dimensional version of the reference process $X_{\mathbb{F}}$), with density (9). Furthermore, consider a STCFMPP Ψ for which we want to have $\text{supp}(M_i) = S_i$ for some $S_i = S_i(X_i, T_i, L_i) \subseteq \mathcal{T}$, $i = 1, \dots, N$. Conditioning on Ψ_G , the auxiliary marks and $(X_i, T_i, L_i) = (x_i, t_i, l_i)$, $i = 1, \dots, n$, it then clearly follows that $f_{(x_1,t_1,l_1),\dots,(x_n,t_n,l_n)}^{\mathbb{F}}(f_1, \dots, f_n) = 0$ if, for any $i = 1, \dots, n$, it holds that $f_i \in \mathbb{F} \setminus \{f \in \mathbb{F} : \text{supp}(f) = S_i\}$. Note that often a natural choice for the supports is $S_i = [T_i, (T_i + L_i) \wedge T^*)$, $L_i \geq 0$, $i = 1, \dots, N$.

4.1.2 Pair correlation functionals

Pair correlation functions are valuable tools for studying second order dependence properties of spatial(-temporal) point processes and, as we shall see, they play a similar role for (ST)CFMPPs.

In light of Proposition 1, when λ and $\rho^{(2)}$ exist, we may define the *pair correlation functional* of a (ST)CFMPP Ψ by

$$\begin{aligned} g_\Psi((g_1, l_1, f_1), (g_2, l_2, f_2)) &= \frac{\rho^{(2)}((g_1, l_1, f_1), (g_2, l_2, f_2))}{\lambda(g_1, l_1, f_1)\lambda(g_2, l_2, f_2)} \\ &= \frac{f_{(g_1, l_1), (g_2, l_2)}^{\mathbb{F}}(f_1, f_2) f_{g_1, g_2}^{\mathbb{A}}(l_1, l_2)}{f_{(g_1, l_1)}^{\mathbb{F}}(f_1) f_{g_1}^{\mathbb{A}}(l_1) f_{(g_2, l_2)}^{\mathbb{F}}(f_2) f_{g_2}^{\mathbb{A}}(l_2)} \frac{\rho_G^{(2)}(g_1, g_2)}{\lambda_G(g_1)\lambda_G(g_2)} \\ &= \frac{f_{(g_1, l_1), (g_2, l_2)}^{\mathbb{F}}(f_1, f_2) f_{g_1, g_2}^{\mathbb{A}}(l_1, l_2)}{f_{(g_1, l_1)}^{\mathbb{F}}(f_1) f_{g_1}^{\mathbb{A}}(l_1) f_{(g_2, l_2)}^{\mathbb{F}}(f_2) f_{g_2}^{\mathbb{A}}(l_2)} g_G(g_1, g_2). \end{aligned} \quad (14)$$

We refer to $g_G(g_1, g_2)$ as the *ground pair correlation function*. Note that when $\mathbb{G} = \mathbb{X}$, $g_G(\cdot)$ is the usual pair correlation function, as can be found in e.g. [Baddeley et al., 2000, Stoyan et al., 1995], and when $\mathbb{G} = \mathbb{X} \times \mathbb{T}$, $g_G(\cdot)$ is a spatio-temporal pair correlation function, as defined in [Cronie and van Lieshout, 2013, Gabriel and Diggle, 2009, Møller and Ghorbani, 2012].

4.2 Campbell and Palm measures

We next turn to the Palm measures of a (ST)CFMPP Ψ and, as usual, they are defined via the Campbell measures.

For any $R \in \Sigma_{\mathcal{N}_{\mathbb{Y}}}$ and bounded $A = G \times H \in \mathcal{B}(\mathbb{Y})$, we define

$$\mathcal{C}^!(A \times R) = \mathcal{C}^!(G \times H \times R) = \mathbb{E} \left[\sum_{(g, l, f) \in \Psi \cap (G \times H)} \mathbf{1}\{\Psi \setminus \{(g, l, f)\} \in R\} \right],$$

which we can extend to a measure on $\mathcal{B}(\mathbb{Y}) \otimes \Sigma_{\mathcal{N}_{\mathbb{Y}}}$. We refer to this measure as the *reduced Campbell measure* and since $\mu(A) = \mathcal{C}^!(A \times \mathcal{N}_{\mathbb{Y}})$ we see that $\mathcal{C}^! \ll \mu$, with Radon-Nikodym derivative $P^{!(g, l, f)}$. Hereby

$$\begin{aligned} \mathcal{C}^!(A \times R) &= \int_{G \times H} \mathcal{C}^!(d(g, l, f) \times R) = \int_G \int_H P^{!(g, l, f)}(R) \mu(d(g, l, f)) \\ &= \int_G \int_H P^{!(g, l, f)}(R) \lambda(g, l, f) \nu(d(g, l, f)) \end{aligned}$$

Definition 12. *It is possible to choose a regular version of the family*

$$\{P^{!(g, l, f)}(R) : (g, l, f) \in \mathbb{Y}, R \in \Sigma_{\mathcal{N}_{\mathbb{Y}}}\},$$

the reduced Palm measures, such that $P^{!(g, l, f)}(R)$ is a measurable functional for fixed $R \in \Sigma_{\mathcal{N}_{\mathbb{Y}}}$ and a measure on $\Sigma_{\mathcal{N}_{\mathbb{Y}}}$ for fixed $(g, l, f) \in \mathbb{Y}$.

We often write $\mathbb{P}^{!(g, l, f)}$ for the $P^{!(g, l, f)}$ -reversely induced probability measure on (Ω, \mathcal{F}) and $\mathbb{E}^{!(g, l, f)}[\cdot]$ for the associated expectation. Heuristically we interpret $P^{!(g, l, f)}(R)$ as the conditional probability of the event $\{\Psi \setminus \{(g, l, f)\} \in R\}$, given $\Psi \cap \{(g, l, f)\} \neq \emptyset$.

The (non-reduced) *Palm measures* may be defined through their reduced counterparts as

$$P^{(g,l,f)}(R) = P^{!(g,l,f)}(\{\varphi \in \mathcal{N}_{\mathbb{Y}} : \varphi + \delta_{(g,l,f)} \in R\}), \quad R \in \Sigma_{\mathcal{N}_{\mathbb{Y}}},$$

and in relation hereto, the *Campbell measure* of Ψ is given by $\mathcal{C}(A \times R) = \mathbb{E}[\Psi(A)\mathbf{1}_R(\Psi)]$. Note that when Ψ is stationary, $P^{!(g,l,f)}(\cdot) = P^{!(0,l,f)}(\cdot)$ and $P^{(g,l,f)}(\cdot) = P^{(0,l,f)}(\cdot)$ for all $(g,l,f) \in \mathbb{Y}$.

In accordance with [Van Lieshout, 2006], we also define the *H-marked reduced Palm measure*

$$P_H^{!g}(R) = \frac{1}{[\nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}}](H)} \int_H P^{!(g,l,f)}(R) [\nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}}](d(l,f)), \quad H \times R \in \mathcal{B}(\mathbb{M}) \otimes \Sigma_{\mathcal{N}_{\mathbb{Y}}}, \quad (15)$$

and we denote the corresponding expectation by $\mathbb{E}_H^{!g}[\cdot]$. Note that the choice of name for $P_H^{!g}(\cdot)$ may be a bit misleading since it does not necessarily define a Palm distribution in the true sense. One case, however, where this is indeed the case is when Ψ has a common mark distribution (see Section 5.1).

We next consider the *Campbell-Mecke formula* and the *reduced Campbell-Mecke formula*, which are central tools in the theory of point processes.

Theorem 1. *For any measurable functional $h : \mathbb{G} \times \mathbb{A} \times \mathbb{F} \times \mathcal{N}_{\mathbb{Y}} \rightarrow [0, \infty)$, the reduced Campbell-Mecke formula is given by*

$$\begin{aligned} \mathbb{E} \left[\sum_{(g,l,f) \in \Psi} h(g,l,f, \Psi \setminus \{(g,l,f)\}) \right] &= \int_{\mathbb{G} \times \mathbb{M}} \int_{\mathcal{N}_{\mathbb{Y}}} h(g,l,f,\varphi) P^{!(g,l,f)}(d\varphi) \mu(d(g,l,f)) \\ &= \int_{\mathbb{G} \times \mathbb{M}} \mathbb{E}^{!(g,l,f)} [h(g,l,f, \Psi)] \lambda(g,l,f) \nu(d(g,l,f)) \end{aligned}$$

and the Campbell-Mecke formula is given by

$$\mathbb{E} \left[\sum_{(g,l,f) \in \Psi} h(g,l,f, \Psi) \right] = \int_{\mathbb{G} \times \mathbb{M}} \mathbb{E}^{(g,l,f)} [h(g,l,f, \Psi + \delta_{(g,l,f)})] \lambda(g,l,f) \nu(d(g,l,f))$$

(with the left hand sides being infinite if and only if the right hand sides are infinite).

Proof. The proof is standard (see e.g. [Van Lieshout, 2000, Theorem 1.5]). First we show that the above expressions hold for $h(g,l,f,\varphi) = \mathbf{1}\{(g,l,f) \in G, \varphi \in R\}$, $G \times R \in \mathcal{B}(\mathbb{G}) \times \Sigma_{\mathcal{N}_{\mathbb{Y}}}$. The rest of the proof follows from standard arguments using linear combinations and monotone limits. \square

4.3 Marked Papangelou conditional intensities

Another important set of point process statistics, here defined in the context of STCFMPPs, are the (marked) Papangelou kernels and the (marked) Papangelou conditional intensities. We will see that the former, in particular, will play an important role in the statistical inference framework constructed here.

By assuming that, for fixed bounded $A = (B \times C) \times (D \times E) \in \mathcal{B}(\mathbb{Y})$, $\mathcal{C}^!(A \times \cdot)$ is absolutely continuous with respect to the distribution $P(\cdot)$ of Ψ on $\Sigma_{\mathcal{N}_{\mathbb{Y}}}$, we find that $\mathcal{C}^!(A \times R) = \int_R \Lambda(A; \varphi) P(d\varphi)$ for some kernel $\{\Lambda(A; \varphi) : A \in \mathcal{B}(\mathbb{Y}), \varphi \in \mathcal{N}_{\mathbb{Y}}\}$. This kernel may be extended to the $\mathcal{B}(\mathbb{Y})$ -setting,

the so-called *Papangelou kernel*. If we assume that $\Lambda(\cdot; \varphi) \ll \ell \otimes \nu_{\mathbb{A}}$, through Fubini's theorem we obtain

$$\mathcal{C}^!(A \times R) = \int_R \int_{(B \times C) \times D} \lambda_E(g, l; \varphi) dg \nu_{\mathbb{A}}(dl) P(d\varphi) = \int_{(B \times C) \times D} \mathbb{E}[\mathbf{1}_R(\Psi) \lambda_E(g, l; \Psi)] dg \nu_{\mathbb{A}}(dl)$$

for $R \in \Sigma_{\mathcal{N}_{\mathbb{Y}}}$.

Definition 13. We refer to $\lambda_E(g, l; \Psi)$, $E \in \mathcal{B}(\mathbb{F})$, as the E -functional marked Papangelou conditional intensity.

Here $\lambda_E(g, l; \Psi) dg \nu_{\mathbb{A}}(dl)$ may be viewed as the conditional probability of finding a point of Ψ located in the infinitesimal set $dg \subseteq \mathbb{G}$ with mark in $dl \times E \in \mathcal{B}(\mathbb{M})$, given all points of Ψ located outside dg .

By additionally assuming that all functional marked Papangelou conditional intensities are absolutely continuous with respect to $\nu_{\mathbb{F}}$, i.e. assuming that $\mathcal{C}^!(A \times \cdot) \ll \nu$, we obtain as Radon-Nikodym derivatives the classical *Papangelou conditional intensities* $\lambda(g, l, f; \Psi)$ which may also be defined, in integral terms, as the non-negative measurable functionals satisfying the *Georgii-Nguyen-Zessin* formula, i.e.

$$\mathbb{E} \left[\sum_{(z, l, f) \in Y} h(g, l, f, \Psi \setminus \{(g, l, f)\}) \right] = \int_{\mathbb{G} \times \mathbb{M}} \mathbb{E} [h(g, l, f, \Psi) \lambda(g, l, f; \Psi)] \nu(d(g, l, f))$$

for any non-negative measurable functional $h : (\mathbb{G} \times \mathbb{M}) \times \mathcal{N}_{\mathbb{Y}} \rightarrow [0, \infty)$. Here, heuristically, $\lambda(g, l, f; \Psi) \nu(d(g, l, f))$ is interpreted as the conditional probability of finding a point of Ψ in the infinitesimal region $d(g, l, f) \subseteq \mathbb{G} \times \mathbb{M}$, given the configuration elsewhere, $\Psi \cap (d(g, l, f))^c$. Note here that

$$\mathbb{E} [h(g, l, f, \Psi) \lambda(g, l, f; \Psi)] \stackrel{a.e.}{=} \mathbb{E}^{!(g, l, f)} [h(g, l, f, \Psi)] \lambda(g, l, f) \quad (16)$$

and $\mathbb{E}[\lambda(g, l, f; \Psi)] = \lambda(g, l, f)$.

5 Mark structures

We now provide some examples of explicit mark structures, which may be considered when constructing (ST)CFMPP models. Note that

- multivariate (ST)CFMPPs (Sections 2.6, 2.7 and 5.3),
- usual marked (spatio-temporal) point processes (Section 3.1),
- spatio-temporal geostatistical marking (Section 3.2),

have already been covered previously in the text.

5.1 Independent marks and common marginal mark distributions

In light of Proposition 1, whenever

$$P_{(g,l)}^{\mathbb{F}}(E) = \nu_{\mathbb{F}}(E), \quad E \in \mathcal{B}(\mathbb{F}),$$

so that $f_{(g,l)}^{\mathbb{F}}(\cdot) \equiv 1$, we say that Ψ has a *common (marginal functional) mark distribution*. Recalling the reference process $X_{\mathbb{F}}$ in (3), which has $\nu_{\mathbb{F}}$ as distribution, we see that under a common marginal functional mark distribution, $M_i \stackrel{d}{=} X_{\mathbb{F}}$ for all $i = 1, \dots, N$. Note that this is a univariate property. One instance where this automatically holds is when Ψ is stationary.

