Université d'Ottawa · University of Ottawa

POINT PROCESSES: DISTRIBUTIONS, PARTIAL ORDERS AND COMPENSATORS

By
Mathieu Plante, B.Sc.
December 2001

A Thesis
submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy in Mathematics[1]

[1]The Ph.D. Program is a joint program with Carleton University, administered by the Ottawa-Carleton Institute of Mathematics and Statistics

0-612-67984-5

Dedication

À la mémoire de Jeanette Cadieux (1916-1993).

Acknowledgements

My doctoral undertaking was made under exceptionally favourable circumstances, with considerable resources at my disposal. I would like to express my most profound gratitude to my supervisors, Dr. André Dabrowski and Dr. Gail Ivanoff, for their diligence, encouragements and insights; to Dr. Ely Merzbach, for patiently explaining one of his articles to me; to the secretarial and administrative staff of the Univerisity of Ottawa's Department of Mathematics and Statistics, for providing an extremely pleasant working environment and for bearing my absentmindedness with kindness and good humour; to the Universty of Ottawa itself, for providing me with scholarships and paid employment; to my colleagues Dean Slonowsky, Gilles Lamothe, Cynthia Bocci, Sarah Sumner, Marie-Claude Bonneau and many others, for their help in technical and mathematical matters; and, last but not least, to my family and friends, for all the moral support they gave me throughout this journey.

Abstract

Jacod ([23]) established that the compensator of a simple point process on $\mathbf{R}_+$, when taken with respect to the point process' internal history, exists as an essentially unique predictable increasing process which determines the point process' distribution. In the present thesis, we endeavour to infer other distributional properties of a point process from its compensator. Specifically, regarding point processes on $\mathbf{R}_+$, we show that the compensator, under appropriate assumptions, (i) determines a sequence of "locally Cox" point processes of discrete support which approximate the original point process' distribution, (ii) determines the stochastic order of two point process distributions with respect to three known partial orders on a certain space of point process realizations, and (iii) determines the association of a point process under any one of the same three partial orders. For the purposes of points (ii) and (iii), we develop a tool called a "representation map", which enables one to infer important distributional properties of random elements of a partially ordered Polish space by "representing" these elements as random sequences of $\bar{\mathbf{R}}_+^\infty$. Regarding point processes on the quadrant $\mathbf{R}_+^2 := [0, \infty) \times [0, \infty)$, we define the compensator as a *family* of compensators on $\mathbf{R}_+$ induced by the planar point process, and show that, under the assumption of strict simplicity and mild regularity conditions, this family exists, is essentially unique, and characterizes the planar point process' distribution - thus generalizing Jacod's result. As a subsidiary result, we develop a regenerative form for the compensator of the *non-simple*, marked point process on $\mathbf{R}_+$, generalizing Jacod's formula ([23]: Proposition 3.1) for the compensator of the simple, marked point process.

Contents

Chapter 1

Introduction

A thesis on the topic of point processes - which aims to study several distinct types of "point processes" - could hardly achieve coherence without a preliminary discussion of the kinds of point processes it purports to investigate. It is therefore important, prior to any mention of results, to make explicit the various meanings that the phrase "point process" shall take at different points of this thesis.

In its most general form, a point process on a topological space E is a random distribution of points on that space. If E is endowed with a (partial) order, the point process may also be viewed as a random (counting) measure on E, i.e. as an E-indexed stochastic process taking values in the set of extended nonnegative integers, or as a random cumulative distribution function on E with integral "jump points". To the reader interested in the general theory of point processes (and other random measures) on abstract topological spaces, we recommend [28].

The present thesis restricts its scope of investigation to point processes either on the half-line $\mathbf{R}_+ := [0, \infty)$ or on the quadrant $\mathbf{R}_+^2 := [0, \infty) \times [0, \infty)$. Both $\mathbf{R}_+$ and $\mathbf{R}_+^2$ are endowed with natural partial orders which allow three characterizations of point processes on those spaces - as random measures, as stochastic processes and as random cumulative distribution functions - to yield interesting results. Accordingly, a preliminary division of this document would group parts pertaining to $\mathbf{R}_+$-indexed point processes (namely: Parts I and II) on the one hand, and the part concerned with point processes on $\mathbf{R}_+^2$ (i.e. Part III) on the other.

<u>Point Processes on $\mathbf{R}_+$</u>

The most general type of $\mathbf{R}_+$-indexed point process studied in this document is found in Chapter 3, where it is generically referred to as "the random measure μ". This point process is *marked*, *non-simple* and possibly *explosive* - characteristics which $\mathbf{R}_+$-indexed point processes examined in the other chapters do not share. Let us briefly explain how Chapter 3's "random measure μ" is constructed, so as to identify the differences which distinguish it from the other $\mathbf{R}_+$-indexed point processes encountered in Chapters 2 and 5 of Parts I and II (Chapter 4 does not pertain directly to point processes; we shall explain its *raison d'être* below).

As Section 3.1 explains in greater detail, the "random measure μ" is one which counts random pairs (T_n, X_n) of the product space $(0, \infty) \times E$, where E is assumed to be a Lusin topological space (see Appendix D). More precisely, we assume the existence of a sequence of positive, extended real random variables

$$T_1, T_2, \ldots : (\Omega, \mathcal{F}, P) \longrightarrow (\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+))$$

(where $\mathcal{B}(\bar{\mathbf{R}}_+)$ denotes the Borel class of the Alexandroff compactification $\bar{\mathbf{R}}_+ := [0, \infty]$ of $\mathbf{R}_+$) ordered in such a fashion that for any $\omega \in \Omega$,

i) $0 < T_1(\omega) \leq T_2(\omega) \leq \ldots \leq T_\infty(\omega) := \sup_{n \in \mathbf{N}} T_n(\omega) \leq \infty$, and

ii) for any $n \in \mathbf{N}$, if $T_n(\omega) < \infty$ then $T_n(\omega) < T_\infty(\omega)$.

In other words, the T_n's are increasing in n and must not attain their supremum if the latter is finite. As for the random elements $X_1, X_2, \ldots$, we first compactify the Lusin topological space E into $\bar{E}$, where $\bar{E} := E \cup \{\Delta\}$ is obtained by adjoining an element $\Delta \notin E$ to E. Denoting the Borel classes of E, $\bar{E}$ and $(0, \infty)$ by $\mathcal{E}$, $\bar{\mathcal{E}}$ and $\mathcal{B}((0, \infty))$ respectively, we let

$$X_n : (\Omega, \mathcal{F}, P) \longrightarrow (\bar{E}, \bar{\mathcal{E}})$$

be such that for any $\omega \in \Omega$, $X_n(\omega) = \Delta$ if and only if $T_n(\omega) = \infty$. Finally, for any $\omega \in \Omega$ and for any set B of the product σ-algebra $\mathcal{B}((0,\infty)) \otimes \mathcal{E}$, we let

$$\mu(\omega, B) := |\{n \in \mathbf{N} : (T_n(\omega), X_n(\omega)) \in B\}|.$$

A few terminological clarifications are now in order.

- If $0 < T_1(\omega) < T_2(\omega) < \ldots < T_n(\omega) < \ldots$ holds for all $\omega \in \Omega$, then μ is called a *simple point process*; otherwise μ is said to be *non-simple.*
- If $T_\infty(\omega) = \infty$ for all $\omega \in \Omega$, then μ is called a *non-explosive point process*; otherwise μ is said to be *explosive.*
- If $E \equiv \mathbf{R}_+$ and $X_n \equiv T_n$ for all $n \in \mathbf{N}$, whereby μ is identified with a random measure on $(0,\infty)$, then μ is called an *unmarked point process*; otherwise, μ is said to be *marked* with *mark space E* and *marks* $X_n, n \in \mathbf{N}$.

In contrast to Chapter 3's "random measure μ", which is constructed in full generality as a (possibly) explosive, marked and non-simple point process, the point processes under consideration in Chapters 2 and 5 are all simple, unmarked and non-explosive. When such restrictions are applied to a given point process, we view the latter as a *random element* of a certain topological space of *measures* on $\mathbf{R}_+$. Let us specify this space at once. The symbol $\mathcal{N}$, when used in Parts I and II (it assumes a different meaning in Part III), denotes the space of all fixed, locally finite counting measures on $\mathbf{R}_+$. A measure ν on $\mathbf{R}_+$ thus belongs to $\mathcal{N}$ if and only if ν can be written as $\nu(\cdot) := \sum_{x\in I} n_x\delta_x(\cdot)$, where I is a finite or countable subset of $\mathbf{R}_+$ with no accumulation points and where, for any $x \in I$, $n_x \in \mathbf{N}$ and δ_x denotes the Dirac measure at point $x \in \mathbf{R}_+$. As explained in Section 2.1, $\mathcal{N}$ can be topologized as a complete and separable metric space by the so-called "metric of vague convergence", which gives rise to a Borel class denoted by $\mathcal{B}(\mathcal{N})$. The set $\mathcal{N}_0 \in \mathcal{B}(\mathcal{N})$ consists of all $\nu \in \mathcal{N}$ which do not charge $\{0\}$ and which do not give any point more than unit mass; that is to say, $\nu \in \mathcal{N}_0$ if and only if $0 \notin I$ and $n_x = 1$ for all $x \in I$. To emphasize their nature as $\mathcal{N}_0$-valued random elements of $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$, simple, non-explosive, unmarked point processes are usually denoted by

$$N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0)),$$

where $\mathcal{B}(\mathcal{N}_0) \equiv \mathcal{B}(\mathcal{N}) \cap \mathcal{N}_0$ denotes the restriction of $\mathcal{B}(\mathcal{N})$ to $\mathcal{N}_0$. When we wish to stress that a point process N on $\mathbf{R}_+$ is a random element of the Polish space $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$, as we do in Chapter 5, we sometimes write

$$N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N})),$$

but even then it is to be understood that $N(\omega) \in \mathcal{N}_0$ for all $\omega \in \Omega$.

Point Processes on $\mathbf{R}_+^2$

Such point processes are the subject of Part III. In this part, and only therein, the symbol $\mathcal{N}$ denotes the space of all measures μ on $\mathbf{R}_+^2$ which can be written as $\mu := \sum_{i\in\mathbf{N}} \delta_{(x_i,y_i)}$ (where, for any $(x, y) \in \mathbf{R}_+^2$, $\delta_{(x,y)}$ denotes the Dirac measure at point (x, y)), and satisfy:

i) $\mu(\{0,0\}) = 0$, and

ii) the sequence $\{(x_i, y_i)\}_{i=1}^{\infty}$ has at most one accumulation point in $\mathbf{R}_+^2$, and if (x, y) is such an accumulation point, then $x_i < x$ and $y_i < y$ for all $i \in \mathbf{N}$.

In this context the space $\mathcal{N}$ is not topologized, but it is endowed with a σ-algebra $\mathcal{F}(\mathcal{N})$ generated by a certain class of "evaluation maps" (see Section 6.2). We then define a point process on $\mathbf{R}_+^2$ as a random element

$$N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{F}(\mathcal{N})).$$

In fairness, it should be mentioned that Part III makes use, as did Maziotto and Merzbach [42] for a similar purpose, of a topological space $(\mathcal{L}, \mathcal{B}(\mathcal{L}))$ of "lower layers" on $\mathbf{R}_+^2$ (see Section 6.1) and then defines a "point process on $\mathcal{L}$" as a random element of a space $(\mathcal{N}_{\mathcal{L}}, \mathcal{F}(\mathcal{N}_{\mathcal{L}}))$ of counting measures on $\mathcal{L}$ (see Section 6.2). Spaces $(\mathcal{L}, \mathcal{B}(\mathcal{L}))$ and $(\mathcal{N}_{\mathcal{L}}, \mathcal{F}(\mathcal{N}_{\mathcal{L}}))$ are contrived for the sole purpose of inferring distributional properties of point processes on $\mathbf{R}_+^2$, and we therefore do not regard "point processes on $(\mathcal{L}, \mathcal{B}(\mathcal{L}))$" as objects of independent interest.

The expression "point process" having been set in all of its possible meanings, we may now address the substantial issues of this thesis.

Compensators, Distributions and Partial Orders of $\mathbf{R}_+$-indexed Point Processes

It has been known for quite some time (see [23]) that a simple, unmarked point process N on $\mathbf{R}_+$ is uniquely determined, in distribution, by the compensator of N taken with respect to N's internal history. This result is conceptually related to our own results of point process approximation, stochastic order (coupling) and association. These refer to the same notional framework and also use the compensator as a criterion. We shall therefore explain the classical result in greater detail prior to unveiling our own. At this point, the reader may review martingale theoretic facts and definitions invoked in this thesis by consulting Appendix B.3.

Let $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$ be a point process, which we also view as an increasing $\mathbf{R}_+$-indexed stochastic process $\{N_t : t \geq 0\}$. The *internal history* of N is the increasing family $\{\mathcal{F}_t^N\}_{t \geq 0}$ of $\mathbf{R}_+$-indexed sub-σ-fields of $\mathcal{F}$ which, for any $t \in \mathbf{R}_+$, satisfy $\mathcal{F}_t^N := \sigma(N_s : s \leq t)$. By Appendix B.3 , there exists an increasing stochastic process $N^\pi : \Omega \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ called the *compensator* of N, which is predictable with respect to $\{\mathcal{F}_t\}_{t \geq 0}$ and such that $\{N_t - N_t^\pi : t \geq 0\}$ is a martingale. The compensator N^π is unique up to evanescence. Now fix the distribution $P \circ N^{-1}$ of N and consider the "canonical" point process

$$\begin{array}{rccc} \chi : & (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0), P \circ N^{-1}) & \longrightarrow & (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0)) \\ & \mu & \longmapsto & \mu \end{array}$$

N and χ share the same distribution. By what precedes, χ admits a compensator $\Lambda : \mathcal{N}_0 \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ with respect to its own internal history. We call this compensator the *path compensator* of N. This terminology is non-standard and we adopt it for the sake of convenience. It should be noted, in this instance, that for any $\mu \in \mathcal{N}_0$ which is fixed in advance, the path $t \mapsto \Lambda(\mu, t)$ is a *deterministic* function of t; however, for any $t \in \mathbf{R}_+$, if $\mu, \nu \in \mathcal{N}_0$ are such that $\mu([0, s]) = \nu([0, s])$ for all $s \leq t$, then $\Lambda(\mu, s) = \Lambda(\nu, s)$ for all $s \leq t$, regardless of the possibility that $\mu([0, u]) \neq \nu((0, u])$

for some $u > t$. This is because the internal history of χ up to t does not "see" the realization of χ_u.

The classical result published by Jacod ([23]: Theorem 3.4) ensures that if $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$ and $\widetilde{N} : (\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P}) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$ are two point processes having the same path compensator $\Lambda : \mathcal{N}_0 \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$, then $P \circ N^{-1}$ and $\widetilde{P} \circ \widetilde{N}^{-1}$ coincide on $\mathcal{B}(\mathcal{N}_0)$. Although Jacod's theorem - which, incidently, applies to all simple, marked and possibly explosive point processes - is formulated in terms of the *compensator* (i.e. if $N : (\Omega, \mathcal{F}) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$ and $\mathcal{F}$-measures P and $\widetilde{P}$ generate the same compensator of N with respect to the internal history $\{\mathcal{F}_t^N\}_{t \geq 0}$, then P and $\widetilde{P}$ coincide on the σ-algebra $\bigvee_{t \in \mathbf{R}_+} \mathcal{F}_t^N$), we state the results of Chapters 2 and 5 in terms of the *path compensator*, because the latter lends itself more easily to manipulation since its measure argument is deterministic.

What are these results, exactly?

Chapter 2's **Theorem 2.3.3** states that if N is a point process whose path compensator Λ satisfies two continuity conditions (given in terms of the measure and time arguments), then Λ determines a sequence $\{N^n\}_{n \in \mathbf{N}}$ of point processes which converge in law to N, which are such that for every $n \in \mathbf{N}$, N^n takes support on the dyadic set $\{\frac{1}{2^n}, \frac{2}{2^n}, \ldots\}$, and such that for any $j \in \mathbf{N}$, the random variable $N^n_{\frac{j}{2^n}} - N^n_{\frac{j-1}{2^n}}$ is Bernoulli distributed with parameter

$$1 - \exp\left\{ - \left[\Lambda\left(N^n|_{[0, \frac{j-1}{2^n}]}, \frac{j}{2^n} \right) - \Lambda\left(N^n|_{[0, \frac{j-1}{2^n}]}, \frac{j-1}{2^n} \right) \right] \right\}.$$

Intuitive interpretations of this result are discussed in the introduction to Chapter 2. Acting in conjunction with Theorem 2.3.3, **Theorem 2.5.2** rigorously validates the common method of approximating a point process' finite dimensional distributions by inverse time changes, provided certain hypotheses on path compensator regularity (whose importance is discussed) do hold. The path compensator thus guarantees a point process' approximability in this case.

Chapter 4, as has been previously mentioned, does not focus on point processes. Rather, it is intended to provide the reader with an overview of results concerning the *stochastic orderings* and *association* of probability measures on partially ordered Polish spaces (POP-spaces). The known results - which predominate - are mostly drawn or adapted from [29] in the case of stochastic orderings, and from [39] in the case of association. The two theories (stochasic orderings and association) are presented in parallel on account of their remarkable analogies. The new results (to our knowledge) of Chapter 4 - namely: **Theorems 4.2.7**, **4.3.9** and **4.3.11** - concern the characterization of stochastic orderings or association on abstract POP-space via *representation maps*, which are defined in Section 4.3.

Chapter 5 applies the results of Chapter 4 to distributions of point processes viewed as random elements of the Polish space $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$. The tools constructed in Chapter 4 enable us to develop path compensator-based criteria for the stochastic ordering and association of point processes with respect to three common partial orders on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$. It turns out the path compensator only needs to be "tested" over certain classes of sets which themselves characterize the stochastic ordering or association of point processes with respect to one of the three partial orders. **Theorems 5.2.2** (for stochastic orderings) and **5.3.7** (for association) constitute the main results to that effect. Section 5.3 defines the *self-exciting property* for point processes in a manner which, unlike the definition given by Kwieciński and Szekli in [37], applies to any arbitrary partial order on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$ (see **Definition 5.3.5** and **Remark 5.3.8**). In Section 5.4, two criteria for point process association (namely: the self-exciting property and something we call the monotone kernel property) are compared in strength. **Theorems 5.4.1** and **5.4.3** each establish, for one particular order, the dominance of one criterion over the other under varying hypotheses, while **Example 5.4.2** illustrates the failure of the self-exciting property to imply the monotone kernel property for the remaining partial order.

At last, Chapter 3 - which the reader will recall as dealing with the possibly non-simple, explosive or marked "random measure μ" - puts forth a so-called "regenerative form" for the compensator, or *dual predictable projection*, of this point process. For simple point processes, the "regenerative form" of the compensator, as defined by Brémaud [7], expresses the compensator at a certain point in terms of the distribution of the (future) $(n+1)^{\mathrm{st}}$ jump conditioned on the first n jumps. This regenerative form, in the simple case, allows one to compute the joint distributions of the jump points (see [23] on p. 243) and hence, the distribution of the point process itself. By providing a regenerative form for the compensator of non-simple point processes (see **Theorem 3.3.2**), we hope to facilitate the characterization of large classes of non-simple point processes by their compensators. As noted in the chapter introduction, we essentially follow Jacod's approach for the compensator of simple processes [24], but provide substantially more details.

A Compensator Scheme for $\mathbf{R}^2_+$-indexed Point Processes

We now turn from partially ordered point processes to point processes with a partially ordered index set, namely: $\mathbf{R}^2_+$. Theorems 2.1 and 3.4 of [23] ensure that for every simple point process N on $\mathbf{R}_+$, the compensator Λ of N - taken with respect to N's internal history - is such that

i) it exists;

ii) it is unique up to indistinguishability;

iii) it characterizes N's distribution.

Several attempts have been made to derive compensators for $\mathbf{R}^2_+$-indexed simple point processes which replicate these three features. As explained in the introduction to Chapter 6, none of these attempts, by which putative compensators were defined as *stochastic processes* on $\mathbf{R}^2_+$, could demonstrably succeed on all three counts. In Chapter 6, we argue in favor of a *family* of $\mathbf{R}_+$-indexed compensators induced by a given point process on $\mathbf{R}^2_+$. We show that when the latter is *strictly simple*, this family, which always exists and is unique up to indistinguishability, characterizes the $\mathbf{R}^2_+$-indexed point process' distribution. Features (i), (ii) and (iii) are thus replicated for strictly simple planar point processes by our proposed scheme.

The main result of Chapter 6 is **Theorem 6.4.1**, which establishes the existence of an essentially unique and distribution-characterizing compensator family for the strictly simple planar point process. Other intermediate results of interest include **Theorems 6.2.18** (characterization of the $\mathbf{R}_+^2$-indexed point process by its associated point process on $\mathcal{L}$ (the space of lower layers)) and **6.3.5** (characterization of the point process on $\mathbf{R}_+^2$ by a family of $\mathbf{R}_+$-indexed point processes), as well as **Remark 6.2.8** (characterization of the associated point process on $\mathcal{L}$ by the original point process on $\mathbf{R}_+^2$) and **Corollary 6.3.8** (determination of the strictly simple point process on $\mathbf{R}_+^2$ by a family of compensators).

Part I

Distributions and Compensators

Chapter 2

The Approximation Theorem

It is well known that the homogeneous Poisson point process on $\mathbf{R}_+$ can be approximated, in distribution, by a sequence of homogeneous Bernoulli point processes of discrete support (see, for example, [41], p. 114). The present chapter seeks to extend this result to point processes with dependent increments. More precisely, let N be a simple point process on $\mathbf{R}_+$ with "path compensator" Λ, where Λ is defined on the space $\mathcal{N}_0 \times \mathbf{R}_+$ and $\mathcal{N}_0$ denotes the set of simple, non-explosive integer measures on $(0, \infty)$ (see (2) on p. 17). The main result of this chapter shall state that if Λ satisfies:

1. $\Lambda(\mu, \cdot) : \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ is continuous for all $\mu \in \mathcal{N}_0$;

2. $\Lambda(\cdot, t) : \mathcal{N}_0 \longrightarrow \mathbf{R}_+$ is continuous for all $t \in \mathbf{R}_+$,

then there exists a sequence $\{N^n\}_{n=1}^{\infty}$ of point processes which converges in law to N, and such that, for any $n \in \mathbf{N}$, the realizations of N^n have support on the dyadic set $\{\frac{1}{2^n}, \frac{2}{2^n}, \ldots\}$. Moreover, for every $n \in \mathbf{N}$, $j \in \mathbf{N} \cup \{0\}$, the random variable $N^n_{\frac{j+1}{2^n}} - N^n_{\frac{j}{2^n}}$ will assume, when conditioned on the realization of $N^n \mid_{(0,\frac{j}{2^n}]}$, a Bernoulli distribution with parameter

$$1 - \exp\left\{-\left[\Lambda(N^n \mid_{(0,\frac{j}{2^n}]}, \frac{j+1}{2^n}) - \Lambda(N^n \mid_{(0,\frac{j}{2^n}]}, \frac{j}{2^n})\right]\right\}.$$

From a heuristic perspective, the process $\{N^n_t - N^n_{\frac{j}{2^n}}\}_{\{\frac{j}{2^n} \le t \le \frac{j+1}{2^n}\}}$ will behave approximately like a Cox process with intensity $\{\Lambda_t - \Lambda_{\frac{j}{2^n}}\}_{\{\frac{j}{2^n} \le t \le \frac{j+1}{2^n}\}}$, thus supporting Karr's intuitive observation ([30]: p. 64) that "a point process with a continuous compensator is locally and conditionally a Poisson process in the sense that its mean and variance are equal".

There exists, at present, a sizeable corpus of literature pertaining to the question of how the convergence in law of a sequence of compensators entails the convergence of the corresponding sequence of point processes. This question has preoccupied, amongst others, Brown [8], Kabanov et al. [27], Kurtz [35] and Jacod, who appears to have given a definitive criterion for such a convergence to take place ([25]: Théorème 4.1). Interesting as these results are in relating the convergence of point processes to that of their compensators, they provide little assistance regarding our problem of establishing the convergence $N^n \Longrightarrow N$ of approximating processes N^n (which will later be properly defined), because they assume prior knowledge of the distributions of the approximating processes - or of their compensators - and it is precisely these distributions we seek to determine here. To our knowledge, McDonald [41] is the only one who has studied, in the context of "discretization" of continuous-time processes, a problem of the type we are now considering.

The chapter comprises five sections, the first three of which are technical in nature. Section 2.1 formally endows the aforementioned set $\mathcal{N}_0$ with the so-called "metric of vague convergence" and characterizes convergence in this metric. Section 2.2 defines the "path compensator" Λ and relates it to the more classical notion of compensator of a point process (see Appendix C.3). Section 2.3 consists of two lemmas and the main theorem (Theorem 2.3.3). Section 2.4 illustrates Theorem 2.3 by applying it to a Cox process with a discretely generated initial σ-field, and uses a result of [41] to ascertain independently that convergence does indeed take place. Section 2.5, at last, invokes Theorem 2.3.3 to legitimize a common method of estimating numerically the distribution that a point process assumes over a certain Borel subset of $\mathbf{R}_+$, and explains why it is important for the path compensator to satisfy the stated hypotheses of continuity for such a method to be valid.

2.1 Topological Framework

First consider the space $S := (0, \infty]$ endowed with the metric ρ defined by the equality:

$$\rho(x, y) := \mid e^{-x} - e^{-y} \mid,$$

where we take that $e^{-\infty} := 0$ by convention.

The Borel sets $\mathcal{S}$ correspond to the "usual" Borel sets on $(0, \infty)$, and admit the singleton $\{\infty\}$ as measurable.

For any $m \in \mathbf{N} \cup \{\infty\}$, we let $S^m := \bigotimes_{i=1}^m S$ and define the metric ρ^m as

$$\rho^m(x, y) := \sum_{i=1}^{m} \frac{\rho(x_i, y_i)}{2^i}$$

for any $x = (x_1, ..., x_m)$, $y = (y_1, ...y_m) \in S^m$.

Note that ρ^m metrizes ρ-convergence coordinatewise in S^m.

If $m < \infty$ let

$$J_m := \{(x_1, ..., x_m) \in S^m : 0 < x_1 < ... < x_m\}$$

and also define

$$J := \{(x_1, x_2, ...) \in S^\infty : \forall i \in \mathbf{N}, x_i < \infty \Rightarrow x_i < x_{i+1}, \text{ and } \lim_{n \to \infty} x_n = \infty\}.$$

It is readily inferred that $J_m \in \mathcal{S}^m$ (J_m is open) and $J \in \mathcal{S}^\infty$. Thus, J_m (resp. J) may be regarded as metric space in its own right when endowed with the metric ρ^m (resp. ρ^∞) restricted to J_m (resp. J). The resulting Borel class will be noted $\mathcal{B}(J_m)$ (resp. $\mathcal{B}(J)$).

Now consider the space $\mathcal{N}$ of Borel measures μ on $\mathbf{R}_+ := [0, \infty)$ such that, for any $t \in \mathbf{R}_+$, $\mu([0, t]) \in \mathbf{N} \cup \{0\}$. $\mathcal{N}$ can be metrized by the metric d of vague convergence, defined as follows. For any $\mu \in \mathcal{N}$, $m \in \mathbf{N}$, let $\mu^m := \mu \mid_{[0,m]}$. One writes, for $\mu, \nu \in \mathcal{N}$,

$$d(\mu, \nu) := \sum_{m=1}^{\infty} \frac{L(\mu^m, \nu^m)}{2^m(1 + L(\mu^m, \nu^m))}$$

where, for any $m \in \mathbf{N}$,

$$L(\mu^m, \nu^m) := \inf\{h > 0 : \mu^m([0, x-h]) - h \leq \nu^m([0, x]) \leq \mu^m([0, x+h]) + h \ \forall \ x \in \mathbf{R}_+\}.$$

(L measures the "distance" between two paths $t \mapsto \mu([0,t])$ and $t \mapsto \nu([0,t])$, for measures μ and ν of bounded support; in [18], L is called the *Lévy metric*). The space $\mathcal{N}$ is complete and separable under d (see [18]: Theorems 4 and 5). Convergence in the metric d is characterized as follows (this is easily established): for any $\mu \in \mathcal{N}$, $n \in \mathbf{N} \cup \{0\}$ let

$$\tau_n(\mu) := \inf\{t \in \mathbf{R}_+ : \mu([0,t]) \geq n\}$$

($\tau_n(\mu)$ is commonly described as the n^{th} jump point of μ). Then, for any sequence $\mu, \mu_1, \mu_2, \ldots$ in $\mathcal{N}$ we have that

$$d(\mu_m, \mu) \longrightarrow 0 \Longleftrightarrow \rho(\tau_n(\mu_m), \tau_n(\mu)) \longrightarrow 0 \ \forall \ n \in \mathbf{N} \cup \{0\}$$

as $m \to \infty$. A subset of $\mathcal{N}$ of special importance is the (measurable) subset $\mathcal{N}_0$ of simple measures. More precisely,

$$\mathcal{N}_0 := \{\mu \in \mathcal{N} : 0 =: \tau_0(\mu) < \tau_1(\mu) < \tau_2(\mu) < \ldots\}.$$

The Borel class of $\mathcal{N}$ will be denoted by $\mathcal{B}(\mathcal{N})$. The restriction of $\mathcal{B}(\mathcal{N})$ to $\mathcal{N}_0$ will be denoted by $\mathcal{B}(\mathcal{N}_0)$.

Given the characterization of vague convergence on $\mathcal{N}$ by the convergence of the jump points, the next result is not surprising at all.

Proposition 2.1.1 *The spaces $(\mathcal{N}_0, d)$ and (J, ρ^∞) are homeomorphic.*

Proof Consider the map

$$\begin{aligned} \Phi : \ & \mathcal{N}_0 \longrightarrow J \\ & \mu \longmapsto (\tau_1(\mu), \tau_2(\mu), \ldots) \end{aligned}$$

where, for any $\mu \in \mathcal{N}_0, n \in \mathbf{N}$,

$$\tau_n(\mu) := \inf\{t \in \mathbf{R}_+ : \mu([0,t]) \geq n\}$$

($\inf \emptyset := \infty$ by convention). The map Φ is evidently bijective, as every sequence $\{t_n\}_{n=1}^{\infty}$ in $(0,\infty]$ which is increasing and such that, for any $n \in \mathbf{N}$, $t_n < \infty \Longrightarrow t_n < t_{n+1}$ determines one and only one element of $\mathcal{N}_0$. Finally, Φ is bicontinuous by virtue of vague convergence being characterized by convergence of the jump points. □

Throughout the sequel, the symbol $\longrightarrow_{\mathcal{L}}$ shall denote convergence in distribution.

Corollary 2.1.2 *Let $N, N^1, N^2, \ldots : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_0,\ \mathcal{B}(\mathcal{N}_0))$ be point processes on $\mathbf{R}_+$. Let $n \to \infty$. $N^n \longrightarrow_{\mathcal{L}} N$ with respect to the vague topology if and only if $\Phi(N^n) \longrightarrow_{\mathcal{L}} \Phi(N)$ with respect to the topology induced by the metric ρ.*

Proof If $h : (S_1, \mathcal{S}_1) \longrightarrow (S_2, \mathcal{S}_2)$ is a continuous mapping between spaces S_1 and S_2, then, for probability measures $P, P_1, P_2, \ldots$ on $(S_1, \mathcal{S}_1)$, $P_n \Longrightarrow P$ as $n \to \infty$ implies $P_n \circ h^{-1} \Longrightarrow P \circ h^{-1}$ on $(S_2, \mathcal{S}_2)$ (see [4]: Theorem 5.1). If, in addition, h is a homeomorphism, then $P_n \circ h^{-1} \Longrightarrow P \circ h^{-1}$ on $(S_2, \mathcal{S}_2)$ implies $P_n \circ h^{-1} \circ h = P_n \Longrightarrow P = P \circ h^{-1} \circ h$ on $(S_1, \mathcal{S}_1)$. The proof is complete by letting $(S_1, \mathcal{S}_1) := (\mathcal{N}_0,\ \mathcal{B}(\mathcal{N}_0))$, $(S_2, \mathcal{S}_2) := (J, \mathcal{B}(J))$ and $h := \Phi$. □

2.2 Choice of Filtration and Path Compensator

Let $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_0,\ \mathcal{B}(\mathcal{N}_0))$ be a point process on $\mathbf{R}_+$ and let $\{\mathcal{F}_t^N\}_{t\geq 0}$ be N's internal history, i.e. for any $t \in \mathbf{R}_+$, $\mathcal{F}_t^N := \sigma(N_s : s \leq t)$. Suppose $\mathcal{F}_0$ is a sub-σ-field of $\mathcal{F}$, and consider a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ given by $\mathcal{F}_t := \mathcal{F}_0 \vee \mathcal{F}_t^N$ for all $t \in \mathbf{R}_+$. If $\{T_n\}_{n=1}^{\infty}$ represents the increasing sequence of N's jump points and, for every $n \in \mathbf{N}$, $\mathcal{F}_{T_n}$ denotes the σ-field $\sigma(A \cap \{T_n \leq t\} : t \in \mathbf{R}_+, A \in \mathcal{F}_t)$ associated to T_n by the augmented history $\{\mathcal{F}_t\}_{t\geq 0}$, then the random measure $\nu : \Omega \times \mathbf{R}_+ \to \mathbf{R}_+$ given by

$$\nu(\omega, dt) := \sum_{n=0}^{\infty} \frac{dP\left[T_{n+1} - T_n \leq t - T_n \mid \mathcal{F}_{T_n}\right](\omega)}{P\left[T_{n+1} - T_n > t - T_n \mid \mathcal{F}_{T_n}\right](\omega)} \mathbf{1}_{\{T_n(\omega) < t \leq T_{n+1}(\omega)\}}$$

defines a version of N's compensator with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$, provided that regular conditional probabilities are employed (see [23]: Proposition 3.1).

The following proposition will enable us to compute the random measure ν with respect to *realized* values of the T_n's and an arbitrary random variable π defined at $t = 0$ (see (2.2) on p. 17). This will apply, in particular, to Cox processes (see Appendix C.3).

Proposition 2.2.1 *If* $\mathcal{F}_0 = \sigma(\pi)$ *for some parameter* $\pi : (\Omega, \mathcal{F}, P) \longrightarrow (\mathbf{R}, \mathcal{B}(\mathbf{R}))$, *then, for any* $n \in \mathbf{N}$,

$$\mathcal{F}_{T_n} = \mathcal{F}^N_{T_n} \vee \mathcal{F}_0 = \sigma(T_1, ..., T_n) \vee \sigma(\pi) = \sigma(\pi, T_1, ..., T_n).$$

Proof That $\sigma(T_1, ..., T_n) = \mathcal{F}^N_{T_n}$ follows from [7] (Theorem A2.T30(i)). We therefore only need to ensure that $\mathcal{F}_{T_n} = \mathcal{F}^N_{T_n} \vee \mathcal{F}_0$. The inclusion $\mathcal{F}_{T_n} \supseteq \mathcal{F}^N_{T_n} \vee \mathcal{F}_0$ is obvious. For the reverse inclusion, consider the $\Omega \times \mathbf{R}_+$-valued process $\tilde{N} = (\tilde{N}_t)_{t\geq 0}$ defined by

$$\begin{array}{rccl} \tilde{N}_t : & (\Omega, \mathcal{F}, P) & \longrightarrow & (\Omega \times \mathbf{R}_+, \mathcal{F}_0 \times \mathcal{B}(\mathbf{R}_+)) \\ & \omega & \longmapsto & (\omega, N_t(\omega)). \end{array}$$

For any $t \in \mathbf{R}_+$, $\sigma(N_t) \vee \mathcal{F}_0 \subseteq \sigma(\tilde{N}_t) \Longrightarrow \mathcal{F}_t \subseteq \mathcal{F}^{\tilde{N}}_t$, where $\{\mathcal{F}^{\tilde{N}}_t\}_{t\geq 0}$ denotes the internal history of $\tilde{N}$. But this implies that $\mathcal{F}_{T_n} \subseteq \mathcal{F}^{\tilde{N}}_{T_n}$. By [7] (Theorem A2.T28), we have $\mathcal{F}^{\tilde{N}}_{T_n} = \sigma(\tilde{N}_{s\wedge T_n}, s \geq 0)$. Furthermore, for any $s \geq 0, A \in \mathcal{F}_0, B \in \mathcal{B}(\mathbf{R}_+)$,

$$\tilde{N}^{-1}_{s\wedge T_n}(A \times B) = \{\omega \in \Omega : \omega \in A, N_{s\wedge T_n}(\omega) \in B\} = A \cap N^{-1}_{s\wedge T_n}(B).$$

Now, if $B = [0, m]$ for some $m \in \mathbf{N}$, then

$$\begin{aligned} N^{-1}_{s\wedge T_n}(B) &= \{\omega \in \Omega : N_{s\wedge T_n(\omega)}(\omega) \leq m\} \\ &= \{\omega \in \Omega : N_s(\omega) \leq m \text{ or } N_{T_n}(\omega) \leq m\} \\ &= \{\omega \in \Omega : N_s(\omega) \leq m\} \cup \{\omega \in \Omega : N_{T_n}(\omega) \leq m\} \\ &= \left\{ \begin{array}{ll} \Omega & \text{if } n \leq m \\ [T_m \geq s] & \text{if } n > m \end{array} \right\} \in \sigma(T_1, ..., T_n), \end{aligned}$$

whence $\tilde{N}^{-1}_{s\wedge T_n}(A \times B) \in \mathcal{F}_0 \vee \sigma\{T_1, ..., T_n\} = \mathcal{F}_0 \vee \mathcal{F}^N_{T_n}$ for all $s \geq 0$, $A \in \mathcal{F}_0$, $B \in \mathcal{B}(\mathbf{R}_+)$; this entails $\mathcal{F}_{T_n} \subseteq \mathcal{F}^{\tilde{N}}_{T_n} \subseteq \mathcal{F}^N_{T_n} \vee \mathcal{F}_0$. □

We henceforth restrict our attention to the case where the parameter $\pi : (\Omega, \mathcal{F}, P) \longrightarrow (\mathbf{R}, \mathcal{B}(\mathbf{R}))$ has a discrete range, say S. For any $n \in \mathbf{N} \cup \{0\}$, $s \in S$, and $x_1, ..., x_n \in \mathbf{R}_+$ with $0 < x_1 < ... < x_n$, let

$$F_{n+1}(\cdot; s; x_1, ..., x_n) := P\left[T_{n+1} - T_n \leq \cdot \mid \pi = s; T_1 = x_1, ..., T_n = x_n\right]$$

be a regular version of the conditional probability distribution, and define the *path compensator* $\Lambda : \mathcal{N}_0 \times \mathbf{R} \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ of N *with initial parameter* π as

$$\Lambda(\mu, s, dt) := \sum_{n \geq 0} \frac{dF_{n+1}(t - \tau_n(\mu); s; \tau_1(\mu), ... \tau_n(\mu))}{1 - F_{n+1}(t - \tau_n(\mu); s; \tau_1(\mu), ... \tau_n(\mu))} \mathbf{1}_{\{\tau_n(\mu) < t \leq \tau_{n+1}(\mu)\}}. \tag{1}$$

One may observe, in view of the last proposition and the discussion preceding it, that for any $\omega \in \Omega$ and $t \in \mathbf{R}_+$, $\Lambda(N(\omega), \pi(\omega), t) := \nu(\omega, t)$ defines a version ν of N's compensator with respect to the filtration $\{\mathcal{F}_t\}_{t \geq 0} := \{\mathcal{F}_t^N \vee \sigma(\pi)\}_{t \geq 0}$. Qualitatively, ν and Λ differ in that ν is random in its first argument, whereas Λ, being computed from *known* conditional probability distributions, is deterministic. When $\sigma(\pi) = \{\emptyset, \Omega\}$ is trivial, the middle argument is suppressed and, for regular versions

$$F_{n+1}(\cdot; x_1, ..., x_n) := P\left[T_{n+1} - T_n \leq \cdot \mid T_1 = x_1, ..., T_n = x_n\right]$$

of the probability distributions, $\Lambda : \mathcal{N}_0 \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ defined as

$$\Lambda(\mu, dt) := \sum_{n \geq 0} \frac{dF_{n+1}(t - \tau_n(\mu); \tau_1(\mu), ... \tau_n(\mu))}{1 - F_{n+1}(t - \tau_n(\mu); \tau_1(\mu), ... \tau_n(\mu))} \mathbf{1}_{\{\tau_n(\mu) < t \leq \tau_{n+1}(\mu)\}} \tag{2}$$

is simply called the *path compensator* of N. Our motivation for presenting the compensator in this fashion will become clearer in the next section. The filtration $\{\mathcal{F}_t\}_{t \geq 0}$ and path compensator Λ - with or without initial parameter - shall be retained for the remainder of this chapter.