Corollary 1. *If Ψ is stationary then all marks (L_i, M_i) , $i = 1, \dots, N$, have the same marginal distribution $\nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}}$.*

Proof. Assuming that Ψ is stationary, we have that $\mu_G(G) = \lambda_G \ell(G)$, $\lambda_G > 0$. It now follows that $\nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}}$ is the uniquely determined probability measure on $(\mathbb{M}, \mathcal{B}(\mathbb{M}))$ which satisfies $\mu(A) = \lambda_G \ell(G) [\nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}}](H)$, $A = G \times H \in B_{\mathbb{Y}}$ [Schneider and Weil, 2008, Thm 3.5.1]. This implies that $\lambda(g, l, f) = \lambda_G$, i.e. that $f_{(g,l)}^{\mathbb{F}}(f) \equiv 1$ and $f_g^{\mathbb{A}}(l) \equiv 1$. \square

Turning next to the multivariate distributions of the functional marks, we may impose different levels of independence.

Definition 14. *Whenever the regular probability distributions (11) satisfy*

$$P_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(E_1 \times \dots \times E_n) = \prod_{i=1}^n P_{(g_i, l_i)}^{\mathbb{F}}(E_i), \quad E_1, \dots, E_n \in \mathcal{B}(\mathbb{F}),$$

for any $n \geq 1$, we say that Ψ has independent functional marks.

When Ψ both has independent functional marks and a common marginal functional mark distribution, we say that Ψ has randomly labelled functional marks, i.e., for any $n \geq 1$ we have that $P_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(E_1 \times \dots \times E_n) = \prod_{i=1}^n \nu_{\mathbb{F}}(E_i)$, $E_1, \dots, E_n \in \mathcal{B}(\mathbb{F})$.

Here, through Proposition 1, we find that the pair correlation functional (14) satisfies

$$g_{\Psi}((g_1, l_1, f_1), (g_2, l_2, f_2)) = \frac{f_{g_1, g_2}^{\mathbb{A}}(l_1, l_2)}{f_{g_1}^{\mathbb{A}}(l_1) f_{g_2}^{\mathbb{A}}(l_2)} g_G(g_1, g_2).$$

Note that when Ψ has randomly labelled functional marks, M_1, \dots, M_N are independent copies of the reference process $X_{\mathbb{F}}$ in (3). Furthermore, given that Ψ has independent functional marks, by additionally assuming that the regular probabilities (10) satisfy

$$P_{g_1, \dots, g_n}^{\mathbb{A}}(D_1 \times \dots \times D_n) = \prod_{i=1}^n P_{g_i}^{\mathbb{A}}(D_i), \quad D_1, \dots, D_n \in \mathcal{B}(\mathbb{A})$$

for any $n \geq 1$, we retrieve the classical definition of independent marking [Daley and Vere-Jones, 2003, Definition 6.4.III] and consequently that of random labelling.

Remark 4. *A weaker form of independent functional marking, conditionally independent functional marking, may be obtained by assuming that*

$$P_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(E_1 \times \dots \times E_n) = \prod_{i=1}^n F_{(g_i, l_i), \dots, (g_n, l_n)}^{\mathbb{F}}(E_i), \quad E_1, \dots, E_n \in \mathcal{B}(\mathbb{F}),$$

for any $n \geq 1$ and some family $\{F_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(E) : (g_1, l_1), \dots, (g_n, l_n) \in \mathbb{G} \times \mathbb{A}, E \in \mathcal{B}(\mathbb{F})\}$ of regular probability distributions.

5.2 Functional mark reference measures and finite-dimensional distributions

Recall from Proposition 1 the probability measures $P_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(\cdot)$ on $(\mathbb{F}^n, \mathcal{B}(\mathbb{F}^n))$, $n \geq 1$. By assigning a specific structure to each $P_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(\cdot)$, or equivalently to $\nu_{\mathbb{F}}$ and $f_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(\cdot)$, we can determine what type of functional marks we may obtain. So far we have not discussed specific choices for either of these components.

To indicate a few such choices, we start here by looking at how one could obtain deterministic marks under the current setup. We then proceed to considering the case where $\nu_{\mathbb{F}}$ is given by Wiener measure (see e.g. [Mörters and Peres, 2010, Chapter 1] or [Skorohod, 1967, p. 2]) on $(\mathbb{F}, \mathcal{B}(\mathbb{F}))$. We stress that in the latter case, most of the ideas indicated may very well be applied to, say, Poisson random measures (see e.g. [Klebaner, 2005, Jacod and Shiryaev, 1987]) or some other Lévy process/semi-martingale generated random measure on $(\mathbb{F}, \mathcal{B}(\mathbb{F}))$ (see e.g. [Jacod and Shiryaev, 1987, Skorohod, 1967]). Note e.g. that in the Poisson case one would be able to generate multivariate functional marks given by multivariate Poisson processes, a construction similarly to [Crété et al., 2013]. This could likely be the required setup for Examples 4 and 5 in Section 3.5. In order to keep the discussion below fairly compact, we will always assume that we may choose a suitable filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{\mathbb{T}}, \mathbb{P})$ under which the constructions can be made.

5.2.1 Point mass reference measures and deterministic functional marks

Consider some pre-defined deterministic function $f^* : \mathbb{G} \times \mathbb{A} \times \mathcal{T} \rightarrow \mathbb{R}$ such that all marginal functions $f_{(g, l)}^* = \{f^*(g, l, t)\}_{t \in \mathcal{T}}$, $(g, l) \in \mathbb{G} \times \mathbb{A}$, are càdlàg functions. If we set

$$P_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(E_1 \times \dots \times E_n) = \int_{E_1} \dots \int_{E_n} \prod_{i=1}^n P_{(g_i, l_i)}^{\mathbb{F}}(df_i) = \int_{E_1} \dots \int_{E_n} \prod_{i=1}^n \delta_{f_{(g_i, l_i)}^*}(df_i)$$

for any $n \geq 1$ and $E_1, \dots, E_n \in \mathcal{B}(\mathbb{F})$ or, in connection to Lemma 1 and Proposition 1, under some arbitrary reference measure $\nu_{\mathbb{F}}$,

$$f_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(f_1, \dots, f_n) = \delta_{(f_{(g_1, l_1)}^*, \dots, f_{(g_n, l_n)}^*)}(f_1, \dots, f_n),$$

we may create a (ST)CFMPP with deterministic marks $M_i(t) = f^*(g_i, l_i, t)$, $i = 1, \dots, N$. Note that this is the case when we consider e.g. the LISTA functions or the classical growth-interaction process.

5.2.2 Wiener reference measures

Since for many applications it may be desirable to let each M_i follow some diffusion process for $t \in \text{supp}(M_i)$, we here consider a setup constructed by choosing $\nu_{\mathbb{F}}$ as a Wiener measure (see e.g. [Mörters and Peres, 2010, Chapter 1] or [Skorohod, 1967, p. 2]). More specifically, let the reference process (3) be given by a standard Brownian motion $X_{\mathbb{F}} = W = \{W(t)\}_{t \in \mathcal{T}}$. This process is generated by Wiener measure $W_{\mathbb{F}}$ on $C_{\mathcal{T}}(\mathbb{R}) = \{f : \mathcal{T} \rightarrow \mathbb{R} : f \text{ continuous}\}$. Noting that $C_{\mathcal{T}}(\mathbb{R}) \in \mathcal{B}(\mathbb{F})$, consider the reference measure $\nu_{\mathbb{F}}(E) = W_{\mathbb{F}}(E \cap C_{\mathcal{T}}(\mathbb{R}))$, $E \in \mathcal{B}(\mathbb{F})$, i.e. the measure assigning probability 0 to discontinuous sample paths and standard Brownian motion

probabilities to sample paths in $C_{\mathcal{T}}(\mathbb{R})$. It now follows that

$$\begin{aligned} P_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(E_1 \times \dots \times E_n) &= \int_{E_1 \times \dots \times E_n} f_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(f_1, \dots, f_n) \prod_{i=1}^n \nu_{\mathbb{F}}(df_i) \\ &= \int_{E_1 \cap C_{\mathcal{T}}(\mathbb{R}) \times \dots \times E_n \cap C_{\mathcal{T}}(\mathbb{R})} f_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(f_1, \dots, f_n) \prod_{i=1}^n W_{\mathbb{F}}(df_i) \end{aligned}$$

for $E_1, \dots, E_n \in \mathcal{B}(\mathbb{F})$, with $P_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(\mathbb{F} \times \dots \times \mathbb{F}) = 1$.

We may next ask ourselves the adequate question how one would obtain explicit forms for the densities $f_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(\cdot)$. To give an indication, assume that, conditionally on Ψ_G and the auxiliary marks, we want to have $(M_1(t), \dots, M_n(t))$ given by, say, an n -dimensional diffusion process $(Y_1(t), \dots, Y_n(t))$, $t \in \mathcal{T}$. Then, under certain conditions, e.g. the Girsanov theorem (see e.g. [Klebaner, 2005, Jacod and Shiryaev, 1987]) and the Cameron-Martin theorem [Mörters and Peres, 2010] give rise to explicit expressions for $f_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(\cdot)$. Furthermore, changing the support of each M_i to some interval $C_i \subseteq \mathcal{T}$ can be obtained by multiplying the density by $\delta_{\Gamma_i}(f)$, where $\Gamma_i = \{f \in \mathbb{F} : \sup\{\text{supp}(f)\} = C_i\}$, $i = 1, \dots, n$, and/or by applying time-change/stopping results to $(Y_1(t), \dots, Y_n(t))$ before applying e.g. Girsanov's theorem. We note that such a setup would be the underlying construction for the extensions discussed in Section 3.4.

5.2.3 Finite-dimensional distributions of the functional marks

As a distribution on the function space $(\mathbb{F}^n, \mathcal{B}(\mathbb{F}^n))$, each $P_{(x_1, t_1, l_1), \dots, (x_n, t_n, l_n)}^{\mathbb{F}}(\cdot)$ is an abstract and non-tractable object, despite that we sometimes may be able to explicitly define its density $f_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(\cdot)$ with respect to some reference measure $\nu_{\mathbb{F}}^n$. Hence, for all practical and mathematically explicit purposes we turn to the *finite-dimensional distributions* (fidis). For an informative discussion on fidis for càdlàg processes, see [Silvestrov, 2004, 1.6.2].

Let Ψ be a STCFMPP and, conditionally on Ψ_G and the auxiliary marks, assume that we have $\{(X_i, T_i, L_i)\}_{i \in I} = \{(x_i, t_i, l_i)\}_{i \in I}$, $I = \{1, \dots, n\}$, and let

$$M_I = \{M_I(t)\}_{t \in \mathcal{T}} = \{(M_1(t), \dots, M_n(t)) | \{(X_j, T_j, L_j) = (x_j, t_j, l_j)\}_{j=1}^n\}_{t \in \mathcal{T}}.$$

It follows that $P_{(x_1, t_1, l_1), \dots, (x_n, t_n, l_n)}^{\mathbb{F}}(\cdot)$, the distribution of M_I on $(\mathbb{F}^n, \mathcal{B}(\mathbb{F}^n))$, is uniquely determined by the fidis of M_I [Silvestrov, 2004, Lemma 1.6.1.],

$$\begin{aligned} P_{M_I} &= \{P_{M_I(S_k)}(A) : k \geq 1, S_k = \{s_1, \dots, s_k\} \subseteq \mathcal{T}, A \in \mathcal{B}(\mathbb{R}^{n \times k})\} \\ &= \{\mathbb{P}((M_I(s_1), \dots, M_I(s_k)) \in A) : k \geq 1, s_1, \dots, s_k \in \mathcal{T}, A \in \mathcal{B}(\mathbb{R}^{n \times k})\}. \end{aligned}$$

Here we may also choose the sets A as products of half-open intervals $(-\infty, u_{ij}] \subseteq \mathbb{R}$, $i = 1, \dots, n$, $j = 1, \dots, k$, where each (u_{1j}, \dots, u_{nj}) , $j = 1, \dots, k$, is a continuity point of the distribution function corresponding to $P_{M_I(S_k)}(\cdot)$. Although we here have considered the STCFMPP case, the CFMPP case is analogous. We now see that, conditionally on Ψ_G (and thereby N) and the auxiliary marks, it follows that $\{M_i\}_{i=1}^N$ is completely determined by the collection $\bigcup_{I \in \mathcal{P}_N} P_{M_I}$.

If in addition $P_{M_I(S_k)}$ is absolutely continuous with respect to $\ell_n^k = \ell_1^{nk}$, the density of $P_{M_I(S_k)}$ will be denoted by

$$f_{(g_1, l_1), \dots, (g_n, l_n)}^{s_1, \dots, s_k}(u_1, \dots, u_n) = f_{(g_1, l_1), \dots, (g_n, l_n)}^{s_1, \dots, s_k} \begin{pmatrix} u_{11} & \dots & u_{n1} \\ \vdots & \ddots & \vdots \\ u_{1k} & \dots & u_{nk} \end{pmatrix}, \quad (17)$$

where $u_{ij} \in \mathbb{R}$, $i = 1, \dots, n$, $j = 1, \dots, k$.

5.2.4 Markovian functional marks

In many cases it may be of interest to let the functional marks be given by Markov processes. This is e.g. the case when each mark is given by some diffusion process.

Recall that, conditionally on Ψ and the auxiliary marks,

$$M_I = \{M_I(t)\}_{t \in \mathcal{T}} = \{(M_i(t))_{i \in I}\}_{t \in \mathcal{T}} \quad (18)$$

for any index set $I \in \mathcal{P}_N$. Define $\mathcal{F}_t^{M_I} = \sigma\{M_I(s)^{-1}(A) \in \mathcal{F} : s \in \mathcal{T} \cap [0, t], A \in \mathcal{B}(\mathbb{R})^{|I|}\}$ and assume that the underlying filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{\mathbb{T}} = \{\mathcal{F}_t\}_{t \in \mathbb{T}}, \mathbb{P})$ satisfies $\mathcal{F}_t^{M_I} \subseteq \mathcal{F}_t$ for any $I \in \mathcal{P}_N$. We say that Ψ has *Markovian marks* if each M_I , $I \in \mathcal{P}_N$, constitutes a Markov process, i.e., for $s \leq t$,

$$\mathbb{P}(M_I(t) \in A | \mathcal{F}_s) = \mathbb{P}(M_I(t) \in A | M_I(s)), \quad A \in \mathcal{B}(\mathbb{R})^{|I|}.$$

Here we refer to $P_{t,s}^{M_I}(A; M_I(s)) = \mathbb{P}(M_I(t) \in A | M_I(s))$ as the M_I -transition probabilities and when there exist transition densities $p_{t,s}^{M_I}(u_t; u_s)$, $u_t, u_s \in \mathbb{R}^n$, with respect to ℓ_n , $n = |I|$, we find that the densities (17) become

$$f_{(g_1, l_1), \dots, (g_n, l_n)}^{s_1}(u_1) \prod_{i=2}^k p_{s_i, s_{i-1}}^{M_I}(u_i; u_{i-1}), \quad u_1, \dots, u_k \in \mathbb{R}^n. \quad (19)$$

5.3 Auxiliary reference measures and multivariate (ST)CFMPPs

Turning next to the auxiliary mark space $(\mathbb{A}, \mathcal{B}(\mathbb{A}))$ and its reference measure $\nu_{\mathbb{A}}$, which we have assumed to be locally finite (the local finiteness allows us to appropriately construct kernels/regular probabilities [Daley and Vere-Jones, 2008, Exercise 9.1.16]), recall from Section 2.2.3 that we either let $\mathbb{A} = \mathbb{A}_d = \{1, \dots, k_{\mathbb{A}}\}$, $\mathbb{A} = \mathbb{A}_c \subseteq \mathbb{R}^{m_{\mathbb{A}}}$, or $\mathbb{A} = \mathbb{A}_d \times \mathbb{A}_c$, where $k_{\mathbb{A}}, m_{\mathbb{A}} \geq 1$.

Note that in the first case, each mark $L_i \in \mathbb{A}_d$, $i = 1, \dots, N$, is a discrete random variable and we may write $\Psi_i = \sum_{(g, l, f) \in \Psi \cap \mathbb{G} \times \{i\} \times \mathbb{F}} \delta_{(g, f)} = \sum_{(g, f) \in \Psi_i} \delta_{(g, f)}$ for the restriction of Ψ to the auxiliary mark set $\{i\}$, $i = 1, \dots, k_{\mathbb{A}}$. For this space we employ some finite measure $\nu_{\mathbb{A}_d}$ as reference measure, which implies that the regular probabilities (10) take the form

$$P_{g_1, \dots, g_n}^{\mathbb{A}}(D_1 \times \dots \times D_n) = \sum_{D_1 \cap \mathbb{A}_d} \dots \sum_{D_n \cap \mathbb{A}_d} f_{g_1, \dots, g_n}^{\mathbb{A}}(l_1, \dots, l_n) \nu_{\mathbb{A}_d}(l_1) \dots \nu_{\mathbb{A}_d}(l_n),$$

so that each $f_{g_1, \dots, g_n}^{\mathbb{A}}(l_1, \dots, l_n)$ becomes a probability mass function/discrete probability density function. On this space the most natural choice for $\nu_{\mathbb{A}_d}$ is the counting measure $\nu_{\mathbb{A}_d}(\cdot) = \sum_{j=1}^{k_{\mathbb{A}}} \delta_j(\cdot)$. At times we want to let the intensity functional of Ψ_i be given by $\lambda_i(g, f) = \lambda(g, i, f) \nu_{\mathbb{A}}(i)$. Then, when $\nu_{\mathbb{A}_d}$ is the counting measure, $\nu_{\mathbb{A}}(i) = 1$ and $\lambda_i(g, f) = f_{(g, l)}^{\mathbb{F}}(f) f_g^{\mathbb{A}}(i) \lambda_G(g)$. Note that if we do not assume that the auxiliary marks are spatially dependent, we may simply let $P_{g_1, \dots, g_n}^{\mathbb{A}}(\cdot) = \nu_{\mathbb{A}}^n(\cdot)$, or equivalently $f_{g_1, \dots, g_n}^{\mathbb{A}}(\cdot) \equiv 1$, $n \geq 1$.

Turning to the second alternative, each mark $L_i \in \mathbb{A}_c$ becomes a (possibly) continuous $m_{\mathbb{A}}$ -dimensional random variable. Here some care should be taken. In most cases the natural candidate for the reference measure $\nu_{\mathbb{A}_c}$ would be Lebesgue measure $\ell_{m_{\mathbb{A}}}$ on $(\mathbb{A}_c, \mathcal{B}(\mathbb{A}_c))$, whereby $f_{g_1, \dots, g_n}^{\mathbb{A}}(l_1, \dots, l_n)$ would become a probability density function in the usual sense. However, at times one must require that $\nu_{\mathbb{A}_c}$ is a finite measure, i.e. $\nu_{\mathbb{A}_c}(\mathbb{A}_c) < \infty$. This is e.g. the case when we treat densities of finite (ST)CFMPPs with respect to Poisson processes (see Section 6.4). If \mathbb{A}_c

is a bounded subset of $\mathbb{R}^{m_{\mathbb{A}}}$, Lebesgue measure will suffice and we may rescale it into the uniform probability measure $\ell_{m_{\mathbb{A}}}(\cdot)/\ell_{m_{\mathbb{A}}}(\mathbb{A})$. On the other hand, when \mathbb{A}_c is not bounded, we choose $\nu_{\mathbb{A}_c}$ as some probability distribution on $(\mathbb{A}_c, \mathcal{B}(\mathbb{A}_c))$, to which the desired distributions of the L_i 's are absolutely continuous.

In the last scenario, where $\mathbb{A} = \mathbb{A}_d \times \mathbb{A}_c$, we simply let the reference measure be given by $\nu_{\mathbb{A}}(\cdot) = [\nu_{\mathbb{A}_d} \otimes \nu_{\mathbb{A}_c}](\cdot)$, the product measure of the two measures defined on the two spaces \mathbb{A}_d and \mathbb{A}_c . Here each auxiliary mark has the form $L_i = (L_{i1}, L_{i2}) \in \mathbb{A}_d \times \mathbb{A}_c$. The discrete random variable L_{i1} indicates which type $1, \dots, k_{\mathbb{A}}$ the i th point belongs to, whereas L_{i2} has the purpose of, say, controlling the functional mark(s). Hereby, given that $(L_{i1}, L_{i2}) = (l_{i1}, l_{i2})$, $i = 1, \dots, n$, the functional conditional densities (9) take the form

$$f_{(g_1, l_{12}), \dots, (g_n, l_{n2})}^{\mathbb{F}}(\cdot; l_{11}, \dots, l_{n1}), \quad g_1, \dots, g_n \in \mathbb{G}. \quad (20)$$

Hence, given the spatial(-temporal) locations $g_1, \dots, g_n \in \mathbb{G}$ and $l_{12}, \dots, l_{n2} \in \mathbb{A}_c$, the functional mark distributions may still vary, depending on which type each point $i = 1, \dots, n$ is assigned. Note that the L_{i2} 's may be treated as random parameter vectors which control e.g. the supports of the marks.

On the other hand, it may also be the case that $\Psi_G = (\Psi_G^1, \dots, \Psi_G^{k_{\mathbb{A}}})$ is a multivariate spatial point process on e.g. $\mathbb{G} = \mathbb{X}$, for instance a multivariate Cox process. As such, it may be treated as the marked point process $\Psi_G = \{(X_i, L_{i1})\}_{i=1}^N$, where L_{i1} indicates which $\Psi_G^1, \dots, \Psi_G^{k_{\mathbb{A}}}$ a point belongs to. By additionally assigning functional marks and auxiliary marks $\{L_{i2}\}_{i=1}^N$ to Ψ_G , with the latter possibly controlling the former through (20), we obtain a CFMPP $\Psi = \{(X_i, L_{i1}, L_{i2}, M_i)\}_{i=1}^N$ where the L_{i1} 's govern both the ground process and the functional marks.

Remark 5. *Since each L_{i2} is a random variable, which works as a parameter in a stochastic process, the current setup of (ST)CFMPPs connects naturally to a Bayesian stochastic process framework.*

5.4 Intensity-dependent marks

A step forward in the marking of stationary unmarked point processes is to allow the distributions of the marks to be dependent on the local intensity, as suggested by [Ho and Stoyan, 2008] and [Myllymäki and Penttinen, 2009] in the context of stationary log Gaussian Cox processes [Møller et al., 1998, Møller and Waagepetersen, 2004]. This intensity-dependent marking assumes conditional independence to hold for the marks, given the random intensity. Heuristically, these models allow the marks to be large (small) in areas of low point intensity and small (large) in areas of high intensity. For log Gaussian Cox process, intensity-dependent marking leads to a correlation of the marks which is affected by the second-order property of the unmarked Cox process. The setup in [Myllymäki and Penttinen, 2009] developed new marking models of such a generality that not only the mean of the mark distribution but also its variance is affected by the local intensity, and these models have been employed for the marking of log Gaussian Cox processes. In this context it is interesting to test for mark independence and for dependence between marks and locations [Grabarnik et al., 2011, Schlather et al., 2004].

In the current STCFMPP context we may extend these ideas further.

Definition 15. *A STCFMPP Ψ with ground intensity $\lambda_G(x, t)$ is said to have intensity-dependent marks if, conditionally on Ψ_G and the auxiliary marks, the functional marks are given by $M_i(t) = \lambda_G(X_i, t)$, $t \in \mathcal{T}$, $i = 1, \dots, N$.*

Note that this falls in the category of deterministic marks (see Section 5.2.1) and the corresponding point masses on $(\mathbb{F}^n, \mathcal{B}(\mathbb{F}^n))$, $n \geq 1$, may or may not depend on the auxiliary marks (recall the discussion in Section 5.3). Moreover, as in the above mentioned references, larger marks indicate where and when there is high intensity.

6 Specific classes of (ST)CFMPPs

Having defined a general structure for (ST)CFMPPs, we next turn to considering different specific constructions. Since by the notion of a spatio-temporal point process is often meant a temporally grounded point process, we look closer at temporally grounded (ST)CFMPPs and as a result obtain a definition of functional marked conditional intensities [Cox and Isham, 1980, Daley and Vere-Jones, 2003, Vere-Jones, 2009]. Furthermore, we look closer at finite (ST)CFMPPs, and then in particular Markov (ST)CFMPPs [Daley and Vere-Jones, 2003, Van Lieshout, 2000]. However, we start by defining (spatio-temporal) càdlàg functional marked Poisson and Cox processes [Daley and Vere-Jones, 2003, Van Lieshout, 2000, Møller and Waagepetersen, 2004, Stoyan et al., 1995].

6.1 Poisson processes

Poisson processes, the most well known point process models, are the benchmark/reference models for representing lack of spatial interaction.

Given a locally finite measure μ on $\mathcal{B}(\mathbb{Y})$, we let a (*spatio-temporal*) *càdlàg functional marked ((ST)CFM) Poisson process* Ψ , with intensity measure μ , be defined as a Poisson process on \mathbb{Y} . In other words, for any disjoint $A_1, \dots, A_n \in \mathcal{B}(\mathbb{Y})$, $n \geq 1$, the random variables $\Psi(A_1), \dots, \Psi(A_n)$ are independent and Poisson distributed with means $\mu(A_i)$, $i = 1, \dots, n$, provided A_i is bounded. When Ψ has an intensity functional $\lambda(\cdot)$, i.e. when the intensity measure in (6) satisfies $\mu(A) = \int_A \lambda(g, l, f) \nu(d(g, l, f))$, through Proposition 1 it follows that

$$\rho^{(n)}((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) = \prod_{i=1}^n \lambda(g_i, l_i, f_i) = \prod_{i=1}^n f_{(g_i, l_i)}^{\mathbb{F}}(f_i) f_{g_i}^{\mathbb{A}}(l_i) \lambda_G(g_i),$$

whereby the pair correlation functional satisfies $g_{\Psi}((g_1, l_1, f_1), (g_2, l_2, f_2)) = 1$. We note also that through (16), since for all $(g, l, f) \in \mathbb{Y}$ the Palm measures satisfy $P^{! (g, l, f)}(\cdot) = P(\cdot)$, its Papangelou conditional intensity satisfies $\lambda(g, l, f; \Psi) = \lambda(g, l, f)$. When Ψ is stationary, due to Corollary 1, Ψ becomes randomly labelled and $\rho^{(n)}((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) = \lambda_G^n > 0$.

6.1.1 Ground Poisson processes

We next relax the Poisson process assumption slightly to only concern the ground process. More specifically, we say that a (ST)CFMPP Ψ is a (*ST*)CFM *ground Poisson process* if its ground process Ψ_G constitutes a simple Poisson process on \mathbb{G} . Note that in light of Proposition 1,

$$\rho^{(n)}((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) = f_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(f_1, \dots, f_n) f_{g_1, \dots, g_n}^{\mathbb{A}}(l_1, \dots, l_n) \prod_{i=1}^n \lambda_G(g_i),$$

where $\lambda_G(\cdot)$ is the ground intensity function. Note that here $g_G(g_1, g_2) \equiv 1$, whereby the pair correlation functional satisfies $g_{\Psi}((g_1, l_1, f_1), (g_2, l_2, f_2)) = f_{g_1, g_2}^{\mathbb{A}}(l_1, l_2) / (f_{g_1}^{\mathbb{A}}(l_1) f_{g_2}^{\mathbb{A}}(l_2))$ if Ψ has independent functional marks and $g_{\Psi}((g_1, l_1, f_1), (g_2, l_2, f_2)) = 1$ if Ψ is independently marked.

6.2 Cox processes

We here consider Cox processes (see e.g. [Stoyan et al., 1995, p. 154]) in the current context of càdlàg functional marking. These are common and interesting models for spatial clustering. Recall from (7) that μ_G is the ground intensity measure of a (ST)CFMPP.