2.3 Discrete Approximation Theorem

As previously announced, this section comprises two lemmas and the main theorem. In the context of notation introduced in Section 2.1, the lemmas provide a criterion for the convergence in law of J-valued random elements of $(S^\infty, \mathcal{S}^\infty)$. Loosely paraphrased, the main theorem states that if a point process has appropriately continuous conditional interarrival distributions - or, equivalently, a path compensator Λ satisfying analogous continuity requirements - then it may be approximated in law by a sequence of discrete binomial processes N^n with the property that for any $j \in \mathbf{N}$, the random variable $N^n_{\frac{j+1}{2^n}} - N^n_{\frac{j}{2^n}}$ assumes a Bernoulli distribution with parameter $\Lambda_{\frac{j+1}{2^n}} - \Lambda_{\frac{j}{2^n}}$.

Lemma 2.3.1 *Let* $N = (T_1, T_2, \ldots)$, $N^1 = (T_1^1, T_2^1, \ldots)$, $N^2 = (T_1^2, T_2^2, \ldots)$, $\ldots : (\Omega, \mathcal{F}, P) \longrightarrow (S^\infty, \mathcal{S}^\infty)$ *be* J*-valued random elements. For any* $n \in \mathbf{N}$ *and* $x_1, \ldots, x_m \in S$ *with* $x_1 < \ldots < x_m$*, let*

$$G_1(\cdot) := P[T_1 \leq \cdot],\ G_1^n(\cdot) := P[T_1^n \leq \cdot],$$

$$G_{m+1}(\cdot; x_1, \ldots, x_m) := P[T_{m+1} \leq \cdot \mid T_1 = x_1, \ldots, T_m = x_m]$$

and

$$G_{m+1}^n(\cdot; x_1, \ldots, x_m) := P[T_{m+1}^n \leq \cdot \mid T_1^n = x_1, \ldots, T_m^n = x_m]$$

be regular versions of the conditional probability distributions. Suppose that

1. *for any* $m \in \mathbf{N} \cup \{0\}$, $x \in S$, $(x_1, \ldots, x_m) \in J_m$,

   a) $G_{m+1}(x; \cdot, \ldots, \cdot)$ *is continuous over* S^m;

   b) $G_{m+1}(\cdot; x_1, \ldots, x_m)$ *is continuous over* $S \setminus \{\infty\}$;

2. $G_1^n \longrightarrow G_1$ *pointwise as* $n \to \infty$ *and, moreover, for any* $x \in S$, $(x_1, \ldots, x_m) \in J_m$*, for any sequence* $\{(y_1^n, \ldots, y_m^n)\}_{n=1}^\infty$ *in* J_m *with limit* $(x_1, \ldots, x_m)$*, the condition*

$$\lim_{n \to \infty} \mid G_{m+1}^n(x; y_1^n, \ldots, y_m^n) - G_{m+1}(x; y_1^n, \ldots y_m^n) \mid = 0$$

   holds.

Then $N^n \longrightarrow_{\mathcal{L}} N$ *as* $n \to \infty$ *(with respect to the topology induced by* ρ^∞ *on* $J \subset S^\infty$*).*

Proof Since $N, N^1, N^2, \ldots$ are random elements of $(S^\infty, \mathcal{S}^\infty)$, we may use the fact that "for a countable product of separable spaces, the finite-dimensional sets form a convergence-determining class" (see [4]: Exercise no. 7, pg. 22). It thus suffices to establish that for any $m \in \mathbf{N}$,

$$(T_1^n, \ldots, T_m^n) \longrightarrow_{\mathcal{L}} (T_1, \ldots, T_m)$$

on $(S^m, \mathcal{S}^m)$ as $n \to \infty$. Fix $m \in \mathbf{N}$ and consider the class $\mathcal{U}_m \subset \mathcal{S}^m$ consisting of sets of the form

$$(a_1, b_1] \times \ldots \times (a_m, b_m]$$

where, for $i = 1, \ldots, m$, $0 \leq a_i < b_i \leq \infty$. Observe that (i) $\mathcal{U}_m$ is closed under the formation of finite intersections, and (ii) for every $(x_1, \ldots, x_m) \in S^m$ and $\epsilon > 0$, there exists $A \in \mathcal{U}_m$ such that $(x_1, \ldots, x_m) \in A^\circ \subseteq A \subseteq B_\epsilon((x_1, \ldots, x_m))$, the latter set denoting the ρ^m-open ball of radius ϵ around $(x_1, \ldots, x_m)$. Since S^m is separable, Corollary 1, pg. 14 of [4] ensures that the condition:

$$P \circ (T_1^n, \ldots, T_m^n)^{-1}(A) \longrightarrow P \circ (T_1, \ldots, T_m)^{-1}(A) \tag{3}$$

as $n \to \infty$ for all $A \in \mathcal{U}_m$ is sufficient to entail $(T_1^n, \ldots, T_m^n) \longrightarrow_{\mathcal{L}} (T_1, \ldots, T_m)$. Now a straightforward inclusion-exclusion argument may be invoked to show that (3) holds if and only if for any $x_1, \ldots, x_m \in S$,

$$P[T_1^n \leq x_1, \ldots, T_m^n \leq x_m] \longrightarrow P[T_1 \leq x_1, \ldots, T_m \leq x_m] \tag{4}$$

as $n \to \infty$. We shall prove (4) by induction over $m \in \mathbf{N}$. For $m = 1$, it is a consequence of the lemma's hypothesis 1.(a) that

$$P[T_1^n \leq x] = G_1^n(x) \longrightarrow G_1(x) = P[T_1 \leq x]$$

as $n \to \infty$ for any $x \in S$, whence (4) is trivially verified in this case. Suppose now that (4) is verified for all integers not exceeding some $m \in \mathbf{N}$. Let $x_1, \ldots, x_m \in S$, and consider, for some $n \in \mathbf{N}$, the expression:

$$\begin{aligned}
&\left|P[T_1^n \le x_1, \ldots, T_{m+1}^n \le x_{m+1}] - P[T_1 \le x_1, \ldots, T_{m+1} \le x_{m+1}]\right| \\
&= \Big| \int_{\{\otimes_{i=1}^m (0,x_i]\}\cap J_m} P[T_{m+1}^n \le x_{m+1} \mid T_1^n = y_1, \ldots, T_m^n = y_m] dP_{T_1^n,\ldots,T_m^n}(y_1,\ldots,y_m) \\
&\quad - \int_{\{\otimes_{i=1}^m (0,x_i]\}\cap J_m} P[T_{m+1} \le x_{m+1} \mid T_1 = y_1, \ldots, T_m = y_m] dP_{T_1,\ldots,T_m}(y_1,\ldots,y_m) \Big| \\
&= \Big| \int_{J_m} I_{\otimes_{i=1}^m (0,x_i]} G_{m+1}^n(x_{m+1}; y_1, \ldots, y_m) dP_{T_1^n,\ldots,T_m^n}(y_1,\ldots,y_m) \\
&\qquad - \int_{J_m} I_{\otimes_{i=1}^m (0,x_i]} G_{m+1}(x_{m+1}; y_1, \ldots, y_m) dP_{T_1,\ldots,T_m}(y_1,\ldots,y_m) \Big| \\
&= \Big| \int_{[0,1]} x dP_{T_1^n,\ldots,T_m^n} \circ [I_{\otimes_{i=1}^m (0,x_i]} G_{m+1}^n(x_{m+1}; \cdot, \ldots, \cdot)]^{-1}(x) \\
&\qquad - \int_{[0,1]} x dP_{T_1,\ldots,T_m} \circ [I_{\otimes_{i=1}^m (0,x_i]} G_{m+1}(x_{m+1}; \cdot, \ldots, \cdot)]^{-1}(x) \Big| .
\end{aligned}$$

Since the identity map is continuous on $[0,1]$, the last expression will tend to 0 once it is shown that

$$\begin{aligned}
&P_{T_1^n,\ldots,T_m^n} \circ [I_{\otimes_{i=1}^m (0,x_i]} G_{m+1}^n(x_{m+1}; \cdot, \ldots, \cdot)]^{-1} \\
&\qquad \Longrightarrow P_{T_1,\ldots,T_m} \circ [I_{\otimes_{i=1}^m (0,x_i]} G_{m+1}(x_{m+1}; \cdot, \ldots, \cdot)]^{-1}
\end{aligned}$$

as $n \to \infty$. Consider the set

$$\begin{aligned}
E = \{(y_1, \ldots, y_m) \in J_m : &\, I_{\otimes_{i=1}^m (0,x_i]}(y_1^n, \ldots, y_m^n) \cdot G_{m+1}^n(x_{m+1}; y_1^n, \ldots, y_m^n) \\
&\not\to I_{\otimes_{i=1}^m (0,x_i]}(y_1, \ldots, y_m) \cdot G_{m+1}(x_{m+1}; y_1, \ldots, y_m) \text{ as } n \to \infty, \\
&\text{for some sequence } \{(y_1^n, \ldots, y_m^n)\}_{n=1}^\infty \text{ with limit } (y_1, \ldots, y_m)\}.
\end{aligned}$$

The separability of $\mathbf{R}_+$ ensures the $\mathcal{S}^m$-measurability of E ([4]: pg.226). That $P_{T_1^n,\ldots,T_m^n} \Longrightarrow P_{T_1,\ldots,T_m}$ is a direct consequence of the induction hypothesis. It therefore suffices, in view of Theorem 5.5 of [4] to prove that $P_{T_1,\ldots,T_m}(E) = 0$. For any sequence $\{(y_1^n, \ldots, y_m^n)\}_{n=1}^\infty$ in J_m with limit $(y_1, \ldots, y_m) \in J_m$,

$$\begin{aligned}
&| G_{m+1}^n(x_{m+1}; y_1^n, \ldots, y_m^n) - G_{m+1}(x_{m+1}; y_1, \ldots, y_m) | \\
&\qquad \le \left|G_{m+1}^n(x_{m+1}; y_1^n, \ldots, y_m^n) - G_{m+1}(x_{m+1}; y_1^n, \ldots, y_m^n)\right| \\
&\qquad\quad + \left|G_{m+1}(x_{m+1}; y_1^n, \ldots, y_m^n) - G_{m+1}(x_{m+1}; y_1, \ldots, y_m)\right|
\end{aligned}$$

for all $n \in \mathbf{N}$. As $n \to \infty$, the first and second term of the right-hand-side tend to 0 by the statements's conditions (2) and (1a) respectively. Thus E may be written

$$E = \{(y_1, \ldots, y_m) \in J_m : I_{\otimes_{i=1}^m (0,x_i]}(y_1^n, \ldots, y_m^n) \not\to I_{\otimes_{i=1}^m (0,x_i]}(y_1, \ldots, y_m) \\ \text{as } n \to \infty, \text{ for some sequence } \{(y_1^n, \ldots, y_m^n)\}_{n=1}^\infty \text{ with limit } (y_1, \ldots, y_m)\}.$$

Observe, however, that if $(y_1, \ldots, y_m) \in E$, then $y_i = x_i < \infty$ for some $i \in 1, \ldots, m$. That $P_{T_1,\ldots,T_m}(E) = 0$ would therefore follow from a claim that for any $a < \infty$ and $j \in \mathbf{N}$, $P[T_j = a] = 0$. Proving the latter is easily done by induction over $j \in \mathbf{N}$. Indeed, for $j = 1$, the result holds by hypothesis (1b) of the statement. If it also holds for an arbitrary j, we then have

$$\begin{aligned} P[T_{j+1} = a] &= P[T_{j+1} \le a] - P[T_{j+1} < a] \\ &= \left| \int_{J_j} G_{j+1}(a; u_1, \ldots, u_j) dP_{T_1,\ldots,T_j}(u_1, \ldots, u_j) \right. \\ &\quad \left. - \int_{J_j} G_{j+1}(a-; u_1, \ldots, u_j) dP_{T_1,\ldots,T_j}(u_1, \ldots, u_j) dP_{T_1,\ldots,T_j}(u_1, \ldots, u_j) \right| \\ &\le \int_{J_j} |G_{j+1}(a; u_1, \ldots, u_j) - G_{j+1}(a-; u_1, \ldots, u_j)| \, dP_{T_1,\ldots,T_j}(u_1, \ldots, u_j) \\ &= 0 \text{ by hypothesis (1b).} \end{aligned}$$

We conclude that $P_{T_1,\ldots,T_m}(E) = 0$, completing our proof. □

Lemma 2.3.1 may be easily generalized by further conditioning the distributions of the T_n's on values taken by an arbitrary parameter $\pi : (\Omega, \mathcal{F}, P) \longrightarrow (\mathbf{R}, \mathcal{B}(\mathbf{R}))$. This is required in order to accommodate the full generality of the filtration described in Section 2.2.

Lemma 2.3.2 *Let* $N = (T_1, T_2, \ldots)$, $N^1 = (T_1^1, T_2^1, \ldots)$, $N^2 = (T_1^2, T_2^2, \ldots)$, $\ldots$: $(\Omega, \mathcal{F}, P) \longrightarrow (S^\infty, \mathcal{S}^\infty)$ *be J-valued random elements. For any* $n \in \mathbf{N}$, $x_1, \ldots, x_m \in S$ *with* $x_1 < \ldots < x_m$, *and* $s \in \mathbf{R}$ *let*

$$\tilde{G}_1(\cdot; s) := P[T_1 \le \cdot \mid \pi = s], \; \tilde{G}_1^n(\cdot; s) := P[T_1^n \le \cdot \mid \pi = s],$$

$$\tilde{G}_{m+1}(\cdot; s; x_1, \ldots, x_m) := P[T_{m+1} \le \cdot \mid \pi = s; T_1 = x_1, \ldots, T_m = x_m]$$

and

$$\tilde{G}^n_{m+1}(\cdot; s; x_1, \ldots, x_m) := P[T^n_{m+1} \leq \cdot \mid \pi = s; T^n_1 = x_1, \ldots, T^n_m = x_m]$$

be regular versions of the conditional probability distributions. Suppose that

1. *for any* $m \in \mathbf{N} \cup \{0\}$, $s \in \mathbf{R}$, $x \in S$, *and* $(x_1, \ldots, x_m) \in J_m$,

   a) $\tilde{G}_{m+1}(x; s; \cdot, \ldots, \cdot)$ *is continuous over* S^m;

   b) $\tilde{G}_{m+1}(\cdot; s; x_1, \ldots, x_m)$ *is continuous over* $S \setminus \{\infty\}$;

2. $\tilde{G}^n_1(\cdot; s) \longrightarrow \tilde{G}_1(\cdot, s)$ *pointwise as* $n \to \infty$ *and, moreover, for any* $x \in S$, $(x_1, \ldots, x_m) \in J_m$, *for any sequence* $\{(y^n_1, \ldots, y^n_m)\}_{n=1}^{\infty}$ *in* J_m *with limit* $(x_1, \ldots, x_m)$, *the condition*

$$\lim_{n\to\infty} \mid \tilde{G}^n_{m+1}(x; s; y^n_1, \ldots, y^n_m) - \tilde{G}_{m+1}(x; s; y^n_1, \ldots y^n_m) \mid = 0$$

   holds.

Then $N^n \longrightarrow_{\mathcal{L}} N$ *as* $n \to \infty$.

Proof By Lemma 2.3.1, we already know that for any $s \in \mathbf{R}$, $P_{N^n|\pi=s} \Longrightarrow P_{N|\pi=s}$ as $n \to \infty$. Viewing $P_{N^n|\pi=s}$, $P_{N|\pi=s}$ as distributions over the metric space $(J, \mathcal{B}(J))$, we have that for any bounded and continuous function $f : (J, \mathcal{B}(J)) \longrightarrow (\mathbf{R}, \mathcal{B}(\mathbf{R}))$,

$$\int_J f(u) dP_{N^n}(u) = \int_{\mathbf{R}} \int_J f(u) dP_{N^n|\pi=s}(u) dP_\pi(s) \longrightarrow \int_{\mathbf{R}} \int_J f(u) dP_{N|\pi=s}(u) dP_\pi(s)$$

as $n \to \infty$, since the (uniformly bounded) inner integral converges. But since the limiting expression equals $\int_J f(u) dP_N(u)$, the convergence $N^n \longrightarrow_{\mathcal{L}} N$ ensues. □

Our main result may now be introduced and proved. The idea which underlies this theorem is that of local time changes. A point process N with a pathwise continuous compensator and such that $N_\infty = \infty$ a.s. may always be expressed (stochastically) as the transformation of a Standard Poisson Process ([36]: Lemma 3.2), and this transformation is typically obtained by substituting random exponential values into pseudo-inverse compensator arguments ([37]: pp. 1216-1217). Here, the half-line is partitioned into intervals of size $\frac{1}{2^n}$, and it is the increments of the approximating process N^n which are determined over these intervals via local time changes. The role of the array $\{\Pi_{\{p,i\}} : p \in I, i \in \mathbf{N} \cup \{0\}\}$ of independent versions of the Standard Poisson Process is to generate independent exponential random variables of varying parameter which are required in order to effect these time changes.

An illustrative example shall be provided in the next section. The symbol θ denotes the null measure on $\mathbf{R}_+$.

Theorem 2.3.3 *Let $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$ be a point process on $(0, \infty)$ and let $\pi : (\Omega, \mathcal{F}, P) \longrightarrow (\mathbf{R}, \mathcal{B}(\mathbf{R}))$ be a random variable with discrete range I. Let $\Lambda : \mathcal{N}_0 \times \mathbf{R} \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ denote N's path compensator with initial parameter π (see Section 2.2); furthermore, assume the existence of a denumerable array $\{\Pi_{\{p,i\}} : p \in I, i \in \mathbf{N} \cup \{0\}\}$ of independent versions of the Standard Poisson Process, which is itself independent of (N, π). For any $p \in I$, suppose Λ satisfies the two conditions*

1. $\Lambda(\mu, p, \cdot) : \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ *is continuous for all* $\mu \in \mathcal{N}_0$, *and*

2. $\Lambda(\cdot, p, t) : \mathcal{N}_0 \longrightarrow \mathbf{R}_+$ *is continuous for all* $t \in \mathbf{R}_+$.

For any $n \in \mathbf{N}$ let $Q_n := \{0; \frac{1}{2^n}; \frac{2}{2^n}; \ldots\}$ and let $N^n : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$ be a point process with support in Q_n, specified as follows: if $\omega \in \Omega$ let

$$\begin{aligned}
&\Delta N^n(\omega, 0) := 0; \\
&\Delta N^n(\omega, \tfrac{1}{2^n}) := \Pi_{\{\pi(\omega), N^n(\omega, \frac{1}{2^n}-)\}}\left(\omega, \Lambda(\theta, \pi(\omega), \tfrac{1}{2^n})\right) \wedge 1; \\
&\Delta N^n(\omega, \tfrac{2}{2^n}) := \Big[\Pi_{\{\pi(\omega), N^n(\omega, \frac{2}{2^n}-)\}}(\omega, \Lambda(N^n |_{(0,\frac{1}{2^n}]}, \pi(\omega), \tfrac{2}{2^n})) \\
&\qquad\qquad - \Pi_{\{\pi(\omega), N^n(\omega, \frac{2}{2^n}-)\}}(\omega, \Lambda(N^n |_{(0,\frac{1}{2^n}]}, \pi(\omega), \tfrac{1}{2^n}))\Big] \wedge 1; \\
&\vdots \\
&\Delta N^n(\omega, \tfrac{j}{2^n}) := \Big[\Pi_{\{\pi(\omega), N^n(\omega, \frac{j}{2^n}-)\}}(\omega, \Lambda(N^n |_{(0,\frac{j-1}{2^n}]}, \pi(\omega), \tfrac{j}{2^n})) \\
&\qquad\qquad - \Pi_{\{\pi(\omega), N^n(\omega, \frac{j}{2^n}-)\}}(\omega, \Lambda(N^n |_{(0,\frac{j-1}{2^n}]}, \pi(\omega), \tfrac{j-1}{2^n}))\Big] \wedge 1; \\
&\vdots
\end{aligned}$$

Then $N^n \longrightarrow_{\mathcal{L}} N$ as $n \to \infty$.

Proof We first make an observation regarding continuity conditions (1) and (2) imposed on Λ. For any $m \in \mathbf{N} \cup \{0\}$, $p \in I$, and $x_1, \ldots, x_m \in S$ with $x_1 < \ldots < x_m$, let

$$G_{m+1}(\cdot; p; x_1, \ldots, x_m) := P[\tau_{m+1}(N) \leq \cdot \mid \pi = p; \tau_1(N) = x_1, \ldots, \tau_m(N) = x_m]$$

be a regular version of the conditional probability, Conditions (1) and (2) are equivalent to conditions that for any $x > 0$ and $p \in I$,

1'. $G_{m+1}(\cdot; p; x_1, \ldots, x_m)$ is continuous over $(0, \infty)$ and

2'. $G_{m+1}(x; p; \cdot, \ldots, \cdot)$ is continuous in its m arguments.

Conditions (1) and (1') are equivalent by virtue of the distribution-to-hazard function formula ([23]: Lemma 3.5); this formula relies on the expression for Λ that was developed in Section 2.2. The equivalence of (2) and (2') results for the characterization of convergence in $\mathcal{N}_0$ by the convergence of the jump points: for any sequence $\{\mu_m\}_{m=1}^{\infty}$ in $\mathcal{N}_0$, μ_m converges to some $\mu \in \mathcal{N}_0$ if and only if for any $n \in \mathbf{N}$, $\tau_n(\mu_m)$ converges to $\tau_n(\mu)$ (see Section 2.1). The expression for Λ appearing in Section 2.2, which involves the distributions of the type $G_{m+1}(x; p; \cdot, \ldots, \cdot)$, thus makes clear that conditions (2) and (2') are equivalent, at least for "good" versions of the $G_{m+1}(x; p; \cdot, \ldots, \cdot)$'s.

We now check that the randomly indexed objects $\Pi_{\{\pi,i\}}$, where $i \in \mathbf{N} \cup \{0\}$, have the desired measurability and distribution properties. For any $A \in \mathcal{B}(\mathcal{N}_0)$, $\{\Pi_{\{\pi,i\}} \in A\} = \bigcup_{p \in I} [\{\pi = p\} \cap \Pi_{\{p,i\}} \in A\}]$, whence the $\Pi_{\{\pi,i\}}$'s are measurable. Moreover,

$$\begin{aligned} P[\{\Pi_{\{\pi,i\}} \in A\}] &= \textstyle\sum_{p\in I} P[\{\pi = p\} \cap \{\Pi_{\{\pi,i\}} \in A\}] \\ &= \textstyle\sum_{p\in I} P[\{\pi = p\}] \cdot P[\Pi_{\{p,i\}} \in A] && \text{(by independence)} \\ &= \textstyle\sum_{p\in I} P[\{\pi = p\}] \cdot P[\Pi_{\{p_0,i\}} \in A] && \text{(for some } p_0 \in I) \\ &= P[\Pi_{\{p_0,i\}} \in A], \end{aligned}$$

so that, for any $p_0 \in I$, the randomly indexed $\Pi_{\{\pi,i\}}$ is equal to $\Pi_{\{p_0,i\}}$ in distribution. It is also important to note that the $\Pi_{\{\pi,i\}}$'s are independent when i varies:

Claim: For each $k \in \mathbf{N}$, $\{\pi, \Pi_{\{\pi,1\}}, \ldots, \Pi_{\{\pi,k\}}\}$ is an array of independent random elements.

Proof of Claim: Observe that the class consisting of $\emptyset$ and sets of the form $\{m\} \times A_1 \times \ldots \times A_k$, for some $m \in I$ and $A_1, \ldots, A_k \in \mathcal{B}(\mathcal{N}_0)$, is a π-system generating the product σ-field $\{\mathcal{B}(\mathbf{R}) \cap \pi(\Omega)\} \times \{\mathcal{B}(\mathcal{N}_0)\}^k$. For any such set we have

$$\begin{aligned}
&P[\pi = m, \Pi_{\{\pi,1\}} \in A_1, \ldots, \Pi_{\{\pi,k\}} \in A_k] \\
&= P[\pi = m, \Pi_{\{m,1\}} \in A_1, \ldots, \Pi_{\{m,k\}} \in A_k] \\
&= P[\pi = m] \cdot P[\Pi_{\{m,1\}} \in A_1] \cdot \ldots \cdot P[\Pi_{\{m,k\}} \in A_k] && \text{(by independence)} \\
&= P[\pi = m] \cdot P[\Pi_{\{\pi,1\}} \in A_1] \cdot \ldots \cdot P[\Pi_{\{\pi,k\}} \in A_k] && \text{(by the argument preceding this claim).}
\end{aligned}$$

The Claim follows. //

For the remainder of the proof, choose $n \in \mathbf{N}$, $p \in I$ and $x_1, \ldots, x_m \in S$ with $x_1 < \ldots < x_m$, and consider

$$G^n_{m+1}(\cdot; p; x_1, \ldots, x_m) := P[\tau_{m+1}(N^n) \le \cdot \mid \pi = p; \tau_1(N^n) = x_1, \ldots, \tau_m(N^n) = x_m]$$

as previously defined. Clearly, our characterization of elements of $\mathcal{N}_0$ by their jump sequences enables us to assume that, whenever $x_m = \infty$,

$$G_{m+1}(\cdot; p; x_1; \ldots; x_m) = \mathbf{1}_{[\cdot = \infty]} = G^n_{m+1}(\cdot; p; x_1, \ldots, x_m).$$

By Corollary 2.1.2, it suffices to prove that the G^n_{m+1}, G_m satisfy conditions (1) and (2) imposed by Lemma 2.3.2 on $\widetilde{G}^n_{m+1}$, $\widetilde{G}_{m+1}$ respectively; a cursory check will confirm this is the case when $x_m = \infty$. If, on the other hand, we have $0 < x_1 < \ldots < x_m < \infty$, then we must determine $G^n_{m+1}(\cdot; p; x_1; \ldots; x_m)$ directly. If $x_1, \ldots, x_m$ do not all belong to Q_n, then we may arbitrarily set

$$G^n_{m+1}(\cdot; p; x_1; \ldots; x_m) := G_{m+1}(\cdot; p; x_1; \ldots, x_m)$$

because the jumps of N^n are always in Q_n. If $x_1, \ldots, x_m \in Q_n$, then for any $x > x_m$ let

$$\underline{x}^n := \sup\left\{\frac{i}{2^n} : i \in \mathbf{N}, \frac{i}{2^n} \le x\right\}.$$

For $x > x_m$, we have:

$$\begin{aligned} & 1 - G^n_{m+1}(x; p; x_1, \ldots, x_m) \\ & = P[\tau_{m+1}(N^n) > x \mid \pi = p; \tau_1(N^n) = x_1, \ldots, \tau_m(N^n) = x_m] \end{aligned}$$

$$\begin{aligned} &= P[N^n(x_m, \underline{x}^n] = 0 \mid \pi = p; \tau_1(N^n) = x_1, \ldots, \tau_m(N^n) = x_m] \\ &= P[\Pi_{\{\pi,m\}}(\Lambda(\phi_m(x_1, \ldots, x_m), \pi, x_m), \Lambda(\phi_m(x_1, \ldots, x_m), \pi, \underline{x}^n)]) = 0 \\ &\qquad\qquad \mid \pi = p; \tau_1(N^n) = x_1; \ldots; \tau_m(N^n) = x_m] \\ &\qquad \text{(where } \phi_m(x_1, \ldots, x_m) \text{ is the measure } \textstyle\sum_{i=1}^m \delta_{x_i} \text{ on } (0, \infty)) \\ &= P[\Pi_{\{\pi,m\}}(\Lambda(\phi_m(x_1, \ldots, x_m), p, x_m), \Lambda(\phi_m(x_1, \ldots, x_m), p, \underline{x}^n)]) = 0 \\ &\qquad\qquad \mid \pi = p; \tau_1(N^n) = x_1; \ldots; \tau_m(N^n) = x_m]. \end{aligned}$$

Observe now that $(\pi, \tau_1(N^n), \ldots, \tau_m(N^n))$ is $\sigma\left\{\pi, \Pi_{\{\pi,1\}}, \ldots, \Pi_{\{\pi,m-1\}}\right\}$-measurable: indeed, for any $p_0 \in I$, $b \in \mathbf{R}_+$ and $i \in \{1, \ldots, m\}$ we have

$$\begin{aligned} [\pi = p_0, \tau_i(N^n) \le b] = \big[\Pi_{\{p_0,i-1\}}(&\Lambda(\phi_{i-1}(\tau_1(N^n), \ldots, \tau_{i-1}(N^n)), p_0, \tau_{i-1}(N^n)), \\ &\Lambda(\phi_{i-1}(\tau_1(N^n), \ldots, \tau_{i-1}(N^n)), p_0, \underline{b}^n)] \ge 1\big]. \end{aligned}$$

Since $\pi, \Pi_{\{\pi,1\}}, \ldots, \Pi_{\{\pi,m\}}$ are independent by the Claim, we deduce

$$\begin{aligned} & 1 - G^n_{m+1}(x; p; x_1, \ldots, x_m) \\ &= P[\Pi_{\{\pi,m\}}(\Lambda(\phi_m(x_1, \ldots, x_m), p, x_m), \Lambda(\phi_m(x_1, \ldots, x_m), p, \underline{x}^n)]) = 0] \\ &= P[\Pi_{\{p,m\}}(\Lambda(\phi_m(x_1, \ldots, x_m), p, x_m), \Lambda(\phi_m(x_1, \ldots, x_m), p, \underline{x}^n)]) = 0] \\ &\qquad (\Pi_{\{\pi,m\}} \text{ and } \Pi_{\{p,m\}} \text{ having the same distribution}) \\ &= \exp\{-[\Lambda(\phi(x_1, \ldots, x_m), p, \underline{x}^n) - \Lambda(\phi(x_1, \ldots, x_m), p, x_m)]\} \\ &= \exp\{-\ln[1 - F_{m+1}(\underline{x}^n - x_m; p; x_1, \ldots, x_m)]\} \\ &\qquad \text{(distribution-to-hazard function formula ([23]: Lemma 3.5))} \\ &= 1 - F_{m+1}(\underline{x}^n - x_m; p; x_1, \ldots, x_m) \\ &= 1 - G_{m+1}(\underline{x}^n; p; x_1, \ldots, x_m). \end{aligned}$$

We may therefore write

$$\begin{aligned} G^n_{m+1}(x; p; x_1, \ldots, x_m) = G_{m+1}(\underline{x}^n; p; x_1, \ldots, x_m) \cdot \textstyle\prod_{i=1}^m \mathbf{1}_{Q_n}(x_i) \\ + G_{m+1}(x; p; x_1, \ldots, x_m) \cdot \{\vee_{i=1}^m \mathbf{1}_{Q_n^c}(x_i)\}. \end{aligned}$$

The condition (1) that Lemma 2.3.2 imposed on $\widetilde{G}^n_{m+1}$ and $\widetilde{G}_{m+1}$ is already satisfied by G^n_{m+1} and G_{m+1}, respectively, by virtue of the present theorem's hypothesis (assumptions (1) and (2)). To verify that condition (2) of Lemma 2.3.2 also prevails, let $\{(y_1^n, \ldots, y_m^n)\}_{n=1}^{\infty}$ be a sequence in J_m converging to $(x_1, \ldots, x_m)$ (remember that $x_1 < \ldots < x_m < \infty$). For any $x > 0$, if $x = \infty$ we certainly - and trivially - have that

$$\limsup_{n \to \infty} \mid G_{m+1}^n(x; p; y_1^n, \ldots, y_m^n) - G_{m+1}(x; p; y_1^n, \ldots, y_m^n) \mid = 0.$$

If, by contrast, $x \in (0, \infty)$, then

$$\begin{aligned}
&\limsup_{n\to\infty} \mid G_{m+1}^n(x; p; x_1, \ldots, x_m) - G_{m+1}(x; p : y_1^n, \ldots, y_m^n) \mid \\
&\quad \leq \limsup_{n\to\infty} G_{m+1}(x; p; y_1^n, \ldots, y_m^n) - G_{m+1}(\underline{x}^n; p; y_1^n, \ldots, y_m^n) \\
&\quad \leq \limsup_{n\to\infty} G_{m+1}(x; p; y_1^n, \ldots, y_m^n) - G_{m+1}(\underline{x}^k; p; y_1^n; \ldots, y_m^n) \\
&\qquad\qquad \text{(for some fixed } k \in \mathbf{N},\ \{\underline{x}^n\}_{n=1}^{\infty} \text{ being increasing)} \\
&\quad = G_{m+1}(x; p; x_1; \ldots, x_m) - G_{m+1}(\underline{x}^k; p; x_1, \ldots, x_m) \\
&\qquad\qquad \text{(by assumption (1) of the theorem's statement)}
\end{aligned}$$

Under assumption (2) of the hypothesis, however, this expression can be made arbitrarily small by letting $k \to \infty$. Thus, by Corollary 2.1.2 and Lemma 2.3.2, $N^n \longrightarrow_{\mathcal{L}} N$ as $n \to \infty$. □

2.4 Application to Mixtures of Poisson Processes

An example illustrates how the sequence $\{N^n\}_{n=1}^{\infty}$ of approximating processes of Theorem 2.3.3 converges to the limiting distribution N. We shall consider a Cox point process $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$ whose "initial information" is generated by a parameter $\pi : (\Omega, \mathcal{F}, P) \longrightarrow (\mathbf{R}, \mathcal{B}(\mathbf{R}))$ which takes a discrete range I of realizations, and we shall verify independently that $N^n \longrightarrow_{\mathcal{L}} N$ by direct computation. The point process N shall first be specified via its path compensator.

Suppose N has a path compensator $\Lambda : \mathcal{N}_0 \times \mathbf{R} \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ with initial parameter π such that $\forall \mu \in \mathcal{N}_0, s \in I, t \in \mathbf{R}_+$,

$$\Lambda(\mu, s, t) := \begin{cases} r_s \cdot t & \text{if } s \in I \\ 0 & \text{otherwise} \end{cases},$$

where $r_s \in (0, \infty)$ is some constant depending on $s \in I$ (note that there must exist a point process having such a path compensator by [23]). Since π takes its values in I with certainty, we suspect that N is in fact a Poisson process conditioned on π because, given $\pi = s$, it has a linear compensator with slope r_s. Let us check that this indeed is the case.

Given the event $[\pi = s]$ and some $n \in \mathbf{N}$, it is readily seen from the formula in the statement of Theorem 2.3.3, that N^n is distributed as a "discrete homogeneous Bernoulli process of rate r_s" in the sense of McDonald ([41]: pg. 111).

More precisely, by Proposition 4.2.2 of [41], conditionally on $[\pi = s]$ we have:

1. The interarrival times $\{\tau_m(N^n) - \tau_{m-1}(N^n)\}_{m=1}^{\infty}$ of N^n are independent and geometrically distributed with parameter $1 - e^{\frac{r_s}{2^n}}$;
2. N^n has independent increments;
3. N^n has stationary increments;
4. $\forall t \in \mathbf{R}, j \in \mathbf{N} \cup \{0\}$, $P[N^n(t) = j \mid \pi = s] = \left(\begin{array}{c} [2^n t] \\ j \end{array} \right) (\gamma_s^n)^j (1 - \gamma_s^n)^{[2^n t]-j}$

where $[\cdot]$ denotes the integer part and where $\gamma_s^n := \exp\{-\frac{r_s}{2^n}\}$.
If the limiting process N is a mixed Poisson process with initial paramenter π, then we expect, for every finite collection $(u_1, v_1], \ldots, (u_m, v_m]$ of disjoint, bounded subintervals of $(0, \infty)$ and for corresponding nonnegative, integer constants $k_1, \ldots, k_m$, to obtain the equality

$$P[N_{v_1} - N_{u_1} = k_1, \ldots, N_{v_m} - N_{u_m} = k_m] = \sum_{s \in I} \left\{ \prod_{i=1}^{m} \frac{[r_s(v_i - u_i)]^{k_i}}{k_i!} e^{-r_s(v_i - u_i)} \right\} P[\pi = s],$$

and such relations would uniquely characterize the distribution of N. But observe that for finite $n \in \mathbf{N}$,

$$P[N_{v_1}^n - N_{u_1}^n = k_1, \ldots, N_{v_m}^n - N_{u_m}^n = k_m] = \sum_{s \in S} \left\{ \prod_{i=1}^{m} \left(\begin{array}{c} [2^n (v_i - u_i)] \\ k_i \end{array} \right) (\gamma_s^n)^{k_i} (1 - \gamma_s^n)^{2^n (v_i - u_i) - k_i} \right\} P[\pi = s],$$

and this expression, by a remark from [41] (pg. 114) coupled with a bounded convergence argument, tends to

$$\sum_{s\in I}\left\{\prod_{i=1}^{m}\frac{[r_s(v_i-u_i)]^{k_i}}{k_i!}e^{-r_s(v_i-u_i)}\right\}P[\pi=s]$$

as $n\to\infty$, so that N is indeed the expected mixed Poisson process with initial parameter π.

2.5 Simulating Fidi's: a Validation of Numerically Computed Time Changes

A common method of "constructing" a certain point process distribution via a given path compensator consists of using this path compensator to effect a time change transformation on a version of the Standard Poisson Process. More precisely, let $N:(\Omega,\mathcal{F},P)\longrightarrow(\mathcal{N}_0,\mathcal{B}(\mathcal{N}_0))$ be a point process with a path compensator $\Lambda:\mathcal{N}_0\times\mathbf{R}_+\longrightarrow\mathbf{R}_+$ which is such that, for any $\mu\in\mathcal{N}_0$,

1. $\lim_{t\to\infty}\Lambda(\mu,t)=\infty$, and
2. $t\mapsto\Lambda(\mu,t)$ is continuous;

additionally, let $\Pi:(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{P})\longrightarrow(\mathcal{N}_0,\mathcal{B}(\mathcal{N}_0))$ be a Standard Poisson Process on $\mathbf{R}_+$. For any $t_1,\ldots,t_m$ such that for any $i\in\{1,\ldots,m-1\}$, $t_i<\infty\implies 0<t_i<t_{i+1}$, let $\phi_m(t_1,\ldots,t_m)$ be the $\mathcal{N}_0$-measure such that $\tau_1(\phi_m(t_1,\ldots,t_m))=t_1,\ldots,\tau_m(\phi_m(t_1,\ldots,t_m))=t_m$ and $\tau_{m+1}(\phi_m(t_1,\ldots,t_m))=\infty$, and make the convention to denote the null measure on $\mathbf{R}_+$ by $\phi_0(\emptyset)$, where "$\emptyset$" represents an empty list of arguments. Define the "time change transform" $\Gamma:(\mathcal{N}_0,\mathcal{B}(\mathcal{N}_0))\longrightarrow(\mathcal{N}_0,\mathcal{B}(\mathcal{N}_0))$ via its image jump points as follows: for any measure $\mu\in\mathcal{N}_0$, let

$s_0:=0;$
$s_1:=\inf\{t>0:\Lambda(\phi_0(\emptyset),(0,t])\geq\tau_1(\mu)\};$

$s_2 := s_1 + \inf\{t > 0 : \Lambda(\phi_1(s_1), (s_1, s_1 + t]) \geq \tau_2(\mu) - \tau_1(\mu)\}$;

$\vdots$

$s_{m+1} := s_m + \inf\{t > 0 : \Lambda(\phi_m(s_1, \ldots, s_m), (s_m, s_m + t]) \geq \tau_{m+1}(\mu) - \tau_m(\mu)\}$

$\vdots$

and put $\tau_m(\Gamma(\mu)) := s_m$ for $m \in \mathbf{N}$. Proposition 3.1 of [36] states that, under such circumstances, $N =_{\text{st}} \Gamma(\Pi)$ (where $=_{\text{st}}$ denotes equality in distribution).

Now suppose that, for some nonempty, bounded Borel sets $B_1, \ldots, B_m \subseteq \mathbf{R}_+$, we would be interested in finding the distribution of the random vector $(N(B_1), \ldots, N(B_m))$. Except for some very simple (and trivial) cases, determining such a distribution in exact form would constitute a daunting combinatorial problem in practice - all conceivable arguments would have to take the (random) locations of the sequentially dependent jump points of N into account.

In the face of such difficulties, we would be tempted to *approximate* the distribution of $(N(B_1), \ldots, N(B_m))$ by means of *simulations*. Since we assume from the outset that N's compensator Λ satisfies (1) and (2), the above construction of $\Gamma(\Pi) =_{\text{st}} N$ suggests a way of specifying the desired simulation algorithm. Suppose our software provides a pseudo-random exponential generator **rand**(). For any m intervals $I_1, \ldots, I_m \subseteq \mathbf{R}_+$ we would estimate $P[N(B_1) \in I_1, \ldots, N(B_m) \in I_m]$ by averaging the number of times the following procedure returns the value 1:

Procedure INDICATOR
begin %{procedure}%
    $i := 0$;
    $i_1 := 0, \ldots, i_m := 0$;
    $T_0 := 0$;
    $T := 0$;
    $E := 0$;
    while $T \leq \sup \bigcup_{k=1}^{m} B_k$ do
        begin %{begin while}%

```
        i := i + 1;
        E := rand();
        T := T_{i-1} + inf{t ∈ R_+ : Λ((φ_{i-1}(T_1, ..., T_{i-1}), (T_{i-1}, t]) ≥ E};
        T_i := T;
        for j := 1 to m
            if T_i ∈ B_j then
                i_j := i_j + 1;
        end;        %{end while}%
    if i_1 ∈ I_1, ..., i_m ∈ I_m then
        return 1 else
        return 0.
end     %{procedure}%
```

Here, we assume the pseudo-random exponential generator **rand()** simulates the distribution of any given interarrival time of the Standard Poisson Process.

There is only one caveat arising from such an approach in practice. Because computers can only generate *discrete* values, the values of **rand()** are bound to be discrete - in other words, **rand()** can only *approximate* an exponential distribution, without ever being truly exponential. At this point, the reader may wonder what all the fuss is about. After all, doesn't any kind of numerical computation involve round-off errors which can be kept within a certain margin of tolerance, provided the computational device being used preserves enough digits?

This is not quite true in our case. The reason for this is that the errors introduced can have a cumulative effect if the path compensator Λ is sensitive to perturbations of its measure argument. The procedure INDICATOR produces a value T_1 which simulates $\tau_1(N)$, a value T_2 which simulates $\tau_2(N)$, and so on. However, since **rand()** is not truly exponentially distributed, T_1 will not be equal in law to $\tau_1(N)$ - in fact, T_1's distribution will admit a discrete support that $\tau_1(N)$'s distribution may not even charge. This is not the whole story. The procedure then uses **rand()** again, in conjunction with T_1, to produce the value T_2. Since T_1 is not distributed exactly as $\tau_1(N)$ is, the obtained value of T_2 will reflect the errors of both **rand()** *and* T_1. Similarly, for any $k \in \mathbf{N}$, the obtained value of T_k will reflect errors inherent in **rand()**, $T_1, \ldots, T_{k-1}$. To see how such effects may lead to a fundamentally erroneous approximation N's distribution, consider the following simple example:

Example 2.5.1 *Suppose N has a path compensator Λ which is specified in the following fashion. For any (possibly empty) list of strictly increasing values $t_1 < t_2 < \ldots < t_m$ define the increasing, continuous path $t \mapsto \Lambda(\phi(t_1, \ldots, t_m), t)$ by writing*

$$\Lambda(\phi(t_1, \ldots, t_m), t) := \begin{cases} 2t & \text{if } t_i \in \mathbf{Q} \text{ for some } i \in \{1, \ldots, m\} \\ t & \text{otherwise} \end{cases}.$$

Since **rand()** *can only produce rational values, T_1 would be a rational number in this case; from the way Λ is specified, the distributions $T_2 - T_1$, $T_3 - T_2 \ldots$ would approximate an exponential distribution with parameter 2. However, since the countable set $\mathbf{Q}$ is not charged by the exponentially distributed $\tau_1(N)$, the true interarrival times $\tau_2(N) - \tau_1(N)$, $\tau_3(N) - \tau_2(N) \ldots$ would actually be exponentially distributed with parameter 1.*

Fortunately, if, in addition to (1) and (2), we assume that Λ is continuous in its *measure* argument for t fixed, we will obtain a valid approximation of N's distribution. In fact, we shall now show that if this extra continuity requirement is satisfied by Λ, then for any $n \in \mathbf{N}$, an algorithmic approximation which rounds the value of **rand()** to the next multiple of $\frac{1}{2^n}$, is equal in distribution to the approximation N^n produced by Theorem 2.3.3.

Let $n \in \mathbf{N}$. For any $x \in \mathbf{R}_+$ let $\overline{x}^n := \inf\{\frac{i}{2^n} : i \in \mathbf{N}, \frac{i}{2^n} \geq x\}$. We define $\Gamma^n : (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0)) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$ almost the same way we defined Γ, except that we round the values of $s_1, s_2, \ldots$ up to the next integer multiple of $\frac{1}{2^n}$. For any measure $\mu \in \mathcal{N}_0$, let

$s_0 := 0;$
$s_1 := \inf\{t > 0 : \Lambda(\phi_0(\emptyset), (0,t]) \geq \tau_1(\mu)\};$
$s_2 := \overline{s_1}^n + \inf\{t > 0 : \Lambda(\phi_1(\overline{s_1}^n), (\overline{s_1}^n, \overline{s_1}^n + t]) \geq \tau_2(\mu) - \tau_1(\mu)\};$
$\vdots$
$s_{m+1} := \overline{s_m}^n + \inf\{t > 0 : \Lambda(\phi_m(\overline{s_1}^n, \ldots, \overline{s_m}^n), (\overline{s_m}^n, \overline{s_m}^n + t]) \geq \tau_{m+1}(\mu) - \tau_m(\mu)\}$
$\vdots$

and put $\tau_m(\Gamma(\mu)) := \overline{s_m}^n$ for $m \in \mathbf{N}$.

Theorem 2.5.2 *Let* $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$ *be a point process whose path compensator* $\Lambda : \mathcal{N}_0 \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ *satisfies:*

1. $\lim_{t\to\infty} \Lambda(\mu, t) = \infty \; \forall \mu \in \mathcal{N}_0$*;*
2. $t \mapsto \Lambda(\mu, t)$ *is continuous* $\forall \mu \in \mathcal{N}_0$*;*
3. $\mu \mapsto \Lambda(\mu, t)$ *is continuous* $\forall t \in \mathbf{R}_+$*.*

For any $n \in N$ *let* N^n *be defined as in Theorem 2.3.3. If* $\Pi : (\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P}) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$ *is a version of the Standard Poisson Process, then* $N^n =_{st} \Gamma^n(\Pi)$*.*

Proof In order to conform to the set-up of Theorem 2.3.3 (applied to the case where there is no parameter π), we start with a sequence $\Pi_0, \Pi_1, \Pi_2, \ldots : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$ of independent versions of the Standard Poisson Process, which are also independent of $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$. We also let $\theta = \phi_0(\emptyset)$ denote the null measure on $\mathbf{R}_+$.

Let $n \in \mathbf{N}$. To simplify notation, we shall denote the jump points $\tau_1(N^n), \tau_2(N^n), \ldots$ by $\tau_1, \tau_2, \ldots$ and $\tau_1(\Gamma^n(\Pi)), \tau_2(\Gamma^n(\Pi)), \ldots$ by $T_1, T_2, \ldots$, respectively. We shall prove that

i) for any $m \in \mathbf{N}$, $P[\tau_1 > \frac{m}{2^n}] = \widetilde{P}[T_1 > \frac{m}{2^n}]$, and

ii) for any $m_1, \ldots, m_k \in \mathbf{N}$ such that $m_1 < m_2 < \ldots < m_k$ and any $m > m_k$, $P\left[\tau_{k+1} > \frac{m}{2^n} | \tau_1 = \frac{m_1}{2^n}, \ldots \tau_k = \frac{m_k}{2^n}\right] = \widetilde{P}\left[T_{k+1} > \frac{m}{2^n} | T_1 = \frac{m_1}{2^n}, \ldots, T_k = \frac{m_k}{2^n}\right]$.

It will follow that $(\tau_1, \tau_2, \ldots) =_{\text{st}} (T_1, T_2, \ldots)$, whence $N^n =_{\text{st}} \Gamma^n(\Pi)$.

For (i): Choose $m \in \mathbf{N}$. We have

$$\begin{aligned}
&\widetilde{P}\left[T_1 > \tfrac{m}{2^n}\right] \\
&= \widetilde{P}\left[\Lambda\left(\phi_0(\emptyset), \left(0, \tfrac{m}{2^n}\right]\right) < \tau_1(\Pi)\right] \\
&= \exp\left\{-\Lambda\left(\phi_0(\emptyset), \tfrac{m}{2^n}\right)\right\} \quad (\tau_1(\Pi) \text{ is exponential with parameter } 1) \\
&= \exp\left\{-\Lambda\left(\theta, \tfrac{m}{2^n}\right)\right\} \\
&= P\left[\tau_1(\Pi_0) > \Lambda\left(\theta, \tfrac{m}{2^n}\right)\right] \quad (\tau_1(\Pi_0) \text{ is exponential with parameter } 1) \\
&= P\left[\Pi_0\left(\omega, \Lambda\left(\theta, \tfrac{m}{2^n}\right)\right) = 0\right] \\
&= P\left[\Pi_0\left(\omega, \Lambda\left(\theta, \tfrac{1}{2^n}\right)\right) = 0,\right. \\
&\qquad \Pi_0\left(\omega, \Lambda\left(\theta, \tfrac{2}{2^n}\right)\right) - \Pi_0\left(\omega, \Lambda\left(\theta, \tfrac{1}{2^n}\right)\right) = 0, \\
&\qquad \vdots \\
&\qquad \left.\Pi_0\left(\omega, \Lambda\left(\theta, \tfrac{m}{2^n}\right)\right) - \Pi_0\left(\omega, \Lambda\left(\theta, \tfrac{m-1}{2^n}\right)\right) = 0\right] \\
&= P\left[\Delta N^n_{\frac{1}{2^n}} = 0, \Delta N^n_{\frac{2}{2^n}} = 0, \ldots, \Delta N^n_{\frac{m}{2^n}} = 0\right] \\
&= P\left[\tau_1 > \tfrac{m}{2^n}\right].
\end{aligned}$$

For (ii): Let $m_1, \ldots, m_k, m$ be as hypothesized and recall the definition of $\overline{x}^n$ from p. 32. We have

$$\begin{aligned}
&\widetilde{P}\left[T_{k+1} > \tfrac{m}{2^n} \middle| T_1 = \tfrac{m_1}{2^n}, \ldots, T_k = \tfrac{m_k}{2^n}\right] \\
&= \widetilde{P}\left[T_k + \overline{\inf\{t > 0 : \Lambda\left(\phi_k(T_1, \ldots, T_k), (T_k, T_k + t]\right) \geq \tau_{k+1}(\Pi) - \tau_k(\Pi)\}}^n > \tfrac{m}{2^n}\right. \\
&\qquad\qquad \left.\mid T_1 = \tfrac{m_1}{2^n}, \ldots, T_k = \tfrac{m_k}{2^n}\right] \\
&= \widetilde{P}\left[T_k + \inf\{t > 0 : \Lambda\left(\phi_k(T_1, \ldots, T_k), (T_k, T_k + t]\right) \geq \tau_{k+1}(\Pi) - \tau_k(\Pi)\} > \tfrac{m}{2^n}\right. \\
&\qquad\qquad \left.\mid T_1 = \tfrac{m_1}{2^n}, \ldots, T_k = \tfrac{m_k}{2^n}\right] \\
&= \widetilde{P}\left[\inf\left\{t > 0 : \Lambda\left(\phi_k\left(\tfrac{m_1}{2^n}, \ldots, \tfrac{m_k}{2^n}\right), \left(\tfrac{m_k}{2^n}, \tfrac{m_k}{2^n} + t\right]\right) \geq \tau_{k+1}(\Pi) - \tau_k(\Pi)\right\}\right. \\
&\qquad\qquad \left.> \tfrac{m-m_k}{2^n} \mid T_1 = \tfrac{m_1}{2^n}, \ldots, T_k = \tfrac{m_k}{2^n}\right].
\end{aligned}$$

Now $\tau_{k+1}(\Pi) - \tau_k(\Pi)$ is, of course, independent of $(\tau_1(\Pi), \ldots, \tau_k(\Pi))$; since $(T_1, \ldots, T_k)$ is $\sigma(\tau_1(\Pi), \ldots, \tau_k(\Pi))$-measurable, as may be checked by induction on k, it follows that the last quantity is equal to

$$\begin{aligned}
&\widetilde{P}\left[\inf\left\{t > 0 : \Lambda\left(\phi_k\left(\tfrac{m_1}{2^n}, \ldots, \tfrac{m_k}{2^n}\right), \left(\tfrac{m_k}{2^n}, \tfrac{m_k}{2^n} + t\right]\right) \geq \tau_{k+1}(\Pi) - \tau_k(\Pi)\right\} > \tfrac{m-m_k}{2^n}\right] \\
&= \widetilde{P}\left[\Lambda\left(\phi_k\left(\tfrac{m_1}{2^n}, \ldots, \tfrac{m_k}{2^n}\right), \left(\tfrac{m_k}{2^n}, \tfrac{m_k}{2^n} + \tfrac{m-m_k}{2^n}\right]\right) < \tau_{k+1}(\Pi) - \tau_k(\Pi)\right] \\
&= \exp\left\{-\Lambda\left(\phi_k\left(\tfrac{m_1}{2^n}, \ldots, \tfrac{m_k}{2^n}\right), \left(\tfrac{m_k}{2^n}, \tfrac{m}{2^n}\right]\right)\right\}
\end{aligned}$$

$(\tau_{k+1}(\Pi) - \tau_k(\Pi)$ is exponential with parameter 1).

Likewise,

$$
\begin{aligned}
&P\left[\tau_{k+1} > \frac{m}{2^n} \mid \tau_1 = \frac{m_1}{2^n}, \dots, \tau_k = \frac{m_k}{2^n}\right] \\
&= P\left[\Delta N^n_{\frac{m_k+1}{2^n}} = 0, \Delta N^n_{\frac{m_k+2}{2^n}} = 0, \dots, \Delta N^n_{\frac{m}{2^n}} = 0 \mid \tau_1 = \tfrac{m_1}{2^n}, \dots, \tau_k = \tfrac{m_k}{2^n}\right] \\
&= P\left[\Pi_k\left(\omega, \Lambda\left(\phi_k(\tau_1,\dots,\tau_k), \tfrac{m_k+1}{2^n}\right)\right) - \Pi_k\left(\omega, \Lambda\left(\phi_k(\tau_1,\dots,\tau_k), \tfrac{m_k}{2^n}\right)\right) = 0,\right. \\
&\qquad \Pi_k\left(\omega, \Lambda\left(\phi_k(\tau_1,\dots,\tau_k), \tfrac{m_k+2}{2^n}\right)\right) - \Pi_k\left(\omega, \Lambda\left(\phi_k(\tau_1,\dots,\tau_k), \tfrac{m_k+1}{2^n}\right)\right) = 0, \\
&\qquad \vdots \\
&\qquad \Pi_k\left(\omega, \Lambda\left(\phi_k(\tau_1,\dots,\tau_k), \tfrac{m}{2^n}\right)\right) - \Pi_k\left(\omega, \Lambda\left(\phi_k(\tau_1,\dots,\tau_k), \tfrac{m-1}{2^n}\right)\right) = 0, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \left.\mid \tau_1 = \tfrac{m_1}{2^n}, \dots, \tau_k = \tfrac{m_k}{2^n}\right] \\
&= P\left[\Pi_k\left(\omega, \Lambda\left(\phi_k(\tau_1,\dots,\tau_k), \tfrac{m}{2^n}\right)\right) - \Pi_k\left(\omega, \Lambda\left(\phi_k(\tau_1,\dots,\tau_k), \tfrac{m_k}{2^n}\right)\right) = 0\right. \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \left.\mid \tau_1 = \tfrac{m_1}{2^n}, \dots, \tau_k = \tfrac{m_k}{2^n}\right].
\end{aligned}
$$

It may be shown by induction on k that $(\tau_1, \dots, \tau_k)$ is $\sigma(\Pi_0, \dots, \Pi_{k-1})$-measurable (this was done in the proof of Theorem 2.3.3); since Π_k is independent of $(\Pi_0, \dots, \Pi_{k-1})$, we obtain that the last expression is equal to

$$
\begin{aligned}
&P\left[\Pi_k\left(\omega, \Lambda\left(\phi_k(\tfrac{m_1}{2^n}, \dots, \tfrac{m_k}{2^n}), \tfrac{m}{2^n}\right)\right) - \Pi_k\left(\omega, \Lambda\left(\phi_k(\tfrac{m_1}{2^n}, \dots, \tfrac{m_k}{2^n}), \tfrac{m_k}{2^n}\right)\right) = 0\right] \\
&= \exp\left[-\left\{\Lambda\left(\phi_k\left(\tfrac{m_1}{2^n}, \dots, \tfrac{m_k}{2^n}\right), \tfrac{m}{2^n}\right) - \Lambda\left(\phi_k\left(\tfrac{m_1}{2^n}, \dots, \tfrac{m_k}{2^n}\right), \tfrac{m_k}{2^n}\right)\right\}\right] \\
&= \exp\left\{-\Lambda\left(\phi_k\left(\tfrac{m_1}{2^n}, \dots, \tfrac{m_k}{2^n}\right), \left(\tfrac{m_k}{2^n}, \tfrac{m}{2^n}\right]\right)\right\},
\end{aligned}
$$

so that $\widetilde{P}\left[T_{k+1} > \frac{m}{2^n} | T_1 = \frac{m_1}{2^n}, \dots, T_k = \frac{m_k}{2^n}\right] = P\left[\tau_{k+1} > \frac{m}{2^n} \mid \tau_1 = \frac{m_1}{2^n}, \dots, \tau_k = \frac{m_k}{2^n}\right]$.
□

Chapter 3

Compensator of the Non-simple Point Process

The object of the present chapter is to determine a regenerative form for the compensator of a non-simple point process indexed by the non-negative reals, with marks in a Lusin topological space (see Appendix D). This extends the result of Jacod ([23]), who determined the regenerative form for the compensator of a simple point process. We essentially emulate Jacod's approach ([24]; pp. 83-86), while providing more details. Lemma 3.3.1 is simply assumed by Jacod ([24]), and our existence proof for Theorem 3.3.2 is much more detailed than Jacod's proof of Proposition 3.41 ([24]).

3.1 The Random Measure μ

Let E be a Lusin topological space, i.e. a Borel subset of a compact metric space (the Lusin property is required to ensure the existence of certain stochastic kernels, as will become apparent later). It is possible to add a point $\Delta \notin E$ to E in a manner which topologizes the set $\bar{E} := E \cup \{\Delta\}$ as a Lusin space as well, and so that the topology of E relative to $\bar{E}$ corresponds to E's original topology (see Appendix D). The Borel classes of E and $\bar{E}$ shall be denoted by $\mathcal{E}$ and $\bar{\mathcal{E}}$ respectively. We consider a sequence

$$T_1, T_2, \ldots : (\Omega, \mathcal{F}, P) \longrightarrow (\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+))$$

of random variables which satisfy

$$0 \equiv T_0 < T_1 \le T_2 \le \ldots \le T_\infty := \sup_{n \in \mathbf{N}} T_n \le T_{\infty+1} := \infty$$

and such that, for any $n \in \mathbf{N}$,

$$T_n < \infty \Longrightarrow T_n < T_\infty.$$

To every T_n, $n \in \mathbf{N}$, we associate an $\bar{E}$-valued random variable

$$X_n : (\Omega, \mathcal{F}, P) \longrightarrow (\bar{E}, \bar{\mathcal{E}})$$

such that

$$X_n = \Delta \Longleftrightarrow T_n = \infty.$$

The random measure μ on $(0, \infty) \times E$ is defined by putting

$$\mu(\omega; dt, dx) := \sum_{n \in \mathbf{N}} \delta_{(T_n(\omega), X_n(\omega))}(dt, dx) \cdot \mathbf{1}_{[T_n < \infty]}.$$

Note that for each $\omega \in \Omega$, $t \in \mathbf{R}_+$ and $B \in \mathcal{E}$, $\mu(\omega; (0, t] \times B)$ counts the number of pairs (T_n, X_n) such that $0 < T_n \le t$ and $X_n \in B$.

For any $t \in \mathbf{R}_+$ we put

$$\mathcal{G}_t := \sigma(\mu((0, s] \times B) : s \le t, B \in \mathcal{E})$$

and define, for some sub-σ-field $\mathcal{F}_0$ of $\mathcal{F}$, the filtration $\{\mathcal{F}_t\}_{t \ge 0}$ via the relation

$$\mathcal{F}_t := \mathcal{F}_0 \vee \mathcal{G}_t.$$

Because E is separable (every Lusin topological space is homeomorphic to a Borel subset of $[0,1]^{\mathbf{N}}$ by Dellacherie and Meyer ([13], pp. 48-49)), its topology admits a countable base $\{G_n\}_{n=1}^{\infty}$, which implies that the filtration $\{\mathcal{F}_t\}_{t \ge 0}$ coincides with the internal history of the $(\Omega, \mathcal{F}_0) \times (\bar{\mathbf{R}}_+^{\infty}, \mathcal{B}(\bar{\mathbf{R}}_+^{\infty}))$-valued process $Y = (Y_t)_{t \ge 0}$ defined by the relation

$$Y_t(\omega) := (\omega, \mu((0, t] \times G_1), \mu((0, t] \times G_2), \ldots)$$

for all $(\omega, t) \in \Omega \times \mathbf{R}_+$. Since this ensures that the hypotheses of Theorem A2.T28 of [7] are satisfied, it follows that the σ-field $\mathcal{F}_T$ associated to any $\{\mathcal{F}\}_{t\geq 0}$-stopping time T satisfies the equalities

$$\mathcal{F}_T = \sigma(Y_{T\wedge t} : t \geq 0) = \mathcal{F}_0 \vee \sigma(\mu((0, T \wedge t]) \times B : t \geq 0, B \in \mathcal{E}).$$

In particular, for any $n \in \mathbf{N}$,

$$\mathcal{F}_{T_n} = \mathcal{F}_0 \vee \sigma(T_i, X_i : 1 \leq i \leq n).$$

Moreover, since $\mathcal{F}_{T_\infty} = \mathcal{F}_0 \vee \sigma(\mu((0, T_\infty \wedge t]) \times B : t \geq 0, B \in \mathcal{E})$, one ensures

$$\bigvee_{n\in\mathbf{N}} \mathcal{F}_{T_n} = \mathcal{F}_{T_\infty}$$

by checking that for any $B \in \mathcal{E}$ and $t \in \mathbf{R}_+$, $\mu((0, T_\infty \wedge t] \times B) = \sup_{n\in\mathbf{N}} \mu((0, T_n \wedge t] \times B)$.

3.2 The Predictable Processes

In this section, we recall a conventional definition of predictability, along with the special representation of predictable processes inferred from the filtration $\{\mathcal{F}_t\}_{t\geq 0}$'s structure. We first make an observation pertaining to that very structure.

Proposition 3.2.1 *For any $n \in \mathbf{N}_0 = \mathbf{N} \cup \{0\}$, $t \in \mathbf{R}_+$ and $A \in \mathcal{F}_t$, there exists a set $A_n \in \mathcal{F}_{T_n}$ such that*

$$A \cap \{t < T_{n+1}\} = A_n \cap \{t < T_{n+1}\}.$$

Proof Let us first adopt the convention that for any set $A \in \mathcal{F}$ and any measurable element $X : (\Omega, \mathcal{F}) \longrightarrow (\bar{E}, \bar{\mathcal{E}})$, $X \mid_A$ denotes the restriction

$$\begin{array}{rccl} X \mid_A : & (A, \mathcal{F} \cap A) & \longrightarrow & (\bar{E}, \bar{\mathcal{E}}) \\ & \omega & \longmapsto & X(\omega) \end{array}.$$

Now, for any $n \in \mathbf{N}_0$ and $t \in \mathbf{R}_+$,

$$\begin{aligned} \mathcal{F}_t \cap \{t < T_{n+1}\} &= [\sigma(\mu((0, s] \times B)) : s \leq t, B \in \mathcal{E}) \vee \mathcal{F}_0] \cap \{t < T_{n+1}\} \\ &= \left[\sigma(T_i \cdot \mathbf{1}_{[T_i < t]}, X_i \mid_{\{T_i < t\}} : 1 \leq 1 \leq n) \vee \mathcal{F}_0\right] \cap \{t < T_{n+1}\} \\ &\subseteq [\sigma(T_i, X_i : 1 \leq i \leq n) \vee \mathcal{F}_0] \cap \{t < T_{n+1}\} \\ &= \mathcal{F}_{T_n} \cap \{t < T_{n+1}\}. \ \square \end{aligned}$$

Let us now define the predictable σ-algebra on $\Omega \times \mathbf{R}_+ \times E$.

Definition 3.2.2 *The predictable σ-algebra on $\Omega \times \mathbf{R}_+ \times E$ is the sub-σ-algebra of $\mathcal{F} \otimes \mathcal{B}(\mathbf{R}_+) \otimes \mathcal{E}$ generated by the sets*

i) $A \times \{0\} \times B$, *where* $A \in \mathcal{F}_0$ *and* $B \in \mathcal{E}$, *and*

ii) $]\!]0, T]\!] \times B$, *where T is an $\{\mathcal{F}_t\}_{t\geq 0}$-stopping time and $B \in \mathcal{E}$ (recall that* $]\!]0, T]\!] \equiv \{(\omega, t) \in \Omega \times \mathbf{R}_+ : 0 < t \leq T(\omega)\}$*).*

It is denoted by $\mathcal{P}$. A stochastic process on $\mathbf{R}_+ \times E$ (i.e, one indexed by $\mathbf{R}_+ \times E$) is said to be predictable if it is measurable with respect to $\mathcal{P}$.

The representation of predictable processes may now be established: the restriction of a predictable process to a stochastic interval of the form $]\!]T_n, T_{n+1}]\!]$ is $\mathcal{F}_{T_n} \otimes \mathcal{B}(\mathbf{R}_+) \otimes \mathcal{E}$-measurable. This fact is invoked in the subsequent analysis.

Proposition 3.2.3 *Let $X = (X_{(t,x)} : t \in \mathbf{R}_+, x \in E)$ be an $\bar{\mathbf{R}}_+$-valued stochastic process on $\mathbf{R}_+ \times E$. X is predictable if and only if the process $X_0 = (X_{(0,x)})_{x\in E}$ is $\mathcal{F}_0 \otimes \mathcal{E}$-measurable, and for all $n \in \mathbf{N} \cup \{0, \infty\}$, there exists an $\mathcal{F}_{T_n} \otimes \mathcal{B}(\mathbf{R}_+) \otimes \mathcal{E}$-measurable process $X^n = (X^n_{(t,x)} : t \in \mathbf{R}_+, x \in E)$ such that*

$$X \equiv X_0 \cdot \mathbf{1}_{[0]\times E} + \sum_{n\in\mathbf{N}_0} X^n \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E} + X^\infty \cdot \mathbf{1}_{[T_\infty,\infty[\times E}.$$

Proof For sufficiency, we need to establish the predictability of processes of the form

$$X_0 \cdot \mathbf{1}_{[0]\times E} + \sum_{n\in\mathbf{N}_0} X^n \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E} + X^\infty \cdot \mathbf{1}_{[T_\infty,\infty[\times E},$$

where, for any $n \in \mathbf{N} \cup \{0, \infty\}$, X^n is $\mathcal{F}_{T_n} \otimes \mathcal{B}(\mathbf{R}_+) \otimes \mathcal{E}$-measurable. First observe that for any $a \in \mathbf{R}_+$, $\{X_0 \cdot \mathbf{1}_{[0]\times E} > a\} = \{X_0 > a\} \times \{0\}$. As it can be verified that $\mathcal{P}$ includes all sets of the form $U \times \{0\}$, where $U \in \mathcal{F}_0 \otimes \mathcal{E}$, it follows that $X_0 \cdot \mathbf{1}_{[0]\times E}$ is predictable. To show that $X^n \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}$ and $X^\infty \cdot \mathbf{1}_{[T_\infty,\infty[\times E}$ are predictable, for any $n \in \mathbf{N} \cup \{0, \infty\}$ consider the class

$$\mathcal{C}_n := \{A_n \times [0, t] \times B : A_n \in \mathcal{F}_{T_n}, t \in \mathbf{R}_+, B \in \mathcal{E}\}.$$

$\mathcal{C}_n$ is a π-system which generates $\mathcal{F}_{T_n} \otimes \mathcal{B}(\mathbf{R}_+) \otimes \mathcal{E}$ (observe that $\mathcal{F}_{T_0} \equiv \mathcal{F}_0$). But for any member $A_n \times [0,t] \times B \in \mathcal{C}_n$ (where $n \in \mathbf{N}$),

$$\mathbf{1}_{A_n \times [0,t] \times B} \cdot \mathbf{1}_{]T_n, T_{n+1}] \times E} = \mathbf{1}_{]T_{n\,A_n}, T_{n+1} \wedge t] \times B}$$

is predictable (note that for any stopping time T and set $A \in \mathcal{F}_T$, $T_A = T$ on A and $T_A = \infty$ on A^c) . Also, for any $A_\infty \times [0, t_\infty] \times B \in \mathcal{C}_\infty$,

$$\mathbf{1}_{A_\infty \times [0,t] \times B} \cdot \mathbf{1}_{[T_\infty, \infty[\times E} = \mathbf{1}_{[T_{\infty\,A_\infty}, t] \times B}.$$

But $A_\infty \in \mathcal{F}_{T_\infty} = \bigvee_{n \in \mathbf{N}} \mathcal{F}_{T_n}$ and $\bigvee_{n \in \mathbf{N}} \mathcal{F}_{T_n} = \mathcal{F}_{T_\infty^-}$, by III.T35 of [12] in conjunction with our current hypothesis that for any finite n, $T_n < T_\infty$ unless $T_n = \infty$. Thus, $T_{\infty A_\infty}$ is a predictable stopping time, which entails the predictability of $\mathbf{1}_{[T_{\infty A_\infty}, t] \times B}$. Sufficiency is now proven.

For necessity, suppose X is a predictable process. It follows from a monotone class argument applied to the generators of $\mathcal{P}$ (the sets (i) and (ii) of Definition 3.2.2 constitute a π-system) that if a process Y is predictable, then Y is adapted, i.e. for any $t \in \mathbf{R}_+$, $Y_t \equiv (Y_{(t,x)})_{x \in E}$ is $\mathcal{F}_t \otimes \mathcal{E}$-measurable. Thus, $X_0 = (X_{(0,x)})_{x \in E}$ is $\mathcal{F}_0 \otimes \mathcal{E}$-measurable. As for the existence of the desired X^n - which is trivially obtained if X's support is limited to $\{0\} \times E$ - one may restrict X, via a monotone class argument on the generators of $\mathcal{P}$, to a process of the form

$$X \equiv \mathbf{1}_{A \times (u, \infty) \times B},$$

where $u \in \mathbf{R}_+$, $A \in \mathcal{F}_u$ and $B \in \mathcal{E}$ (note that $A \times (u, \infty) \times B =]u_A, \infty[\times B$). If $n = \infty$, we may take $X^\infty := X$ because X is $\mathcal{F}_\infty \otimes \mathcal{B}(\mathbf{R}_+) \otimes \mathcal{E}$-measurable and

$$\mathcal{F}_\infty = \bigvee_{t \in \mathbf{R}_+} \mathcal{F}_t = \mathcal{F}_0 \vee \bigvee_{t \in \mathbf{R}_+} \mathcal{G}_t = \mathcal{F}_0 \vee \bigvee_{n \in \mathbf{N}} \sigma(T_1, X_1, \ldots, T_n, X_n) = \bigvee_{n \in \mathbf{N}} \mathcal{F}_{T_n} = \mathcal{F}_{T_\infty}.$$

If n is finite, then for any $q \in \{0, 1, \ldots, n-1\}$ let $A_q \in \mathcal{F}_{T_q}$ be such that

$$A \cap \{u < T_{q+1}\} = A_q \cap \{u < T_{q+1}\}$$

(such an A_q exists by Proposition 3.2.1). We claim that

$$X^n := \sum_{q<n} \mathbf{1}_{[A_q \cap \{T_q \le u < T_{q+1}\}] \times (u,\infty) \times B} + \mathbf{1}_{[A_n \cap \{T_n \le u\}] \times (u,\infty) \times B}$$

satisfies the theorem's statement for all $n \in \mathbf{N}_0$.