Definition 16. *Given a locally finite random measure Λ_G on \mathbb{G} , a (ST)CFMPP Ψ is called a (spatio-temporal) càdlàg functional marked ((ST)CFM) Cox process (directed by Λ_G) if the ground process Ψ_G constitutes a Λ_G -directed Cox process on \mathbb{G} . In other words, conditionally on Λ_G , Ψ_G is a Poisson process with $\mu_G = \Lambda_G$.*

We note that, conditionally on Λ_G , Ψ becomes a (ST)CFMPP ground Poisson process. Hence, a more suitable name for Ψ would possibly be (ST)CFM ground Cox process. Assume next that the locally finite random measure $\Lambda_G(G) = \int_G \Lambda(g)dg$ is generated by an a.s. non-negative random field $\Lambda = \{\Lambda(g)\}_{g \in \mathbb{G}}$, which consequently must be a.s. locally integrable. When $\mathbb{G} = \mathbb{X}$ (CFM Cox process), we may write $\Lambda = \{\Lambda(x)\}_{x \in \mathbb{X}}$ and when $\mathbb{G} = \mathbb{X} \times \mathbb{T}$ (STCFM Cox process) we may write $\Lambda = \{\Lambda(x, t)\}_{(x, t) \in \mathbb{X} \times \mathcal{T}}$. For a (ST)CFM Cox process, in light of Proposition 1, the n th product density is given by [Daley and Vere-Jones, 2003, Chapter 6.2.]

$$\rho^{(n)}((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) = f_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(f_1, \dots, f_n) f_{g_1, \dots, g_n}^{\mathbb{A}}(l_1, \dots, l_n) \prod_{i=1}^n \mathbb{E}[\Lambda_G(g_i)],$$

whereby $g_G(\cdot) \equiv 1$ and its pair correlation functional becomes

$$g_{\Psi}((g_1, l_1, f_1), (g_2, l_2, f_2)) = \frac{f_{(g_1, l_1), (g_2, l_2)}^{\mathbb{F}}(f_1, f_2) f_{g_1, g_2}^{\mathbb{A}}(l_1, l_2)}{f_{(g_1, l_1)}^{\mathbb{F}}(f_1) f_{g_1}^{\mathbb{A}}(l_1) f_{(g_2, l_2)}^{\mathbb{F}}(f_2) f_{g_2}^{\mathbb{A}}(l_2)}.$$

When Ψ is a (ST)CFM Cox process with spatio-temporal geostatistical marking (recall Definition 7), i.e. $M_i(t) = Z_{X_i}(t)$ for some spatio-temporal random field $Z = \{Z_x(t)\}_{(x, t) \in \mathbb{X} \times \mathcal{T}}$, we may connect random fields and point processes simultaneously in two different ways; the driving random field Λ 'from underneath' and a random field Z 'from above'. This structure is simplified when we consider intensity dependent marks (Section 5.4). In the current context this translates into the following definition.

Definition 17. *A STCFM Cox process Ψ with random intensity field $\Lambda = \{\Lambda(x, t)\}_{(x, t) \in \mathbb{X} \times \mathcal{T}}$ is said to have intensity-dependent marks if, conditionally on Ψ_G and the random field Λ , the functional marks are given by $M_i(t) = \Lambda(X_i, t)$, $t \in \mathcal{T}$, $i = 1, \dots, N$.*

6.3 Temporally grounded STCFMPPs and conditional intensities

The product densities are extremely useful tools and e.g. they allow for the development of an array of different tools/statistics, useful for performing statistical inference of different kinds. There is, however, one case which allows one to take a step further in this development and that is when a STCFMPP Ψ is temporally grounded. More specifically, recall from Definition 5 that when a STCFMPP Ψ is temporally grounded we may treat it as a temporal point process $\Psi_{\mathbb{T}} = \{T_i\}_{i=1}^N$ on \mathbb{T} with marks $\{(X_i, L_i, M_i)\}_{i=1}^N$. This allows one to exploit the natural ordering of \mathbb{T} and thus exploit the more general theory of temporal stochastic processes, in particular that of cumulative processes (see e.g. [Daley and Vere-Jones, 2003, 2008]).

It should be stressed that when Ψ is a CFMPP and $\mathbb{X} \subseteq \mathbb{R}$, by construction, the ground process $\Psi_G = \{X_i\}_{i=1}^N \subseteq \mathbb{R}$ may also be treated as a temporal point process with marks $\{(L_i, M_i)\}_{i=1}^N$. Naturally the ideas presented below still hold under this setup.

6.3.1 The temporal ground product densities

We first look closer at the behaviour of $\rho_G^{(n)}$ when Ψ is temporally grounded and for this purpose we consider $G_1, \dots, G_n = (G_1^S \times G_1^T), \dots, (G_n^S \times G_n^T) \in \mathcal{B}(\mathbb{G}) = \mathcal{B}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{T})$.

Since Ψ is temporally grounded, Ψ_G may be treated as the temporal point process $\Psi_{\mathbb{T}} = \{T_i\}_{i=1}^N$ on \mathbb{T} , with associated \mathbb{X} -valued marks $\{X_i\}_{i=1}^N$. Then, under the assumption of local finiteness, we may extend

$$\alpha_{G,T}^{(n)}(G_1^T \times \dots \times G_n^T) = \alpha_G^{(n)}((\mathbb{X} \times G_1^T) \times \dots \times (\mathbb{X} \times G_n^T))$$

to become the n th *temporal ground factorial moment measure* of $\Psi_{\mathbb{T}}$. Consequently, when we additionally assume that $\alpha_{G,T}^{(n)} \ll \ell_1^n$, where the Radon-Nikodym derivative $\rho_{G,T}^{(n)}$ is referred to as the n th *temporal ground product density*, we may disintegrate $\alpha_G^{(n)}$ with respect to some family $\{P_{t_1, \dots, t_n}^{G,S}(\cdot) : t_1, \dots, t_n \in \mathbb{T}\}$ of regular conditional probability distributions on $\mathcal{B}(\mathbb{X}^n)$. When we further assume that $P_{t_1, \dots, t_n}^{G,S} \ll \ell_d^n$, with density $f_{t_1, \dots, t_n}^{G,S}(x_1, \dots, x_n)$, we find that

$$\alpha_G^{(n)}(G_1 \times \dots \times G_n) = \int_{G_1 \times \dots \times G_n} f_{t_1, \dots, t_n}^{G,S}(x_1, \dots, x_n) \rho_{G,T}^{(n)}(t_1, \dots, t_n) \prod_{i=1}^n dx_i dt_i$$

and we write $\lambda_{G,T}(t) = \rho_{G,T}^{(1)}(t)$ for the corresponding intensity. Note that when $\rho_G^{(n)}$ exists, $\rho_G^{(n)}((x_1, t_1), \dots, (x_n, t_n)) = f_{t_1, \dots, t_n}^{G,S}(x_1, \dots, x_n) \rho_{G,T}^{(n)}(t_1, \dots, t_n)$.

Remark 6. When on the other hand Ψ is spatially grounded, Ψ_G may be described as a spatial point process $\Psi_{\mathbb{X}} = \{X_i\}_{i=1}^N$ with \mathbb{T} -valued temporal marks $\{T_i\}_{i=1}^N$. Then, under analogous assumptions and in an identical fashion, we may derive the spatial ground product density $\rho_{G,S}^{(n)}(x_1, \dots, x_n)$ and the corresponding family $f_{x_1, \dots, x_n}^{G,T}(t_1, \dots, t_n)$ of densities for which the product constitutes $\rho_G^{(n)}((x_1, t_1), \dots, (x_n, t_n))$, provided it exists. However, at least when $d \geq 2$ (which is a natural assumption), there is no natural ordering to be gained, whereby the temporal component is absorbed into the auxiliary mark and we end up considering the usual CFMPP case.

It should be pointed out at this stage that not all STCFMPPs are temporally grounded. E.g. when $\mathbb{G} = \mathbb{X} \times \mathbb{T} = \mathbb{R}^2 \times [0, \infty)$, in the case of a Poisson process, we must require that $\alpha_{G,T}^{(1)}(C) = \int_C \lambda_{G,T}(t) dt = \mu_G(\mathbb{R}^2 \times C) < \infty$ for bounded $C \in \mathcal{B}(\mathbb{T})$ in order for $\Psi_{\mathbb{T}}$ to be a well-defined point process on \mathbb{T} . When Ψ is a finite STCFMPP this follows automatically.

6.3.2 Cumulative STCFMPPs and conditional intensities

Assume now that $\mathbb{T} \subseteq [0, \infty)$ and consider some suitable filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{\mathbb{T}} = \{\mathcal{F}_t\}_{t \in \mathbb{T}}, \mathbb{P})$. Note that in this section we essentially follow the structure and notation of [Daley and Vere-Jones, 2008, Chapter 14].

Define $(\mathbb{M}_{\mathbb{T}}, \mathcal{B}(\mathbb{M}_{\mathbb{T}}), \nu_{\mathbb{M}_{\mathbb{T}}}) = (\mathbb{X} \times \mathbb{A} \times \mathbb{F}, \mathcal{B}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{A}) \otimes \mathcal{B}(\mathbb{F}), \ell_d \otimes \nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}})$ so that (X_i, L_i, M_i) takes values in $(\mathbb{M}_{\mathbb{T}}, \mathcal{B}(\mathbb{M}_{\mathbb{T}}), \nu_{\mathbb{M}_{\mathbb{T}}})$, and define

$$\Psi_{\mathbb{T}}(C, K) = \Psi(B \times C \times D \times E)$$

for $C \in \mathcal{B}(\mathbb{T})$ and $K = B \times D \times E \in \mathcal{B}(\mathbb{M}_{\mathbb{T}})$ which gives us the *cumulative process* $\Psi_{\mathbb{T}} : \mathbb{T} \times \mathcal{B}(\mathbb{M}_{\mathbb{T}}) \times \Omega \rightarrow [0, \infty]$,

$$\Psi_{\mathbb{T}}(t, K) = \Psi_{\mathbb{T}}([0, t], K) = \sum_{i=1}^N \mathbf{1}_K(X_i, L_i, M_i) \mathbf{1}_{(-\infty, t]}(T_i), \quad t \in \mathbb{T}. \quad (21)$$

We note that this, in fact, is a cumulative process in the sense of [Daley and Vere-Jones, 2008, p. 378] since $\Psi_{\mathbb{T}}(t, K)$ is monotonically increasing in t , for fixed K , and a locally finite measure on $\mathcal{B}(\mathbb{M}_{\mathbb{T}})$ for fixed t . Consider further the history/filtration

$$\mathcal{H}_t = \sigma(\{\Psi_{\mathbb{T}}((s, t], K) : 0 < s \leq t, K \in \mathcal{B}(\mathbb{M}_{\mathbb{T}})\}) \subseteq \mathcal{F}_t \quad (22)$$

and note that $\Psi_{\mathbb{T}}(t, K)$ becomes progressively measurable with respect to $\mathcal{F}_{\mathbb{T}}$ [Daley and Vere-Jones, 2003, (A3.3.2)] and the so-called mark-predictable σ -algebra [Daley and Vere-Jones, 2003, p. 379] is coarser than $\mathcal{F}_{\mathbb{T}}$. Hence, since $\Psi_{\mathbb{T}}(t, K)$ is adapted to $\mathcal{F}_{\mathbb{T}}$, we have a setup well enough specified to accommodate all purposes/constructions considered below.

Consider next the *compensator* $A(t, K)$ of $\Psi_{\mathbb{T}}$, which is the unique (mark-)predictable cumulative process $A(t, K)$ such that, for each $K \in \mathcal{B}(\mathbb{M}_{\mathbb{T}})$, the process $\Psi_{\mathbb{T}}(t, K) - A(t, K)$ is a martingale with respect to $\mathcal{F}_{\mathbb{T}}$ [Daley and Vere-Jones, 2008, Definition 14.2.III.]. Under the current setup, such an $A(t, K)$ exists uniquely and through it we may define a *conditional intensity* [Daley and Vere-Jones, 2008, Definition 14.3.I.]. The conditional intensity is an $\mathcal{F}_{\mathbb{T}}$ -adapted process $\lambda^*(t, (x, l, f); \omega)$ (we emphasise this by writing out the ω) which is measurable with respect to $\mathcal{B}(\mathbb{T}) \otimes \mathcal{B}(\mathbb{M}_{\mathbb{T}}) \otimes \mathcal{F}$ and satisfying

$$A(t, K; \omega) = \int_{(0, t] \times K} \lambda^*(u, (x, l, f); \omega) du \nu_{\mathbb{M}_{\mathbb{T}}}(d(x, l, f))$$

a.s. for all $t \in \mathbb{T} \setminus \{0\}$ and $K \in \mathcal{B}(\mathbb{M}_{\mathbb{T}})$. However, in order to explicitly define λ^* we have to consider the Campbell measure $\mathcal{C}_{\Psi_{\mathbb{T}}}$ on $\mathbb{T} \times \mathbb{M}_{\mathbb{T}} \times \Omega$, which is defined through the relations

$$\mathcal{C}_{\Psi_{\mathbb{T}}}(C \times K \times F) = \int_F \Psi_{\mathbb{T}}(C, K; \omega) \mathbb{P}(d\omega), \quad F \in \mathcal{F},$$

for bounded $C \in \mathcal{B}(\mathbb{T})$ and $K \in \mathcal{B}(\mathbb{M}_{\mathbb{T}})$. By assuming that $\mathcal{C}_{\Psi_{\mathbb{T}}}(\cdot) \ll \ell_1 \otimes \nu_{\mathbb{M}_{\mathbb{T}}} \otimes \mathbb{P}$, we find that a predictable version of λ^* is given by the corresponding Radon-Nikodym derivative [Daley and Vere-Jones, 2008, Proposition 14.3.II.]

$$\lambda^*(u, (x, l, f); \omega) = \frac{d\mathcal{C}_{\Psi_{\mathbb{T}}}(t, K; \omega)}{d(\ell_1 \otimes \nu_{\mathbb{M}_{\mathbb{T}}} \otimes \mathbb{P})}$$

and this version coincides (except possibly on a $[\ell_1 \otimes \nu_{\mathbb{M}_{\mathbb{T}}} \otimes \mathbb{P}]$ -null set) with any other conditional intensity defined through the integral equation above. Hereby the conditional intensity of the ground process $\Psi_{\mathbb{T}}^G$ is given by

$$\lambda_G^*(t; \omega) = \int_{\mathbb{M}_{\mathbb{T}}} \lambda^*(t, (x, l, f); \omega) \nu_{\mathbb{M}_{\mathbb{T}}}(d(x, l, f)).$$

Lemma 2. *The conditional intensity is given by*

$$\lambda^*(t, (x, l, f); \omega) = f_{(x, t, l)}^{\mathbb{F}}(f; \omega) f_{(x, t)}^{\mathbb{A}}(l; \omega) f_t^{G, S}(x; \omega) \lambda_G^*(t; \omega),$$

with $\mathbb{E}[\lambda^*(t, (x, l, f))] = \lambda(t, x, l, f) = f_{(x, t, l)}^{\mathbb{F}}(f) f_{(x, t)}^{\mathbb{A}}(l) f_t^{G, S}(x) \lambda_G^*(t)$, and

$$\lambda_K^*(t; \omega) = \int_K \lambda^*(t, (x, l, f); \omega) [\ell \otimes \nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}}](d(x, l, f))$$

is the conditional intensity of $\Psi_{\mathbb{T}}(t, K)$.