Clearly, X^n as defined is $\mathcal{F}_{T_n} \otimes \mathcal{B}(\mathbf{R}_+) \otimes \mathcal{E}$-measurable, so there remains to check that

$$X \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E} = X^n \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}$$

for all $n \in \mathbf{N}_0$. Fix $n \in \mathbf{N}_0$, $\omega \in \Omega$, $t \in \mathbf{R}_+$ and $x \in E$. If

$$X \cdot \mathbf{1}_{]T_n,T_{n+1}]}(\omega, t, x) = \mathbf{1}_{\{[A\times(u,\infty)]\cap]T_n,T_{n+1}]\}\times B}(\omega, t, x) = 1,$$

then $\omega \in A$, $u \vee T_n(\omega) < t \leq T_{n+1}(\omega)$ and $x \in B$. Observe that $u < T_{n+1}(\omega)$. Let $q^* := \inf\{q \in \mathbf{N}_0 : u < T_{q+1}\}$. We have $\omega \in A \cap \{T_{q^*} \leq u < T_{q^*+1}\} = A_{q^*} \cap \{T_{q^*} \leq u < T_{q^*+1}\}$. If $q^* = n$, then $\omega \in A_n \cap \{T_n \leq u\}$, $t > u$, $x \in B$ and $T_n(\omega) < t \leq T_{n+1}(\omega)$, which implies

$$X^n \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E} = \mathbf{1}_{[A_n\cap\{T_n\leq u\}]\times(u,\infty)\times B}(\omega, t, x) \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}(\omega, t, x) = 1.$$

If $q^* < n$, then $\omega \in A_{q^*} \cap \{T_{q^*} \leq u < T_{q^*+1}\}$, $u < t$, $x \in B$ and $T_n(\omega) < t \leq T_{n+1}(\omega)$, which implies

$$X^n \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E} = \mathbf{1}_{\left[A_{q^*}\cap\{T_{q^*}\leq u<T_{q^*+1}\}\right]\times(u,\infty)\times B}(\omega, t, x) \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}(\omega, t, x) = 1.$$

Therefore,

$$X \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E} = 1 \Longrightarrow X^n \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E} = 1.$$

If $X \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}(\omega, t, x) = 0$, then either $x \notin B$, $(\omega, t) \notin]T_n, T_{n+1}]$ or $(\omega, t) \notin A \times (u, \infty)$. If $x \notin B$ or $(\omega, t) \notin]T_n, T_{n+1}]$, then $X^n \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}(\omega, t, x) = 0$. If $(\omega, t) \notin A\times(u, \infty)$, then either $\omega \notin A$ or $t \leq u$. If $t \leq u$ then $X^n\cdot\mathbf{1}_{]T_n,T_{n+1}]\times E}(\omega, t, x) = 0$. Having eliminated all other possibilities, there only remains the case when $\omega \notin A$ but $u \vee T_n(\omega) < t \leq T_{n+1}(\omega)$. In this case, there exists $q^* \leq n$ such that $T_{q^*}(\omega) \leq u < T_{q^*+1}(\omega)$, but since $A \cap \{T_{q^*} \leq u < T_{q^*+1}\} = A_q \cap \{T_{q^*} \leq u < T_{q^*+1}\}$, $\omega \notin A \Longrightarrow \omega \notin A_q \Longrightarrow X^n \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}(\omega, t, x) = 0$. We thus have $X \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}(\omega, t, x) = X^n \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}(\omega, t, x)$. □

3.3 The Dual Predictable Projection

A technical lemma will precede the introduction of a dual predictable projection ν for μ. The lemma may appear somewhat trivial, but it nevertheless justifies one crucial line in the argument which proves the theorem that follows it.

Lemma 3.3.1 *Let $(E_1, \mathcal{E}_1)$, $(E_2, \mathcal{E}_2)$ be measurable spaces, E_2 a Lusin space. If $X_1 : (\Omega, \mathcal{F}, P) \longrightarrow (E_1, \mathcal{E}_1)$ and $X_2 : (\Omega, \mathcal{F}, P) \longrightarrow (E_2, \mathcal{E}_2)$ are random elements and*

$$f : E_1 \times E_2 \longrightarrow (\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+))$$

is a Borel function, then

$$\mathbf{E}[f(X_1, X_2)] = \mathbf{E}\left[\int_{E_2} f(X_1, x_2) dP[X_2 = x_2 | X_1]\right].$$

Proof The regular conditional probability $P[X_2 \in \cdot | X_1]$ exists because, E_2 being Lusin, it is isomorphic to a Borel subset of $\mathbf{R}$ (see [13], pp. 48-49 and Theorem 6.6.5 of [1]). By a monotone class argument it suffices to consider f of the form

$$\begin{array}{rccc} f: & E_1 \times E_2 & \longrightarrow & \bar{\mathbf{R}}_+ \\ & (x, y) & \longmapsto & \mathbf{1}_{A_1 \times A_2}(x, y) \end{array},$$

where $A_1 \in \mathcal{E}_1$ and $A_2 \in \mathcal{E}_2$. We then have

$$\begin{aligned} \mathbf{E}\left[f(X_1, X_2)\right] &= \mathbf{E}\left[\mathbf{1}_{\{X_1 \in A_1\}} \mathbf{1}_{\{X_2 \in A_2\}}\right] \\ &= \mathbf{E}\left[\mathbf{1}_{\{X_1 \in A_1\}} \mathbf{E}\left[\mathbf{1}_{\{X_2 \in A_2\}} | X_1\right]\right] \\ &= \mathbf{E}\left[\mathbf{1}_{\{X_1 \in A_1\}} P\left[X_2 \in A_2 | X_1\right]\right] \\ &= \mathbf{E}\left[\mathbf{1}_{\{X_1 \in A_1\}} \int_{E_2} \mathbf{1}_{A_2}(x_2) dP\left[X_2 = x_2 | X_1\right]\right] \text{ (definition of the integral)} \\ &= \mathbf{E}\left[\int_{E_2} \mathbf{1}_{A_1}(X_1) \mathbf{1}_{A_2}(x_2) dP\left[X_2 = x_2 | X_1\right]\right] \\ &= \mathbf{E}\left[\int_{E_2} f(X_1, x_2) dP[X_2 = x_2 | X_1]\right] \quad \square \end{aligned}$$

The sought regenerative form of the dual predictable projection ν of μ is now in sight. As was previously mentioned, the proof of the following result is considerably longer than that of Jacod's Proposition 3.41 ([24]). It is not trivial - at least to the present author's mind - that for any $n \in \mathbf{N}_0$, any "random set" $A \times B \in \mathcal{F} \otimes \mathcal{B}(\mathbf{R}_+) \otimes \mathcal{E}$ with sections $A(\omega) \times B$, $\omega \in \Omega$ (where $B \in \mathcal{E}$), and any conditional probability ν_n of (T_{n+1}, X_{n+1}) with respect to $\mathcal{F}_{T_n}$, the equality

$$\mathbf{E}[\nu_n(A(\omega) \times B)] = P[T_{n+1}(\omega) \in A(\omega), X_{n+1} \in B]$$

would necessarily hold almost surely; in fact, it may not even hold in general. That it holds when $A =]T_n, \infty[$ is implicit in Jacod's argument ([24]), but we would like to highlight the fact that it is the specific structure of the set $]T_n, \infty[$ - i.e. its open sections, and thus the possibility of approximating its ν_n-measure "from below" using *deterministic* intervals - which warrants this property. The following theorem is the promised generalization of Jacod's regenerative form for the compensator of a simple process ([24]: Proposition 3.41), which now includes the compensator of a non-simple point process.

Theorem 3.3.2 *For any $n \in \mathbf{N}_0$, $i \in \mathbf{N}$ let*

$$\nu_n^{n+i}(\cdot) := P\left[(T_{n+i}, X_{n+i}) \in \cdot\,; T_{n+i} = T_{n+1} | \mathcal{F}_{T_n}\right]$$

denote a regular version of the conditional probability. The random measure ν on $(0, \infty) \times E$ given by

$$\nu(dt, dx) := \sum_{n \in \mathbf{N}_0} \left(\sum_{i \in \mathbf{N}} \frac{\nu_n^{n+i}(dt, dx)}{\nu_n^{n+1}([t, \infty] \times \bar{E})} \right) \cdot \mathbf{1}_{\{T_n < t \leq T_{n+1}\}}$$

constitutes a version of the dual predictable projection (i.e. compensator) of μ; that is:

i) *for any $B \in \mathcal{E}$, the map $(\omega, t, x) \mapsto \nu(\omega, (0, t] \times B)$ is $\mathcal{P}$-measurable, and*

ii) *for any $\bar{\mathbf{R}}_+$-valued, $\mathcal{P}$-measurable process X,*

$$\mathbf{E}\left[\int_{(0,\infty)\times E} X(t, x)\mu(dt, dx)\right] = \mathbf{E}\left[\int_{(0,\infty)\times E} X(t, x)\nu(dt, dx)\right].$$

Proof The regular conditional probabilities ν_n^{n+i} exist because, as a consequence of Tychonoff's theorem, the topological product of finitely many Lusin spaces is Lusin, so that $(0,\infty] \times \bar{E} \times [0,1]$ is Lusin. To prove (i), observe that for any $B \in \mathcal{E}$, $n \in \mathbf{N}_0$ and $(t,x) \in \mathbf{R}_+ \times E$,

$$\nu\left((0,t] \times B\right) \cdot \mathbf{1}_{]T_n,T_{n+1}] \times E} \\ = \left[\left(\sum_{p=0}^{n-1}\sum_{i=1}^{\infty}\int_{T_p}^{T_{p+1}} \frac{\nu_p^{p+i}(du,B)}{\nu_p^{p+1}\left([u,\infty] \times \bar{E}\right)}\right) + \sum_{j=1}^{\infty}\int_{T_n}^{t} \frac{\nu_n^{n+j}(du,B)}{\nu_n^{n+1}\left([u,\infty] \times \bar{E}\right)}\right] \cdot \mathbf{1}_{]T_n,T_{n+1}] \times E}.$$

Since the map

$$(\omega,t,x) \longmapsto \left(\sum_{p=0}^{n-1}\sum_{i=1}^{\infty}\int_{T_p}^{T_{p+1}} \frac{\nu_p^{p+i}(du,B)}{\nu_p^{p+1}\left([u,\infty] \times \bar{E}\right)}\right) + \left(\sum_{j=1}^{\infty}\int_{T_n}^{t} \frac{\nu_n^{n+j}(du,B)}{\nu_n^{n+1}\left([u,\infty] \times \bar{E}\right)}\right)$$

is $\mathcal{F}_{T_n} \otimes \mathcal{B}(\mathbf{R}_+) \otimes \mathcal{E}$-measurable and ν charges neither $[0] \times E$ nor $[T_\infty,\infty[\times E$, the map $(\omega,t,x) \mapsto \nu\left(\omega,(0,t] \times B\right)$ must be predictable by Proposition 3.2.3. To prove (ii), observe that sets of the form $A \times (u,\infty) \times B$, where $u \in \mathbf{R}_+$, $A \in \mathcal{F}_u$ and $B \in \mathcal{E}$, generate the class $\mathcal{P} \cap (]0,\infty[\times E)$ and constitute a π-system. Again using the fact that ν as written charges neither $[0] \times E$ nor $[T_\infty,\infty[\times E$, (ii) will be proven once we ascertain that

$$\mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{A\times(u,\infty)\times B}\, d\mu\right] = \mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{A\times(u,\infty)\times B}\, d\nu\right]$$

for all such sets. But this equality will itself hold once we show that for any $n \in \mathbf{N}_0$,

$$\mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{A\times(u,\infty)\times B} \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}\, d\mu\right] = \\ \mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{A\times(u,\infty)\times B} \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}\, d\nu\right].$$

Choose $n \in \mathbf{N}_0$, $u \in \mathbf{R}_+$, $A \in \mathcal{F}_u$ and $B \in \mathcal{E}$. The left-hand-side of the preceding equality is decomposed as

$$\mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{A\times(u,\infty)\times B} \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}\, d\mu\right] \\ = \sum_{i=1}^{\infty} \mathbf{E}\left[\mathbf{1}_A(\omega) \cdot \mathbf{1}_{\{T_{n+i}=T_{n+1}>T_n\}} \cdot \mathbf{1}_{\{u<T_{n+i}(\omega)<\infty\}} \cdot \mathbf{1}_B(X_{n+i}(\omega))\right].$$

As for the right-hand-side, observe that, by virtue of Proposition 3.2.3, there exists an $\mathcal{F}_{T_n} \otimes \mathcal{B}(\mathbf{R}_+) \otimes \mathcal{E}$-measurable process X^n such that

$$\mathbf{1}_{A\times(u,\infty)\times B} \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E} = X^n \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}, \tag{5}$$

and X^n may even be assigned, using the proof of Proposition 3.2.1, the explicit formulation

$$X^n(\omega,t,x) \equiv \sum_{q<n} \mathbf{1}_{[A_q\cap\{T_q\leq u<T_{q+1}\}]\times(u,\infty)\times B}(\omega,t,x) + \mathbf{1}_{[A_n\cap\{T_n\leq u\}]\times(u,\infty)\times B}(\omega,t,x),$$

where, for all $q \leq n$, $A_q \in \mathcal{F}_{T_q}$ is such that

$$A \cap \{u < T_{q+1}\} = A_q \cap \{u < T_{q+1}\}.$$

We therefore have, by (5) and the definition of ν in the statement of the theorem, that

$$\begin{aligned}
&\mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{A\times(u,\infty)\times B} \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E} d\nu\right] \\
&= \mathbf{E}\left[\int_{(0,\infty)\times E} X^n \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E} d\nu\right] \\
&= \mathbf{E}\left[\int_E \int_{T_n}^{T_{n+1}} X^n \cdot \sum_{i=1}^{\infty} \frac{\nu_n^{n+i}(dt,dx)}{\nu_n^{n+1}([t,\infty]\times\bar{E})}\right] \\
&= \sum_{i=1}^{\infty} \mathbf{E}\left[\int_{(0,\infty]} \int_{(0,\infty)\times E} \mathbf{1}_{]T_n,s]} \cdot X^n \right. \\
&\qquad \left. \cdot \frac{\nu_n^{n+i}(dt,dx)}{\nu_n^{n+1}([t,\infty]\times\bar{E})} \cdot \nu_n^{n+1}(ds\times\bar{E})\right] \qquad \text{(by Lemma 3.3.1)} \\
&= \sum_{i=1}^{\infty} \mathbf{E}\left[\int_{(0,\infty]} \int_{(0,\infty]\times E} \mathbf{1}_{\{T_n<t<\infty\}} \cdot \mathbf{1}_{\{t\leq s\}} \right. \\
&\qquad \left. \cdot \frac{X^n(t,x)}{\nu_n^{n+1}([t,\infty]\times\bar{E})} \nu_n^{n+i}(dt,dx)\nu_n^{n+1}(ds\times\bar{E})\right] \\
&= \sum_{i=1}^{\infty} \mathbf{E}\left[\int_{(0,\infty]\times E} \mathbf{1}_{\{T_n<t<\infty\}} \frac{X^n(t,x)}{\nu_n^{n+1}([t,\infty]\times\bar{E})} \right. \\
&\qquad \left. \cdot \left(\int_{(0,\infty]} \mathbf{1}_{\{t\leq s\}} \nu_n^{n+1}(ds\times\bar{E})\right) \nu_n^{n+i}(dt,dx)\right] \qquad \text{(Fubini)}
\end{aligned}$$

$$= \sum_{i=1}^{\infty} \mathbf{E}\left[\int_{(0,\infty]\times E} \mathbf{1}_{\{T_n<t<\infty\}} \frac{X^n(t,x)}{\nu_n^{n+1}([t,\infty]\times\bar{E})} \cdot \nu_n^{n+1}([t,\infty]\times\bar{E})\nu_n^{n+i}(dt,dx)\right]$$
$$= \sum_{i=1}^{\infty} \mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{]T_n,\infty[\times E} X^n(t,x)\nu_n^{n+i}(dt,dx)\right].$$

Fix $i \in \mathbf{N}$ to consider only one of the summands. We have

$$\mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{]T_n,\infty[\times E} X^n(t,x)\nu_n^{n+i}(dt,dx)\right]$$
$$= \sum_{q<n} \mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{]T_n,\infty[\times E}\cdot \mathbf{1}_{[A_q\cap\{T_q\le u<T_{q+1}\}]\times(u,\infty)\times B}\nu_n^{n+i}(dt,dx)\right]$$
$$+ \mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{]T_n,\infty[\times E}\cdot \mathbf{1}_{[A_n\cap\{T_n\le u\}]\times(u,\infty)\times B}\nu_n^{n+i}(dt,dx)\right].$$

For any $q < n$, we approximate the sections of $]T_n,\infty[$ "from below" and invoke monotone convergence to obtain

$$\mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{]T_n,\infty[\times E}\cdot \mathbf{1}_{[A_q\cap\{T_q\le u<T_{q+1}\}]\times(u,\infty)\times B}\nu_n^{n+i}(dt,dx)\right]$$
$$= \mathbf{E}\left[\int_{(0,\infty)\times E} \sup_{m\in\mathbf{N}} \sum_{p\in\mathbf{N}_0}\sum_{j=1}^{m}\left(\mathbf{1}_{\{p+\frac{j-1}{m}\le T_n<p+\frac{j}{m}\}\times(p+\frac{j}{m},\infty)\times E}\right.\right.$$
$$\left.\left.\cdot\mathbf{1}_{[A_q\cap\{T_q\le u<T_{q+1}\}]\times(u,\infty)\times B}\right)\cdot\nu_n^{n+i}(dt,dx)\right]$$
$$= \sup_{m\in\mathbf{N}} \sum_{p\in\mathbf{N}_0}\sum_{j=1}^{m} \mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{[\{p+\frac{j-1}{m}\le T_n<p+\frac{j}{m}\}\cap A_q\cap\{T_q\le u<T_{q+1}\}]\times(u\vee(p+\frac{j}{m}),\infty)\times B}\right.$$
$$\left.\cdot\nu_n^{n+i}(dt,dx)\right]$$
$$= \sup_{m\in\mathbf{N}} \sum_{p\in\mathbf{N}_0}\sum_{j=1}^{m} \mathbf{E}\left[\mathbf{1}_{\{p+\frac{j-1}{m}\le T_n<p+\frac{j}{m}\}\cap A_q\cap\{T_q\le u<T_{q+1}\}}\cdot\int_{(u\vee(p+\frac{j}{m}),\infty)\times B}\nu_n^{n+i}(dt,dx)\right]$$
$$= \sup_{m\in\mathbf{N}} \sum_{p\in\mathbf{N}_0}\sum_{j=1}^{m} \mathbf{E}\left[\mathbf{1}_{\{p+\frac{j-1}{m}\le T_n<p+\frac{j}{m}\}\cap A_q\cap\{T_q\le u<T_{q+1}\}}\right.$$
$$\left.\cdot P\left[u\vee\left(p+\frac{j}{m}\right)<T_{n+i}=T_{n+1}<\infty, X_{n+i}\in B|\mathcal{F}_{T_n}\right]\right]$$

$$= \sup_{m\in\mathbf{N}} \sum_{p\in\mathbf{N}_0} \sum_{j=1}^{m} P\left[\left\{p+\frac{j-1}{m} \le T_n < p+\frac{j}{m}\right\} \cap A_q \cap \{T_q \le u < T_{q+1}\}\right.$$
$$\left. \cap \left\{u \vee \left(p+\frac{j}{m}\right) < T_{n+i} = T_{n+1} < \infty\right\} \cap \{X_{n+i} \in B\}\right]$$
$$= P\left[A_q \cap \{T_q \le u < T_{q+1}\} \cap \{T_n < T_{n+i} = T_{n+1} < \infty\} \cap \{X_{n+i} \in B\}\right]$$
$$= P\left[A \cap \{T_q \le u < T_{q+1}\} \cap \{T_n < T_{n+i} = T_{n+1} < \infty\} \cap \{X_{n+i} \in B\}\right],$$

and moreover,

$$\mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{]T_n,\infty[\times E} \cdot \mathbf{1}_{[A_n\cap\{T_n\le u\}]\times(u,\infty)\times B}\nu_n^{n+i}(dt,dx)\right]$$
$$= \mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{[A_n\cap\{T_n\le u\}]\times(u,\infty)\times B}\nu_n^{n+i}(dt,dx)\right]$$
$$= \mathbf{E}\left[\mathbf{1}_{A_n\cap\{T_n\le u\}} \cdot \int_{(u,\infty)\times B} \nu_n^{n+i}(dt,dx)\right]$$
$$= \mathbf{E}\left[\mathbf{1}_{A_n\cap\{T_n\le u\}} \cdot P\left[u < T_{n+i} = T_{n+1} < \infty; X_{n+i} \in B|\mathcal{F}_{T_n}\right]\right]$$
$$= P\left[A_n \cap \{T_n \le u\} \cap \{u < T_{n+i} = T_{n+1} < \infty; X_{n+i} \in B\}\right]$$
$$= P\left[A_n \cap \{T_n \le u < T_{n+1}\} \cap \{u \vee T_n < T_{n+i} = T_{n+1} < \infty\} \cap \{X_{n+i} \in B\}\right]$$
$$= P\left[A \cap \{T_n \le u\} \cap \{u \vee T_n < T_{n+i} = T_{n+1} < \infty\} \cap \{X_{n+i} \in B\}\right].$$

Summing these expressions over all $i \in \mathbf{N}$, we obtain:

$$\mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{A\times(u,\infty)\times B} \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E}d\nu\right]$$
$$= \sum_{i=1}^{\infty}\left(\sum_{q<n} P[A \cap \{T_q \le u < T_{q+1}\} \cap \{T_n < T_{n+i} = T_{n+1} < \infty\} \cap \{X_{n+i} \in B\}]\right.$$
$$\left. + P[A \cap \{T_n \le u\} \cap \{u \vee T_n < T_{n+i} = T_{n+1} < \infty\} \cap \{X_{n+i} \in B\}]\right)$$
$$= \sum_{i=1}^{\infty}\sum_{q<n} P[A \cap \{T_q \le u < T_{q+1}\} \cap \{u \vee T_n < T_{n+i} = T_{n+1} < \infty\} \cap \{X_{n+i} \in B\}]$$
$$+ P[A \cap \{T_n \le u\} \cap \{u \vee T_n < T_{n+i} = T_{n+1} < \infty\} \cap \{X_{n+i} \in B\}]$$
$$= \sum_{i=1}^{\infty} P[A \cap \{u \vee T_n < T_{n+i} = T_{n+1} < \infty\} \cap \{X_{n+i} \in B\}]$$
$$= \sum_{i=1}^{\infty} \mathbf{E}[\mathbf{1}_A \cdot \mathbf{1}_{\{T_{n+i}=T_{n+1}>T_n\}} \cdot \mathbf{1}_{\{u<T_{n+i}<\infty\}} \cdot \mathbf{1}_B(X_{n+i})]$$

$$= \mathbf{E}\left[\int_{(0,\infty)\times E} \mathbf{1}_{A\times(u,\infty)\times B} \cdot \mathbf{1}_{]T_n,T_{n+1}]\times E} d\mu\right]. \ \square$$

Part II

Point Process Representations, Couplings and Association

Chapter 4

POP-spaces, Stochastic Order and Association

As its title suggests, this chapter is an exposition on essential notions of partially ordered Polish spaces, stochastic orderings of pairs of probability measures, and association of single probability measures on such spaces.

Nachbin's classical study of the interplay between topology and partial orders on abstract spaces [44], as well as a result attributed to Strassen ([48]: Theorem 11) on the existence of probability measures on product spaces with stochastically ordered marginals, led Kamae, Krengel and O'Brien [29], among others, to investigate the *stochastic orderings* of probability measures on partially ordered Polish spaces - in particular, on countable products of partially ordered Polish spaces - and to identify conditions under which such orderings do hold.

Association emerged as a dependence property of collections of random variables. First introduced by Esary, Proschan and Walkup [17], this property found immediate applications in reliability theory [16]. As central limit theorems were later obtained for associated random variables (see, for example, [45], [10] and [11]), probabilists sought to generalize the concept of association to abstract partially ordered Polish spaces. Lindqvist [39] adapted techniques used by Kamae et al. [29] in their study of stochastic orderings, to derive important results on the association of probability measures on partially ordered Polish spaces, and on countable products thereof. As a result, the frameworks of the two theories (stochastic orderings and association) have now much in common, as we shall show in this chapter.

Not surprisingly, the two sources from which we will most often quote are [29] and [39] - the former as regards stochastic orderings, the latter in relation to association. The bulk of the quoted results are taken from these sources. Results involving representation maps (**Theorems 4.2.7**, **4.3.9** and **4.3.11**) are, to our knowledge, new. These maps reduce the question of orderings on a generic space to a corresponding question on $\bar{\mathbf{R}}_+^\infty$.

We first define partially ordered Polish spaces and related concepts in Section 4.1, while reserving the treatment of stochastic orderings and association for Sections 4.2 and 4.3 respectively. A deliberate attempt will be made to highlight the striking similarities existing between criteria for stochastic orderings of pairs of probability measures on one hand, and criteria for the association of a single probability measure on the other.

In Chapter 5, we will apply some of these results to point processes on the half-line, which will then be viewed as random elements of the partially ordered Polish space $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$ if locally integer-valued, nonexplosive measures on $\mathbf{R}_+$.

4.1 Partially Ordered Polish Spaces

A *Polish* space is a complete separable metric space. A Polish space S endowed with the Borel σ-field $\mathcal{S}$ generated by the metric of S is noted $(S, \mathcal{S})$. Given a Polish space $(S, \mathcal{S})$, a *closed partial order* $\prec$ on $(S, \mathcal{S})$ is a binary relation satisfying the following properties:

1. reflexivity: $x \prec x \;\; \forall\, x \in S$;

2. transitivity: $x \prec y$ and $y \prec z$ together imply $x \prec z \;\; \forall\, x, y, z \in S$;
3. antisymmetry: $x \prec y$ and $y \prec x$ together imply $x = y \;\; \forall\, x, y, z \in S$;
4. closure: if $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ are two sequences in S converging to x and y respectively and, moreover, $x_n \prec y_n \;\; \forall\, n \in \mathbf{N}$, then $x \prec y$.

The foregoing equips us to define partially ordered Polish spaces (POP-spaces) and the related concept of increasing (resp. decreasing) sets:

Definition 4.1.1 *Let $(S, \mathcal{S})$ be a Polish space and let $\prec$ be a closed partial order on S.*

1. *The triplet $(S, \mathcal{S}, \prec)$ is called a partially ordered Polish space (POP-space for short).*
2. *A set $A \subseteq S$ is said to be $\prec$-increasing (or simply increasing when the order is unambiguous) if, for any $x \in A$, $y \in S$, $x \prec y \implies y \in A$. A is said to be ($\prec$-)decreasing if it is the complement of a ($\prec$-)increasing set. If $A \subseteq S$ is arbitrary, we note*

   $$\mathrm{Inc}(A) := \{y \in S : \exists\, x \in A \ni x \prec y\}$$

   and

   $$\mathrm{Dec}(A) := \{y \in S : \exists\, x \in A \ni y \prec x.\}$$

   Obviously, Inc(A) *is increasing by the transitivity of $\prec$, and* Dec(A) *is a decreasing set because its complement is increasing. If $A = \{x\}$ is a singleton, we write C_x for* Inc(A) *and D_x for* Dec(A).
3. *An increasing set $A \subseteq S$ is said to be compact generated if there exists a compact set $K \subseteq S$ such that $A =$ Inc(K); a decreasing set B is said to be compact generated if there exists a compact set $\widetilde{K} \subseteq S$ such that $B =$ Dec$(\widetilde{K})$.*

Remark 4.1.2 *If $A \subseteq S$ is compact generated (as an increasing or decreasing set), the sequential criterion for compactness may be invoked along with the closure of the partial order $\prec$ to show that A is closed, and thus measurable. This fact will be used later.*

As POP-spaces have been defined, we may now introduce a special variety of POP-spaces, namely, *product* POP-spaces:

Definition 4.1.3 *Let $(S_1, \mathcal{S}_1, \prec_1)$, $(S_2, \mathcal{S}_2, \prec_2)$,... be POP-spaces endowed with metrics $d_1, d_2, \ldots$ respectively. Let $n \in \mathbf{N} \cup \{\infty\}$. We define the product partially ordered Polish space $(S^{(n)}, \mathcal{S}^{(n)}, \prec^n)$ as follows:*

1. *$S^{(n)} := \bigotimes_{i=1}^n S_i$ is endowed with the product metric d^n defined by*

$$d^n(x, y) := \sum_{i=1}^{n} \frac{d_i(x_i, y_i)}{2^i[1 + d_i(x_i, y_i)]}$$

   for any $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ in $S^{(n)}$;

2. *$\mathcal{S}^{(n)} = \bigotimes_{i=1}^n \mathcal{S}_i$ is the product Borel σ-field generated by the product metric d^n;*

3. *For any $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n) \in S^{(n)}$, $\prec^n$ satisfies:*

$$(x_1, \ldots, x_n) \prec^n (y_1, \ldots, y_n)$$

   if and only if $x_i \prec_i y_i$ for all $i = 1, \ldots, n$. Observe that $\prec^n$ is a closed partial order on account of the fact that the $\prec_i$'s are closed partial orders, and that convergence in $(S^{(n)}, d^n)$ corresponds to coordinate convergence.

A technical, but desirable property for a POP-space to have is that of *normal orderliness*. This property will be invoked, in particular, to ensure that association is preserved under weak convergence of probability measures (see Theorem 4.3.3).

Let $(S, \mathcal{S}, \prec)$ be a POP-space endowed with a metric d. For any pair A, $B \subseteq S$ of nonempty sets let

$$d(A, B) := \inf\{d(u, v) : u \in A, b \in B\}.$$

Definition 4.1.4 *The POP-space $(S, \mathcal{S}, \prec)$ is said to be normally ordered if there exists a metric d on $(S, \mathcal{S})$ such that*

(N1) *$d(D_x, C_y) = 0$ implies $y \prec x$ for all $x, y \in S$;*

(N2) *$d(D_x, C_z) \leq d(D_x, C_y) + d(D_y, C_z)$ for all $x, y, z \in S$.*

"Normal orderliness" has been defined a great many different ways in the literature (see, for example, [44], [49] and [39]). Our definition is stronger that Lindqvist's ([39]), but conditions (N1) and (N2) have been used by Lindqvist ([39]: Theorem 5.2) to establish the following important result - which he phrases differently, in accordance with his own definition:

Theorem 4.1.5 *If $(S_1, \mathcal{S}_1, \prec_1)$, $(S_2, \mathcal{S}_2, \prec_2)$,... are normally ordered POP-spaces, then the product POP-space $(S^{(\infty)}, \mathcal{S}^{(\infty)}, \prec^{\infty})$ is normally ordered.*

At last, we consider two normally ordered POP-spaces, one of which we shall use frequently. Let $\bar{\mathbf{R}}_+ := [0, \infty]$ be endowed with the metric d defined by

$$d(x, y) := \mid e^{-x} - e^{-y} \mid$$

for any $x, y \in \bar{\mathbf{R}}_+$, where $e^{-\infty} \equiv 0$ by convention. The Borel sets $\mathcal{B}(\bar{\mathbf{R}}_+)$ of $\bar{\mathbf{R}}_+$ correspond to the "usual" Borel sets on $[0, \infty)$ and admit the singleton $\{\infty\}$ as measurable. It is clear that $(\bar{\mathbf{R}}_+, d)$ is a Polish space, and that the usual partial order $\leq$ is closed on $(\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+))$. It is also clear, by inspection, that conditions (N1) and (N2) are met by the POP-space $(\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+), \leq)$, so that $(\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+), \leq)$ is normally ordered. By Theorem 4.1.5, the POP-space $(\bar{\mathbf{R}}_+^{\infty}, \mathcal{B}(\bar{\mathbf{R}}_+^{\infty}), \leq^{\infty})$ is also normally ordered. We shall later use $(\bar{\mathbf{R}}_+^{\infty}, \mathcal{B}(\bar{\mathbf{R}}_+^{\infty}), \leq^{\infty})$ to define representation maps for closed partial orders on abstract Polish spaces (see Definition 4.3.8). For the sake of legibility and future reference, let us summarize these points as follows:

Remark 4.1.6 *The POP-spaces $(\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+), \leq)$ and $(\bar{\mathbf{R}}_+^{\infty}, \mathcal{B}(\bar{\mathbf{R}}_+^{\infty}), \leq^{\infty})$ are normally ordered.*

4.2 Stochastic Orderings of Probability Measures on POP-spaces

This section summarizes and adapts results extracted from [29]. Our presentation is designed to highlight the close analogies between criteria that ensure a coupling relation on the one hand, and criteria which secure association on the other. The current section deals with the stochastic ordering (i.e. coupling) of *pairs* of probability measures, whereas its successor will address association of *single* probability measures.

Let $(S, \mathcal{S}, \prec)$ be an arbitrary POP-space.

Definition 4.2.1 *Let P_1 and P_2 be probability measures on $(S, \mathcal{S}, \prec)$.*

1. *We say that P_1 is stochastically smaller than P_2 (with respect to $\prec$ if the order is ambiguous) if, for any $\prec$-increasing set $A \in \mathcal{S}$,*

$$P_1(A) \leq P_2(A).$$

We denote this relationship $P_1 \prec P_2$, and say that P_1 and P_2 satisfy a coupling relation, *or a* coupling *for short.*

2. *If $X_1, X_2 : (\Omega, \mathcal{F}, P) \longrightarrow (S, \mathcal{S}, \prec)$ are two random elements of S such that $P \circ X_1^{-1} \prec P \circ X_2^{-1}$, we say that X_1 is stochastically smaller than X_2, and denote this relationship*

$$X_1 \prec_{\text{st}} X_2.$$

As random elements cannot be confused with probability measures, we also say, in this case, that X_1 and X_2 satisfy a coupling (relation).

It turns out that not all $\prec$-increasing sets A need be substituted in the relation $P_1(A) \leq P_2(A)$ for a coupling to be proven. Indeed,

Theorem 4.2.2 *Let P_1 and P_2 be probability measures on $(S, \mathcal{S}, \prec)$. The following statements are equivalent:*

1. *$P_1 \prec P_2$;*
2. *$P_1(A) \leq P_2(A)$ for all closed, $\prec$-increasing sets $A \subseteq S$;*
3. *$P_1(A) \leq P_2(A)$ for all $\prec$-increasing, compact generated sets $A \subseteq S$.*

Proof The equivalence of (1) and (2) is proved in [29] (Theorem 1). As (3) is a special case of (2), only the implication (3)$\Longrightarrow$(2) requires an explanation. Suppose (3) holds and let $A \subseteq S$ be closed and increasing. Since $(S, \mathcal{S})$ is Polish, for any $\epsilon > 0$ there exists a compact set $K \subseteq A$ such that $P_1(A \setminus K) < \epsilon$ ([4]: Theorems 1.1 and 1.4). But as $\text{Inc}(K) \subseteq A$ is compact generated, we obtain $P_1(A) \leq P_1(\text{Inc}(K)) + \epsilon \leq P_2(\text{Inc}(K)) + \epsilon \leq P_2(A) + \epsilon$, and since ϵ is arbitrary, (2) follows. □

The partial order $\prec$ is closed, i.e. for any sequences $\{x_n\}_{n=1}^{\infty}$, $\{y_n\}_{n=1}^{\infty}$ in S converging to x and y respectively, if $x_n \prec y_n$ for all $n \in \mathbf{N}$, then $x \prec y$. This property is mirrored by pairs of componentwise coupled weakly convergent sequences of probability measures. More precisely, according to [29] (Proposition 3),

Theorem 4.2.3 *Let $\{P_n\}_{n=1}^{\infty}$, $\{Q_n\}_{n=1}^{\infty}$ be two sequences of probability measures on $(S, \mathcal{S}, \prec)$ converging weakly to probability measures P and Q respectively. If $P_n \prec Q_n$ for all $n \in \mathbf{N}$, then $P \prec Q$.*

Now turning to product POP-spaces, let us state how the coupling of two probability measures on a product POP-space relates to the couplings of their respective marginals. According to Proposition 2 of [29], one measure is stochastically smaller than another on a product POP-space if and only if the marginals of the former are stochastically smaller than the marginals of the latter on the projection spaces. In mathematical notation, this is expressed as follows:

Theorem 4.2.4 *Let $(S^{(\infty)}, \mathcal{S}^{(\infty)}, \prec^{\infty})$ be a countable product of POP-spaces $(S_1, \mathcal{S}_1, \prec_1)$, $(S_2, \mathcal{S}_2, \prec_2)$,... For any $n \in \mathbf{N}$ let*

$$\begin{array}{rccc} \pi_n: & S^{(\infty)} & \longrightarrow & S^{(n)} \\ & (x_1, x_2, \ldots) & \longmapsto & (x_1, \ldots, x_n) \end{array}$$

denote the projection map associated to the first n coordinates. If P and Q are two probability measures on $(S^{(\infty)}, \mathcal{S}^{(\infty)})$, then $P \prec Q$ if and only if $P \circ \pi_n^{-1} \prec Q \circ \pi_n^{-1}$ on the product space $(S^{(n)}, \mathcal{S}^{(n)}, \prec^n)$ for all $n \in \mathbf{N}$.

Remark 4.2.5 *Kamae et al. [29] only proved sufficiency in Theorem 4.2.4, but necessity follows from this simple observation: for any $n \in \mathbf{N}$, if $A \in \mathcal{S}^{(n)}$ is $\prec^n$-increasing, then $A \times S_{n+1} \times S_{n+2} \times \ldots$ is both $\prec^{\infty}$-increasing and $\mathcal{S}^{(\infty)}$-measurable.*

Theorem 4.2.4 characterizes couplings of probability measures on an infinite product of POP-spaces via corresponding couplings of the finite dimensional marginals on the projection POP-spaces. Its import, however, is more theoretical than practical, because it provides no criterion for ensuring the couplings of finite dimensional marginals in the first place. One such criterion, which constitutes the so-called "discrete-time comparison theorem" in the terminology of [29], uses stochastic kernels; stochastic kernels - or transition probabilities, as they are sometimes called - are handier in practice because probability laws on product spaces tend to be more often specified via their stochastic kernels than via their joint distributions. In Theorem 4.2.4's notation, Kamae et al. ([29]: Theorem 2) state:

Theorem 4.2.6 *For any $n \in \mathbf{N}$, if*

$$P_n^{n+1}, Q_n^{n+1} : S^{(n)} \times \mathcal{S}_{n+1} \longmapsto [0,1]$$

denote the unique stochastic kernels in $S^{(n+1)} = S^n \times S_{n+1}$ such that, for any $A^{(n)} \in \mathcal{S}^{(n)}$, $A_{n+1} \in \mathcal{S}_{n+1}$,

$$\int_{A^{(n)}} P_n^{n+1}(x_1, \ldots, x_n, A_{n+1}) dP \circ \pi_n^{-1}(x_1, \ldots, x_n) = P \circ \pi_{n+1}^{-1}(A^{(n)} \times A_{n+1})$$

and

$$\int_{A^{(n)}} Q_n^{n+1}(x_1, \ldots, x_n, A_{n+1}) dQ \circ \pi_n^{-1}(x_1, \ldots, x_n) = Q \circ \pi_{n+1}^{-1}(A^{(n)} \times A_{n+1}),$$

and if

1. *$P \circ \pi_1^{-1} \prec_1 Q \circ \pi_1^{-1}$ and*

2. *for any $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n) \in S^{(n)}$ such that $(x_1, \ldots, x_n) \prec^n (y_1, \ldots, y_n)$,*

$$P_n^{n+1}(x_1, \ldots, x_n, \cdot) \prec_n Q_n^{n+1}(y_1, \ldots, y_n, \cdot)$$

   as probability measures,

then $P \prec^\infty Q$ on $(S^{(\infty)}, \mathcal{S}^{(\infty)}, \prec^\infty)$.