Proof. By [Daley and Vere-Jones, 2008, Proposition 14.3.II.], there exist regular probabilities $P_t^{\mathbb{M}_T}(\cdot)$ on $\mathcal{B}(\mathbb{M}_T)$ which are absolutely continuous with respect to $\nu_{\mathbb{M}_T}$ with densities $f_{\mathbb{M}_T}(x, l, f|t; \omega)$ such that $\lambda^*(t, (x, l, f)) = f_{\mathbb{M}_T}(x, l, f|t; \omega) \lambda_G^*(t; \omega)$. By applying Lemma 1 we obtain the first statement. Note further that since $\mathbb{E}[A(t, K)] = \mathbb{E}[\Psi_T(t, K)]$ it follows that $\mathbb{E}[\lambda^*(t, (x, l, f))] = \lambda(t, x, l, f)$, the intensity functional of Ψ . The last part follows from [Daley and Vere-Jones, 2008, p. 392]. \square

Note here that λ^* depends on Ψ only through the history \mathcal{H}_t and not the whole realisation (emphasised by ω) so we consequently may write $\lambda^*(\cdot; \mathcal{H}_t)$ or simply $\lambda^*(\cdot)$, which is more common notation in point process literature. Provided that the limit exists, there exists a version of λ^* such that

$$\lambda^*(t, x, l, f) = \lim_{\Delta \downarrow 0} \frac{\mathbb{E}[\Psi(B_{\mathbb{X}}[x, \Delta] \times [t, t + \Delta] \times B_{\mathbb{A}}[l, \Delta] \times B_{\mathbb{F}}[f, \Delta]) | \mathcal{H}_t]}{\ell_d(B_{\mathbb{X}}[x, \Delta]) \ell_1([t, t + \Delta]) \nu_{\mathbb{A}}(B_{\mathbb{A}}[l, \Delta]) \nu_{\mathbb{F}}(B_{\mathbb{F}}[f, \Delta])},$$

where e.g. $B_{\mathbb{F}}[f, \Delta]$ denotes the closed ball which is centred around $f \in \mathbb{F}$ with $d_{\mathbb{F}}$ -radius Δ .

Note also that extensions where $\mathbb{T} = \mathbb{R}$ are possible [Daley and Vere-Jones, 2008, p. 394]. However, since in most applications it is natural to assume that the temporal starting point is 0, we do not choose to consider such a setup.

To give an example, recall Section 5.3 and assume that $\mathbb{A} = \mathbb{A}_c = \mathbb{R}$. By [Daley and Vere-Jones, 2003, Lemma 6.4.VI.] we may construct a STCFMPP Ψ where $\{(T_i, L_i)\}_{i=1}^N$ constitutes a Compound Poisson process. This is done by letting Ψ_T be a Poisson process with intensity $\lambda_{G,T}(t)$ and $f_{(x,t)}^{\mathbb{A}}(l)$ be the density of the underlying jump size kernel $F(\cdot|\cdot)$. Note that in this construction we allow for the jump sizes to be location dependent.

6.3.3 Total temporal evolution and Markovian functional marks

Here we may extend the concept of Markovian functional marks a bit. Let $\mathbb{T} = \mathcal{T}$ and let $I_t = \{i : T_i \leq t\}$, $t \in \mathbb{T}$, where $I_t = \bigcup_{s \leq t} I_s$. Letting $I = I_t$ in (18), we obtain $M_{I_t}(s) = (M_i(s))_{i \in I_t}$, $s \in \mathbb{T}$, i.e. the multivariate conditional mark process of all point observed by time $t \in \mathbb{T}$. Note also that when we assume that $\inf\{\text{supp}(M_i)\} = T_i$, then $I_t = \bigcup_{s \leq t} \{i : s \in \text{supp}(M_i)\}$. Recall now the history \mathcal{H}_t in (22) and define $\mathcal{F}_t^M = \sigma\{M_{I_t}(s)^{-1}(A) \in \mathcal{F} : s \in \mathbb{T} \cap [0, t], A \in \mathcal{B}(\mathbb{R})^{|I_s|}\}$ so that, if one were to study the simultaneous temporal evolution of the points and the marks, it should hold that

$$\mathcal{F}_t^* = \mathcal{H}_t \bigvee \mathcal{F}_t^M \subseteq \mathcal{F}_t, \quad t \in \mathbb{T},$$

for the underlying filtration $\mathcal{F}_{\mathbb{T}} = \{\mathcal{F}_t\}_{t \in \mathbb{T}}$. Hereby we may study the simultaneous evolution of $\Psi_{\mathbb{T}}$ (or $\Psi_{\mathbb{T}}^G$) and M_{I_t} , i.e. the *total temporal evolution*. Note here that when Ψ has Markovian marks, it follows that

$$\mathbb{P}(M_{I_t}(t) \in A | \mathcal{F}_s) = \mathbb{P}(M_{I_t}(t) \in A | M_{I_t}(s))$$

for $s \leq t$ and $A \in \mathcal{B}(\mathbb{R}^n)$.

6.4 Finite (ST)CFMPPs

Recall the definitions of finite CFMPPs and finite STCFMPPs given in Definitions 3 and 4, respectively. Following [Daley and Vere-Jones, 2003, Chapter 5.3], we find that the distribution P of a finite (ST)CFMPP Ψ is completely specified by its *Janossy measures*, $J_n(A_1 \times \cdots \times A_n)$,

$A_1, \dots, A_n \in \mathcal{B}(\mathbb{Y})$, $n \geq 0$, where $\sum_{n=0}^{\infty} \frac{1}{n!} J_n(\mathbb{Y}^n) = 1$. We assume that these have symmetric densities with respect to the n -fold products ν^n of the reference measure in (2), i.e.

$$J_n(A_1 \times \dots \times A_n) = \int_{A_1 \times \dots \times A_n} j_n((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) \prod_{i=1}^n \nu(d(g_i, l_i, f_i)),$$

and we refer to these densities $j_n(\cdot)$ as the *Janossy densities*. Here \mathbb{Y}^0 denotes an ideal point such that $\mathbb{Y}^0 \times \mathbb{Y} = \mathbb{Y} \times \mathbb{Y}^0 = \mathbb{Y}$ [Daley and Vere-Jones, 2003, Proposition 5.3.II.]. The interpretation is that $J_n(d(g_1, l_1, f_1) \times \dots \times d(g_n, l_n, f_n))$ gives the probability of Ψ having *exactly* n marked points in the infinitesimal ground-mark regions $d(g_1, l_1, f_1), \dots, d(g_n, l_n, f_n) \subseteq \mathbb{Y}$ and no points anywhere else. We note here that by construction the ground process Ψ_G is also a finite point process, on \mathbb{G} , and its Janossy densities will be denoted by j_n^G , $n \geq 0$.

Recalling the interpretation of $\rho^{(n)}$, it quickly becomes clear that there is a connection between the product densities and j_n . The fundamental difference between the two is that j_n gives the infinitesimal probability of having exactly n points at the specified marked locations and no points anywhere else.

Lemma 3. *For each $n \geq 1$, recalling the conditional densities (8) and (9), the Janossy densities satisfy*

$$j_n((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) = f_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(f_1, \dots, f_n) f_{g_1, \dots, g_n}^{\mathbb{A}}(l_1, \dots, l_n) j_n^G(g_1, \dots, g_n).$$

Proof. We here see that the existence of all j_n , $n \geq 1$, implies that Ψ is *regular* in the sense of [Daley and Vere-Jones, 2003, p. 247]. Hence, we find that Ψ_G is regular in its own right and for any $n \geq 1$, given $\Psi_G = \{g_1, \dots, g_n\}$, the conditional distribution of the marks has density $f_{g_1, \dots, g_n}^{\mathbb{M}}(m_1, \dots, m_n)$ on $\mathbb{M}^n = (\mathbb{A} \times \mathbb{F})^n$, with respect to $(\nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}})^n$ [Daley and Vere-Jones, 2003, Proposition 7.3.I.]. Hereby an application of Lemma 1 gives the desired result. \square

In the current setup, for any bounded $G \in \mathcal{B}(\mathbb{G})$ and any $n \geq 1$, the *local Janossy measure* $J_n(\cdot | G \times \mathbb{M})$, i.e. the Janossy measure of $\Psi \cap (G \times \mathbb{M})$, has a well-defined probability density $j_n((g_1, l_1, f_1), \dots, (g_n, l_n, f_n) | G \times \mathbb{M})$ with respect to ν^n , the *local Janossy density* [Daley and Vere-Jones, 2003, p. 247]. We further call Ψ *totally finite* if \mathbb{G} is bounded and we see that when Ψ is totally finite, $j_n^G(g_1, \dots, g_n) = j_n^G(g_1, \dots, g_n | G \times \mathbb{M})$ may be treated as a local Janossy density.

Remark 7. *This construction is equivalent to specifying Ψ through a) a family of discrete probability distributions $p_n = \mathbb{P}(\Psi(\mathbb{Y}) = n) = \mathbb{P}(N = n)$, $n \in \mathbb{N}$, and b) a family of symmetric probability measures $\Pi_n(A_1 \times \dots \times A_n) = \frac{J_n(A_1 \times \dots \times A_n)}{p_n n!}$, $n \geq 1$ (so that $J_0(\mathbb{Y}^0) = p_0$) [Daley and Vere-Jones, 2003, Proposition 5.3.II.]. Hereby the density of Π_n with respect to ν^n is $\pi_n(\cdot) = \frac{j_n(\cdot)}{p_n n!}$ so that, conditionally on $N = n$, $\pi_n(\cdot)$ determines the multivariate distribution of the points of Ψ in \mathbb{Y} .*

6.4.1 Densities with respect to Poisson processes

Recalling Section 5.3, assume now that the auxiliary reference measure $\nu_{\mathbb{A}}$ is finite on $\mathcal{B}(\mathbb{A})$. This implies that $\nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}}$ is a finite measure on $\mathcal{B}(\mathbb{M}) = \mathcal{B}(\mathbb{A}) \otimes \mathcal{B}(\mathbb{F})$. Note that $[\nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}}](\mathbb{A} \times \mathbb{F}) = \nu_{\mathbb{A}}(\mathbb{A}) \nu_{\mathbb{F}}(\mathbb{F}) = \nu_{\mathbb{A}}(\mathbb{A})$ and if $\nu_{\mathbb{A}}$ is given by a probability measure, $[\nu_{\mathbb{A}} \otimes \nu_{\mathbb{F}}](\mathbb{A} \times \mathbb{F}) = 1$.

Let further Ψ^* be a finite (ST)CFM Poisson process with ground intensity function λ_* , finite ground intensity measure μ_G^* , i.e. $\mu_G^*(\mathbb{G}) = \int_{\mathbb{G}} \lambda_*(g) dg < \infty$, and intensity functional $\lambda_*(g, l, f) = f_{(g, l)}^{\mathbb{F}}(f) f_g^{\mathbb{A}}(l) \lambda_*(g)$. Denote the distribution of Ψ^* on $\Sigma_{\mathcal{N}_{\mathbb{Y}}}$ by P_* .

Assume next that Ψ is a finite (ST)CFMPP with distribution P on $\Sigma_{\mathcal{N}_{\mathbb{Y}}}$ and assume that $P \ll P_*$, where the corresponding density will be denoted by $p_{\Psi}(\varphi)$, $\varphi \in \mathcal{N}_{\mathbb{Y}}$. We may equivalently and conveniently consider the supports of P and P_* , whereby we may define the density as $p_{\Psi} : \mathcal{N}^f \rightarrow [0, \infty)$, where $\mathcal{N}^f = \{\{(g_1, l_1, f_1), \dots, (g_n, l_n, f_n)\} \subseteq \mathbb{Y} : n < \infty\}$ is the collection of all finite subsets of \mathbb{Y} (see e.g. [Daley and Vere-Jones, 2008, Van Lieshout, 2000, Møller and Waagepetersen, 2004]). Hereby (see e.g. [Van Lieshout, 2000, Chapter 2.3] and [Møller and Waagepetersen, 2004, Chapter 6.6.1]) the density $p_{\Psi}(\cdot)$ satisfies $\int_{\mathcal{N}^f} p_{\Psi}(\mathbf{y}) P_*(d\mathbf{y}) = 1$ and

$$\begin{aligned} \mathbb{P}(\Psi \in R) &= \sum_{n=0}^{\infty} \frac{e^{-\mu_G^*(\mathbb{G})\nu_{\mathbb{A}}(\mathbb{A})\nu_{\mathbb{F}}(\mathbb{M})}}{n!} \int_{\mathbb{Y}^n} \mathbf{1}\{\{(g_1, l_1, f_1), \dots, (g_n, l_n, f_n)\} \in R\} \times \\ &\quad \times p_{\Psi}((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) \prod_{i=1}^n f_{(g_i, l_i)}^{\mathbb{F}}(f_i) f_{g_i}^{\mathbb{A}}(l_i) \lambda_*(g_i) \nu(d(g_i, l_i, f_i)) \end{aligned}$$

for $R \in \Sigma_{\mathcal{N}_{\mathbb{Y}}}$, where

$$\begin{aligned} p_{\Psi}((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) &= e^{\mu_G^*(\mathbb{G})\nu_{\mathbb{A}}(\mathbb{A})\nu_{\mathbb{F}}(\mathbb{M})} j_n((g_1, l_1, f_1), \dots, (g_n, l_n, f_n)) \\ &= e^{\mu_G^*(\mathbb{G})\nu_{\mathbb{A}}(\mathbb{A})\nu_{\mathbb{F}}(\mathbb{M})} f_{(g_1, l_1), \dots, (g_n, l_n)}^{\mathbb{F}}(f_1, \dots, f_n) f_{g_1, \dots, g_n}^{\mathbb{A}}(l_1, \dots, l_n) j_n^G(g_1, \dots, g_n) \end{aligned} \quad (23)$$

through Lemma 3. Note also that the densities satisfy $\mathbb{E}[p_{\Psi}(\Psi \cup \{(g_1, l_1, f_1), \dots, (g_n, l_n, f_n)\})] = \rho^{(n)}((g_1, l_1, u_1), \dots, (g_n, l_n, u_n))$ [Møller and Waagepetersen, 2007, (16)].