We conclude this section with a result that relates a coupling relation on an arbitrary POP-space to an induced relation on the space $(\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty), \leq^\infty)$. Our motivation for doing this is to complete the remarkable parallel to be drawn from criteria which ensure coupling relations on one hand and criteria which establish the property of association on the other. The instrument we invoke here - a representation map - will be formally defined in the next section.

Theorem 4.2.7 *Let $(S, \mathcal{S}, \prec)$ be a POP-space. Suppose there exists an injective, measurable map*

$$\Phi_\prec : (S, \mathcal{S}) \longrightarrow (\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty))$$

such that for any $x, y \in S$, $x \prec y$ if and only if $\Phi_\prec(x) \leq^\infty \Phi_\prec(y)$. Then, for any two probability measures P_1, P_2 on $(S, \mathcal{S})$, $P_1 \prec P_2$ holds if and only if $P_1 \circ \Phi_\prec^{-1} \leq^\infty P_2 \circ \Phi_\prec^{-1}$ holds on $(\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty))$.

Proof For sufficiency, suppose $P_1 \prec P_2$ on $(S, \mathcal{S})$, and let $A \in \mathcal{B}(\bar{\mathbf{R}}_+^\infty)$ be a $\leq^\infty$-increasing set. Then $\Phi_\prec^{-1}(A)$ is $\prec$-increasing in S. We therefore obtain $P_1 \circ \Phi_\prec^{-1}(A) = P_1(\Phi_\prec^{-1}(A)) \leq P_2(\Phi_\prec^{-1}(A)) = P_2 \circ \Phi_\prec^{-1}(A)$, and so $P_1 \circ \Phi_\prec^{-1} \leq^\infty P_2 \circ \Phi_\prec^{-1}$. For necessity, suppose $A \in \mathcal{S}$ is $\prec$-increasing. As $(S, \mathcal{S})$ and $(\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty))$ are Polish and $\Phi_\prec$ is injective, $\Phi_\prec(A) \in \mathcal{B}(\bar{\mathbf{R}}_+^\infty)$ by Kuratowski's theorem ([33]: Corollary 2 of Theorem 13.1.9). Also, for any $\epsilon > 0$, Billingsley ([4]: Theorems 1.1 and 1.4) warrants the existence of a compact set $K \subseteq \Phi_\prec(A)$ such that $P_1 \circ \Phi_\prec^{-1}(A \setminus K) < \epsilon$. Observe, then, that $\Phi_\prec^{-1}(\mathrm{Inc}(K)) \subseteq A$, as $\Phi_\prec^{-1}$ preserves orders and A is increasing. It follows that $P_1(A) \leq P_1 \circ \Phi_\prec^{-1}(\mathrm{Inc}(K)) + \epsilon \leq P_2 \circ \Phi_\prec^{-1}(\mathrm{Inc}(K)) + \epsilon \leq P_2(A) + \epsilon$. As ϵ is arbitrary, the conclusion follows. □

Remark 4.2.8 *Sufficiency does not require that $\Phi_\prec$ be injective in the proof of Theorem 4.2.7. In general, if $(S_1, \mathcal{S}_1, \prec_1)$ and $(S_2, \mathcal{S}_2, \prec_2)$ are POP-spaces and*

$$\phi : (S_1, \mathcal{S}_1, \prec_1) \longrightarrow (S_2, \mathcal{S}_2, \prec_2)$$

is measurable and increasing (meaning that $\forall\, x, y \in S_1$, $x \prec_1 y \Longrightarrow \phi(x) \prec_2 \phi(y)$), the coupling $P_1 \prec_1 P_2$ of two probability measures on $(S_1, \mathcal{S}_1, \prec_1)$ entails the coupling $P_1 \circ \phi^{-1} \prec_2 P_2 \circ \phi^{-1}$ of the corresponding probability measures induced by ϕ on $(S_2, \mathcal{S}_2)$.

Remark 4.2.9 *If, in the statement of Theorem 4.2.7, the property that* $x \prec y$ *if and only if* $\Phi_{\prec}(x) \leq^{\infty} \Phi_{\prec}(y)$ *is replaced by the property that* $x \prec y$ *if and only if* $\Phi_{\prec}(y) \leq^{\infty} \Phi_{\prec}(x)$*, then* $P_1 \prec P_2$ *holds if and only if* $P_2 \circ \Phi_{\prec}^{-1} \leq^{\infty} P_1 \circ \Phi_{\prec}^{-1}$ *holds on* $(\bar{\mathbf{R}}_+^{\infty}, \mathcal{B}(\bar{\mathbf{R}}_+^{\infty}))$*. That is, when* $\Phi_{\prec}$ *defines* $\prec$ *"in reverse", the pre-image of a* $\leq^{\infty}$*-increasing set is a* $\prec$*-decreasing set.*

4.3 Association of Probability Measures on POP-spaces

This section relies primarily on [39]. When unspecified, $(S, \mathcal{S}, \prec)$ shall denote an arbitrary POP-space. Let us first define the property of association:

Definition 4.3.1 *Let* P *be a probability measure on* $(S, \mathcal{S}, \prec)$.

1. P *is said to be associated (or associated* $(\prec)$ *when the order is ambiguous) if, for any pair* A, B *of measurable,* $\prec$*-increasing subsets of* S,

$$P(A \cap B) \geq P(A) \cdot P(B).$$

   This relationship holds if and only if

$$P(\widetilde{A} \cap \widetilde{B}) \geq P(\widetilde{A}) \cdot P(\widetilde{B})$$

   for any pair of $\prec$*-decreasing sets* $\widetilde{A}$, $\widetilde{B} \in \mathcal{S}$.

2. *A random element* $X : (\Omega, \mathcal{F}, P) \longrightarrow (S, \mathcal{S}, \prec)$ *of* S *is said to be associated (or associated* $(\prec)$ *when the order is ambiguous) if its distribution* $P \circ X^{-1}$ *is associated* $(\prec)$.

Theorem 3.1 of [39] includes a practical characterization of association which closely reflects Theorem 4.2.2's characterization of stochastic order:

Theorem 4.3.2 *Let* P *be a probability measure on a POP-space* $(S, \mathcal{S}, \prec)$*. The following statements are equivalent:*

1. *P is associated ($\prec$).*

2. *$P(C_1 \cap C_2) \geq P(C_1) \cdot P(C_2)$ for all pairs C_1, C_2 of $\prec$-increasing (resp. $\prec$-decreasing), closed subsets of S.*

3. *$P(G_1 \cap G_2) \geq P(G_1) \cdot P(G_2)$ for all pairs G_1, G_2 of $\prec$-increasing (resp. $\prec$-decreasing) compact generated subsets of S (see Remark 4.1.2).*

The usefulness of Theorem 4.3.2 resides in (2) being a special case of (1) and (3) being a special case of (2).

Recall Theorems 4.2.3 and 4.2.4. They stated, respectively, that coupling relations are preserved by weak limits, and that coupling relations on product spaces are characterized by coupling relations on the projection spaces. We ask ourselves if and when similar results could hold for association, namely:

1. When is association preserved by weak limits of associated probability measures? - and -

2. When is a probability measure associated on a countable product of POP-spaces?

It is here that the generality of Section 4.2 cannot be replicated entirely. The property of normal ordeliness is necessary to answer either question. Insofar as weak limits on abstract POP-spaces are concerned, we have the following theorem from [39] (Theorem 3.5):

Theorem 4.3.3 *If $(S, \mathcal{S}, \prec)$ is* normally ordered *and $\{P_n\}_{n=1}^{\infty}$ is a sequence of associated probability measures on $(S, \mathcal{S}, \prec)$ converging weakly to a probability measure P, then P is associated.*

Similarly, on a normally ordered *product* POP-space, a probability measure is associated if and only if its finite dimensional marginals are associated on the projection spaces. Theorem 5.1 of [39] may be stated as follows:

Theorem 4.3.4 *Let* $(S^{(\infty)}, \mathcal{S}^{(\infty)}, \prec^{\infty})$ *be a countable product of POP-spaces* $(S_1, \mathcal{S}_1, \prec_1)$, $(S_2, \mathcal{S}_2, \prec_2)$,... *For any* $n \in \mathbf{N}$ *let*

$$\begin{array}{rccc} \pi_n : & S^{(\infty)} & \longrightarrow & S^{(n)} \\ & (x_1, x_2, \ldots) & \longmapsto & (x_1, \ldots, x_n) \end{array}$$

denote the projection map associated to the first n coordinates. If $(S^{(\infty)}, \mathcal{S}^{(\infty)}, \prec^{\infty})$ *is normally ordered, then a probability measure P on* $(S^{(\infty)}, \mathcal{S}^{(\infty)}, \prec^{\infty})$ *is associated if and only if* $P \circ \pi_n^{-1}$ *is associated on the product space* $(S^{(n)}, \mathcal{S}^{(n)}, \prec^n)$ *for all* $n \in \mathbf{N}$.

In particular, by virtue of Remark 4.1.6,

Remark 4.3.5 *A probability measure P on* $(\bar{\mathbf{R}}_+^{\infty}, \mathcal{B}(\bar{\mathbf{R}}_+^{\infty}), \leq^{\infty})$ *is associated if and only if* $P \circ \pi_n^{-1}$ *is associated on* $(\bar{\mathbf{R}}_+^n, \mathcal{B}(\bar{\mathbf{R}}_+^n), \leq^n)$ *for all* $n \in \mathbf{N}$.

We have thus characterized the association of a probability measure on a normally ordered product POP-space as the association of its finite dimensional marginals. Still lacking is a criterion guaranteeing association on the product space - a certain characteristic of the measure which could be verified by way of computation. With Theorem 4.2.6 in hindsight, the reader may suspect this criterion will involve stochastic kernels. Define the *monotone kernel property* as follows:

Definition 4.3.6 *Let* $(S^{(\infty)}, \mathcal{S}^{(\infty)}, \prec^{\infty})$ *be a countable product of POP-spaces* $(S_1, \mathcal{S}_1, \prec_1)$, $(S_2, \mathcal{S}_2, \prec_2)$,..., *and let P be a probability measure on* $(S^{(\infty)}, \mathcal{S}^{(\infty)}, \prec^{\infty})$. *For any* $n \in \mathbf{N}$, *let*

$$\begin{array}{rccc} \pi_n : & S^{(\infty)} & \longrightarrow & S^{(n)} \\ & (x_1, x_2, \ldots) & \longmapsto & (x_1, \ldots, x_n) \end{array}$$

denote the projection map associated to the first n coordinates, and let

$$P_n^{n+1} : S^{(n)} \times \mathcal{S}_{n+1} \longmapsto [0,1]$$

denote the unique stochastic kernel in $S^{(n+1)} = S^{(n)} \times S_{n+1}$ *such that, for any* $A^{(n)} \in \mathcal{S}^{(n)}$, $A_{n+1} \in \mathcal{S}_{n+1}$,

$$\int_{A^{(n)}} P_n^{n+1}(x_1, \ldots, x_n, A_{n+1}) dP \circ \pi_n^{-1}(x_1, \ldots, x_n) = P \circ \pi_{n+1}^{-1}(A^{(n)} \times A_{n+1}).$$

If, for any $n \in \mathbf{N}$,

1. $P_n^{n+1}(x_1, \dots, x_n, \cdot)$ *is associated* $(\prec_{n+1})$ *for all* $(x_1, \dots, x_n) \in S^{(n)}$, *and*

2. *for any* $(x_1, \dots, x_n)$, $(y_1, \dots, y_n) \in S^{(n)}$, $(x_1, \dots, x_n) \prec^n (y_1, \dots, y_n)$ *implies* $P_n^{n+1}(x_1, \dots, x_n, \cdot) \prec_{n+1} P_n^{n+1}(y_1, \dots, y_n, \cdot)$ *as probability measures,*

then P *is said to satisfy the monotone kernel property, or, equivalently, is said to exhibit monotone kernels.*

This leads to the following criterion for association, which is a direct consequence of [39] (Theorem 4.1):

Theorem 4.3.7 *Let* P *be a probability measure on a countable product* $(S^{(\infty)}, \mathcal{S}^{(\infty)}, \prec^{\infty})$ *of POP-spaces. If* P *satisfies the monotone kernel property, then* P *is associated.*

It should be noted that the monotone kernel property is strictly stonger than association in general ([39]: p. 120), so that Theorem 4.3.7 does not have a converse.

It must also be mentioned that in Lindqvist's terminology, a "monotone stochastic kernel" is one which satisfies (2) of Definition 4.3.6, whereas an "associated stochastic kernel" is one which satisfies (1). We took the liberty of combining the two properties into one definition on account of the fact that $(\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty), \leq^\infty)$ will be the only product POP-space ever encountered in this thesis, and that (1) is automatically satisfied by $(\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty), \leq^\infty)$ for all $n \in \mathbf{N}$, as $\bar{\mathbf{R}}_+$ is totally ordered by $\leq$ (probability measures are always associated with respect to total orders).

The monotone kernel property is one of the easiest ways of ensuring the association of a probability measure on $(\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty), \leq^\infty)$. In this regard, $\bar{\mathbf{R}}_+^\infty$ is a privileged POP-space, one for which association may easily be verified. On more abstract POP-spaces, however, the criterion of kernel monotonicity is seldom available (the spaces in question may not even be product spaces). It is therefore of interest to try to take advantage of $\bar{\mathbf{R}}_+^\infty$'s simple structure by attempting to characterize the association of a probability measure on an abstract POP-space via the association of a related probability measure on $(\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty), \leq^\infty)$. Whence the appropriateness of a tool called *representation map* - our own creation.

Definition 4.3.8 *Let $(S, \mathcal{S})$ be a Polish space, and $\prec$ a closed partial order on $(S, \mathcal{S})$. A representation map $\Phi_\prec$ for $\prec$ is an injective, measurable map*

$$\Phi_\prec : (S, \mathcal{S}) \longrightarrow (\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty))$$

such that either

1. $x \prec y \iff \Phi_\prec(x) \leq^\infty \Phi_\prec(y)$ *for all* $x, y \in S$, *or*

2. $x \prec y \iff \Phi_\prec(y) \leq^\infty \Phi_\prec(x)$ *for all* $x, y \in S$.

If (1) holds, then $\Phi_\prec$ is said to be order-preserving; if (2) holds, it is said to be order-reversing.

The usefulness of such maps becomes apparent with this result:

Theorem 4.3.9 *Let P be a probability measure on a POP-space $(S, \mathcal{S}, \prec)$. If there exists a representation map*

$$\Phi_\prec : (S, \mathcal{S}) \longrightarrow (\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty))$$

for $\prec$, then P is associated $(\prec)$ if and only if $P \circ \Phi_\prec^{-1}$ associated $(\leq^\infty)$.

Proof For necessity, let us assume that P is associated $(\prec)$. If $A,\ B \in \mathcal{B}(\bar{\mathbf{R}}_+^\infty)$ are two $\leq^\infty$-increasing sets, then $\Phi_\prec^{-1}(A)$ and $\Phi_\prec^{-1}(B)$ are either both $\prec$-increasing or both $\prec$-decreasing, depending on whether $\Phi_\prec$ satisfies (1) or (2) of Definition 4.3.8. Therefore,

$$\begin{aligned} P \circ \Phi_\prec^{-1}(A \cap B) &= P[\Phi_\prec^{-1}(A) \cap \Phi_\prec^{-1}(B)] \\ &\geq P[\Phi_\prec^{-1}(A)] \cdot P[\Phi_\prec^{-1}(B)] \\ &= P \circ \Phi_\prec^{-1}(A) \cdot P \circ \Phi_\prec^{-1}(B) \end{aligned}$$

and it follows that $P \circ \Phi_\prec^{-1}$ is associated $(\leq^\infty)$. For sufficiency, suppose that $P \circ \Phi_\prec^{-1}$ is associated $(\leq^\infty)$, and let $A,\ B \in \mathcal{S}$ be two $\prec$-increasing sets. As $(S, \mathcal{S})$ and $(\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty))$ are two Polish spaces and $\Phi_\prec$ is both measurable and injective, Kuratowski's theorem ([33]: Corollary 2 of Theorem 13.1.9) ensures that the image sets

$\Phi_{\prec}(A)$ and $\Phi_{\prec}(B)$ are in $\mathcal{B}(\bar{\mathbf{R}}_{+}^{\infty})$. Also, the fact that $(\bar{\mathbf{R}}_{+}^{\infty}, \mathcal{B}(\bar{\mathbf{R}}_{+}^{\infty}))$ is Polish implies, by Theorems 1.1 and 1.4 of [4], that, for any $\epsilon > 0$, there exists compact sets $K_1 \subseteq \Phi_{\prec}(A)$ and $K_2 \subseteq \Phi_{\prec}(B)$ such that

$$[P \circ \Phi_{\prec}^{-1}(\Phi_{\prec}(A) \setminus K_1)] \vee [P \circ \Phi_{\prec}^{-1}(\Phi_{\prec}(B) \setminus K_2)] < \epsilon.$$

Pick ϵ, K_1 and K_2 accordingly and consider $G_1 := \mathrm{Inc}(K_1)$ and $G_2 := \mathrm{Inc}(K_2)$ if $\Phi_{\prec}$ is order-preserving, $G_1 := \mathrm{Dec}(K_1)$ and $G_2 := \mathrm{Dec}(K_2)$ if $\Phi_{\prec}$ is order-reversing. By Theorem 4.3.2 we know that

$$P \circ \Phi_{\prec}^{-1}(G_1 \cap G_2) \geq P \circ \Phi_{\prec}^{-1}(G_2) \cdot P \circ \Phi_{\prec}^{-1}(G_2),$$

and we also observe that $\Phi_{\prec}^{-1}(G_1) \subseteq A$ and $\Phi_{\prec}^{-1}(G_2) \subseteq B$ as $\Phi_{\prec}$ is a representation map. It follows that

$$\begin{aligned} P(A \cap B) &\geq P[\Phi_{\prec}^{-1}(G_1) \cap \Phi_{\prec}^{-1}(G_2)] \\ &= P \circ \Phi_{\prec}^{-1}(G_1 \cap G_2) \\ &\geq P \circ \Phi_{\prec}^{-1}(G_2) \cdot P \circ \Phi_{\prec}^{-1}(G_2) \\ &\geq [P(A) - \epsilon] \cdot [P(B) - \epsilon] \\ &\geq P(A) \cdot P(B) - 2\epsilon + \epsilon^2. \end{aligned}$$

As ϵ is arbitrary, we obtain $P(A \cap B) \geq P(A) \cdot P(B)$, which results in P being associated $(\prec)$. □

Remark 4.3.10 *Sufficiency does not require that $\Phi_{\prec}$ be injective in the proof of Theorem 4.3.9. In general, if if $(S_1, \mathcal{S}_1, \prec_1)$ and $(S_2, \mathcal{S}_2, \prec_2)$ are POP-spaces and*

$$\phi : (S_1, \mathcal{S}_1, \prec_1) \longrightarrow (S_2, \mathcal{S}_2, \prec_2)$$

is measurable and monotone (meaning that $\forall\, x, y \in S_1$, either $x \prec_1 y \Longrightarrow \phi(x) \prec_2 \phi(y)$ or $x \prec_1 y \Longrightarrow \phi(y) \prec_2 \phi(x)$), then the association of a probability measure P on $(S_1, \mathcal{S}_1, \prec_1)$ entails the association of the probability $P \circ \phi^{-1}$ induced by ϕ on $(S_2, \mathcal{S}_2, \prec_2)$.

At last, we tackle the problem of preserving association under the weak convergence of associated probability measures on abstract POP-spaces. This will be done by exhibiting representation maps for the closed partial orders considered, which are continuous almost surely in the limiting distribution.

Theorem 4.3.11 *Let $P_1, P_2, \ldots$ be probability measures on a POP-space $(S, \mathcal{S}, \prec)$ converging weakly to a probability measure P. If P_n is associated for all n and there exists a representation map*

$$\Phi_\prec : (S, \mathcal{S}) \longrightarrow (\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty))$$

for $\prec$ such that the set

$$\{x \in S : \Phi_\prec(x) \text{ is discontinuous at } x\}$$

is P-null, then P is also associated.

Proof Assuming the P_n's are associated and such a $\Phi_\prec$ exists, $P_n \circ \Phi_\prec^{-1}$ is associated ($\leq^\infty$) for all $n \in \mathbf{N}$ by Theorem 4.3.9. The convergence $P_n \Longrightarrow P$, coupled with the continuity of $\Phi_\prec$ everywhere outside a set of P-measure 0, entails that $P_n \circ \Phi_\prec^{-1} \Longrightarrow P \circ \Phi_\prec^{-1}$ on $(\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty))$ as $n \to \infty$ ([4]: Theorem 5.1). But since $(\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty), \leq^\infty)$ is normally ordered, $P \circ \Phi_\prec^{-1}$ must be associated ($\leq^\infty$) on $(\bar{\mathbf{R}}_+^\infty, \mathcal{B}(\bar{\mathbf{R}}_+^\infty))$. This implies, by Theorem 4.3.9, that P is associated on $(S, \mathcal{S}, \prec)$. $\square$.

Chapter 5

Couplings and Association of Point Processes

In this chapter we investigate various criteria which ensure couplings and association of point processes when viewed as random elements of the metric space $(\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$, which is itself embedded in the Polish space $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$. Three closed partial orders on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$ will be considered, along with corresponding representation maps.

The stochastic ordering of point processes finds wide application, including the study of "repairable sytems with schemes of planned replacements" (see [38], as well as their references). On the other hand, Burton and Waymire have established that certain stationary random measures (including point processes) which are associated satisfy "classical scaling limits" ([9]: Theorem 4.1). As we shall see, the stochastic ordering or association of point processes is often related to the stochastic ordering or association of random variables which characterize these point processes (interarrival times, for example). See the examples of Kwieciński and Szekli [37]. We may thus state unambiguously that there exists an independent interest for both stochastic orderings and association of point processes.

Section 5.1 defines, in the manner of Kwieciński and Szekli [37], the three partial orders $\prec_{\mathcal{N}}$, $\prec_{\mathcal{D}}$ and $\prec_{\infty}$ on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$, and associates representation maps to each of the three (representation maps were discussed in Section 4.3). Section 5.2 provides criteria on the path compensators of two point processes which ensure that the latter are stochastically ordered with respect to one of the three aforementioned partial orders. This section uses background material introduced in Section 4.2.

In a similar fashion, Section 5.3 develops criteria on the path compensator of one point process which ensure the latter is associated with respect to one of the three partial orders; in passing, it redefines a unified self-exciting property that Kwieciński and Szekli [37] had formalized separately for the three orders. Section 5.3 refers to results in Section 4.3.

Finally, as an application, Section 5.4 attempts to replicate, from the angle of monotone kernels, Kwieciński and Szekli's result ([37]: Theorem 4.2) that a point process which is self-exciting with respect to one of the three partial orders, is automatically associated with respect to that order.

5.1 Orders on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$

Recall that $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$ is the topological space consisting of all measures μ on $(\mathbf{R}_+, \mathcal{B}(\mathbf{R}_+))$ such that, for any $t \in \mathbf{R}_+$, $\mu([0,t]) \in \mathbf{N} \cup \{0\}$. This space is made complete and separable by the metric of vague convergence, which ensures that a sequence $\{\mu_n\}_{n=1}^{\infty}$ in $\mathcal{N}$ converges to some $\mu \in \mathcal{N}$ if and only if for any $m \in \mathbf{N}$, $\tau_m(\mu_n) \longrightarrow \tau_m(\mu)$ in $\bar{\mathbf{R}}_+$ as $n \to \infty$. The three closed partial orders on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$ that we shall define and represent are $\prec_{\mathcal{N}}$, $\prec_{\mathcal{D}}$ and $\prec_{\infty}$.

$\prec_{\mathcal{N}}$ is commonly called the thinning order. For any μ, $\nu \in \mathcal{N}$, one writes $\mu \prec_{\mathcal{N}} \nu$ if $\mu(B) \leq \nu(B)$ for any bounded Borel set $B \subseteq \mathbf{R}_+$. $\prec_{\mathcal{N}}$ is obviously a reflexive, transitive and antisymmetric binary relation, and its closure has been established by Kallenberg ([28]: 15.7); it is thus a bona fide closed partial order on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$.

The first representation map $\Phi^{(1)}_{\prec_{\mathcal{N}}}$ we shall define for $\prec_{\mathcal{N}}$ involves a countable class of relatively open intervals of $\mathbf{R}_+$. Let $\{r_n\}_{n=1}^{\infty}$ be an enumeration of the nonnegative rationals, define the class

$$\mathcal{T}^{(1)} := \{[0, r_n) : n \in \mathbf{N}\} \cup \{(r_l, r_m) : m, l \in \mathbf{N}\},$$

and let $\{C_n\}_{n=1}^{\infty}$ be an enumeration of the elements of $\mathcal{T}^{(1)}$. We claim:
Proposition 5.1.1

$$\begin{aligned} \Phi^{(1)}_{\prec_{\mathcal{N}}} : \mathcal{N} &\longrightarrow \bar{\mathbf{R}}_+^{\infty} \\ \mu &\longmapsto (\mu(C_1), \mu(C_2), \ldots) \end{aligned}$$

defines an order-preserving representation map for $\prec_{\mathcal{N}}$.

Proof Let us first show that $\Phi^{(1)}_{\prec_{\mathcal{N}}}$ is injective. Observe that the class $\mathcal{T}^{(1)}$ is a π-system generating the Borel σ-field over $\mathbf{R}_+$, and that both μ and ν are σ-finite over $\mathcal{T}^{(1)}$. If $\Phi^{(1)}_{\prec_{\mathcal{N}}}(\mu) = \Phi^{(1)}_{\prec_{\mathcal{N}}}(\nu)$, then μ and ν agree on $\mathcal{T}^{(1)}$, and must therefore agree on $\sigma(\mathcal{T}^{(1)}) = \mathcal{B}(\mathbf{R}_+)$ by Theorem 10.3 of [5]. As for $\Phi^{(1)}_{\prec_{\mathcal{N}}}$ being order-preserving, it is clear, by the definition of $\prec_{\mathcal{N}}$, that $\mu \prec_{\mathcal{N}} \nu \Longrightarrow \Phi^{(1)}_{\prec_{\mathcal{N}}}(\mu) \leq^{\infty} \Phi^{(1)}_{\prec_{\mathcal{N}}}(\nu)$. Conversely, if $\Phi^{(1)}_{\prec_{\mathcal{N}}}(\mu) \leq^{\infty} \Phi^{(1)}_{\prec_{\mathcal{N}}}(\nu)$, one may invoke the regularity of locally finite measures on bounded open subsets of a metric space ([4]: Theorem 1.1), along with the fact that both μ and ν are integer-valued, to show that for any bounded Borel set $B \subseteq \mathbf{R}_+$, there exist a finite, disjoint collection $B_{i_1}, \ldots, B_{i_n}$ of members of $\mathcal{T}^{(1)}$ such that $B \subseteq \bigcup_{j=1}^{n} B_{i_j}$, $\sum_{j=1}^{n} \mu(B_{i_j}) = \mu(B)$ and $\sum_{j=1}^{n} \nu(B_{i_j}) = \nu(B)$. In this event, the inequality $\Phi^{(1)}_{\prec_{\mathcal{N}}}(\mu) \leq^{\infty} \Phi^{(1)}_{\prec_{\mathcal{N}}}(\nu)$ entails $\mu(B) = \sum_{j=1}^{n} \mu(B_{i_j}) \leq \sum_{j=1}^{n} \nu(B_{i_j}) = \nu(B)$; as B is arbitrary, the relation $\mu \prec_{\mathcal{N}} \nu$ ensues. □

Another representation map $\Phi^{(1)}_{\prec_{\mathcal{N}}}$ for $\prec_{\mathcal{N}}$ may be obtained by modifying the "dissecting class" $\mathcal{T}^{(1)}$. Using the same enumeration $\{r_n\}_{n=1}^{\infty}$ of the nonnegative rationals, define the class

$$\mathcal{T}^{(2)} := \{[0, r_n] : n \in \mathbf{N}\} \cup \{(r_l, r_m] : m, l \in \mathbf{N}\},$$

along with the corresponding map

$$\begin{aligned} \Phi^{(2)}_{\prec_{\mathcal{N}}} : \mathcal{N} &\longrightarrow \bar{\mathbf{R}}_+^{\infty} \\ \mu &\longmapsto (\mu(\widetilde{C}_1), \mu(\widetilde{C}_2), \ldots), \end{aligned}$$

where $\{\widetilde{C}_n\}_{n=1}^{\infty}$ is an enumeration of members of $\mathcal{T}^{(2)}$.

Proposition 5.1.2 *$\Phi^{(2)}_{\prec_{\mathcal{N}}}$ is an order-preserving representation map for $\prec_{\mathcal{N}}$.*

Proof As every coordinate function of $\Phi^{(2)}_{\prec_{\mathcal{N}}}$ is the limit of a decreasing sequence of coordinate functions of $\Phi^{(1)}_{\prec_{\mathcal{N}}}$, it is clear that $\Phi^{(2)}_{\prec_{\mathcal{N}}}$ is measurable. $\Phi^{(2)}_{\prec_{\mathcal{N}}}$ may also be shown to be injective and order-preserving by repeating the argument used for $\Phi^{(1)}_{\prec_{\mathcal{N}}}$, substituting "$\mathcal{T}^{(2)}$" for every occurrence of "$\mathcal{T}^{(1)}$".

Remark 5.1.3 *Neither $\Phi^{(1)}_{\prec_{\mathcal{N}}}$ nor $\Phi^{(2)}_{\prec_{\mathcal{N}}}$ are continuous in the metric of vague convergence. To show $\Phi^{(1)}_{\prec_{\mathcal{N}}}$ is not continuous, observe that $[0,1)$ belongs to $\mathcal{T}^{(1)}$. Then $\mu_n := \delta_{1-\frac{1}{n}} \longrightarrow \mu := \delta_1$ in the metric of vague convergence as $n \to \infty$, but $\mu_n((0,1]) = 1 \not\longrightarrow \mu((0,1]) = 0$, so $\Phi^{(1)}_{\prec_{\mathcal{N}}}$ is not continuous. A similar reasoning may be used to show $\Phi^{(2)}_{\prec_{\mathcal{N}}}$ is discontinuous at measures which charge the left endpoints of members of $\mathcal{T}^{(2)}$.*

Let us now turn our attention to the partial order $\prec_{\mathcal{D}}$, commonly referred to as the Skorokhod order. Let $\mu, \nu \in \mathcal{N}$. We write $\mu \prec_{\mathcal{D}} \nu$ if, for any $t \in \mathbf{R}_+$, $\mu([0,t]) \leq \nu([0,t])$. As the intervals $[0,t]$ constitute a π-system which generates $\mathcal{B}(\mathbf{R}_+)$, it is clear that the binary relation $\prec_{\mathcal{D}}$ is antisymmetric in addition to being reflexive and transitive. Another characterization of the relation $\mu \prec_{\mathcal{D}} \nu$ is to say that $\tau_n(\mu) \geq \tau_n(\nu)$ for all $n \in \mathbf{N}$. Thence, the characterization of vague convergence by convergence of the jump points enables one to conclude that the order $\prec_{\mathcal{D}}$ is closed.

Two representation maps, one continuous and one not continuous, will be constructed for $\prec_{\mathcal{D}}$. The continuous representation involves the jump point of the measure argument; the non-continuous one involves intervals of the form $[0,t]$.

Proposition 5.1.4 *The map*

$$\begin{aligned} \Phi^{(1)}_{\prec_{\mathcal{D}}} : \mathcal{N} &\longrightarrow \bar{\mathbf{R}}_+^{\infty} \\ \mu &\longmapsto (\tau_1(\mu), \tau_2(\mu), \ldots) \end{aligned}$$

is a continuous, order-reversing representation map for $\prec_{\mathcal{D}}$.

Proof A measure of $\mathcal{N}$ is uniquely determined by its jump points, so $\Phi^{(1)}_{\prec_{\mathcal{D}}}$ is injective. $\Phi^{(1)}_{\prec_{\mathcal{D}}}$ is also continuous (in its coordinate functions), because the vague convergence of $\mathcal{N}$-measures is equivalent to the convergence of their jump points. At last, $\Phi^{(1)}_{\prec_{\mathcal{D}}}$ is order-reversing as a result of our characterization of $\prec_{\mathcal{D}}$ by the very same jump points. □

To define the second representation map for $\prec_{\mathcal{D}}$, retain our previous enumeration $\{r_n\}_{n=1}^{\infty}$ of nonnegative rationals.

Proposition 5.1.5 *The map*

$$\begin{aligned} \Phi^{(2)}_{\prec_{\mathcal{D}}} : \mathcal{N} &\longrightarrow \bar{\mathbf{R}}_{+}^{\infty} \\ \mu &\longmapsto (\mu([0,r_1]), \mu([0,r_2]), \ldots) \end{aligned}$$

is an order-preserving representation map for $\prec_{\mathcal{D}}$.

Proof In addition to being measurable (see the proof of Proposition 5.1.2), $\Phi^{(2)}_{\prec_{\mathcal{D}}}$ is injective because the class of intervals of the form $[0,r_n]$ is a π-system generating $\mathcal{B}(\mathbf{R}_+)$ over which any $\mu \in \mathcal{N}$ is σ-finite. It also results from a simple approximation argument that $\Phi^{(2)}_{\prec_{\mathcal{D}}}$ is order-preserving. □

Remark 5.1.6 *As in Remark 5.1.3, one can show that* $\Phi^{(2)}_{\prec_{\mathcal{D}}}$ *is not continuous in the metric of vague convergence.*

The last partial order to be studied is $\prec_\infty$, the "infinity" order. For any $\mu, \nu \in \mathcal{N}$, write $\mu \prec_\infty \nu$ if $\tau_n(\mu) - \tau_{n-1}(\mu) \geq \tau_n(\nu) - \tau_{n-1}(\nu)$ for all $n \in \mathbf{N}$. As the inter-jump distances of $\mathcal{N}$-measures determine the jump points themselves, it is clear that $\prec_\infty$ is antisymmetric, as well as reflexive and transitive. It is also closed by virtue of the characterization of vague convergence by convergence of the jump points. As we did for $\prec_{\mathcal{D}}$, we introduce two representation maps for $\prec_\infty$, one continuous, the other not continuous.

Proposition 5.1.7 *The map*

$$\begin{aligned} \Phi^{(1)}_{\prec_\infty} : \mathcal{N} &\longrightarrow \bar{\mathbf{R}}_{+}^{\infty} \\ \mu &\longmapsto (X_1(\mu), X_2(\mu), X_3(\mu), \ldots) \end{aligned}$$

where, for any $n \in \mathbf{N}$,

$$X_n(\mu) := \begin{cases} \tau_n(\mu) - \tau_{n-1}(\mu) & \textit{if } \tau_{n-1}(\mu) < \infty \\ \infty & \textit{otherwise} \end{cases},$$

is a continuous, order-reversing representation map for $\prec_\infty$.

Proof An argument similar to that of the proof of Proposition 5.1.4 may be invoked. □

At last, the non-continuous representation map $\Phi^{(2)}_{\prec_\infty}$ for $\prec_\infty$ will involve a class of intervals which, in contrast with previously used classes, will depend on the measure argument of $\Phi^{(2)}_{\prec_\infty}$. As usual, retain the enumeration $\{r_n\}_{n=1}^\infty$ of nonnegative rationals. For any $\mu \in \mathcal{N}$ define the class $\mathcal{I}^\mu$ as

$$\mathcal{I}^\mu := \{(\tau_n(\mu), \tau_n(\mu) + r_m] : m, n \in \mathbf{N}\} \cup \{[0, r_l] : l \in \mathbf{N}\},$$

and enumerate it as $\{I_n^\mu\}_{n=1}^\infty$.

Proposition 5.1.8 *The map*

$$\begin{aligned} \Phi^{(2)}_{\prec_\infty} : \mathcal{N} &\longrightarrow \bar{\mathbf{R}}_+^\infty \\ \mu &\longmapsto (\mu(I_1^\mu), \mu(I_2^\mu), \ldots) \end{aligned}$$

is an order-preserving representation map for $\prec_\infty$.

Proof Let us first show that $\Phi^{(2)}_{\prec_\infty}$ is measurable. For any $l, m, n \in \mathbf{N}$, the set $\{\mu \in \mathcal{N} : \mu((\tau_n(\mu), \tau_n(\mu) + r_m]) \geq l\}$ is equal to $\{\mu \in \mathcal{N} : \tau_{n+l}(\mu) - \tau_n(\mu) \leq r_m\}$; likewise, $\{\mu \in \mathcal{N} : \mu([0, r_m]) \geq l\}$ is equal to $\{\mu \in \mathcal{N} : \tau_l(\mu) \leq r_m\}$. As the jump points are continuous functions of their measures, the coordinate functions of $\Phi^{(2)}_{\prec_\infty}$ must be measurable. $\Phi^{(2)}_{\prec_\infty}$ is injective because, for any $\mu \in \mathcal{N}$, $n \in \mathbf{N} \cup \{0\}$, the distance $\tau_n(\mu) - \tau_{n+1}(\mu)$ can be recovered from the values $\mu((\tau_n(\mu), \tau_n(\mu) + r_l])$ by letting the index $l \in \mathbf{N}$ vary; similarly, the point $\tau_1(\mu)$ can be situated using the values $\mu([0, r_l])$ and letting l vary. Finally, as regards the order-preserving character of $\Phi^{(2)}_{\prec_\infty}$, it is immediate that if $\mu \prec_\infty \nu$ in $\mathcal{N}$, then $\Phi^{(2)}_{\prec_\infty}(\mu) \leq^\infty \Phi^{(2)}_{\prec_\infty}(\nu)$. Conversely, if $\Phi^{(2)}_{\prec_\infty}(\mu) \leq^\infty \Phi^{(2)}_{\prec_\infty}(\nu)$, then one must have $\tau_1(\mu) \geq \tau_1(\nu)$ from $\mu([0, r_l]) \leq \nu([0, r_l])$ for all $l \in \mathbf{N}$; $\tau_2(\mu) - \tau_1(\mu) \geq \tau_2(\nu) - \tau_1(\nu)$ from $\mu((\tau_1(\mu), \tau_1(\mu) + r_l]) \leq \nu((\tau_1(\nu), \tau_1(\nu) + r_l])$ for all $l \in \mathbf{N}, \ldots$ and so on to obtain $\mu \prec_\infty \nu$. □

Remark 5.1.9 *It may be shown that $\Phi^{(2)}_{\prec_\infty}$ is discontinuous at measures μ which charge the right endpoint of a member of $\mathcal{I}^\mu$. The argument is similar to the one which was developed in Remark 5.1.3.*

This completes our survey of closed partial orders on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$. The reader should take note of the concluding remarks.

Remark 5.1.10 *As we have exhibited continuous representation maps for closed partial orders $\prec_{\mathcal{D}}$ and $\prec_\infty$, Theorem 4.3.11 secures the following result: let $\{P_n\}$ be a sequence of probability measures on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$ which are associated which converge weakly to a probability measure P.*

1. *If $\prec \in \{\prec_{\mathcal{D}}, \prec_\infty\}$ and P_n is associated $(\prec)$ for all $n \in \mathbf{N}$, then P is associated $(\prec)$.*
2. *If P_n is associated $(\prec_{\mathcal{N}})$ for all $n \in \mathbf{N}$ and the set*

$$\{\mu \in \mathcal{N} : \Phi^{(2)}_{\prec_{\mathcal{N}}} \text{ is discontinuous at } \mu\}$$

is P-null, then P is associated $(\prec_{\mathcal{N}})$.

Remark 5.1.11 *It is clear that $\prec_{\mathcal{D}}$ is weaker than either $\prec_{\mathcal{N}}$ or $\prec_\infty$ in the sense that if $\mu \prec_{\mathcal{N}} \nu$ or $\mu \prec_\infty \nu$ in $\mathcal{N}$, then $\mu \prec_{\mathcal{D}} \nu$. The orders $\prec_{\mathcal{N}}$ and $\prec_\infty$, however, are not comparable: it is easy to construct examples which show that $\mu \prec_{\mathcal{N}} \nu$ does not imply $\mu \prec_\infty \nu$ and $\mu \prec_\infty \nu$ does not imply $\mu \prec_{\mathcal{N}} \nu$ in general.*

5.2 Compensator Criteria for Point Process Coupling

In light of the previous section, Theorems 4.2.4 and 4.2.7 entail the following result as pertains to the closed partial orders $\prec_{\mathcal{N}}$, $\prec_{\mathcal{D}}$ and $\prec_\infty$ on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$ (recall that point process relizations are by definition $\mathcal{N}_0$-valued and therefore do not charge $\{0\}$).

Theorem 5.2.1 *Let $N_1, N_2 : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ be two point processes. These three statements hold:*

1. $N_1 \prec_{\mathcal{N}_{st}} N_2 \iff \Phi^{(1)}_{\prec_{\mathcal{N}}}(N_1) \leq^{\infty}_{st} \Phi^{(1)}_{\prec_{\mathcal{N}}}(N_2) \iff \Phi^{(2)}_{\prec_{\mathcal{N}}}(N_1) \leq^{\infty}_{st} \Phi^{(2)}_{\prec_{\mathcal{N}}}(N_2)$ *if and only if*

$$(N_1((x_1,y_1]),\ldots,N_1((x_m,y_m])) \leq^{m}_{st} (N_2((x_1,y_1]),\ldots,N_2((x_m,y_m]))$$

for all $(x_1,y_1),\ldots,(x_m,y_m) \in \mathbf{R}^2_+$.

2. $N_1 \prec_{\mathcal{D}_{st}} N_2 \iff \Phi^{(1)}_{\prec_{\mathcal{D}}}(N_2) \leq^{\infty}_{st} \Phi^{(1)}_{\prec_{\mathcal{D}}}(N_1) \iff \Phi^{(2)}_{\prec_{\mathcal{D}}}(N_1) \leq^{\infty}_{st} \Phi^{(2)}_{\prec_{\mathcal{D}}}(N_2)$ *if and only if*

$$(N_1((0,t_1]),\ldots,N_1((0,t_m])) \leq^{m}_{st} (N_2((0,t_1]),\ldots,N_2((0,t_m]))$$

for all $t_1,\ldots,t_m \in \mathbf{R}_+$.

3. $N_1 \prec_{\infty_{st}} N_2 \iff \Phi^{(1)}_{\prec_{\infty}}(N_2) \leq^{\infty}_{st} \Phi^{(1)}_{\prec_{\infty}}(N_1) \iff \Phi^{(2)}_{\prec_{\infty}}(N_1) \leq^{\infty}_{st} \Phi^{(2)}_{\prec_{\infty}}(N_2)$ *if and only if*

$$\begin{aligned}(N_1((\tau_n(N_1),\tau_n(N_1)+t_1]),\ldots,N_1((\tau_n(N_1),\tau_n(N_1)+t_m]))\\ \leq^{m}_{st} (N_2((\tau_n(N_2),\tau_n(N_2)+t_1]),\ldots,N_2((\tau_n(N_2),\tau_n(N_2)+t_m]))\end{aligned}$$

for all $n \in \mathbf{N}\cup\{0\},\ t_1,\ldots,t_m \in \mathbf{R}_+$.

It may be inferred from results in [36] and [47] that, provided some regularity conditions hold, the *path compensators* of N_1 and N_2 may be tested over the same classes of intervals - taken individually - as those appearing in the preceding theorem to ensure a coupling of the two point processes. More precisely:

Theorem 5.2.2 *Let* $N_1, N_2 : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ *be two point processes with path compensators* $\Lambda_1, \Lambda_2 : \mathcal{N}_0 \times \mathbf{R}_+ \longmapsto \mathbf{R}_+$ *respectively, such that* $N_1(\mathbf{R}_+) = N_2(\mathbf{R}_+) = \infty$ *almost surely and* Λ_1, Λ_2 *are pathwise continuous.*

1. *If* Λ_1 *and* Λ_2 *are pathwise absolutely continuous and* $\Lambda_1(\mu,(s,t]) \leq \Lambda_2(\nu,(s,t])$ *for any* $s,t \in \mathbf{R}_+$, $\mu,\nu \in \mathcal{N}_0$ *such that* $s < t$ *and* $\mu \prec_{\mathcal{N}} \nu$, *then*

$$N_1 \prec_{\mathcal{N}_{st}} N_2;$$

2. *If* $\Lambda_1(\mu,(0,t]) \leq \Lambda_2(\nu,(0,t])$ *for all* $t \in \mathbf{R}_+$ *whenever* $\mu,\nu \in \mathcal{N}_0$ *satisfy* $\mu \prec_{\mathcal{D}} \nu$, *then*

$$N_1 \prec_{\mathcal{D}_{st}} N_2;$$

3. *If* $\Lambda_1(\mu,(\tau_n(\mu),\tau_n(\mu)+t]) \leq \Lambda_2(\nu,(\tau_n(\nu),\tau_n(\nu)+t])$ *for any* $t \in \mathbf{R}_+$, $n \in \mathbf{N} \cup \{0\}$ *and* $\mu,\nu \in \mathcal{N}_0$ *such that* $\mu \prec_\infty \nu$, *then*

$$N_1 \prec_{\infty_{st}} N_2.$$

Proof 1. Rolski and Szekli ([47]: Theorem 2) assert that if λ_1 and λ_2 are (path) densities of Λ_1 and Λ_2 respectively, then the condition $\lambda(\mu,t) \leq \lambda(\nu,t)$, satisfied for all $t > 0$ and $\mu,\nu \in \mathcal{N}_0$ such that $\mu \prec_{\mathcal{N}} \nu$, entails $N_1 \prec_{\mathcal{N}_{\mathrm{st}}} N_2$. But by the current hypothesis we may write $\lambda_1(\mu,t) := \lim_{s\to t^-} \frac{\Lambda_1(\mu,(s,t])}{t-s} \leq \lim_{s\to t^-} \frac{\Lambda_2(\nu,(s,t])}{t-s} =: \lambda_2(\nu,t)$.

2. Follows from [36] (Theorem 3.1 and Remark 3.1), once a few simple notational changes are made. The reader is reminded that two $\mathcal{N}_0$-measures μ and ν satisfy $\mu \prec_{\mathcal{D}} \nu$ if and only if $\tau_n(\mu) \geq \tau_n(\nu)$ for all $n \in \mathbf{N} \cup \{0\}$.

3. Let $\Pi : (\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P}) \longrightarrow (\mathcal{N}_0,\mathcal{B}(\mathcal{N}_0))$ be a Standard Poisson Process on $\mathbf{R}_+$; let θ represent the null measure on $\mathbf{R}_+$ and furthermore, for any $t_1,\ldots,t_m$ such that for any $i \in \{1,\ldots,m-1\}$, $t_i < \infty \Rightarrow 0 < t_i < t_{i+1}$, let $\phi_m(t_1,\ldots,t_m)$ be the $\mathcal{N}_0$-measure such that $\tau_1(\phi_m(t_1,\ldots,t_m)) = t_1,\ldots,\tau_m(\phi_m(t_1,\ldots,t_m)) = t_m$, and $\tau_{m+1}(\phi_m(t_1,\ldots,t_m)) = \infty$. Consider the maps $\Gamma_1,\Gamma_2 : (\mathcal{N}_0,\mathcal{B}(\mathcal{N}_0)) \longrightarrow (\mathcal{N}_0,\mathcal{B}(\mathcal{N}_0))$ defined as follows: for any measure $\mu \in \mathcal{N}_0$, let

$$\begin{aligned}
&s_1 := \inf\{t > 0 : \Lambda_1(\theta,(0,t]) \geq \tau_1(\mu)\};\\
&s_2 := s_1 + \inf\{t > 0 : \Lambda_1(\phi_1(s_1),(s_1,s_1+t]) \geq \tau_2(\mu) - \tau_1(\mu)\};\\
&\vdots\\
&s_{m+1} := s_m + \inf\{t > 0 : \Lambda_1(\phi_m(s_1,\ldots,s_m),(s_m,s_m+t]) \geq \tau_{m+1}(\mu) - \tau_m(\mu)\}\\
&\vdots
\end{aligned}$$

and put $\tau_m(\Gamma_1(\mu)) := s_m$ for $m \in \mathbf{N}$; similarly let

$t_1 := \inf\{t > 0 : \Lambda_2(\theta, (0, t]) \geq \tau_1(\mu)\};$

$t_2 := s_1 + \inf\{t > 0 : \Lambda_2(\phi_1(t_1), (t_1, t_1 + t]) \geq \tau_2(\mu) - \tau_1(\mu)\};$

$\vdots$

$t_{m+1} := s_m + \inf\{t > 0 : \Lambda_2(\phi_m(t_1, \ldots, t_m), (t_m, t_m + t]) \geq \tau_{m+1}(\mu) - \tau_m(\mu)\}$

$\vdots$

and put $\tau_m(\Gamma_2(\mu)) := t_m$ for $m \in \mathbf{N}$.

According to [36] (Proposition 3.1), $N_1 =_{\text{st}} \Gamma_1(\Pi)$ and $N_2 =_{\text{st}} \Gamma_2(\Pi)$. Assuming the current hypothesis (i.e. $\forall\, t > 0, n \in \mathbf{N} \cup \{0\}$, $\Lambda_1(\mu, (\tau_n(\mu), \tau_n(\mu) + t]) \leq \Lambda_2(\mu, (\tau_n(\nu), \tau_n(\nu) + t])$ whenever $\mu \prec_\infty \nu$ in $\mathcal{N}_0$), we shall obtain $N_1 \prec_{\infty_{\text{st}}} N_2$ once we show that $\Gamma_1(\mu) \prec_\infty \Gamma_2(\mu)$ for all $\mu \in \mathcal{N}_0$, or, equivalently, that $\tau_{n+1}(\Gamma_1(\mu)) - \tau_n(\Gamma_1(\mu)) \geq \tau_{n+1}(\Gamma_2(\mu)) - \tau_n(\Gamma_2(\mu))$ for all $\mu \in \mathcal{N}_0$ and all $n \in \mathbf{N} \cup \{0\}$ such that $\tau_n(\Gamma_1(\mu)) < \infty$. This, however, is easily shown by induction over n. □.

We have now developed criteria for the coupling of two point processes which use their respective path compensators. Historically, criteria which ensure the association of individual random elements have generally been crafted to resemble criteria which were already known to ensure the stochastic orderings of pairs of random elements - we have stressed this aspect throughout the present chapter. In the next section, we shall endeavour to develop a criterion for the association of a point process (with respect to one of the three common partial orders) that uses the compensators in a fashion similar to that exhibited in Theorem 5.2.2. The notion of *self-exciting* point process will emerge from this endeavour.

5.3 Compensator Criteria for Point Process Association

Using Theorems 4.3.4 and 4.3.9, the association of a point process on $\mathbf{R}_+$ with respect to $\prec_\mathcal{N}$, $\prec_\mathcal{D}$ or $\prec_\infty$ is characterized as follows:

Theorem 5.3.1 *Let $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ be a point process. These three statements hold:*

1. *N is associated ($\prec_{\mathcal{N}}$) if and only if $\Phi^{(1)}_{\prec_{\mathcal{N}}}(N)$ is associated ($\leq^{\infty}$) if and only if $\Phi^{(2)}_{\prec_{\mathcal{N}}}(N)$ is associated ($\leq^{\infty}$) if and only if*

$$(N((x_1, y_1]), \ldots, N_1((x_m, y_m])) \text{ is associated } (\leq^m)$$

for all $(x_1, y_1), \ldots, (x_m, y_m) \in \mathbf{R}^2_+$.

2. *N is associated ($\prec_{\mathcal{D}}$) if and only if $\Phi^{(1)}_{\prec_{\mathcal{D}}}(N)$ is associated ($\leq^{\infty}$) if and only if $\Phi^{(2)}_{\prec_{\mathcal{D}}}(N)$ is associated ($\leq^{\infty}$) if and only if*

$$(N((0, t_1]), \ldots, N((0, t_m])) \text{ is associated } (\leq^m)$$

for all $t_1, \ldots, t_m \in \mathbf{R}_+$.

3. *N is associated ($\prec_{\infty}$) if and only if $\Phi^{(1)}_{\prec_{\infty}}(N)$ is associated ($\leq^{\infty}$) if and only if $\Phi^{(2)}_{\prec_{\infty}}(N)$ is associated ($\leq^{\infty}$) if and only if*

$$(N((\tau_n(N), \tau_n(N) + t_1]), \ldots, N((\tau_n(N), \tau_n(N) + t_m])) \text{ is associated } (\leq^m)$$

for all $n \in \mathbf{N} \cup \{0\}$, $t_1, \ldots, t_m \in \mathbf{R}_+$.

As previously announced, the aim of this section is to produce criteria on the *path compensator* of N which entail its association of N with respect to $\prec_{\mathcal{N}}$, $\prec_{\mathcal{D}}$ or $\prec_{\infty}$, and which resemble those used in Theorem 5.2.2 to ensure coupling relations.

It is here that we introduce the notion of a *self-exciting* point process. Kwieciński and Szekli (1996) have been the first to formalize such a notion in relation to the partial orders $\prec_{\mathcal{N}}$, $\prec_{\mathcal{D}}$ and $\prec_{\infty}$. Our definition is more general in that it is not restricted to these three orders, and does not impose conditions of absolute continuity on the compensator (see [37]). Intuitively speaking, if $\prec$ is a closed partial order on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$, a point process N on $\mathbf{R}_+$ is self-exciting with respect to $\prec$ if, for any $t > 0$, knowing that the realization of N restricted to $(0, t]$ is "large" in terms of $\prec$, enables one to forecast that the realization on N restricted to (t, ∞) is likely to be "large" as well.

Our definition will hinge on a formal object called an *echelon set.* For any set $A \subseteq \mathcal{N} \times \mathbf{R}_+$, let $A^\mu := \{t \in \mathbf{R}_+ : (\mu, t) \in A\}$.

Definition 5.3.2 *A subset $A_{n,x,y}$ of $\mathcal{N} \times \mathbf{R}_+$, where $n \in \mathbf{N}_0$ and $0 \leq x < y < \infty$, is called an echelon set if, for all $\mu \in \mathcal{N}$,*

$$A^\mu_{n,x,y} = \{t \in \mathbf{R}_+ : (\mu, t) \in A_{n,x,y}\} = (\tau_n(\mu) + x, \tau_n(\mu) + y].$$

Echelon sets are notationally convenient tools used in comparing measures over intervals of the form $A^\mu_{n,x,y} = (\tau_n(\mu) + x, \tau_n(\mu) + y]$. For the purposes of defining the self-exciting property with respect to a certain closed partial order $\prec$ on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$, the echelon sets we shall consider are those which satisfy the property of "$\prec$-concordance:"

Definition 5.3.3 *If $\prec$ is a closed partial order on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$, an echelon set $A_{n,x,y}$ is said to be $\prec$-concordant if, $\forall$ μ, $\nu \in \mathcal{N}_0$ such that $\mu \prec \nu$,*

$$\mu(A^\mu_{n,x,y}) \leq \nu(A^\nu_{n,x,y}).$$

Given a closed partial order $\prec$ on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$ and a measure $\mu \in \mathcal{N}_0$, the value $\mu(A^\mu_{n,x,y})$ of the μ-measure of the section (at μ) of the $\prec$-concordant set $A_{n,x,y}$ constitutes an "indicator" of how "big" μ is in terms of $\prec$.

Echelon sets have been taylored to produce familiar $\prec$-concordant sections when $\prec \in \{\prec_{\mathcal{N}}, \prec_{\mathcal{D}}, \prec_\infty\}$; these sections are of the same forms as the intervals over which a point process was evaluated in the last theorem. Indeed:

Proposition 5.3.4 *Let $A_{n,x,y}$ be an echelon set. Then:*

1. *$A_{n,x,y}$ is $\prec_{\mathcal{N}}$-concordant if and only if $n = 0$;*

2. *$A_{n,x,y}$ is $\prec_{\mathcal{D}}$-concordant if and only if $n = 0$ and $x = 0$;*

3. *$A_{n,x,y}$ is $\prec_\infty$-concordant if and only if $x = 0$.*

Thus,

- $A_{n,x,y}$ *is* $\prec_{\mathcal{N}}$*-concordant* $\Longleftrightarrow A^{\mu}_{n,x,y} = (x, y]\ \forall\ \mu \in \mathcal{N}$;
- $A_{n,x,y}$ *is* $\prec_{\mathcal{D}}$*-concordant* $\Longleftrightarrow A^{\mu}_{n,x,y} = (0, y]\ \forall\ \mu\ \in \mathcal{N}$;
- $A_{n,x,y}$ *is* $\prec_{\infty}$*-concordant* $\Longleftrightarrow A^{\mu}_{n,x,y} = (\tau_n(\mu), \tau_n(\mu) + y]\ \forall\ \mu\ \in \mathcal{N}$.

Proof 1. If $n = 0$, then $\forall\ \mu,\ \nu \in \mathcal{N}_0$ such that $\mu \prec_{\mathcal{N}} \nu$ we have $\mu(A^{\mu}_{0,x,y}) = \mu((x, y]) \leq \nu((x, y]) = \nu(A^{\nu}_{0,x,y})$, which implies that $A_{n,x,y} = A_{0,x,y}$ is $\prec_{\mathcal{N}}$-concordant. If $n \neq 0$, let $\nu \in \mathcal{N}_0$ be such that $\tau_{n+1}(\nu) = \tau_n(\nu) + y + 1$ and $\tau_i(\nu) = \tau_{i-1}(\nu) + (x + y)/2$ for $i \geq n + 2$. Construct $\mu \in \mathcal{N}_0$ such that $\tau_i(\mu) = \tau_i(\nu)$ for $i \in \{1, ..., n - 1\}$ and $\tau_i(\mu) = \tau_{i+1}(\nu)$ for $i \geq n$. Then, clearly, $\mu \prec_{\mathcal{N}} \nu$ but $\mu(A^{\mu}_{n,x,y}) > \nu(A^{\nu}_{n,x,y}) = 0$. Therefore, $A_{n,x,y}$ is not $\prec_{\mathcal{N}}$-concordant if $n \neq 0$.

2. If $n = 0 = x$, then $\forall\ \mu,\ \nu \in \mathcal{N}_0$ such that $\mu \prec_{\mathcal{D}} \nu$ we have $\mu(A^{\mu}_{0,0,y}) = \mu((0, y]) \leq \nu((0, y]) = \nu(A^{\nu}_{0,0,y})$, which implies that $A_{n,x,y} = A_{0,0,y}$ is $\prec_{\mathcal{D}}$-concordant. If $n \neq 0$, construct μ and ν as for (1). Then $\mu \prec_{\mathcal{D}} \nu$ but $\mu(A^{\mu}_{n,x,y}) > \nu(A^{\nu}_{n,x,y})$, so $A_{n,x,y}$ is not $\prec_{\mathcal{D}}$-concordant if $n \neq 0$. Now consider $A_{n,x,y} = A_{0,x,y}$ with $x \neq 0$. Let $\nu \in \mathcal{N}_0$ be such that $\tau_1(\nu) = x/2$ and $\tau_2(\nu) = y + 1$; let $\mu \in \mathcal{N}_0$ be such that $\tau_1(\mu) = (x + y)/2$ and $\tau_i(\mu) = \tau_i(\nu)\ \forall\ i \geq 2$. $A_{0,x,y}$ is not $\prec_{\mathcal{D}}$-concordant because $\mu \prec_{\mathcal{D}} \nu$, while $\mu(A^{\mu}_{0,x,y}) > \nu(A^{\nu}_{0,x,y})$.

3. If $x = 0$, then $\forall\ \mu,\ \nu \in \mathcal{N}_0$ such that $\mu \prec_{\infty} \nu$ we have $\mu(A^{\mu}_{n,0,y}) = \mu((\tau_n(\mu), \tau_n(\mu) + y]) \leq \nu((\tau_n(\nu), \tau_n(\nu) + y]) = \nu(A^{\nu}_{n,0,y})$. Therefore, $A_{n,x,y} = A_{n,0,y}$ is $\prec_{\infty}$-concordant if $x = 0$. If $x \neq 0$, let $\nu \in \mathcal{N}_0$ be such that $\tau_{n+1}(\nu) = [2\tau_n(\nu) + x]/2$, $\tau_{n+2}(\nu) = y + 1$, and $\tau_i(\nu) = \tau_{i-1}(\nu) + 1$ for $i \geq n + 3$; let $\mu \in \mathcal{N}_0$ be such that $\tau_i(\mu) = \tau_i(\nu)$ for $i \in \{1, .., n\}$, $\tau_{n+1}(\mu) = \tau_n(\nu) + (x + y)/2$, and $\tau_i(\mu) = \tau_{i-1}(\mu) + [\tau_i(\nu) - \tau_{i-1}(\nu)]$ for $i \geq n+2$. Then $\mu \prec_{\infty} \nu$ but $\mu(A^{\mu}_{n,x,y}) > \nu(A^{\nu}_{n,x,y})$, whence $A_{n,x,y}$ is not $\prec_{\infty}$-concordant if $x \neq 0$. □

We are now poised to define the self-exciting property formally:

Definition 5.3.5 *Let* $\prec$ *be a closed partial order on* $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$. *A point process* $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ *with path compensator* $\Lambda : \mathcal{N}_0 \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ *is said to* *be self-exciting with respect to $\prec$ if, for any $\prec$-concordant echelon set $A_{n,x,y} \subseteq \mathcal{N} \times \mathbf{R}_+$,*

$$\int_{A^{\mu}_{n,x,y}} \Lambda(\mu, dt) \leq \int_{A^{\nu}_{n,x,y}} \Lambda(\nu, dt)$$

holds whenever μ, $\nu \in \mathcal{N}_0$ satisfy $\mu \prec \nu$.

Remark 5.3.6 *By Proposition 5.3.4, a point process $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ with path compensator $\Lambda : \mathcal{N}_0 \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+ \ldots$*

1. *...is self-exciting with respect to $\prec_{\mathcal{N}}$ if and only if*

$$\Lambda(\mu, (x, y]) \leq \Lambda(\nu, (x, y])$$

   for any $x, y \in \mathbf{R}_+$, $\mu, \nu \in \mathcal{N}_0$ such that $\mu \prec_{\mathcal{N}} \nu$;

2. *...is self-exciting with respect to $\prec_{\mathcal{D}}$ if and only if*

$$\Lambda(\mu, (0, t]) \leq \Lambda(\nu, (0, t])$$

   for any $t \in \mathbf{R}_+$, $\mu, \nu \in \mathcal{N}_0$ such that $\mu \prec_{\mathcal{D}} \nu$;

3. *...is self-exciting with respect to $\prec_{\infty}$ if and only if*

$$\Lambda(\mu, (\tau_n(\mu), \tau_n(\mu) + t]) \leq \Lambda(\nu, (\tau_n(\nu), \tau_n(\nu) + t])$$

   for any $t \in \mathbf{R}_+$, $n \in \mathbf{N} \cup \{0\}$, and $\mu, \nu \in \mathcal{N}_0$ such that $\mu \prec_{\infty} \nu$.

At last, the compensator criterion for association may be stated. It was essentially formulated in [37]. The reader may ascertain that according to Remark 5.3.6, Theorems 5.3.1 and 5.3.7 constitute a direct analogue to Theorems 5.2.1 and 5.2.2.

Theorem 5.3.7 *Let $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ be a point process with a path compensator $\Lambda : \mathcal{N}_0 \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ that is pathwise continuous, and such that, almost surely, $N(\mathbf{R}_+) = \infty$. These three statements hold:*

1. *If Λ is pathwise absolutely continuous and N is self-exciting with respect to $\prec_{\mathcal{N}}$, then N is associated $(\prec_{\mathcal{N}})$.*

2. *If N is self-exciting with respect to $\prec_{\mathcal{D}}$, then N is associated ($\prec_{\mathcal{D}}$).*

3. *If N is self-exciting with respect to $\prec_{\infty}$, then N is associated ($\prec_{\infty}$).*

Proof 1. Kwieciński and Szekli ([37]: Theorem 4.2) have shown that if Λ admits a density $\lambda : \mathcal{N}_0 \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$, then the condition

$$\lambda(\mu, t) \leq \lambda(\nu, t)$$

for all $t > 0$ and all $\mu, \nu \in \mathcal{N}_0$ such that $\mu \prec_{\mathcal{N}} \nu$, entails that N is associated ($\prec_{\mathcal{N}}$). Under the current hypothesis, however, for any such μ, ν and t we may put

$$\lambda(\mu, t) := \lim_{s \to t^-} \frac{\Lambda(\mu, (s, t])}{t - s} \leq \lim_{s \to t^-} \frac{\Lambda(\nu, (s, t])}{t - s} := \lambda(\nu, t).$$

2. This follows immediately from Theorem 4.2 of [37] and Remark 5.3.6.
3. With proper notational changes effected, Kwieciński and Szekli ([37]: Theorem 4.2) assert that if, for any $\mu, \nu \in \mathcal{N}_0$ such that $\mu \prec_{\infty} \nu$, any $n \in \mathbf{N}$ and any $t > 0$, the condition

$$\Lambda(\mu\,|_{[0, \tau_n(\mu)]}, (\tau_n(\mu), \tau_n(\mu) + t]) \leq \Lambda(\nu\,|_{[0, \tau_n(\nu)]}, (\tau_n(\nu), \tau_n(\nu) + t])$$

entails that N is associated ($\prec_{\infty}$). One observes, however, that $\mu \prec_{\infty} \nu$ implies $\mu\,|_{[0, \tau_n(\mu)]} \prec_{\infty} \mu\,|_{[0, \tau_n(\nu)]}$, so that, under the current hypothesis, such a condition is met by Remark 5.3.6. □

Remark 5.3.8 *Our definition of the self-exciting property coincides with that of Kwieciński and Szekli for the cases $\prec = \prec_{\mathcal{N}}$ and $\prec = \prec_{\mathcal{D}}$ when the hypotheses of Theorem 5.3.7 are fulfilled. When $\prec = \prec_{\infty}$, however, our definition is strictly stronger, since renewal processes with increasing failure rates are self-exciting with respect to $\prec_{\infty}$ in Kwieciński and Szekli's definition, but not in ours (see the remark of [37] inserted between Lemma 4.2 and Example 4.2, as opposed to Theorem 3.1 of [46]).*

5.4 The Self-exciting Property and Monotone Kernels as Criteria for Point Process Association

Let $\prec \in \{\prec_{\mathcal{N}}, \prec_{\mathcal{D}}, \prec_{\infty}\}$. Two criteria may be inferred from the material presented in this chapter and the preceding one to ensure that a point process $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ be associated $(\prec)$.

The first criterion is that for some representation map

$$\Phi_{\prec} : (\mathcal{N}, \mathcal{B}(\mathcal{N})) \longrightarrow (\bar{\mathbf{R}}_+^{\infty}, \mathcal{B}(\bar{\mathbf{R}}_+^{\infty}))$$

for $\prec$, the distribution $P \circ \Phi_{\prec}(N)^{-1}$ exhibit monotone kernels. Indeed, in such an event, the random element $\Phi_{\prec}(N)$ is associated $(\leq^{\infty})$ by Theorem 4.3.7, which implies, in turn, that N itself is associated $(\prec)$ by Theorem 4.3.9.

The second criterion is simply that N be self-exciting with respect to $(\prec)$ and satisfy the hypotheses of Theorem 5.3.7 which specifically apply to $(\prec)$.

The question at hand is whether the second criterion implies the first. This question is very broad and - unfortunately - elicits no simple answers. Depending on $\prec$, varying conclusions may be reached.

Easiest is the case $\prec = \prec_{\infty}$, $\Phi_{\prec} = \Phi^{(1)}_{\prec_{\infty}}$, for which a clean result has been obtained:

Theorem 5.4.1 *Let $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ be a point process with a path compensator $\Lambda : \Omega \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ that everywhere admits continuous paths. If N is self-exciting with respect to $\prec_{\infty}$, then $P \circ \Phi^{(1)}_{\prec_{\infty}}(N)^{-1}$ satisfies the monotone kernel property, and N is therefore associated $(\prec_{\infty})$.*

Proof Suppose N is self-exciting with respect to $\prec_{\infty}$. To prove that $P \circ \Phi^{(1)}_{\prec_{\infty}}(N)^{-1}$ satisfies the monotone kernel property, it suffices to show that for all $n \in \mathbf{N}$ and

$x \geq 0$, the inequality

$$\begin{aligned} &P[X_{n+1} > x \mid X_1(N) = X_1(\mu), \ldots, X_n(N) = X_n(\mu)] \\ &\quad \geq P[X_{n+1} > x \mid X_1(N) = X_1(\nu), \ldots, X_n(N) = X_n(\nu)] \end{aligned} \tag{6}$$

whenever $\mu \prec_\infty \nu$ in $\mathcal{N}_0$. If $X_n(\mu) = \infty$, then the left-hand-side of (6) equals 1. If $X_n(\mu) < \infty$, then $X_n(\nu) < \infty$ since $\mu \prec_\infty \nu$. Without loss of generality we may assume $\tau_{n+1}(\mu) = \tau_{n+1}(\nu) = \infty$ without altering the conditional probabilities or the relation $\mu \prec_\infty \nu$; this is because the only information that is provided by the conditioning arguments are the first n jump points of μ and ν respectively, and the very same information would be provided by all measures sharing the same first n jump points. Accordingly,

$$\begin{aligned} &P[X_{n+1} > x \mid X_1(N) = X_1(\mu), \ldots, X_n(N) = X_n(\mu)] \\ &\quad = P[X_{n+1} > x \mid \tau_1(N) = \tau_1(\mu), \ldots, \tau_n(N) = \tau_n(\mu)] \\ &\quad = 1 - P[\tau_{n+1}(N) \leq x + \tau_n(\mu) \mid \tau_1(N) = \tau_1(\mu), \ldots, \tau_n(N) = \tau_n(\mu)] \\ &\quad = 1 - F_{n+1}(x; \tau_1(\mu), \ldots, \tau_n(\mu)) \end{aligned}$$

$$\begin{aligned} &= \exp\{-[\Lambda(\mu; x + \tau_n(\mu)) - \Lambda(\mu; \tau_n(\mu))]\} \\ &\qquad\qquad \text{(by Lemma 3.5 of [23], } \Lambda \text{ being pathwise continuous)} \\ &= \exp\{-\textstyle\int_{\tau_n(\mu)}^{\tau_n(\mu)+x} \Lambda(\mu, dt)\} \\ &\geq \exp\{-\textstyle\int_{\tau_n(\nu)}^{\tau_n(\mu)+x} \Lambda(\nu, dt)\} \qquad (N \text{ being self-exciting w.r.t. } \prec_\infty) \\ &= P[X_{n+1} > x \mid X_1(N) = X_1(\nu), \ldots, X_n(N) = X_n(\nu)]\ . \ \Box \end{aligned}$$

A similar theorem would not be valid for the case $\prec = \prec_{\mathcal{D}}$ and $\Phi_\prec = \Phi^{(1)}_{\prec_{\mathcal{D}}}$, as the following counterexample illustrates:

Example 5.4.2 *Let* $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ *be a point process with a path compensator* $\Lambda : \Omega \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ *specified as follows: for any* $\mu \in \mathcal{N}_0$, $t \in \mathbf{R}_+$, *if* $\tau_1(\mu) \leq 1$ *let*

$$\Lambda(\mu, t) = \begin{cases} t & \text{if } 0 \leq t \leq 3 \\ 3 + \frac{1}{2}(t-3) & \text{if } 3 < t \leq 5 \\ 4 & \text{if } t > 5 \end{cases},$$

otherwise let

$$\Lambda(\mu, t) = \begin{cases} t & \text{if } 0 \leq t \leq 2 \\ 2 + \frac{1}{2}(t-2) & \text{if } 2 < t \leq 4 \\ t - 1 & \text{if } 4 < t \leq 5 \\ 4 & \text{if } t > 5 \end{cases} .$$

Then N is self-exciting with repect to $\prec_{\mathcal{D}}$, but $P \circ \Phi^{(1)}_{\prec_{\mathcal{D}}}{}^{-1}$ does not satisfy the monotone kernel property.

Proof It is obvious that $\Lambda(\mu, t) \leq \Lambda(\nu, t)$ for all $t \in \mathbf{R}_+$ whenever $\mu \prec_{\mathcal{D}} \nu$ in $\mathcal{N}_0$, so N is self-exciting with respect to $\prec_{\mathcal{D}}$ (such a point process exists by Theorem 3.6 of [23]). Let $\mu, \nu \in \mathcal{N}_0$ be such that $\tau_1(\mu) = 2$, $\tau_2(\mu) = 4$, $\tau_3(\mu) = \infty$ and $\tau_1(\nu) = 1$, $\tau_2(\nu) = 4$, $\tau_3(\nu) = \infty$. Then $\mu \prec_{\mathcal{D}} \nu$ and if $P \circ \Phi^{(1)}_{\prec_{\mathcal{D}}}{}^{-1}$ were to satisfy the monotone kernel property, one would expect

$$P[\tau_3(N) > 5 \mid \tau_1(N) = 2, \tau_2(N) = 4] \geq P[\tau_3(N) > 5 \mid \tau_1(N) = 1, \tau_2(N) = 4],$$

but computations reveal:

$$\begin{aligned} P[\tau_3(N) > 5 \mid \tau_1(N) = 2, \tau_2(N) = 4] \\ &= P[\tau_3(N) - \tau_2(N) > 1 \mid \tau_1(N) = 2, \tau_2(N) = 4] \\ &= \exp\{-\textstyle\int_4^5 \Lambda(\mu, dt)\} \\ &= \tfrac{1}{e} < \tfrac{1}{\sqrt{e}} \\ &= \exp\{-\textstyle\int_4^5 \Lambda(\nu, dt)\} \\ &= P[\tau_3(N) - \tau_2(N) > 1 \mid \tau_1(N) = 1, \tau_2(N) = 4] \\ &= P[\tau_3(N) > 5 \mid \tau_1(N) = 1, \tau_2(N) = 4], \end{aligned}$$

and we conclude that $P \circ \Phi^{(1)}_{\prec_{\mathcal{D}}}{}^{-1}$ does not satisfy the monotone kernel property. □

Although we do not have counterexamples to illustrate that N being self-exciting with respect to $\prec_{\mathcal{D}}$ does not necessarily entail that $P \circ \Phi^{(2)}_{\prec_{\mathcal{D}}}{}^{-1}$ exhibits monotone kernels, we suspect this is the case.

Finally, the case $\prec \, = \, \prec_{\mathcal{N}}$ is problematic because known representation maps for the partial order $\prec_{\mathcal{N}}$ - namely, $\Phi^{(1)}_{\prec_{\mathcal{N}}}$ and $\Phi^{(2)}_{\prec_{\mathcal{N}}}$ - produce sequential vectors whose entries cannot be ordered in an intuitively coherent fashion. By that we mean that the image vector's entries, in general, list values taken by a measure over a certain class of subets of $\mathbf{R}_+$, and these subsets cannot be enumerated with an index whose values correspond to the order of appearance of the subsets on the half-line.

If, however, a point process $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ is self-exciting with respect to $\prec_{\mathcal{N}}$ and satisfies the hypotheses of Theorem 2.3.3, then the sequence $\{N^n\}_{n=1}^{\infty}$ of approximating processes whose existence Theorem 2.3.3 guarantees, is such that for any $n \in \mathbf{N}$, N^n admits a representation whose distribution exhibits monotone kernels.

Pick $n \in \mathbf{N}$. Recall that $Q_n = \{0, \frac{1}{2^n}, \frac{2}{2^n}, \ldots\}$ according to Theorem 2.3.3's statement. Define

$$\mathcal{N}_0^n := \{\mu \in \mathcal{N}_0 : \mu(\mathbf{R}_+ \setminus Q_n) = 0\}.$$

It is obvious that $\mathcal{N}_0^n$ is a closed subset of $\mathcal{N}$, and thus $(\mathcal{N}_0^n, \mathcal{B}(\mathcal{N}) \cap \mathcal{N}_0^n, \prec_{\mathcal{N}})$ becomes a POP-space in its own right. Let

$$\begin{array}{rcl} \Phi^n_{\prec_{\mathcal{N}}} : \mathcal{N}_0^n & \longrightarrow & (\bar{\mathbf{R}}_+^{\infty}, \mathcal{B}(\bar{\mathbf{R}}_+^{\infty})) \\ \mu & \longmapsto & (\mu(\{0\}), \mu(\{\frac{1}{2^n}\}), \mu(\{\frac{1}{3^n}\}), \ldots) \end{array} .$$

Evidently, $\Phi^n_{\prec_{\mathcal{N}}}$ constitutes a representation map for the order $\prec_{\mathcal{N}}$ restricted to $\mathcal{N}_0^n$.