6.4.2 Markov (ST)CFMPPs

For (ST)CFMPPs there are many different ways in which one can describe (local) interaction structures. One important specific construction in the current context, when Ψ is finite with density p_{Ψ} , is the class of *Markov (ST)CFMPPs*.

Given a symmetric and reflexive *neighbour relation* \sim on \mathbb{Y} , we say that points $(g_1, l_1, f_1) \in \mathbb{Y}$ and $(g_2, l_2, f_2) \in \mathbb{Y}$ are *neighbours* if $(g_1, l_1, f_1) \sim (g_2, l_2, f_2)$ and we define the *neighbourhood* of (g_1, l_1, f_1) as $\partial(\{(g_1, l_1, f_1)\}) = \{(g, l, f) \in \mathbb{Y} : (g_1, l_1, f_1) \sim (g, l, f)\}$. In addition, we write $\partial_{\Psi}(\{(g_1, l_1, f_1)\}) = \partial(\{(g_1, l_1, f_1)\}) \cap \Psi$.

Definition 18. *Given a symmetric and reflexive relation \sim on \mathbb{Y} , we say that a (ST)CFMPP Ψ is a Markov (ST)CFMPPs if, for all \mathbf{y} such that $p_{\Psi}(\mathbf{y}) > 0$,*

a) $p_{\Psi}(\mathbf{z}) > 0$ for all $\mathbf{z} \subseteq \mathbf{y}$,

b) for all $(g, l, f) \in \mathbb{Y}$, $p_{\Psi}(\mathbf{y} \cup \{(g, l, f)\})/p_{\Psi}(\mathbf{y})$ depends only on (g, l, f) and $\partial_{\Psi}(\{(g_1, l_1, f_1)\})$.

If $(g_1, l_1, f_1) \sim (g_2, l_2, f_2)$ for all $(g_1, l_1, f_1), (g_2, l_2, f_2) \in \mathbf{y}$, we say that \mathbf{y} is a *clique*. By writing $c_{\mathbf{y}}^{\sim} = \{\mathbf{z} \subseteq \mathbf{y} : \mathbf{z} \text{ is a clique under } \sim\}$, the Hammersley-Clifford theorem [Van Lieshout, 2000, (2.7)] states that there is a measurable interaction function $\phi : \mathcal{N}^f \rightarrow [0, \infty)$ such that $p_{\Psi}(\mathbf{y}) = \prod_{\mathbf{z} \in c_{\mathbf{y}}^{\sim}} \phi(\mathbf{z})$ if and only if Ψ is a Markov (ST)CFMPP. This is a more common way of defining a Markov process. We may express b) in the above definition through the Papangelou conditional intensity.

Lemma 4. *A finite (ST)CFMPP with density p_{Ψ} with respect to a Poisson process Ψ^* has a Papangelou conditional intensity, which is given by*

$$\lambda(g, l, f; \mathbf{y}) = \frac{p_{\Psi}(\mathbf{y} \cup \{(g, l, f)\})}{p_{\Psi}(\mathbf{y})}$$

for $(g, l, f) \notin \mathbf{y} \in \mathcal{N}^f$ and $\lambda(g, l, f; \mathbf{y}) = 0$ for $(g, l, f) \in \mathbf{y}$.

Proof. The proof exploits the Campbell-Mecke formula and the Georgii-Nguyen-Zessin formula. It follows the exact steps of [Van Lieshout, 2000, Theorem 1.6], with obvious modifications for the marks. That $\lambda(g, l, f; \mathbf{y}) = 0$ for $(g, l, f) \in \mathbf{y}$ follows since Ψ is simple. \square

For convenience we often write $i \sim j$ if $(g_i, l_i, f_i), (g_j, l_j, f_j) \in \Psi$ are neighbours. Note here that a Markov STCFMPP Ψ may possess several different types of Markovianity:

1. Spatial Markovianity: e.g. $i \sim j$ if $d(X_i, X_j) \leq R$, for some maximum range of interaction $R > 0$.
2. Temporal Markovianity: e.g. $i \sim j$ if $|T_i - T_j| \leq R$, $R > 0$.
3. Spatio-temporal Markovianity: e.g. $i \sim j$ if $(X_j, T_j) \in (X_i, T_i) + C_{R_S}^{R_T}$, $R_S, R_T > 0$ (recall the cylinder sets C_u^v in (1)). For an example, see [Cronie and van Lieshout, 2013].
4. Functional mark Markovianity: see Section 5.2.4.
5. If Ψ is temporally grounded, the cumulative process $\Psi_{\mathbb{T}}$ in (21) may be a Markov process: $\mathbb{P}(\Psi_{\mathbb{T}}(t, K) \in \cdot | \mathcal{F}_s) = P(\Psi_{\mathbb{T}}(t, K) \in \cdot | \Psi_{\mathbb{T}}(s, K)), s < t$.

Note that for a Markov (ST)CFMPP we may define a type of *local independent marking*. For cliques $I_1 = \{i_1, \dots, i_k\}$ and $I_2 = \{1, \dots, n\} \setminus I_1 = \{i_{k+1}, \dots, i_n\}$, conditionally on Ψ_G , we let the distribution $P_{g_1, \dots, g_n}^{\mathbb{M}}(\cdot)$ of the marks $\{(L_i, M_i)\}_{i=1}^n$ on $(\mathbb{M}^n, \mathcal{B}(\mathbb{M}^n))$ satisfy

$$P_{g_1, \dots, g_n}^{\mathbb{M}}(d(m_1, \dots, m_n)) = P_{g_{i_1}, \dots, g_{i_k}}^{\mathbb{M}}(d(m_{i_1}, \dots, m_{i_k})) P_{g_{i_{k+1}}, \dots, g_{i_n}}^{\mathbb{M}}(d(m_{i_{k+1}}, \dots, m_{i_n})),$$

i.e. the marks are independent if the points are not neighbours.

7 Discretely sampled functional marks

In many situations one might be interested in different characteristics of Ψ when the functional marks $\{M_i\}_{i=1}^N$ are only observed at times $S_k = \{s_1, \dots, s_k\} \subseteq \mathcal{T}$, $k \geq 1$, i.e. instead of whole functional marks $M_i = \{M_i(t)\}_{t \in \mathcal{T}}$ we consider multivariate marks of the form $(M_i(s_1), \dots, M_i(s_k))$, $i = 1, \dots, N$. Recall that this e.g. is the case when we define a classical marked spatio-temporal point process $\bar{\Psi} = \{(X_i, T_i, L_i, M_i(0)) : (X_i, T_i) \in \Psi_G\} = \{(X_i, T_i, L_i, \xi_i) : (X_i, T_i) \in \Psi_G\}$ through a (ST)CFMPP Ψ , i.e. $S_1 = \{0\}$.

7.1 Functional mark sampled point process characteristics

We shall see that the expressions below play a crucial role in the statistical inference since they form the basis for e.g. likelihood functions, when the mark-functions are sampled at discrete times s_1, \dots, s_k .

In order to accommodate such a restriction we consider sets

$$E_{U_i}^{S_k} = \{f \in \mathbb{F} : (f(s_1), \dots, f(s_k)) \in U_i \in \mathcal{B}(\mathbb{R}^k)\}, \quad i = 1, \dots, n, \quad (24)$$

which give rise to sets $A_1, \dots, A_n = (G_1 \times (D_1 \times E_{U_1}^{S_k})), \dots, (G_n \times (D_n \times E_{U_n}^{S_k})) \in \mathcal{B}(\mathbb{Y})$. Recalling the reference stochastic process $X_{\mathbb{F}}$ in (3), which has $\nu_{\mathbb{F}}$ as distribution on $\mathcal{B}(\mathbb{F})$, note that

$$\nu_{\mathbb{F}}(E_{U_i}^{S_m}) = \mathbb{P}((X_{\mathbb{F}}(s_1), \dots, X_{\mathbb{F}}(s_k)) \in U_i).$$

Hence, under this discrete sampling of the functional marks, through the sets $E_{U_i}^{S_k}$ most characteristics are automatically transferred over to the fidis of the functional mark processes. Note e.g. that if the reference process $X_{\mathbb{F}}(t) = W(t)$ is, say, a standard Brownian motion, then $\nu_{\mathbb{F}}(E_{U_i}^{S_m})$ can be calculated explicitly as a product of (Gaussian) transition probabilities.

7.1.1 Functional mark sampled product densities

By considering sets of the form (24), we obtain an expression for the product densities under S_k -sampling, the S_k -mark-sampled product densities.

Lemma 5. *Under the assumptions of Proposition 1 and the existence of densities (17), the n th S_k -mark-sampled product density is given by*

$$\begin{aligned} \rho_{s_1, \dots, s_k}^{(n)}((g_1, l_1, u_1), \dots, (g_n, l_n, u_n)) &= \\ &= f_{(g_1, l_1), \dots, (g_n, l_n)}^{s_1, \dots, s_k}(u_1, \dots, u_n) f_{g_1, \dots, g_n}^{\mathbb{A}}(l_1, \dots, l_n) \rho_G^{(n)}(g_1, \dots, g_n), \end{aligned}$$

where $u_1, \dots, u_n \in \mathbb{R}^k$.

Proof. We may define the factorial moment measure with respect to S_k and $U_1, \dots, U_n \in \mathcal{B}(\mathbb{R}^k)$,

$$\alpha_{S_k}^{(n)}((G_1 \times D_1 \times U_1) \times \dots \times (G_n \times D_n \times U_n)) = \alpha^{(n)}(A_1 \times \dots \times A_n),$$

where $A_1, \dots, A_n = (G_1 \times (D_1 \times E_{U_1}^{S_k})), \dots, (G_n \times (D_n \times E_{U_n}^{S_k})) \in \mathcal{B}(\mathbb{Y})$. We note that by setting each $U_i = \mathbb{R}^k$, we obtain $\alpha_{S_k}^{(n)} = \alpha_G^{(n)}$, hence $\alpha_{S_k}^{(n)} \ll \alpha_G^{(n)}$. The family of regular probabilities on $\mathcal{B}(\mathbb{A}) \otimes \mathcal{B}(\mathbb{R}^{n \times k})$, which are the corresponding Radon-Nikodym derivatives, are clearly absolutely continuous with respect to the probability measures $P^{\mathbb{A}, n}$ in (10). Hereby, since $\rho_G^{(n)}$ and densities (8) and (17) exist,

$$\begin{aligned} \alpha_{S_k}^{(n)}((G_1 \times D_1 \times U_1) \times \dots \times (G_n \times D_n \times U_n)) &= \\ &= \int_{G_1 \times D_1 \times U_1} \dots \int_{G_n \times D_n \times U_n} f_{(g_1, l_1), \dots, (g_n, l_n)}^{s_1, \dots, s_k}(u_1, \dots, u_n) f_{g_1, \dots, g_n}^{\mathbb{A}}(l_1, \dots, l_n) \\ &\quad \times \rho_G^{(n)}(g_1, \dots, g_n) \prod_{i=1}^n \nu_{\mathbb{A}}(dl_i) dg_i du_i \end{aligned}$$

and similarly, by directly assuming that $\alpha_{S_k}^{(n)} \ll \ell \otimes \nu_{\mathbb{A}} \otimes \ell_{nk}$, we find that

$$\begin{aligned} \alpha_{S_k}^{(n)}((G_1 \times D_1 \times U_1) \times \dots \times (G_n \times D_n \times U_n)) &= \\ &= \int_{G_1 \times D_1 \times U_1} \dots \int_{G_n \times D_n \times U_n} \rho_{s_1, \dots, s_k}^{(n)}((g_1, l_1, u_1), \dots, (g_n, l_n, u_n)) \prod_{i=1}^n dg_i \nu_{\mathbb{A}}(dl_i) du_i, \end{aligned}$$

which implies that the two integrands are equal a.e. □

As a consequence, the pair correlation functional is given by

$$g_{\Psi}((g_1, l_1, f_1), (g_1, l_2, f_2)) = \frac{f_{(g_1, l_1), (g_1, l_2)}^{s_1, \dots, s_k}(u_1, u_2) f_{g_1, g_2}^{\mathbb{A}}(l_1, l_2)}{f_{(g_1, l_1)}^{s_1, \dots, s_k}(u_1) f_{(g_1, l_2)}^{s_1, \dots, s_k}(u_2) f_{g_1}^{\mathbb{A}}(l_1) f_{g_2}^{\mathbb{A}}(l_2)} g_G(g_1, g_2).$$

7.1.2 Functional mark sampled conditional intensities

Turning next to conditional intensities of Section 6.3, let $U = U_1 \in \mathcal{B}(\mathbb{R}^k)$ in expression (24), recall Lemma 2 and define

$$\begin{aligned}\lambda_{S_k}^*(t, (x, l, U); \omega) &= \int_{E_U^{S_k}} \lambda^*(t, (x, l, f); \omega) \nu_{\mathbb{F}}(df) \\ &= f_{(x,t)}^{\mathbb{A}}(l; \omega) f_t^{G,S}(x; \omega) \lambda_G^*(t; \omega) \int_{E_U^{S_k}} f_{(x,t,l)}^{\mathbb{F}}(f; \omega) \nu_{\mathbb{F}}(df) \\ &= f_{(x,t)}^{\mathbb{A}}(l; \omega) f_t^{G,S}(x; \omega) \lambda_G^*(t; \omega) P_{S_k}^*(U|x, t, l),\end{aligned}$$

which is well-defined since $\lambda_K^*(t; \omega)$, $K \in \mathcal{B}(\mathbb{M}_{\mathbb{T}})$, exists. If additionally the probability measure $P_{S_k}^*(\cdot|x, t, l)$ is absolutely continuous with respect to $\ell_k(\cdot)$, it follows that

$$\lambda_{S_k}^*(t, (x, l, U); \omega) = \int_U f_{(x,t,l)}^{s_1, \dots, s_m}(u; \omega) f_{(x,t)}^{\mathbb{A}}(l; \omega) f_t^{G,S}(x; \omega) \lambda_G^*(t; \omega) du.$$

We refer to the integrand

$$\lambda_{S_k}^*(t, (x, l, u); \omega) = f_{(x,t,l)}^{s_1, \dots, s_m}(u; \omega) f_{(x,t)}^{\mathbb{A}}(l; \omega) f_t^{G,S}(x; \omega) \lambda_G^*(t; \omega) \quad (25)$$

as the S_k -mark-sampled conditional intensity.

7.1.3 Functional mark sampled Papangelou conditional intensities

Recall the E -functional marked Papangelou conditional intensities $\lambda_E(g, l; \Psi)$ in Section 4.3. In a similar fashion, one may define the marked S_k -mark-sampled Papangelou kernel

$$\lambda^{S_k}(g, l, U; \Psi) = \lambda_{E_U^{S_k}}(g, l; \Psi), \quad U \in \mathcal{B}(\mathbb{R}^k),$$

through the set $E_U^{S_k} \in \mathcal{B}(\mathbb{F})$. Note that $\lambda^{S_k}(g, l, U; \Psi)$ may be viewed as the infinitesimal conditional probability of finding a point of Ψ in $dg \subseteq \mathbb{G}$ with auxiliary mark in $dl \subseteq \mathbb{A}$ and functional mark M_i such that $(M_i(s_1), \dots, M_i(s_k)) \in U$, given all points of Ψ located outside dg . Conditionally on the i th point of Ψ having spatial(-temporal) location g and auxiliary mark l , and conditionally on all points $j \neq i$ of Ψ , assume that $\lambda^{S_k}(g, l, \cdot; \Psi)$ is absolutely continuous with respect to the distribution of $(M_i(s_1), \dots, M_i(s_k))$ and assume that this distribution, in turn, is absolutely continuous with respect to ℓ_k . Then the corresponding density (see expression (17)) is given by $f_{(g,l)}^{s_1, \dots, s_m}(u)$, $u \in \mathbb{R}^m$, and consequently

$$\lambda^{S_m}(g, l, U; \Psi) = \int_U \lambda^{S_m}(g, l, du; \Psi) = \int_U f_{(g,l)}^{s_1, \dots, s_m}(u) \lambda^{S_m}(g, l, u; \Psi) du, \quad (26)$$

where we note that $f_{(g,l)}^{s_1, \dots, s_m}(u) \lambda^{S_m}(g, l, u; \Psi) du$ may be interpreted as $\lambda^{S_m}(g, l, U; \Psi)$ with U given by the infinitesimal region $du \subseteq \mathbb{R}^k$.

7.1.4 Functional mark sampled Janossy densities

A further important entity for the likelihood inference is the set of Janossy densities. By combining Lemma 3 and Lemma 5, we find that the n th S_k -sampled Janossy density is given by

$$\begin{aligned}j_n^{S_k}((g_1, l_1, u_1), \dots, (g_n, l_n, u_n)) &= \\ &= f_{(g_1, l_1), \dots, (g_n, l_n)}^{s_1, \dots, s_k}(u_1, \dots, u_n) f_{g_1, \dots, g_n}^{\mathbb{A}}(l_1, \dots, l_n) j_n^G(g_1, \dots, g_n),\end{aligned} \quad (27)$$

where $u_i = (u_{i1}, \dots, u_{ik}) \in \mathbb{R}^k$, $i = 1, \dots, n$. In addition, from (23) we obtain that under the S_k sampling the density of a finite (ST)CFMPP with respect to a Poisson process is given by

$$p_{\Psi}^{S_k}((g_1, l_1, u_1), \dots, (g_n, l_n, u_n)) = e^{\mu_G^*(\mathbb{G})\nu_{\mathbb{A}}(\mathbb{A})\nu_{\mathbb{F}}(\mathbb{M})} j_n^{S_k}((g_1, l_1, u_1), \dots, (g_n, l_n, u_n)). \quad (28)$$

7.2 Statistical inference

We now turn to the statistical inference for (ST)CFMPPs with discretely sampled functional marks. Specifically, we consider a STCFMPP $\Psi = \{(X_i, T_i, L_i, M_i)\}_{i=1}^N$ with distribution P_{θ_0} , which belongs to some parametric family of models $\{P_{\theta} : \theta \in \Theta\}$, and we here indicate how a few estimation schemes may be defined to obtain an estimate $\hat{\theta}$ of θ_0 .

Note here that when we observe all of Ψ , but with functional marks sampled according to $S_k = \{s_1, \dots, s_k\} \subseteq \mathcal{T}$, and wish to find an estimate $\hat{\theta}$, we essentially consider the usual marked spatio-temporal point process setup, where the marked point process under consideration is given by $\bar{\Psi} = \{((X_i, T_i), (L_i, M_i(s_1), \dots, M_i(s_k)))\}_{i=1}^N$. Assume that we observe $\bar{\Psi}$ as $\{((x_i, t_i), (l_i, u_i))\}_{i=1}^n$, $u_i = (u_{i1}, \dots, u_{ik})$, within some compact space-time region $W_S \times W_T \subseteq \mathbb{G}$. Given the imposed probabilistic structures of the ground process and the marking, it may be possible to derive e.g. a (pseudo)likelihood function. The literature on the subject is vast and good accounts can be found in e.g. [Daley and Vere-Jones, 2003, 2008, Diggle, 2003, Gelfand et al., 2010, Illian et al., 2008, Van Lieshout, 2000, Møller and Waagepetersen, 2004]. A quick indication of how one could proceed is as follows.

- When Ψ is assumed to be spatially grounded (see Section 6.3), by consulting e.g. [Daley and Vere-Jones, 2003] and recalling (25), we find that the log-likelihood function is given by

$$\begin{aligned} \log L(\theta) &= \sum_{i=1}^n \log \lambda_{S_k}^*(t_i, (x_i, l_i, u_i)) - \int_{W_T} \int_{W_S} \int_{\mathbb{A}} \int_{\mathbb{R}^k} \lambda_{S_k}^*(t, (x, l, u)) d\nu_{\mathbb{A}}(dl) dx dt \\ &= \sum_{i=1}^n \log f_{(x_i, t_i, l_i)}^{s_1, \dots, s_m}(u_i) + \log f_{(x_i, t_i)}^{\mathbb{A}}(l_i) + \log f_{t_i}^{G, S}(x_i) + \log \lambda_G^*(t_i) \\ &\quad - \int_{W_T} \int_{W_S} f_t^{G, S}(x) \lambda_G^*(t) dx dt. \end{aligned}$$

When we maximise $\log L(\theta)$ with respect to $\theta \in \Theta$, we obtain a maximum likelihood estimate $\hat{\theta}$. This is probably the most well-known likelihood estimation procedure for point processes.

- When Ψ is finite (and not temporally grounded), following e.g. [Van Lieshout, 2000, Chapter 3.7], through expressions (27) and (28), we obtain the likelihood function as

$$\begin{aligned} L(\theta) &= p_{\Psi}^{S_k}((x_1, t_1, l_1, u_1), \dots, (x_n, t_n, l_n, u_n)) \\ &\propto j_N^{S_k}((x_1, t_1, l_1, u_1), \dots, (x_n, t_n, l_n, u_n)) \\ &= f_{(x_1, t_1, l_1), \dots, (x_n, t_n, l_n)}^{s_1, \dots, s_k}(u_1, \dots, u_n) f_{(x_1, t_1), \dots, (x_n, t_n)}^{\mathbb{A}}(l_1, \dots, l_n) j_n^G((x_1, t_1), \dots, (x_n, t_n)). \end{aligned}$$

By maximising $L(\theta)$ or $\log L(\theta)$ with respect to $\theta \in \Theta$, we obtain a maximum likelihood estimate $\hat{\theta}$.

- When Ψ is a Markov STCFMPP as defined in Section 6.4.2, it is completely specified by its Papangelou conditional intensity. Hereby, through e.g. [Van Lieshout, 2000, Chapter 3.8],

by recalling expression (26), we find that the pseudo-likelihood function is given by

$$PL(\theta) = \exp \left\{ - \int_{W_S \times W_T \times \mathbb{A}} \lambda^{S_m}((x, t), l, \mathbb{R}^k; \{(x_i, t_i, l_i, u_i)\}_{i=1}^n) [\ell \otimes \nu_{\mathbb{A}}](d(x, t, l)) \right\} \\ \times \prod_{i=1}^n f_{(x_i, t_i, l_i)}^{s_1, \dots, s_m}(u_i) \lambda^{S_m}(x_i, t_i, l_i, u_i; \{(x_i, t_i, l_i, u_i)\}_{i=1}^n)$$

and upon maximising $PL(\theta)$ or $\log PL(\theta)$ with respect to $\theta \in \Theta$, we obtain a maximum pseudo-likelihood estimate $\hat{\theta}$.

- For a finite Ψ , assume that we are able to estimate the parameters of the ground process and the auxiliary marks in some suitable way. Conditionally on $\{(x_i, t_i, l_i)\}_{i=1}^n$, given some parametric family $\{f(t; \theta, n) = (f_1(t; \theta), \dots, f_n(t; \theta)) : t \in \mathcal{T}, \theta \in \Theta\} \subseteq \mathbb{F}^n$, $n \geq 1$, of L^2 -functions, when we sample the marks according to S_k , with a slight abuse of notation, we obtain the least-squares estimator

$$\begin{aligned} \hat{\theta} &= \arg \min_{\theta \in \Theta} \int_{\mathcal{T}} \sum_{j=1}^k \delta_{s_j}(t) [M_I(t | \Psi_G, \{L_i\}_{i=1}^N) - f(t; \theta, n)]^2 dt \\ &\approx \arg \min_{\theta \in \Theta} \int_{\mathcal{T}} \sum_{j=1}^k \delta_{s_j}(t) [M_I(t | \{(x_i, t_i, l_i)\}_{i=1}^n) - f(t; \theta, n)]^2 dt \\ &= \arg \min_{\theta \in \Theta} \sum_{j=1}^k \sum_{i=1}^n [(u_{ij} - f_i(s_j; \theta, n))^2], \end{aligned}$$

where we recall M_I , $I = \{1, \dots, n\}$, from (18). The second (approximate) equality follows since the observed marks are likely influenced by points falling outside $W_S \times W_T$. Note hereby that this approach does not account for edge effects. To correct for edge effect in a setting as the one above, [Cronie and Särkkä, 2011] suggested a few different approaches. The general idea is to, successively and conditionally on $\{((x_i, t_i), (l_i, u_i))\}_{i=1}^n$, simulate realisations on a torus \mathbb{G} , with $W_S \times W_T \subseteq \mathbb{G}$, from the parametric model under consideration and then use all simulated marked points falling in $\mathbb{G} \setminus (W_S \times W_T)$ as the missing data in $\mathbb{G} \setminus (W_S \times W_T)$. Assuming that there are $n^* \geq 0$ simulated points falling in $\mathbb{G} \setminus (W_S \times W_T)$, the estimator above is adjusted by considering some parametric family $\{f(t; \theta, n + n^*) : t \in \mathcal{T}, \theta \in \Theta\} \subseteq \mathbb{F}^{n+n^*}$ of functions and the estimator $\hat{\theta}$ becomes exact.

Note above that if the M_i 's are Markov processes with existing transition densities, as described in Section 5.2.4, then $f_{(x_1, t_1, l_1), \dots, (x_n, t_n, l_n)}^{s_1, \dots, s_k}(u_1, \dots, u_n)$, $n \geq 1$, are given by products of the transition densities found in expression (19).

7.3 Observable processes, thinning and parameter estimation

The above sampling structure, with spatio-temporally continuously sampled points and discretely sampled marks, may be reasonable in certain situations. However, one may argue that if the functional marks are sampled discretely, also the spatio-temporal ground process and possibly the auxiliary marks should be sampled according to whether we are able to observe them at the sample times $S_k = \{s_1, \dots, s_k\}$. Below we discuss such a setup and indicate how the associated statistical inference could be performed.

In order to provide a mathematical structure which reflects such a sampling scheme, for a STCFMPP Ψ consider the (modified) evaluation functional

$$\pi_t^{\mathbb{F}}[\varphi] = \{(x, t, l, f(t)) : (x, t, l, f) \in \varphi\}, \quad \varphi \in \mathcal{N}_{\mathbb{Y}}, t \in \mathcal{T}.$$

When considering certain types of real data, such as in the sampling described above, it is useful to consider the *observable process*

$$\begin{aligned} \Psi_O^{\mathbb{Y}}(t) &= \pi_t^{\mathbb{F}}[\Psi \cap (\mathbb{X} \times \mathbb{T} \times \mathbb{A} \times \{f \in \mathbb{F} : t \in \text{supp}(f)\})] \\ &= \pi_t^{\mathbb{F}}[\{(X_i, T_i, L_i, M_i) \in \Psi : t \in \text{supp}\{M_i\}\}] \\ &= \{(X_i, T_i, L_i, M_i(t)) : (X_i, T_i, L_i, M_i) \in \Psi, t \in \text{supp}\{M_i\}\}. \end{aligned}$$

Note that this corresponds to *what we observe* at time t , i.e. we treat a point of Ψ as observable/present at time t only if its functional mark is non-zero. The definition is analogous for CFMPPs.

Consider next the scenario where we sample the observable process at times s_1, \dots, s_k . A typical data set which corresponds to this type of sampling could be e.g. a forest stand, where each X_i represents the location of a tree, each T_i its birth time, and each $(M_i(s_1), \dots, M_i(s_k))$ the sizes (e.g. radius or height) of the i th tree, measured at s_1, \dots, s_m . We note that what we in fact are observing is a location-dependent thinning of Ψ , i.e. for some measurable function $p : \mathbb{Y} \rightarrow [0, 1]$, each point (x, t, l, f) of Ψ is retained with the probability $p(x, t, l, f)$. The product densities of a thinned STCFMPP [Daley and Vere-Jones, 2008] are given by

$$\begin{aligned} \rho_{\text{th}}^{(n)}((x_1, t_1, l_1, f_1), \dots, (x_n, t_n, l_n, f_n)) &= \\ &= \rho^{(n)}((x_1, t_1, l_1, f_1), \dots, (x_n, t_n, l_n, f_n)) \prod_{i=1}^n p(x_i, t_i, l_i, f_i). \end{aligned}$$

The correct choice of retention probability function to reflect the nature of the S_k -evaluated observable process is

$$p(x, t, l, f) = p(x, t, l, f; S_k) = \mathbf{1}\{S_k \cap \text{supp}(f) \neq \emptyset\}, \quad S_k = \{s_1, \dots, s_k\},$$

Note that if the support of the i th functional mark of Ψ is given by $\text{supp}(M_i) = [t, t + l)$ when $(X_i, T_i, L_i, M_i) = (x, t, l, f)$, we set $\text{supp}(f) = [t, t + l)$. Hence, we remove all unobserved points which we do not observe at any of the sample times s_1, \dots, s_k .