Theorem 5.4.3 *Let $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ be a point process satisfying the hypotheses of Theorem 2.3.3. If N is self-exciting with respect to $\prec_{\mathcal{N}}$, then the sequence $\{N^n\}_{n=1}^{\infty}$ of approximating processes is such that, for all $n \in \mathbf{N}$, $P \circ \Phi^n_{\prec_{\mathcal{N}}}(N^n)^{-1}$ satisfies the monotone kernel property.*

Proof Borrowing all relevant notation from the statement of Theorem 2.3.3, we suppose the parameter π is, in fact, deterministic, so that we may rewrite the array $\{\Pi_{\{p,i\}} : p \in I, i \in \mathbf{N} \cup \{0\}\}$ as $\{\Pi_i\}_{i \in \mathbf{N}}$ and $\Lambda : \Omega \times \mathbf{R} \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ as $\Lambda : \Omega \times \mathbf{R}_+ \longrightarrow \mathbf{R}_+$, suppressing the middle argument.

Let $m, n \in \mathbf{N}$. Assuming N is self-exciting with respect to $\prec_{\mathcal{N}}$, we must show

$$P\left[\Delta N^n_{\frac{m+1}{2^n}} \mid \Delta N^n_{\frac{1}{2^n}} = \delta_1, \ldots \Delta N^n_{\frac{m}{2^n}} = \delta_m\right] \leq P\left[\Delta N^n_{\frac{m+1}{2^n}} \mid \Delta N^n_{\frac{1}{2^n}} = \delta'_1, \ldots \Delta N^n_{\frac{m}{2^n}} = \delta'_m\right]$$

whenever $\delta_1, \ldots, \delta_m, \delta'_1, \ldots, \delta'_m \in \{0,1\}$ are such that $\delta_i \leq \delta'_i$ for $i = 1, \ldots, m$. Picking the δ_i, δ'_i accordingly, let $\mu, \nu \in \mathcal{N}_0$ be defined by

$$\mu((0,t]) := \sum_{i=1}^{m} \mathbf{1}_{[\frac{i}{2^n} \leq t]} \cdot \delta_i; \qquad \nu((0,t]) := \sum_{i=1}^{m} \mathbf{1}_{[\frac{i}{2^n} \leq t]} \cdot \delta'_i$$

for $t \in \mathbf{R}_+$. Then $\mu \prec_{\mathcal{N}} \nu$. Let $a := \mu(\mathbf{R}_+)$, $b := \nu(\mathbf{R}_+)$. We now have

$$\begin{aligned} &P[\Delta N^n_{\frac{m+1}{2^n}} \mid \Delta N^n_{\frac{1}{2^n}} = \delta_1, \ldots, \Delta N^n_{\frac{m}{2^n}} = \delta_m] \\ &\quad = 1 - P\left\{\Pi_a\left(\left(\Lambda(\mu, \tfrac{m}{2^n}), \Lambda(\mu, \tfrac{m+1}{2^n})\right]\right) = 0\right\} \\ &\quad = 1 - \exp\{-\left[\Lambda(\mu, \tfrac{m+1}{2^n}) - \Lambda(\mu, \tfrac{m}{2^n})\right]\} \\ &\quad \leq 1 - \exp\{-\left[\Lambda(\nu, \tfrac{m+1}{2^n}) - \Lambda(\nu, \tfrac{m}{2^n})\right]\} \\ &\qquad\qquad (N \text{ is self-exciting w.r.t. } \prec_{\mathcal{N}}) \\ &\quad = 1 - P\left\{\Pi_b\left(\left(\Lambda(\nu, \tfrac{m}{2^n}), \Lambda(\nu, \tfrac{m+1}{2^n})\right]\right) = 0\right\} \\ &\quad = P[\Delta N^n_{\frac{m+1}{2^n}} \mid \Delta N^n_{\frac{1}{2^n}} = \delta'_1, \ldots, \Delta N^n_{\frac{m}{2^n}} = \delta'_m], \end{aligned}$$

completing the proof. □

Corollary 5.4.4 *If $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ is a point process which is self-exciting with respect to $\prec_{\mathcal{N}}$ and which satisfies the hypotheses of Theorem 2.3.3, then N is associated $(\prec_{\mathcal{N}})$.*

Proof By Theorems 5.4.3, 4.3.7 and 4.3.9, N^n is associated $(\prec_{\mathcal{N}})$ for all $n \in \mathbf{N}$. But since $N^n \longrightarrow_{\mathcal{L}} N$, by Remark 5.1.10 it suffices to show that the set

$$\{\omega \in \Omega : \Phi^{(2)}_{\prec_{\mathcal{N}}} \text{ is discontinuous at } N(\omega)\}$$

is P-null. Observe, however that the set $\{\mu \in \mathcal{N} : \Phi^{(2)}_{\prec_{\mathcal{N}}} \text{ is discontinuous at } \mu\}$ is contained in the set

$$\bigcup_{r \in \mathbf{Q} \cap \mathbf{R}_+} \{\mu \in \mathcal{N} : \mu(\{r\}) > 0\}.$$

Since Λ is pathwise continuous and N has unit jumps, for any $r \in \mathbf{Q} \cap \mathbf{R}_+$ we have

$$\begin{aligned} P[\omega \in \Omega : N(\omega)(r) > 0] &= \mathbf{E}\{\Delta N_r\} \\ &= \mathbf{E}\{\Delta \Lambda_r\} \\ &= 0 \end{aligned}$$

Since the set of nonnegative rational numbers is denumerable, the set

$$\{\omega \in \Omega : \Phi^{(2)}_{\prec_{\mathcal{N}}} \text{ is discontinuous at } N(\omega)\}$$

is P-null and it follows that N is associated $(\prec_{\mathcal{N}})$. $\square$

Part III

Point Processes on $\mathbf{R}_+^2$

Chapter 6

Planar Point Processes and Flows

According to Jacod ([23]: Theorems 2.1 and 3.4), every simple point process N on $\mathbf{R}_+$ which does not charge 0, and such that $\mathbf{E}[N_t] < \infty \; \forall t \in \mathbf{R}_+$, admits an essentially unique dual predictable projection (compensator) Λ with respect to its internal history; furthermore, Λ characterizes the distribution of N in the sense that no two simple point processes having different distributions admit the same compensator. The essential features of Λ are therefore

i) its existence;

ii) its (essential) uniqueness;

iii) its power to characterize the underlying point process distribution.

Several attempts - none entirely satisfactory - have been made to generalize these three features to (putative) compensators of simple point processes on $\mathbf{R}_+^2$. Points (i) and (ii) follow from the Doob-Meyer decomposition of submartingales of class (D) on $\mathbf{R}_+$. Because every point process is an increasing process and every (locally integrable) increasing process is a (local) submartingale of class (D), some authors (e.g. Brennan [6], Dozzi [15], Gushchin [19]) have sought outright generalizations of the Doob-Meyer decomposition to planar submartingales; others have restricted the scope of their investigation to planar increasing processes (e.g. Merzbach and Zakai [43]), while others still (e.g. Ivanoff [20], Ivanoff and Merzbach [21]) have specifically focused on planar point processes. These Doob-Meyer decompositions vary in form and assumptions, depending on how the authors extend the notions of "past", "predictability" and "compensator" - which are natural and unambiguous on $\mathbf{R}_+$ - to the planar context (competing notions of "past" and "compensator" are discussed in Section 4, while the reader may glance at pp. 58-59 of [19] for an exhaustive list of the many types of "predictability"). A common assumption made by Dozzi ([15]: Proposition 1), Merzbach and Zakai ([43]: Proposition 6) and Gushchin ([19]: Theorems 2, 3, 6 and 9) in stating their decomposition results, is the so-called Cairoli-Walsh - or "F4" - condition on a planar filtration $\{\mathcal{F}_z\}_{z\in\mathbf{R}_+^2}$: for any $x, y, s, t \in \mathbf{R}_+$, σ-fields $\mathcal{F}_{(x+s,y)}$ and $\mathcal{F}_{(x,y+t)}$ are conditionally independent given $\mathcal{F}_{(x,y)}$. This condition, which is quite strong, appears crucial in guaranteeing the existence of regular versions of martingales of the type $\{\mathbf{E}(\mathbf{1}_A|\mathcal{F}_z)\}_{z\in\mathbf{R}_+^2}$ ([2]: Théorème on p. 152), and such versions seem, in turn, to ensure both the existence ([14]: Proposition 5) and uniqueness ([3]: Theorem 2) of the predictable projections necessary to establish the uniqueness of the compensator ([3]: Theorems 3 and 4). Proposition 1.1 of Ivanoff [20] guarantees the existence of a compensator for simple planar point processes without recourse to "F4", but still relies on "F4" for the compensator's uniqueness. At last, Maziotto and Merzbach did prove the existence of a compensator for a point process with linearly ordered jumps ([42]: Theorem 2.4); this chapter, in fact, owes much to their idea of embedding the "lower layers" of a point process in a space which is "larger" than the plane and where these "layers" are totally ordered. Nevertheless, as emerged from a private conversation with one of the authors, it is the compensator's uniqueness as a *stochastic measure* on the predictable sets which was established, rather than its uniqueness as a *process*. We can only conclude, based on these sources, that the existence of a unique Doob-Meyer decomposition has yet to be proven for simple $\mathbf{R}_+^2$-indexed point process in general.

As regards a compensator's power to characterize the distribution of a underlying simple point process on $\mathbf{R}_+^2$, results in this direction have been fragmentary. Ivanoff and Merzbach ([21]: (2), pg. 401) have shown that the so-called *compensator* does *not* characterize the underlying point process distribution in general, but have also shown the so-called **-compensator does* characterize the point process' distribution when this point process is Poisson or Cox ([22]: Theorem 5.3.1). Compensators and *-compensators are discussed in Section 4.

In this chapter, we argue in favor of a *family* of $\mathbf{R}_+$-indexed compensators induced by a planar point process, rather than suggesting a single $\mathbf{R}_+^2$-indexed compensator. We show that, under mild conditions, the three essential features of existence, uniqueness, and power to characterize the underlying distribution, are retained by our proposed scheme.

The approach taken was inspired by Mazziotto and Merzbach [42], and consists of embedding $\mathbf{R}_+^2$ ($\bar{\mathbf{R}}_+^2$, in fact) in the so-called space $(\mathcal{L}, \mathcal{B}(\mathcal{L}))$ of *lower layers* of $\bar{\mathbf{R}}_+^2$. Lower layers are closed subsets of $\bar{\mathbf{R}}_+^2$ which are decreasing with respect to the coordinate partial order $\leq$. It so happens that if μ is an integer valued measure on $\mathbf{R}_+^2$, then for any integer $n \in \mathbf{N}$, the set $L_n^\mu = \overline{\{z \in \bar{\mathbf{R}}_+^2 : \mu(\{z' \in \bar{\mathbf{R}}_+^2 : z' \leq z\}) < n\}}$ is a lower layer; additionally, $L_1^\mu \subseteq L_2^\mu \subseteq L_3^\mu \ldots$, a property which is used repeatedly. The measure μ on $\mathbf{R}_+^2$ determines, and is determined by an associated measure $\tilde{\mu}$ on $\mathcal{L}$ which "counts" the L_i^μ's.

Such a correspondence is explicitly exhibited in Section 2, along with an analogous correspondence which accommodates the measure μ being randomized into a planar point process N. To this effect, **Remark 6.2.8** states that $P \circ N^{-1}$ uniquely determines $P \circ \widetilde{N}^{-1}$, while **Theorem 6.2.18** asserts the converse. Section 2 does not involve any discussion of compensators; this is saved for Section 3.

Section 3 introduces the notion of flow. A flow is an injective, increasing and right-continuous map $f : \bar{\mathbf{R}}_+ \longrightarrow \mathcal{L}$ which "traverses" the space $\mathcal{L}$ in a sense which will be later clarified. Every flow f gives rise to a point process N^f on $\mathbf{R}_+$, which itself induces a compensator Λ^f with respect to its internal history. **Theorem 6.3.5** states that the law $P \circ N^{-1}$ uniquely determines, and is uniquely determined by the family of "flow process" distributions $P \circ {N^f}^{-1}$. **Corollary 6.3.8**, which may be viewed as the ultimate objective of this chapter, guarantees that if N satisfies the property of *strict simplicity*, then $P \circ N^{-1}$ is uniquely determined by the family of linear compensators $\Lambda^f = \{\Lambda_t^f\}_{t\geq 0}$.

In addition to summarizing the main results of this chapter (**Theorem 6.4.1**), Section 4 stresses the advantages of the current approach, in comparison to past attempts of determining a planar point process distribution by the latter's "compensator(s)". Emphasis is put on the differing notions of "past information" which have surfaced in connection with contemporary studies of $\mathbf{R}_+^2$-indexed processes.

At last, in Section 1, which immediately follows, the space $\mathcal{L}$ is topologized and its subsequently invoked properties are outlined.

6.1 The Topological Space $(\mathcal{L}, \mathcal{B}(\mathcal{L}))$

Consider a compact metrization of the extended half-line $\bar{\mathbf{R}}_+ := [0, \infty]$. The extended plane $\bar{\mathbf{R}}_+^2 := [0, \infty] \times [0, \infty]$, endowed with a product metric derived from that of $\bar{\mathbf{R}}_+$, is (locally) compact, Hausdorff and second-countable. In turn, the space $\mathcal{K}$ consisting of all non-empty, compact subsets of $\bar{\mathbf{R}}_+^2$ may then, according to Matheron ([40]: Corollary 1-4-1 and Proposition 1-4-4), be topologized as a locally compact, Hausdorff and second-countable (LCS) space using the Haudorff metric ρ defined by writing

$$\rho(K, K') := \max\left\{\sup_{x\in K} d(x, K'), \sup_{x'\in K'} d(x', K)\right\},$$

where d represents the metric on $\bar{\mathbf{R}}_+^2$ and K, K' are nonempty compact subsets of $\bar{\mathbf{R}}_+^2$. We let $\mathcal{B}(\mathcal{K})$ denote the Borel σ-field on $\mathcal{K}$ induced by ρ.

Consider the coordinate partial order $\leq$ on $\bar{\mathbf{R}}_+^2$ defined via the relation $(x, y) \leq (x', y') \iff x \leq x'$ and $y \leq y'$. An "infimum" $\wedge$ and a "supremum" $\vee$ may naturally be defined on $\bar{\mathbf{R}}_+^2$ by writing

$$(x, y) \wedge (x', y') \equiv (x \wedge x', y \wedge y')$$

and

$$(x, y) \vee (x', y') \equiv (x \vee x', y \vee y'),$$

and it is straightforward to ascertain the continuity of $\bar{\mathbf{R}}^2_+ \times \bar{\mathbf{R}}^2_+ \to \bar{\mathbf{R}}^2_+$-mappings $(z, z') \rightsquigarrow z \wedge z'$ and $(z, z') \rightsquigarrow z \vee z'$.

The space $(\mathcal{L}, \mathcal{B}(\mathcal{L}))$ will appear as a subspace of $(\mathcal{K}, \mathcal{B}(\mathcal{K}))$ and will soon be shown to constitute an LCS space in its own right. The elements of $(\mathcal{L}, \mathcal{B}(\mathcal{L}))$ shall be called the *lower layers* of the extended plane; they are similar in concept, though not identical, to what Maziotto and Merzbach ([42]) have called *sets of separation.*

Definition 6.1.1 *A lower layer of $\bar{\mathbf{R}}^2_+$ is a closed (and hence, compact) subset L of $\bar{\mathbf{R}}^2_+$ which satisfies the following two conditions:*

i) $(0, 0) \in L$;

ii) for any $x, x', y, y' \in \bar{\mathbf{R}}_+$, $(x, y) \in L$ and $(x', y') \leq (x, y)$ together imply $(x', y') \in L$.

In other words, a lower layer is a closed, nonempty and $\leq$-decreasing subset of $\bar{\mathbf{R}}^2_+$. The set of all lower layers of $\bar{\mathbf{R}}^2_+$ is noted $\mathcal{L}$.

The following proposition is a consequence of Matheron's so-called "tractable convergence theorem", which was invoked for a similar purpose by Mazziotto and Merzbach ([42]).

Proposition 6.1.2 *$\mathcal{L}$ is a closed subset of $\mathcal{K}$ in the Hausdorff topology induced by ρ.*

Proof Since $\bar{\mathbf{R}}^2_+$ is compact, Theorems 1-2-2 and 1-4-1 of [40] yield the following characterization for convergence in $\mathcal{K}$: a sequence $\{K_n\}_{n=1}^{\infty}$ in $\mathcal{K}$ converges to some $K \in \mathcal{K}$ if and only if

i) for any $z \in K$ there exists, for every sufficiently large $n \in \mathbf{N}$, a point $z_n \in K_n$ such that $\lim_{n\to\infty} z_n = z$ topologically;

ii) for every subsequence $\{K_{n_k}\}_{k=1}^{\infty}$ of $\{K_n\}_{n=1}^{\infty}$ and every corresponding sequence $\{z_{n_k}\}_{k=1}^{\infty}$ such that $z_{n_k} \in K_{n_k}$ for all $k \in \mathbf{N}$, if $\{z_{n_k}\}_{k=1}^{\infty}$ converges, then $\lim_{k\to\infty} z_{n_k} \in K$.

Let $\{L_n\}_{n=1}^{\infty}$ be a sequence in $\mathcal{L}$ converging to some $L \in \mathcal{K}$. As a consequence of part (ii) of the convergence criterion, L is a closed subset of $\bar{\mathbf{R}}_+^2$ which contains the origin. If $(x, y) \in L$ and some (x', y') satisfies $(x', y') \leq (x, y)$, then, by part (i), there exists a sequence $\{(x_n, y_n)\}_{n=N_{(x,y)}}^{\infty}$ converging to (x, y) with the property that $(x_n, y_n) \in L_n$ for all $n \geq N_{(x,y)}$. For any such n, however, the element $(x', y') \wedge (x_n, y_n) = (x' \wedge x_n, y' \wedge y_n)$ must belong to L_n because L_n is a lower layer. But since the binary operation $\wedge : \bar{\mathbf{R}}_+^2 \times \bar{\mathbf{R}}_+^2 \to \bar{\mathbf{R}}_+^2$ is continuous, $\{(x' \wedge x_n, y' \wedge y_n)\}_{n \geq N_{(x,y)}}$ converges to $(x' \wedge x, y' \wedge y) = (x', y')$, and this limit must be in L by part (ii) of the criterion. L is therefore a *bona fide* lower layer and the closure of $\mathcal{L}$ ensues. □

Corollary 6.1.3 *$\mathcal{L}$ is an LCS topological space. Its Borel class shall be denoted by $\mathcal{B}(\mathcal{L})$.*

The space $(\mathcal{L}, \mathcal{B}(\mathcal{L}))$ may be endowed with the partial order $\leq$ derived from the relation of inclusion: $L \leq L' \iff L \subseteq L'$. This binary relation is obviously reflexive, transitive and antisymmetric. As for the extended plane, a "minimum" $\wedge$ and a "maximum" $\vee$ may be defined - in this case using the set operations of intersection and union:

$$L \wedge L' := L \cap L'$$

and

$$L \vee L' := L \cup L'$$

for any $L, L' \in \mathcal{L}$. The $\mathcal{L} \times \mathcal{L}$ binary operations $(L, L') \rightsquigarrow L \wedge L'$ and $(L, L') \rightsquigarrow L \vee L'$ are also continuous by virtue of the "tractable convergence criterion" used in the proof of Proposition 6.1.2. Moreover, the definitions of $\wedge$ and $\vee$ may be extended in an obvious way to cover lattice infima and suprema of countable subsets of $\mathcal{L}$, and it is clear, by the way the metric ρ is defined, that any decreasing sequence of lower layers will converge topologically to its lattice infimum. A transitive, but not reflexive binary relation $<<$ on $\mathcal{L}$ may be obtained by strengthening $\leq$ as follows: for any $L, L' \in \mathcal{L}$, one writes $L << L'$ if $L \leq L'$ and $L \in \{L'' \in \mathcal{L} : L'' \leq L'\}^o$ (the interior being taken with respect to the Hausdorff topology). An important property of $(\mathcal{L}, \mathcal{B}(\mathcal{L}))$ is "separability from above" with respect to the strengthened partial order $<<$. The expression "separability from above" was introduced by Kurtz ([34]) and subsequently used by Maziotto and Merzbach ([42]), but we have altered its meaning somewhat to signify the property proven to hold here:

Proposition 6.1.4 *There exists a countable and dense subset $\mathcal{L}_0$ of $\mathcal{L} \setminus \{\bar{\mathbf{R}}^2_+\}$ which is closed under finite iterations of $\wedge$, and such that for any $L \in \mathcal{L} \setminus \{\bar{\mathbf{R}}^2_+\}$, there exists a sequence $\{L^{(n)}\}_{n=1}^{\infty}$ in $\mathcal{L}_0$ which converges to L and which satisfies $L << \ldots << L^{(3)} << L^{(2)} << L^{(1)}$.*

Proof For any $n \in \mathbf{N}$, let $\mathcal{C}^{(n)}$ denote the class of all "squares"

$$C_{ij}^{(n)} := \left[\frac{i-1}{2^n}, \frac{i}{2^n}\right] \times \left[\frac{j-1}{2^n}, \frac{j}{2^n}\right], \qquad C_{i\infty} := \left[\frac{i-1}{2^n}, \frac{i}{2^n}\right] \times \{\infty\}$$

and

$$C_{\infty j} := \{\infty\} \times \left[\frac{j-1}{2^n}, \frac{j}{2^n}\right], \qquad i, j \in \mathbf{N},$$

and let $\mathcal{U}_n$ denote the class of all lower layers different from $\bar{\mathbf{R}}^2_+$ which are unions of members of $\mathcal{C}^{(n)}$. Because any lower layer in $\mathcal{U}_n$ must have a finite number of "corners", $\mathcal{U}_n$ is countable. Let $\mathcal{L}_0 := \bigcup_{n \in \mathbf{N}} \mathcal{U}_n$, and let us now show that every element of $L \in \mathcal{L} \setminus \{\bar{\mathbf{R}}^2_+\}$ can be "approximated from above" by members of $\mathcal{L}_0$. Let $L \in \mathcal{L} \setminus \{\bar{\mathbf{R}}^2_+\}$. We construct the approximating sequence $\{L^{(n)}\}_{n=1}^{\infty}$ as follows: for any $n \in \mathbf{N}$, we let $\widetilde{L}^{(n)}$ be the intersection of all members of $\mathcal{U}_n$ which cover L (note that $\widetilde{L}^{(n)} \in \mathcal{U}_n$), and we let $L^{(n)}$ be the smallest member of $\mathcal{U}_n$ which covers $\widetilde{L}^{(n)}$ and whose border does not intersect that of $\widetilde{L}^{(n)}$. It is clear that $L^{(1)}, L^{(2)}, \ldots$ thus defined satisfy $L << \ldots << L^{(3)} << L^{(2)} << L^{(1)}$ and converge to L. □

Separability from above, coupled with the continuity of the binary operation $\wedge$, gives rise to the following interesting property:

Proposition 6.1.5 *Let $L, L' \in \mathcal{L}$ satisfy $L \leq L'$ and $L \neq L'$. There exists a decreasing sequence $\{\widetilde{L}^{(n)}\}_{n=1}^{\infty}$ in $\mathcal{L}$ with lattice infimum L, and with the properties that $L \leq \widetilde{L}^{(n)} \leq L'$ and $L \neq \widetilde{L}^{(n)} \neq L'$ for all $n \in \mathbf{N}$.*

Proof Let $L, L' \in \mathcal{L}$ be as stated, and consider the sequence $\{L^{(n)}\}_{n=1}^{\infty}$ which approximates L from above as described by Proposition 6.1.4. Let us first show there exists an $n_0 \in \mathbf{N}$ such that $L' \not\leq L^{(n_0)}$. If no such n_0 exists, then $L' \leq L^{(n)}$ $\forall n \in \mathbf{N}$, which implies, by the continuity of $\wedge$, that $L' = L' \wedge L^{(n)} \longrightarrow L' \wedge L = L$ as $n \to \infty$, i.e. that $L = L'$, which contradicts the hypothesis. With $n_0 \in \mathbf{N}$ set, observe that $L' \not\leq L^{(n)}$ for all $n \geq n_0$ by transitivity. Let us check that setting $\widetilde{L}^{(n)} := L^{(n_0+n)} \wedge L'$, $n \in \mathbf{N}$ completes the proof. We clearly have $L \leq \widetilde{L}^{(n)} \leq L'$, so there only remains to prove that $L \neq \widetilde{L}^{(n)} \neq L'$ for all $n \in \mathbf{N}$. Fix $n \in \mathbf{N}$. If $\widetilde{L}^{(n)} = L'$, then $L^{(n+n_0)} \cap L' = L'$, which implies $L' \leq L^{(n+n_0)}$, contradicting our choice of n_0. If $L = \widetilde{L}^{(n)}$, then $L^{(n+n_0)} \cap L' \subseteq L$, which implies $(L^{(n+n_0)} \setminus L) \cap L' = \emptyset$ which, at last, entails that $L' \subseteq L$ by the fact that L' is a lower layer. But this contradicts the hypothesis because $L \neq L'$. □

6.2 A Point Process on the Plane and its Associated Process on $\mathcal{L}$

In this section, a point process N on $\mathbf{R}_+^2$ is formally defined, along with a corresponding random measure $\widetilde{N}$ on $\mathcal{L}$. Relations between the respective distributions of N and $\widetilde{N}$ are scrutinized. The salient results of this section are Remark 6.2.8 and Theorem 6.2.18.

For the purposes of the present chapter, $\mathcal{N} = \mathcal{N}(\mathbf{R}_+^2)$ shall denote the space of all point measures $\mu := \sum_{i\in\mathbf{N}} \delta_{z_i}$ on $\mathbf{R}_+^2$ with the restrictions that

i) $\mu(\{0,0\}) = 0$, and

ii) the sequence $\{z_i\}_{i\in\mathbf{N}} = \{(x_i, y_i)\}_{i\in\mathbf{N}}$ has at most one accumulation point in $\mathbf{R}_+^2$, and if $z = (x, y)$ is such an accumulation point, then $x_i < x$ and $y_i < y$ for all $i \in \mathbf{N}$.

Condition (i) is imposed for technical reasons which will become apparent later, whereas condition (ii) ensures the measure μ is uniquely characterized by the values it takes on the sets $[(0,0), z]$, $z \in \mathbf{R}_+^2$. Of course, for any set A belonging to the class $\mathcal{B}(\mathbf{R}_+^2)$, $\mu(A)$ is defined as the cardinality of the set $\{i \in \mathbf{N} : z_i \in A\}$. To simplify notation, for any $z \in \bar{\mathbf{R}}_+^2$ we write

$$\mu_z := \mu([(0,0), z] \cap \mathbf{R}_+^2).$$

Since we will study random elements of $\mathcal{N}$, it is important to endow the space $\mathcal{N}$ with a σ-algebra $\mathcal{F}(\mathcal{N})$ ($\mathcal{N}$ will not be topologized, however). Define the "canonical" σ-algebra $\mathcal{F}(\mathcal{N})$ on $\mathcal{N}$ as the one generated by all the evaluation maps $\phi_z : \mathcal{N} \longrightarrow (\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+))$ which, for any fixed point $z \in \mathbf{R}_+^2$, associate to the measure $\mu \in \mathcal{N}$ the value $\phi_z(\mu) = \mu_z$. Our definition of a point process on $\mathbf{R}_+^2$ is certainly not universal, but is nevertheless convenient:

Definition 6.2.1 *A point process on $\mathbf{R}_+^2$ is a random element $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{F}(\mathcal{N}))$.*

As in the case of point processes on the line, a point process N may be viewed as a collection of $\{N_z\}_{z \in \mathbf{R}_+^2}$ of extended random variables such that for any $z \in \mathbf{R}_+^2$,

$$N_z := N([(0,0), z]),$$

and this notation is consistent with the one expressing μ_z as $\mu([(0,0), z])$, for some $\mu \in \mathcal{N}$.

With $\mu \in \mathcal{N}$ chosen, let us now define, as promised in the chapter introduction, the point measure $\tilde{\mu} := \sum_{i \in \mathbf{N}} \delta_{L_i^\mu}$ on $(\mathcal{L}, \mathcal{B}(\mathcal{L}))$ by constructing the lower layers L_i^μ corresponding to the measure μ as follows: for any $i \in \mathbf{N}$ let

$$L_i^\mu := \overline{\{z \in \bar{\mathbf{R}}_+^2 : \mu_z < i\}}.$$

Let us verify the L_i^μ's are *bona fide* lower layers with the following

Proposition 6.2.2 *For any $i \in \mathbf{N}$, L_i^μ is a lower layer.*

Proof For any $i \in \mathbf{N}$, $(0,0) \in L_i^\mu$ because μ does not charge the origin. Since L_i^μ is also closed in $\bar{\mathbf{R}}_+^2$, it remains to check that for any $z \in L_i^\mu$, $z' \leq z$ implies $z' \in L_i^\mu$. If $z \in L_i^\mu$, there exists a sequence $\{z_n\}_{n=1}^\infty$ converging to z, and such that $\mu_{z_n} < i$ for all $n \in \mathbf{N}$. Furthermore, if $z' \leq z$, then the sequence $\{z' \wedge z_n\}_{n=1}^\infty$ converges to $z' \wedge z = z'$

by the continuity of the binary operation $\wedge$ on $\bar{\mathbf{R}}_+^2 \times \bar{\mathbf{R}}_+^2$. But since $\mu_{z' \wedge z_n} < i$ for all $n \in \mathbf{N}$, z' must be in L_i^μ. □

The following observation is trivial:

Observation 6.2.3 $0 << L_1^\mu \leq L_2^\mu \leq L_3^\mu \leq \ldots$

Evidently, μ uniquely determines $\tilde{\mu}$. As has already been reported by Maziotto and Merzbach ([42]; Theorem 1.4), the converse is also true in the following sense: if κ is a point measure on $\mathcal{L}$ of which it is known beforehand that there exists a measure μ on $\mathbf{R}_+^2$ such that $\kappa = \tilde{\mu}$, then κ uniquely and explicitly determines μ (by "explicitly", we mean that κ determines the values μ_z, $z \in \mathbf{R}_+^2$).

To see this, adopt the following notation: for any $z = (x, y) \in \mathbf{R}_+^2$ define $D_z \in \mathcal{L}$ as

$$D_z := \{(u, v) \in \bar{\mathbf{R}}_+^2 : u \leq x \text{ or } v \leq y\}.$$

D_z has the geometric appearance of an "L-shaped" region which is the union of two perpendicular "bands", each band adjacent to one extended axis. Also, if L is a lower layer, let $\tilde{\mu}_L$ be the value

$$\tilde{\mu}_L := \tilde{\mu}(\{L' \in \mathcal{L} : L' \leq L\}).$$

That κ determines μ in the previously expounded sense is a direct consequence of the following

Remark 6.2.4 *For any $z \in \mathbf{R}_+^2$, $\mu_z = \tilde{\mu}_{D_z}$.*

Proof Pick $z \in \mathbf{R}_+^2$. For any $i \in \mathbf{N}$, $L_i^\mu \leq D_z$

$$\begin{aligned}
&\iff \overline{\{(u, v) \in \bar{\mathbf{R}}_+^2 : \mu_{(u,v)} < i\}} \subseteq D_z \\
&\iff \{(u, v) \in \bar{\mathbf{R}}_+^2 : \mu_{(u,v)} < i\} \subseteq D_z \qquad (D_z \text{ is closed}) \\
&\iff \mu_{z+(\frac{1}{n},\frac{1}{n})} \geq i \quad \forall n \in \mathbf{N} \\
&\iff \mu_z \geq i \qquad \text{(by the way } \mu \text{ is specified).}
\end{aligned}$$

Thus, $\tilde{\mu}_{D_z} = |\{i \in \mathbf{N} : L_i^\mu \leq D_z\}| = |\{i \in \mathbf{N} : \mu_z \geq i\}| = \mu_z$. □

We now focus our attention on the analogous question of determining one measure from the other, but as it pertains to point processes. More precisely, if $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, (\mathcal{N}))$ is a point process on $\mathbf{R}^2_+$, is the associated random measure $\widetilde{N}$ on $\mathcal{L}$ such that the distribution of N uniquely determines that of $\widetilde{N}$, and vice-versa?

Note that it makes no sense to talk about "the distribution" of $\widetilde{N}$ if the random map $\omega \rightsquigarrow \widetilde{N}(\omega)$ has not been properly defined. In this case, an appropriate image space of integer measures on $(\mathcal{L}, \mathcal{B}(\mathcal{L}))$ must be specified. To alleviate notation, define the $\mathcal{L}$-elements $\{(0,0)\}$ and $\bar{\mathbf{R}}^2_+$ as $0_{\mathcal{L}}$ and $\infty_{\mathcal{L}}$ respectively, and for any $L, L' \in \mathcal{L}$ let

$$[L, L'] := \{L'' \in \mathcal{L} : L \leq L'' \leq L'\}.$$

Definition 6.2.5 *Let $\mathcal{N}_{\mathcal{L}}$ denote the space of all point measures $\kappa := \sum_{i \in \mathbf{N}} \delta_{L_i}$ on $\mathcal{L}$ which satisfy*

i) $\kappa(0_{\mathcal{L}}) = 0$, and

ii) the sequence $\{L_i\}_{i \in \mathbf{N}}$ admits at most one accumulation point in $\mathcal{L}$, and if L is such an accumulation point, then $L_i << L$ for all $i \in \mathbf{N}$.

Let $\mathcal{F}(\mathcal{N}_{\mathcal{L}})$ denote the "canonical" σ-algebra on $\mathcal{N}_{\mathcal{L}}$ generated by all evaluation maps $\phi_L : \mathcal{N}_{\mathcal{L}} \longrightarrow \bar{\mathbf{R}}_+$, i.e. those that map the measure $\kappa \in \mathcal{N}_{\mathcal{L}}$ to the value $\phi_L(\kappa) = \kappa_L$. A point process M on $\mathcal{L}$ shall be understood as a random mapping $M : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_{\mathcal{L}}, \mathcal{F}(\mathcal{N}_{\mathcal{L}}))$.

Observe the mapping $\mu \rightsquigarrow \tilde{\mu}$ takes its images in $\mathcal{N}_{\mathcal{L}}$ for $\mu \in \mathcal{N}$. The question of its $\mathcal{F}(\mathcal{N})/\mathcal{F}(\mathcal{N}_{\mathcal{L}})$-measurability naturally arises, and will be settled forthwith:

Proposition 6.2.6 *The map* $\Phi : \begin{array}{ccc} (\mathcal{N}, \mathcal{F}(\mathcal{N})) & \longrightarrow & (\mathcal{N}_{\mathcal{L}}, \mathcal{F}(\mathcal{N}_{\mathcal{L}})) \\ \mu & \longmapsto & \tilde{\mu} \end{array}$ *is measurable.*

Proposition 6.2.6 is a consequence of the next proposition, which itself constitutes a generalization of Remark 6.2.4.