Turning now to the characteristics corresponding to $(\Psi_O^{\mathbb{Y}}(s_1), \dots, \Psi_O^{\mathbb{Y}}(s_1))$, consider first the S_k -marked-sampled product densities in Lemma 5, which are needed since we also sample the functional marks discretely. We find that their thinned versions are given by

$$\begin{aligned} \rho_{S_k \text{th}}^{(n)}((x_1, t_1, l_1, u_1), \dots, (x_n, t_n, l_n, u_n)) &= \\ &= f_{(x_1, t_1, l_1), \dots, (x_n, t_n, l_n)}^{s_1, \dots, s_k}(u_1, \dots, u_n) f_{(x_1, t_1), \dots, (x_n, t_n)}^{\mathbb{A}}(l_1, \dots, l_n) \rho_G^{(n)}((x_1, t_1), \dots, (x_n, t_n)) \\ &\times \prod_{i=1}^n \mathbf{1}\{S_k \cap \text{supp}(f_i) \neq \emptyset\}. \end{aligned}$$

Under the assumptions of Proposition 1, when the product densities of all orders $n \geq 1$ exist,

through [Daley and Vere-Jones, 2003, Lemma 5.4.III.] we find that

$$\begin{aligned}
& j_n^{S_k \text{th}}((x_1, t_1, l_1, u_1), \dots, (x_n, t_n, l_n, u_n)) = \\
& = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} \int_{\mathbb{X} \times \mathbb{T} \times \mathbb{A} \times \mathbb{R}^k} \cdots \int_{\mathbb{X} \times \mathbb{T} \times \mathbb{A} \times \mathbb{R}^k} \times \\
& \times \rho_{S_k \text{th}}^{(n+j)}((x_1, t_1, l_1, u_1), \dots, (x_n, t_i, l_n, u_n), y_1, \dots, y_j) \prod_{i=1}^j [\ell_d \otimes \ell_1 \otimes \nu_{\mathbb{A}} \otimes \ell_k](dy_i) \\
& = \prod_{i=1}^n \mathbf{1}\{S_k \cap \text{supp}(f_i) \neq \emptyset\} f_{(x_1, t_1, l_1), \dots, (x_n, t_n, l_n)}^{s_1, \dots, s_k}(u_1, \dots, u_n) f_{(x_1, t_1), \dots, (x_n, t_n)}^{\mathbb{A}}(l_1, \dots, l_n) \times \\
& \times \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} \int_{\mathbb{G}} \cdots \int_{\mathbb{G}} \rho_G^{(n+j)}((x_1, t_1), \dots, (x_n, t_n), g_1, \dots, g_j) \prod_{i=1}^j dg_i \\
& = f_{(x_1, t_1, l_1), \dots, (x_n, t_n, l_n)}^{s_1, \dots, s_k}(u_1, \dots, u_n) f_{(x_1, t_1), \dots, (x_n, t_n)}^{\mathbb{A}}(l_1, \dots, l_n) j_n^G(g_1, \dots, g_n) \times \\
& \times \prod_{i=1}^n \mathbf{1}\{S_k \cap \text{supp}(f_i) \neq \emptyset\}.
\end{aligned}$$

Hence, the thinned S_k -sampled Papangelou conditional intensity $\lambda^{S_m}(x_i, t_i, l_i, u_i; \Psi)$ is given by $\mathbf{1}\{S_k \cap \text{supp}(f_i) \neq \emptyset\} \lambda^{S_m}(x_i, t_i, l_i, u_i; \Psi)$.

By further integrating out over \mathbb{A} and/or \mathbb{T} in the expressions above, we obtain equivalences for

$$\begin{aligned}
\Psi_O^{\mathbb{X} \times \mathbb{F}}(t) &= \text{proj}_{\mathbb{X} \times \mathbb{F}}(\Psi_O^{\mathbb{Y}}(t)) = \{(X_i, M_i(t)) : (X_i, T_i, L_i, M_i(t)) \in \Psi_O^{\mathbb{Y}}(t)\} \\
&= \{(X_i, M_i(t)) : (X_i, T_i, L_i, M_i) \in \Psi, t \in \text{supp}\{M_i\}\}, \\
\Psi_O^{\mathbb{X} \times \mathbb{A} \times \mathbb{F}}(t) &= \text{proj}_{\mathbb{X} \times \mathbb{A} \times \mathbb{F}}(\Psi_O^{\mathbb{Y}}(t)) = \{(X_i, L_i, M_i(t)) : (X_i, T_i, L_i, M_i) \in \Psi, t \in \text{supp}\{M_i\}\},
\end{aligned}$$

where e.g. $\text{proj}_{\mathbb{X} \times \mathbb{F}}(A)$ denotes the projection of $A \subseteq \mathbb{X} \times \mathbb{T} \times \mathbb{A} \times \mathbb{F}$ onto $\mathbb{X} \times \mathbb{F}$. Note that such a setup might be more realistic in most cases, since e.g. in a forest stand we do not actually observe the birth times T_i , $i = 1, \dots, N$, of the trees.

Acknowledgements

The authors are truly grateful to N.M.M. van Lieshout (CWI, The Netherlands) for feedback, ideas and proofreading. The authors are also grateful to Aila Särkkä (Chalmers University of Technology, Sweden) and Eric Renshaw (University of Strathclyde, U.K.) for ideas, discussions and for introducing us to concepts underlying the field. This research was supported by the Netherlands Organisation for Scientific Research NWO (613.000.809) and the Spanish Ministry of Education and Science (NTN2010-14961).

References

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

- D. Baum and V.V. Kalashnikov. Stochastic models for communication networks with moving customers. *Information Processes*, 1:1–23, 2001.
- P. Billingsley. *Convergence of Probability Measures*. Wiley, 2nd edition, 1999.
- V. Bogachev. *Measure Theory*. Springer, 2007.
- L.B. Collins and N. Cressie. Analysis of spatial point patterns using bundles of product lisa function. *Journal of Agricultural, Biological, and Environmental Statistics*, 6:118–135, 2001.
- C. Comas. Modelling forest regeneration strategies through the development of a spatio-temporal growth interaction model. *Stochastic Environmental Research and Risk Assessment*, 23:1089–1102, 2009.
- C. Comas, P. Delicado, and Mateu J. A second order approach to analyse spatial point patterns with functional marks. *Test*, 20:503–523, 2011.
- D. Cox and V. Isham. *Point processes*. CRC, 1980.
- N.A.C. Cressie. *Statistics for Spatial Data*. Wiley, revised edition, 1993.
- R. Cr  t  , B. Pumo, S. Soubeyrand, F. Didelot, and Caffier V. A continuous time-and-state epidemic model fitted to ordinal categorical data observed on a lattice at discrete times. *Journal of Agricultural, Biological, and Environmental Statistics*, 18:538–555, 2013.
- O. Cronie. Likelihood inference for a functional marked point process with cox-ingersoll-ross process marks. *arXiv*, 2012.
- O. Cronie and A. S  rkk  . Some edge correction methods for marked spatio-temporal point process models. *Computational Statistics & Data Analysis*, 55:2209–2220, 2011.
- O. Cronie and M.N.M. van Lieshout. A J -function for inhomogeneous spatio-temporal point processes. *arXiv*, 2013.
- O. Cronie, K. Nystr  m, and J. Yu. Spatiotemporal modeling of swedish scots pine stands. *Forest Science*, 59:505–516, 2013.
- D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*. Springer Series in Statistics, 2nd edition, 2003.
- D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes: General Theory and Structure*. Springer, 2nd edition, 2008.
- P. Delicado, R. Giraldo, C. Comas, and J. Mateu. Statistics for spatial functional data: some recent contributions. *Environmetrics*, 21:224–239, 2010.
- P. Diggle. *Statistical Analysis of Spatial Point Patterns*. Hodder Education Publishers – Arnold, 2nd edition, 2003.
- S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley-Interscience, 1986.

- E. Gabriel and P. Diggle. Second-order analysis of inhomogeneous spatio-temporal point process data. *Statistica Neerlandica*, 63:43–51, 2009.
- A. Gelfand, M. Fuentes, P. Guttorp, and P. Diggle. *Handbook of Spatial Statistics*. Taylor & Francis, 2010.
- R. Giraldo, P. Delicado, and J. Mateu. Continuous time-varying kriging for spatial prediction of functional data: An environmental application. *Journal of Agricultural, Biological, and Environmental Statistics*, 15:66–82, 2010.
- R. Giraldo, P. Delicado, and J. Mateu. Ordinary kriging for function-valued spatial data. *Environmental and Ecological Statistics*, 18:411–426, 2011.
- P. Grabarnik, M. Myllymaki, and D. Stoyan. Correct testing of mark independence for marked point patterns. *Ecological Modelling*, 222:3888–3894, 2011.
- P.R. Halmos. *Measure Theory*. Springer, 1974.
- L.P. Ho and D. Stoyan. Modelling marked point patterns by intensity-marked cox processes. *Statistics & Probability Letters*, 78:1194–1199, 2008.
- J. Illian, E. Benson, J. Crawford, and H. Staines. *Principal component analysis for spatial point processes – assessing the appropriateness of the approach in an ecological context*, pages 135–150. Springer, 2006.
- J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan. *Statistical Analysis and Modelling of Spatial Point Patterns*. Wiley, 2008.
- J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, 1987.
- A. Karr. *Point Processes and Their Statistical Inference*. CRC, 2nd edition, 1991.
- F.C. Klebaner. *Introduction to Stochastic Calculus with Applications*. Imperial College Press, 2nd edition, 2005.
- M.N.M. van Lieshout. *Markov Point Processes and Their Applications*. Imperial College Press, London, 2000.
- M.N.M. van Lieshout. A J -function for marked point patterns. *Annals of the Institute of Statistical Mathematics*, 58:235–259, 2006.
- J. Mateu, G. Lorenzo, and E. Porcu. Detecting features in spatial point processes with clutter via local indicators of spatial association. *Journal of Computational and Graphical Statistics*, 16: 968–990, 2007.
- J. Mateu, G. Lorenzo, and E. Porcu. Detecting features in spatial point processes with clutter via local indicators of spatial association. *Journal of Computational and Graphical Statistics*, 16: 968–990, 2008.
- J. Møller and M. Ghorbani. Aspects of second-order analysis of structured inhomogeneous spatio-temporal point processes. *Statistica Neerlandica*, 66:472–491, 2012.

- J. Møller and R.P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall/CRC Press, 2004.
- J. Møller and R.P. Waagepetersen. Modern statistics for spatial point processes. *Scandinavian Journal of Statistics*, 34:643–684, 2007.
- J. Møller, A.R. Syversveen, and R.P. Waagepetersen. Log gaussian cox processes. *Scandinavian Journal of Statistics*, 25:451–482, 1998.
- P. Mörters and Y. Peres. *Brownian motion*. Cambridge University Press, 2010.
- M. Myllymäki and A. Penttinen. Conditionally heteroscedastic intensity-dependent marking of log gaussian cox processes. *Statistica Neerlandica*, 63:450–473, 2009.
- Y. Ogata. Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, 50:379–402, 1998.
- J.O. Ramsay and B.W. Silverman. *Applied Functional Data Analysis*. Springer, 2002.
- J.O. Ramsay and B.W. Silverman. *Functional Data Analysis*. Springer, 2nd edition, 2005.
- E. Renshaw and C. Comas. Space-time generation of high intensity patterns using growth-interaction processes. *Statistics and Computing*, 19:423–437, 2009.
- E. Renshaw and A. Särkkä. Gibbs point processes for studying the development of spatial-temporal stochastic processes. *Computational Statistics & Data Analysis*, 36:85–105, 2001.
- E. Renshaw, C. Comas, and J. Mateu. Analysis of forest thinning strategies through the development of space-time growth-interaction simulation models. *Stochastic Environmental Research and Risk Assessment*, 23:275–288, 2009.
- B.D. Ripley. *Spatial Statistics*. Wiley, 1981.
- B.D. Ripley and A.I. Sutherland. Finding spiral structures in images of galaxies. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, 332:477–485, 1990.
- F.J. Rodríguez-Cortés, M. Ghorbani, and J. Mateu. On the expected value and variance for an estimator of the spatio-temporal product density function. *Submitted*, 2014.
- A. Särkkä and E. Renshaw. The analysis of marked point patterns evolving through space and time. *Computational Statistics & Data Analysis*, 51:1698–1718, 2006.
- M. Schlather, Ribeiro Jr., P.J., and P.J. Diggle. Detecting dependence between marks and locations of marked point processes. *Journal of the Royal Statistical Society B*, 66:79–93, 2004.
- R. Schneider and W. Weil. *Stochastic and Integral Geometry*. Springer, 2008.
- F.P. Schoenberg. Consistent parametric estimation of the intensity of a spatial-temporal point process. *JSPI*, 128:79–93, 2004.
- F.P. Schoenberg. *Introduction to Point Processes*. Wiley, 2011.

- F.P. Schoenberg and K.E. Tranbarger. Description of earthquake aftershock sequences using prototype point processes. *Environmetrics*, 19:271–286, 2008.
- S.D. Silvestrov. *Limit Theorems for Randomly Stopped Stochastic Processes*. Springer, 2004.
- A.V. Skorohod. On the densities of probability measures in functional spaces. *Proc. Fifth Berkeley Symp. on Math. Statist. and Prob.*, 2:163–182, 1967.
- D. Stoyan and H. Stoyan. *Fractals, Random Shapes and Point Fields*. Wiley, 1994.
- D. Stoyan, W. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. John Wiley & sons, West Sussex, 2nd edition, 1995.
- K.E. Tranbarger and F.P. Schoenberg. On the computation and application of point process prototypes. *Open Applied Informatics Journal*, 4:1–9, 2010.
- D. Vere-Jones. Some models and procedures for space-time point processes. *Environmental and Ecological Statistics*, 16:173–195, 2009.