Proposition 6.2.7 *For any $\mu \in \mathcal{N}$, if $u_1, \ldots, u_n \in \mathbf{R}^2_+$, then*

$$\tilde{\mu}_{D_{u_1} \wedge \ldots \wedge D_{u_n}} = \inf_{1 \le k \le n} \mu_{u_k}.$$

Proof We first argue that

$$\left| \bigcap_{1 \le k \le n} \{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\} \right| = \inf_{1 \le k \le n} |\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\}|.$$

Let $k^* \in \{1, \ldots, n\}$ be such that

$$|\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_{k^*}}\}| = \inf_{1 \le k \le n} |\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\}|.$$

Since $\bigcap_{1 \le k \le n}\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\} \subseteq \{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_{k^*}}\}$, we certainly have that $|\bigcap_{1 \le k \le n}\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\}| \le |\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_{k^*}}\}|$. On the other hand, if, for some $i_0 \in \mathbf{N}$, $L_{i_0}^\mu \le D_{u_{k^*}}$, then $i_0 \le |\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_{k^*}}\}|$ $= \inf_{1 \le k \le n} |\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\}|$, the inequality holding by virtue of the fact that $L_1^\mu \le L_2^\mu \le \ldots$ Thus, for any $k \in \{1, \ldots, n\}$, $i_0 \le |\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\}|$, which implies $L_{i_0}^\mu \in \{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\}$ (because $L_1^\mu \le L_2^\mu \le \ldots$), i.e. $L_{i_0} \in \bigcap_{1 \le k \le n}\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\}$. We have therefore shown that $\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_{k^*}}\} \subseteq \bigcap_{1 \le k \le n}\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\}$, which entails $|\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_{k^*}}\}| \le |\bigcap_{1 \le k \le n}\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\}|$. Since the reverse inequality has already been established, we have that $|\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_{k^*}}\}|$ $= |\bigcap_{1 \le k \le n}\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\}|$, as claimed. To complete the proof observe: $\tilde{\mu}_{D_{u_1} \wedge \ldots \wedge D_{u_k}} = |\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_1} \wedge \ldots \wedge D_{u_n}\}| = |\bigcap_{1 \le k \le n}\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\}|$ $= \inf_{1 \le k \le n} |\{L_i^\mu : i \in \mathbf{N}, L_i^\mu \le D_{u_k}\}| = \inf_{1 \le k \le n} \tilde{\mu}_{D_{u_k}} = \inf_{1 \le k \le n} \mu_{u_k}$, the last equality following from Remark 6.2.4. □

We now proceed to the

Proof of Proposition 6.2.6 To prove that $\begin{array}{rccc}\Phi : & (\mathcal{N}, \mathcal{F}(\mathcal{N})) & \longrightarrow & (\mathcal{N}_{\mathcal{L}}, \mathcal{F}(\mathcal{N}_{\mathcal{L}})) \\ & \mu & \longmapsto & \tilde{\mu}\end{array}$ is measurable, it suffices to verify that for any $L \in \mathcal{L}$ and $k \in \mathbf{N}$, the set $\Phi^{-1}\{\kappa \in \mathcal{N}_{\mathcal{L}} : \kappa_L \ge k\} = \{\mu \in \mathcal{N} : \tilde{\mu}_L \ge k\}$ belongs to $\mathcal{F}(\mathcal{N})$. For any such L and k, if $L = \infty_{\mathcal{L}}$, this immediately follows from the equalities $\{\mu \in \mathcal{N} : \tilde{\mu}_L \ge k\}$

$= \bigcup_{n\in\mathbf{N}}\{\mu \in \mathcal{N} : \tilde{\mu}_{D_{(n,n)}} \geq k\} = \bigcup_{n\in\mathbf{N}}\{\mu \in \mathcal{N} : \mu_{(n,n)} \geq k\}$, guaranteed by Remark 6.2.4. Suppose, therefore, that $L \neq \infty_{\mathcal{L}}$. Observe that, by the way the elements of $\mathcal{N}_{\mathcal{L}}$ were specified, if $\{\widetilde{L}^{(n)}\}_{n\in\mathbf{N}}$ is a decreasing sequence in $\mathcal{L}$ which converges to L, then $\tilde{\mu}_{\widetilde{L}^{(n)}} \downarrow \tilde{\mu}_L$ as $n \to \infty$, which would imply

$$\{\mu \in \mathcal{N} : \tilde{\mu}_L \geq k\} = \bigcap_{n\in\mathbf{N}}\{\mu \in \mathcal{N} : \tilde{\mu}_{\widetilde{L}^{(n)}} \geq k\}.$$

Construct the sequence $\{\widetilde{L}^{(n)}\}_{n\in\mathbf{N}}$ as follows: let $\{z_m\}_{m\in\mathbf{N}}$ be an enumeration of $\mathbf{Q}_+^2 \setminus L$, and for any $n \in \mathbf{N}$ set $\widetilde{L}^{(n)} := D_{z_1} \wedge \ldots \wedge D_{z_n}$. Clearly, $\widetilde{L}^{(n)} \downarrow L$ as $n \to \infty$. Using Proposition 6.2.7, we thus have $\{\mu \in \mathcal{N} : \tilde{\mu}_L \geq k\} = \bigcap_{n\in\mathbf{N}}\{\mu \in \mathcal{N} : \tilde{\mu}_{\widetilde{L}^{(n)}} \geq k\}$ $= \bigcap_{n\in\mathbf{N}}\{\mu \in \mathcal{N} : \tilde{\mu}_{D_{z_1}\wedge\ldots\wedge D_{z_n}} \geq k\} = \bigcap_{n\in\mathbf{N}}\{\mu \in \mathcal{N} : \inf_{1\leq i\leq n}\mu_{z_i} \geq k\}$ $= \bigcap_{n\in\mathbf{N}}\bigcap_{1\leq i\leq n}\{\mu \in \mathcal{N} : \mu_{z_i} \geq k\} \in \mathcal{F}(\mathcal{N})$. □

At this point, the measurability of the map $\Phi : \begin{array}{ccc}(\mathcal{N}, \mathcal{F}(\mathcal{N})) & \longrightarrow & (\mathcal{N}_{\mathcal{L}}, \mathcal{F}(\mathcal{N}_{\mathcal{L}})) \\ \mu & \longmapsto & \tilde{\mu}\end{array}$ enables us to make the following remark, the first milestone of our program outlined in the chapter introduction:

Remark 6.2.8 *If $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{F}(\mathcal{N}))$ is a point process, then the distribution $P \circ N^{-1}$ of N on $\mathbf{R}_+^2$ uniquely determines the distribution $P \circ \widetilde{N}^{-1}$ of the associated point process $\widetilde{N} = \Phi(N)$ on $\mathcal{L}$.*

We dedicate the rest of this section to proving the "converse" of Remark 6.2.8, namely, that if $M : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_{\mathcal{L}}, \mathcal{F}(\mathcal{N}_{\mathcal{L}}))$ is a point process on $\mathcal{L}$ of which it is known that there exists a point process $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{F}(\mathcal{N}))$ on $\mathbf{R}_+^2$ such that $M = \widetilde{N}$, then the distribution $P \circ M^{-1}$ uniquely determines the distribution $P \circ N^{-1}$. This will be achieved by exhibiting the finite dimensional distributions of N on a determining class of cylinders as transformations of the finite dimensional distributions of $\widetilde{N}$.

The setup is as follows: for any $n \in \mathbf{N}$, we shall consider n^2 distinct elements $z_{(1,1)}, \ldots, z_{(1,n)}, z_{(2,1)}, \ldots z_{(2,n)}, \ldots, z_{(n,1)}, \ldots, z_{(n,n)}$ of $\mathbf{R}_+^2$ arranged in such a fashion that for any $i \in \{1, \ldots, n\}$, $z_{(i,1)}, \ldots z_{(i,n)}$ belongs to a common horizontal line (row), and for any $j \in \mathbf{N}$, $z_{(1,j)}, \ldots, z_{(n,j)}$ belongs to a common vertical line (column). We further require that if $1 \leq i < k \leq n$, points $z_{(i,1)}, \ldots z_{(i,n)}$ lie above points $z_{(k,1)}, \ldots z_{(k,n)}$, and points $z_{(1,i)}, \ldots, z_{(n,i)}$ lie to the left of $z_{(1,k)}, \ldots, z_{(n,k)}$. Also, for simplicity, assume $z_{(n,1)} = (0,0)$; this will imply that points $z_{(1,1)}, \ldots, z_{(n,1)}$ lie on the y-axis, and points $z_{(n,1)}, \ldots, z_{(n,n)}$ lie on the x-axis. For the sake of convenience, we shall call the set $\{z_{(i,j)}\}_{i,j=1}^{n}$ a *planar grid of size n*, or a *grid* for short.

For any set $A \subseteq \mathbf{R}_{+}^{2}$, agree to call an element $z \in A$ *minimal in A* if there exists no $z' \in A \setminus \{z\}$ such that $z' \leq z$. The following procedure is intended to produce an ordered list $U = \{u_1; u_2; \ldots; u_{n^2}\}$ of the grid $\{z_{(1,1)}, \ldots, z_{(n,n)}\}$, along with a corresponding list $\mathcal{L}^U := \{l_1; \ldots; l_{n^2}\}$ of lower layers placed in increasing order. The procedure itself is not deterministic (i.e. it can produce several versions of U and $\mathcal{L}^U$), but we shall later exhibit a concise characterization of all of its possible outputs.

PROCEDURE

Let $S := \{z_{(1,1)}, \ldots z_{(1,n)}, \ldots, z_{(n,1)}, \ldots z_{(n,n)}\}$.

STEP 1 Let $S_1 := \{z \in S : z \text{ is minimal in } S\}$ $\quad (= \{z_{(n,1)}\} = \{(0,0)\})$
Choose $u_1 \in S_1$ and set $l_1 := \bigwedge\{D_z : z \in S_1\}$ $\quad (= D_{z_{(n,1)}} = D_{(0,0)})$.

STEP 2: Let $S_2 := \{z \in S \setminus \{u_1\} : z \text{ is minimal in } S \setminus \{u_1\}\}$
Choose $u_2 \in S_2$ and set $l_2 := \bigwedge\{D_z : z \in S_2\}$.

⋮

STEP k: Let $S_k := \{z \in S \setminus \{u_1, \ldots, u_{k-1}\} : z \text{ is minimal in } S \setminus \{u_1, \ldots, u_{k-1}\}\}$
Choose $u_k \in S_k$ and set $l_k := \bigwedge\{D_z : z \in S_k\}$.

⋮

STEP n²: Let $S_{n^2} := \{z \in S \setminus \{u_1, \ldots, u_{n^2-1}\} : z \text{ is minimal in } S \setminus \{u_1, \ldots, u_{n^2-1}\}\}$ $\quad (= \{z_{(1,n)}\})$
Choose $u_{n^2} \in S_{n^2}$ and set $l_{n^2} := \bigwedge\{D_z : z \in S_k\}$ $\quad (= D_{z_{(1,n)}})$. □

Let us make three observations regarding the ordered sets U and $\mathcal{L}^U$ produced by this procedure.

Observation 6.2.9 *For any $i, j \in \{1, \ldots, n^2\}$, $i < j$ implies $u_j \not\leq u_i$.*

Proof Let $i, j \in \{1, \ldots, n^2\}$ with $i < j$. Let us first show that, for any $z \in S_j$ (see the procedure), either

i) $z \in S_i$, or

ii) $\exists\, v \in S_i$ such that $v \leq z$.

We prove this for $j = i + 1$, the case $j \geq i + 2$ following by induction. Accordingly, assuming $j = i+1$ $z \in S_j \backslash S_i = S_{i+1} \backslash S_i$, we must show $\exists\, v \in S_i$ such that $v \leq z$. Now, $z \in S_{i+1} \Longrightarrow z \in S \backslash \{u_1, \ldots, u_i\} \Longrightarrow z \in S \backslash \{u_1, \ldots, u_{i-1}\}$. If $\nexists\, \tilde{v} \in S \backslash \{u_1, \ldots, u_{i-1}\}$ such that $\tilde{v} \leq z$, then z is minimal in $S \setminus \{u_1, \ldots, u_{i-1}\}$, which implies $z \in S_i$ and thus contradicts the hypothesis. Therefore, $\exists\, \tilde{v} \in S \setminus \{u_1, \ldots, u_{i-1}\}$ such that $\tilde{v} \leq z$; since S_i consists of the minimal elements of $S \setminus \{u_1, \ldots, u_{i-1}\}$, by Zorn's lemma there must exist $v \leq \tilde{v}, v \in S_i$ such that $v \leq z$. Having validated (i) and (ii), we now prove the statement. If $u_j \in S_i$, then the fact that u_i is minimal in $S \setminus \{u_1, \ldots, u_{i-1}\}$ implies $u_j \not\leq u_i$ because $u_j \neq u_i$. If $u_j \notin S_i$, then, by the preliminary argument, $\exists\, v \in S_i$ such that $v \leq u_j$. But then, $u_j \leq u_i$ would imply $v \leq u_i$, which would imply, in turn, that $u_i = v$ owing to the minimality of $u_i \in S \setminus \{u_1, \ldots, u_{i-1}\}$, which would entail, at last, that $u_i = u_j$, a contradiction. Thus, $u_j \not\leq u_i$ if $u_j \notin S_i$. □

Observation 6.2.10 *For any $k \in \{1, \ldots, n^2\}$, $l_k = D_{u_k} \wedge D_{u_{k+1}} \wedge \ldots \wedge D_{u_{n^2}}$. In particular, $l_1 \leq l_2 \leq \ldots \leq l_{n^2}$.*

Proof For any such k the procedure yields $l_k = \bigwedge\{D_z : z \in S_k\}$. Now S_k is included in $S \setminus \{u_1, \ldots, u_{k-1}\} = \{u_k, u_{k+1}, \ldots, u_{n^2}\}$, so it is clear that $l_k \geq D_{u_k} \wedge D_{u_{k+1}} \wedge \ldots \wedge D_{u_{n^2}}$. To establish the reverse inequality, it suffices to prove $l_k \leq D_{u_i}$ for all $i \in \{k, k+1, \ldots, n^2\}$. According to the previous observation, for any such i either

i) $u_i \in S_k$, or

ii) $\exists\, v \in S_k$ such that $v \leq u_i$.

If $u_i \in S_k$, then $l_k = \bigwedge\{D_z : z \in S_k\} \leq D_{u_i}$; if $\exists\, v \in S_k$ such that $v \leq u_i$, then $l_k = \bigwedge\{D_z : z \in S_k\} \leq D_v \leq D_{u_i}$. In either case, $l_k \leq D_{u_i}$, whence $l_k \leq D_{u_k} \wedge D_{u_{k+1}} \wedge \ldots \wedge D_{u_{n^2}}$. □

The third observation is a converse of the first; from these two observations will emerge an elegant characterization of the ordered set U.

Observation 6.2.11 *Let $V = \{v_1; \ldots; v_{n^2}\}$ be an ordered enumeration of $S = \{z_{(1,1)}, \ldots, z_{(1,n)}, \ldots, z_{(n,1)}, \ldots, z_{(n,n)}\}$. If V is such that for any $i, j \in \{1, \ldots, n^2\}$, $i < j$ implies $v_j \not\leq v_i$, then the aforementioned procedure for choosing U could yield $U = \{u_1; \ldots; u_{n^2}\} = \{v_1; \ldots; v_{n^2}\} = V$.*

Proof The criterion $i < j \implies v_j \not\leq v_i$ ensures $v_1 = z_{(n,1)} = u_1$, and thus $\{v_1\}$ is a possible version of $\{u_1\}$. Let $k \in \{1, \ldots, n^2 - 1\}$ and suppose that for any $i \leq k$, $\{v_1; \ldots; v_i\}$ is a possible version of $\{u_1; \ldots; u_i\}$. To show that $\{v_1; \ldots; v_k; v_{k+1}\}$ is a possible version of $\{u_1; \ldots; u_k; u_{k+1}\}$, it suffices to prove that v_{k+1} is minimal in $S \setminus \{v_1, \ldots, v_k\} = S \setminus \{u_1, \ldots, u_k\}$. Suppose this is not the case; that is, $\exists\, v \in S \setminus \{v_1, \ldots, v_k\}$ such that $v \neq v_{k+1}$ and $v \leq v_{k+1}$. But then, $v \in \{v_{k+2}, \ldots, v_{n^2}\}$ and $v \leq v_{k+1}$, contradicting the hypothesis. In conclusion, v_{k+1} is minimal in $S \setminus \{v_1, \ldots, v_k\}$ and the induction argument is complete. □

Let us give a name to the ordered set U produced by the procedure, and let us state once and for all how its order is characterized:

Definition 6.2.12 *An ordered set $U = \{u_1; \ldots; u_{n^2}\}$ is called a labeling of the grid $\{z_{(1,1)}, \ldots, z_{(1,n)}, \ldots, z_{(n,1)}, \ldots, z_{(n,n)}\}$ (a labeling for short) if its elements are those of $\{z_{(1,1)}, \ldots, z_{(1,n)}, \ldots, z_{(n,1)}, \ldots, z_{(n,n)}\}$, and if, for any $i, j \in \{1, \ldots, n^2\}$ with $i < j$, $u_j \not\leq u_i$.*

Remark 6.2.13 *An ordered subset U of $\mathbf{R}_+^2$ is a labeling of $\{z_{(1,1)}, \ldots, z_{(1,n)}, \ldots, z_{(n,1)}, \ldots, z_{(n,n)}\}$ if and only if it constitutes a possible output of the previously stated procedure.*

That an ordered set $\mathcal{L}^U$ of lower layers is constructed in the same time as the labeling U by the procedure, is in no way a happenstance; these two sets are used to relate, in a very special case, the finite dimensional distributions of a point process on $\mathbf{R}_+^2$ to those of its associated point process on $\mathcal{L}$, as the following lemma demonstrates.

Lemma 6.2.14 *Let $U = \{u_1; \ldots; u_{n^2}\}$ be a labeling, and $\mathcal{L}^U = \{l_1; \ldots; l_{n^2}\}$ the ordered set of lower layers associated to U by the procedure. If $k_1, \ldots, k_{n^2}$ are integers such that $0 \le k_1 \le k_2 \le \ldots \le k_{n^2}$, we have that*

$$\{\mu \in \mathcal{N} : \mu_{u_1} \ge k_1, \mu_{u_2} \ge k_2, \ldots, \mu_{u_{n^2}} \ge k_{n^2}\}$$
$$= \{\mu \in \mathcal{N} : \tilde{\mu}_{l_1} \ge k_1, \tilde{\mu}_{l_2} \ge k_2, \ldots, \tilde{\mu}_{l_{n^2}} \ge k_{n^2}\}.$$

Proof By Proposition 6.2.7 and Observation 6.2.10, for any $i \in \{1, \ldots, n^2\}$ we have

$$\begin{aligned}
\{\mu \in \mathcal{N} : \tilde{\mu}_{l_i} \ge k_i\} &= \{\mu \in \mathcal{N} : \tilde{\mu}_{D_{u_i} \wedge \ldots \wedge D_{u_{n^2}}} \ge k_i\} \\
&= \{\mu \in \mathcal{N} : \inf_{i \le j \le n^2} \mu_{u_j} \ge k_i\} \\
&= \bigcap_{i \le j \le n^2} \{\mu \in \mathcal{N} : \mu_{u_j} \ge k_i\} \\
&= \{\mu \in \mathcal{N} : \mu_{u_i} \ge k_i\} \cap \bigcap_{i+1 \le j \le n^2} \{\mu \in \mathcal{N} : \mu_{u_j} \ge k_i\}.
\end{aligned}$$

However, since $k_1 \le k_2 \le \ldots \le k_{n^2}$, it results that

$$\begin{aligned}
&\{\mu \in \mathcal{N} : \tilde{\mu}_{l_1} \ge k_1, \tilde{\mu}_{l_2} \ge k_2, \ldots, \tilde{\mu}_{l_{n^2}} \ge k_{n^2}\} \\
&\quad = \bigcap_{1 \le i \le n^2} \{\mu \in \mathcal{N} : \tilde{\mu}_{l_i} \ge k_i\} \\
&\quad = \bigcap_{1 \le i \le n^2} \left[\{\mu \in \mathcal{N} : \mu_{u_i} \ge k_i\} \cap \bigcap_{i+1 \le j \le k^2} \{\mu \in \mathcal{N} : \mu_{u_j} \ge k_i\} \right] \\
&\quad = \bigcap_{1 \le i \le n^2} \{\mu \in \mathcal{N} : \mu_{u_i} \ge k_i\} \\
&\quad = \{\mu \in \mathcal{N} : \mu_{u_1} \ge k_1, \mu_{u_2} \ge k_2, \ldots, \mu_{u_{n^2}} \ge k_{n^2}\}. \square
\end{aligned}$$

It is not clear, at this point, whether the finite dimensional distributions of a point process N on $\mathbf{R}_+^2$ are entirely characterized by the measures of the reverse images of the sets expressed in the foregoing lemma. We now aim to prove that this is indeed the case. In order to do this, we must introduce a special kind of permutation on labelings, then to ask ourselves whether the resulting ordered sets are still labelings and, should they be, whether they comprise enough cases to characterize the finite dimensional distributions of N.

Definition 6.2.15 *Let $U = \{u_1; \dots; u_{n^2}\}$ be a labeling of the grid $\{z_{(i,j)}\}_{i,j=1}^{n}$. Suppose there exist $k_1, k_2 \in \{1, \dots, n^2\}$ such that $k_1 < k_2$ and $u_{k_1} \not\leq u_{k_2}$ (and $u_{k_2} \not\leq u_{k_1}$, necessarily). Let*

$$\begin{aligned} U^{k_1} &:= \{k \in \{k_1+1, \dots, k_2-1\} : u_{k_1} \leq u_k\}; \\ U_{k_2} &:= \{k \in \{k_1+1, \dots, k_2-1\} : u_k \leq u_{k_2}\}; \\ U_{k_1}^{k_2} &:= \{k \in \{k_1+1, \dots, k_2-1\} : u_{k_1} \not\leq u_k \text{ and } u_k \not\leq u_{k_2}\}. \end{aligned}$$

Let $\{l_1; \dots; l_{|U^{k_1}|}\}$ be an enumeration of the elements of U^{k_1} respecting their order of appearance in U; let $\{m_1; \dots; m_{|U_{k_2}|}\}$ and $\{n_1; \dots; n_{|U_{k_1}^{k_2}|}\}$ be analogous enumerations for U_{k_2} and $U_{k_1}^{k_2}$, respectively.

Define the permutation $\sigma_{k_1}^{k_2}$ on $\{1, \dots, n^2\}$ as follows:

- *if $k \in \{1, \dots, k_1 - 1\}$ set $\sigma_{k_1}^{k_2}(k) := k$;*
- *set* $$\begin{aligned} &\sigma_{k_1}^{k_2}(k_1) := m_1; \\ &\sigma_{k_1}^{k_2}(k_1+1) := m_2; \\ &\vdots \\ &\sigma_{k_1}^{k_2}(k_1+|U_{k_2}|-1) := m_{|U_{k_2}|}; \end{aligned}$$
- *set $\sigma_{k_1}^{k_2}(k_1+|U_{k_2}|) := k_2$;*
- *set* $$\begin{aligned} &\sigma_{k_1}^{k_2}(k_1+|U_{k_2}|+1) := n_1; \\ &\sigma_{k_1}^{k_2}(k_1+|U_{k_2}|+2) := n_2; \\ &\vdots \\ &\sigma_{k_1}^{k_2}(k_1+|U_{k_2}|+|U_{k_1}^{k_2}|) := n_{|U_{k_1}^{k_2}|}; \end{aligned}$$
- *set $\sigma_{k_1}^{k_2}(k_1+|U_{k_2}|+|U_{k_1}^{k_2}|+1) := k_1$;*

- *set* $\sigma_{k_1}^{k_2}(k_1 + |U_{k_2}| + |U_{k_1}^{k_2}| + 2) := l_1;$
  $\sigma_{k_1}^{k_2}(k_1 + |U_{k_2}| + |U_{k_1}^{k_2}| + 3) := l_2;$
  $\vdots$
  $\sigma_{k_1}^{k_2}(k_1 + |U_{k_2}| + |U_{k_1}^{k_2}| + |U^{k_1}| + 1) = \sigma_{k_1}^{k_2}(k_2) := l_{|U^{k_1}|};$
- *if* $k \in \{k_2+1, \ldots, n^2\}$ *set* $\sigma_{k_1}^{k_2}(k) := k.$

Lemma 6.2.16 *Let* $U = \{u_1; \ldots; u_{n^2}\}$ *be a labeling of* $\{z_{(i,j)}\}_{i,j=1}^{n}$. *If* $k_1, k_2 \in \{1, \ldots, n^2\}$ *such that* $k_1 < k_2$ *and* $u_{k_1} \not\leq u_{k_2}$ *(and* $u_{k_2} \not\leq u_{k_1}$*), then* $V := \{v_1; \ldots; v_{n^2}\}$ *defined by writing*

$$v_k := u_{\sigma_{k_1}^{k_2}(k)}, \qquad k = 1, 2, \ldots, n^2,$$

is another labeling of $\{z_{(i,j)}\}_{i,j=1}^{n}$.

Proof As they relate to U, define $U^{k_1}, U_{k_1}^{k_2}$ and U_{k_2}, as well as their respective enumerations, in the same fashion as in Definition 6.2.15. We now proceed to show that for any $k, k' \in \{1, \ldots, n^2\}$ such that $k < k'$, $v_{k'} \not\leq v_k$. Pick any such k and k'. The argument embraces a multitude of cases and subcases.

<u>For $k \in \{1, \ldots, k_1 - 1\}$:</u>

- if $k' \leq k_1 - 1$, then $v_{k'} = u_{k'} \not\leq u_k = v_k$ because U is a labeling;
- if $k' \geq k_1$, then $v_{k'} \not\leq v_k$ because $\{v_{k_1}, \ldots, v_{n^2}\} = \{u_{k_1}, \ldots, u_{k_{n^2}}\}$ (as unordered sets) and U is a labeling.

<u>For $k_1 \leq k \leq k_1 + |U_{k_2}| - 1$:</u>

- if $k' \leq k_1 + |U_{k_2}| - 1$, then $v_{k'} \not\leq v_k$ because $u_{m_1}, \ldots, u_{m_{|U_{k_2}|}}$ appear exactly in that order in U;
- if $k' = k_1 + |U_{k_2}|$, then $v_{k'} \leq v_k \Longleftrightarrow u_{k_2} \leq u_m$ for some $m \in U_{k_2}$; however, since $u_m \leq u_{k_2}$, this would imply $u_m = u_{k_2}$, contradicting the fact that the $z_{(i,j)}$'s are distinct;

- if $k_1 + |U_{k_2}| + 1 \leq k' \leq k_1 + |U_{k_2}| + |U_{k_1}^{k^2}|$, then $v_{k'} \leq v_k \Longleftrightarrow u_n \leq u_m$ for some $n \in U_{k_1}^{k_2}$, $m \in U_{k_2}$; but since $u_m \leq u_{k_2}$, this would imply $u_n \leq u_{k_2}$ by transitivity, which contradicts the definition of $U_{k_1}^{k_2}$;

- if $k' = k_1 + |U_{k_2}| + |U_{k_1}^{k^2}| + 1$, then $v_{k'} \leq v_k$ would imply $u_{k_1} \leq u_m$ for some $m \in U_{k_2}$; but since $u_m \leq u_{k_1}$, we would obtain the contradiction $u_m = u_{k_1}$;

- if $k_1 + |U_{k_2}| + |U_{k_1}^{k^2}| + 2 \leq k' \leq k_2$, then $v_{k'} \leq v_k$ would imply $u_l \leq u_m$ for some $l \in U^{k_1}$ and $m \in U_{k_2}$; but as $u_m \leq u_{k_2}$ and $u_{k_1} \leq u_l$, this would entail the contradiction $u_{k_1} \leq u_{k_2}$;

- if $k' \geq k_2 + 1$, then $v_{k'} \leq v_k$ would imply $u_{k'} \leq u_m$ for some $m \in U_{k_2}$, which contradicts $m < k'$ in this case.

For $k = k_1 + |U_{k_2}|$:

- if $k_1+|U_{k_2}|+1 \leq k' \leq k_1+|U_{k_2}|+|U_{k_1}^{k_2}|$, then $v_{k'} \not\leq v_k$ because $u_n \not\leq u_{k_2}\ \forall\, n \in U_{k_1}^{k_2}$;

- if $k' = k_1 + |U_{k_2}| + |U_{k_1}^{k_2}| + 1$, then $v_{k'} \not\leq v_k$ because $u_{k_1} \not\leq u_{k_2}$;

- if $k_1 + |U_{k_2}| + |U_{k_1}^{k_2}| + 2 \leq k' \leq k_2$, then $v_{k'} \not\leq v_k$ because $u_l \not\leq u_{k_2}\ \forall\, l \in U^{k_1}$ (the opposite would imply $u_{k_1} \leq u_{k_2}$ by transitivity);

- if $k' \geq k_2 + 1$, then $v_{k'} \leq v_k$ would imply $u_{k'} \leq u_{k_2}$, which cannot hold because $k_2 < k'$ in this case.

For $k_1 + |U_{k_2}| + 1 \leq k \leq k_1 + |U_{k_2}| + |U_{k_1}^{k_2}|$:

- if $k' \leq k_1 + |U_{k_2}| + |U_{k_1}^{k_2}|$, then $v_{k'} \not\leq v_k$ because $u_{n_1}, \ldots, u_{n_{|U_{k_1}^{k_2}|}}$ appear exactly in that order in $U_{k_1}^{k_2}$;

- if $k' = k_1 + |U_{k_2}| + |U_{k_1}^{k_2}| + 1$, then $v_{k'} \leq v_k$ would imply $u_{k_1} \leq u_n$ for some $n \in U_{k_1}^{k_2}$, which is a contradiction;

- if $k_1 + |U_{k_2}| + |U_{k_1}^{k_2}| + 2 \leq k' \leq k_2$, then $v_{k'} \leq v_k$ cannot hold because this would imply $u_{k_1} \leq u_l \leq u_n$ for some $n \in U_{k_1}^{k_2}$ and $l \in U^{k_1}$ ($u_{k_1} \not\leq u_n\ \forall\, n \in U_{k_1}^{k_2}$);

- if $k' \geq k_2 + 1$, then $v_{k'} \leq v_k$ would imply $u_{k'} \leq u_n$ for some $n \in U_{k_1}^{k_2}$, which would contradict $n < k'$ in this case.

For $k = k_1 + |U_{k_2}| + |U_{k_1}^{k_2}| + 1$:

- if $k_1 + |U_{k_2}| + |U_{k_1}^{k_2}| + 2 \leq k' \leq k_2$, then $v_{k'} \leq v_k$ would imply $u_l \leq u_{k_1}$ for some $l \in U^{k_1}$; but since $u_{k_1} \leq u_l$, the contradiction $u_l = u_{k_1}$ would ensue in this case, as $l \neq k_1$;

- if $k' \geq k_2 + 1$, then $v_{k'} \leq v_k$ would imply $u_{k'} \leq u_{k_1}$, contradicting $k_1 < k'$ in this case.

For $k_1 + |U_{k_2}| + |U_{k_1}^{k_2}| + 2 \leq k \leq k_2$:

- if $k' \leq k_2$, then $v_{k'} \not\leq v_k$ because $l_1, l_2, \ldots, l_{|U^{k_1}|}$ are listed exaclty in this order in U;

- if $k' \geq k_2 + 1$, then $v_{k'} \leq v_k$ would imply $u_{k'} \leq u_l$ for some $l \in U^{k_1}$, but this cannot happen because $l < k'$ in this case.

As all cases have been exhausted, we conclude that $v_{k'} \not\leq v_k$ if $k > k'$, and hence, that $V = \{v_1; \ldots; v_{n^2}\}$ is a *bona fide* labeling of $\{z_{(i,j)}\}_{i,j=1}^n$. □

Innocuous as it may appear, the following theorem nonetheless constitutes the pivot on which much of this chapter hinges. It does suggest the manner by which labelings will be used in characterizing the finite dimensional distributions of a point process on $\mathbf{R}_+^2$ - although this particular point will be explicitly addressed in the subsequent corollary.

Theorem 6.2.17 *Let $U := \{u_1; \ldots; u_{n^2}\}$ be a labeling of $\{z_{(i,j)}\}_{i,j=1}^n$ and let $\vec{k} := (k_1, \ldots, k_{n^2})$ be a vector of nonnegative integers such that for all $i, j \in \{1, \ldots, n^2\}$, if $u_i \leq u_j$ then $k_i \leq k_j$. There exists a permutation σ on $\{1, \ldots, n^2\}$ such that $V := \{u_{\sigma(1)}; \ldots; u_{\sigma(n^2)}\}$ is a labeling of $\{z_{(i,j)}\}_{i,j=1}^n$ and such that, in addition, $0 \leq k_{\sigma(1)} \leq \ldots \leq k_{\sigma(n^2)}$.*

Proof With U and $\vec{k}$ as given, let us first acknowledge the following fact: if $i < j$ and $k_i > k_j$, then $u_i \not\leq u_j$ and $u_j \not\leq u_i$. Indeed, if $i < j$, then $u_j \not\leq u_i$ by the fact that U is a labeling. Moreover, if $u_i \leq u_j$ were to hold, the hypothesis that $k_i \leq k_j$ would be contradicted.

Given any labeling $W := \{w_1; \ldots; w_{n^2}\}$ and any vector $\vec{h} := (h_1, \ldots, h_{n^2})$ of corresponding integers, define the procedure SWAP($W, \vec{h}$) as follows (this procedure is expressed in pseudo-code; W and $\vec{h}$ are accepted as input and are returned, possibly modified, as output):

Procedure SWAP($W, \vec{h}$)
begin %{procedure}%
    $j := 0$;
    while ($j \leq n^2$ and $\nexists\ i, j \in \{1, \ldots, n^2\}$ such that $i < j$ and $l_i > l_j$) do
        $j := j + 1$;
    if ($j \leq n^2$) then
        begin %{begin if}%
        choose $i \in \{1, \ldots, j-1\}$ such that $l_i > l_j$;
        $w_1 := w_{\sigma_i^j(1)}; \ldots; w_{n^2} := w_{\sigma_i^j(n^2)}$;
        $h_1 := h_{\sigma_i^j(1)}; \ldots; h_{n^2} := h_{\sigma_i^j(n^2)}$;
        end; %{end if}%
end %{procedure}%

In view of Lemma 6.2.16 and the observation we made at the beginning of this proof, it is clear the procedure SWAP($W, \vec{h}$) returns a *bona fide* labeling of $\{z_{(i,j)}\}_{i,j=1}^n$, along with a vector whose entries are a permutation of the procedure's input vector's entries. Starting with labeling U and integer vector $\vec{k}$, we create the following sequence of labelings and corresponding vectors:

$(U^{[1]}, \vec{k}^{[1]}) :=$ SWAP($U, \vec{k}$);
$(U^{[2]}, \vec{k}^{[2]}) :=$ SWAP($U^{[1]}, \vec{k}^{[1]}$);

$(U^{[3]}, \vec{k}^{[3]}) := \text{SWAP}(U^{[2]}, \vec{k}^{[2]});$

$\vdots$

$(U^{[m+1]}, \vec{k}^{[m+1]}) := \text{SWAP}(U^{[m]}, \vec{k}^{[m]});$

$\vdots$

Claim: There exists $M \in \mathbf{N}$ such that $(U^{[m]}, \vec{k}^{[m]}) = (U^{[M]}, \vec{k}^{[M]})$; for all $m \geq M$ (in other words, the procedure stops modifying its input after a finite number of iterations).

Proof of Claim: Recall the lexicographic order $\leq_l$ on $\mathbf{R}^{n^2}$: if $\vec{x} = (x_1, \dots, x_{n^2})$ and $\vec{y} = (y_1, \dots, y_{n^2})$ are vectors of $\mathbf{R}^{n^2}$we declare $\vec{x} \leq_l \vec{y}$ and $\vec{y} \leq_l \vec{x}$ if $\vec{x} = \vec{y}$; otherwise we take $i^* := \inf\{i \in \{1, \dots, n^2\} : x_i \neq y_i\}$ and write $\vec{x} \leq_l \vec{y}$ if $x_{i^*} < y_{i^*}$, $\vec{y} \leq_l \vec{x}$ if $y_{i^*} < x_{i^*}$. Clearly, $\leq_l$ is a total order, and thus a transitive and antisymmetric relation. We argue that

$$\dots \leq_l \vec{k}^{[2]} \leq_l \vec{k}^{[1]} \leq_l \vec{k}.$$

Let $m \in \mathbf{N}$, and suppose, without loss of generality, that $\vec{k}^{[m]} \neq \vec{k}^{[m+1]}$. Writing $\vec{k}^{[m]} = (k_1^{[m]}, \dots, k_{n^2}^{[m]}), \vec{k}^{[m+1]} = (k_1^{[m+1]}, \dots, k_{n^2}^{[m+1]}), U^{[m]} = (u_1^{[m]}; \dots; u_{n^2}^{[m]})$ and $U^{[m+1]} = (u_1^{[m+1]}; \dots; u_{n^2}^{[m+1]})$, there thus exists $i, j \in \{1, \dots, n^2\}$ such that $i < j$, $k_i^{[m]} > k_j^{[m]}$ and $k_1^{[m+1]} = k_{\sigma_i^j(1)}^{[m+1]}, \dots k_{n^2}^{[m+1]} = k_{\sigma_i^j(n^2)}^{[m+1]}$. By the way σ_i^j is defined, it is clear that $k_1^{[m+1]} = k_1^{[m]}, \dots, k_{i-1}^{[m+1]} = k_{i-1}^{[m]}$. We claim $k_i^{[m+1]} < k_i^{[m]}$. Indeed, by the way σ_i^j is defined, either

i) $k_i^{[m+1]} = k_j^{[m]} < k_i^{[m]}$, or

ii) there exists $l \in \{i+1, \dots, j-1\}$ such that $u_l^{[m]} \leq u_j^{[m]}$ and $\sigma_i^j(i) = l$; but then, under the current hypothesis, $k_i^{[m+1]} = k_l^{[m]} \leq k_j^{[m]} < k_i^{[m]}$.

We conclude that $\vec{k}^{[m+1]} \leq_l \vec{k}^{[m]}$, whether $\vec{k}^{[m]} = \vec{k}^{[m+1]}$ or not. If there exists no $M \in \mathbf{N}$ such that $(U^{[m]}, \vec{k}^{[m]}) = (U^{[M]}, \vec{k}^{[M]})$ for all $m \geq M$, then the $\leq_l$-decreasing sequence $\{\vec{k}^{[m]}\}_{m=1}^{\infty}$ must have infinitely many terms, since $\leq_l$ is transitive and antisymmetric. But this contradicts the fact that there are only finitely many permutations of $\{1, \dots, n^2\}$; such an $M \in \mathbf{N}$ therefore exists. //

Let $M \in \mathbf{N}$ be such that $(U^{[m]}, \vec{k}^{[m]}) = (U^{[M]}, \vec{k}^{[M]})$; for all $m \geq M$. The M iterations of the procedure SWAP$(W, \vec{h})$ guarantee that for any $m \geq M$, $0 \leq k_1^{[m]} \leq k_2^{[m]} \leq \ldots \leq k_{n^2}^{[m]}$; moreover, $U^{[m]}$ remains a labeling of $\{z_{(i,j)}\}_{i,j=1}^n$ by virtue of Lemma 6.2.16 acting in conjunction with the hypothesis. The permutation σ therefore exists. □

The next result shall conclude this section; it establishes the long anticipated converse of Remark 6.2.8, namely, the determination of $P \circ N^{-1}$ by $P \circ \tilde{N}^{-1}$.

Theorem 6.2.18 *Let $M : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}_{\mathcal{L}}, \mathcal{F}(\mathcal{N}_{\mathcal{L}}))$ be a point process on $\mathcal{L}$ of which it is known that there exists a point process $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{F}(\mathcal{N}))$ such that $\Phi(N) = \tilde{N} = M$. The distribution $P \circ M^{-1}$ uniquely determines the distribution $P \circ N^{-1}$. In particular, $P \circ N^{-1}$ is uniquely determined by the distributions of random vectors of the form $(\tilde{N}_{l_1}, \tilde{N}_{l_2}, \ldots, \tilde{N}_{l_{n^2}})$, where, for some grid $\{z_{(i,j)}\}_{i,j=1}^n$ and labeling $U := \{u_1; \ldots; u_{n^2}\}$ of $\{z_{(i,j)}\}_{i,j=1}^n$, $l_1 = D_{u_1} \wedge \ldots \wedge D_{u_{n^2}}$, $l_2 = D_{u_2} \wedge \ldots \wedge D_{u_{n^2}}, \ldots, l_{n^2} = D_{u_{n^2}} = D_{z_{(1,n)}}$.*

Proof The σ-field $\mathcal{F}(N)$ is such that $\sigma(N)$ is generated by random vectors of the form $(N_{z_{(1,1)}}, \ldots, N_{z_{(1,n)}}, \ldots, N_{z_{(n,1)}}, \ldots, N_{z_{(n,n)}})$, where $\{z_{(i,j)}\}_{i,j=1}^n$ is a planar grid of size n. Hence, the π-system which consists of the $\mathcal{F}$-sets $[N_{z_{(1,1)}} \geq k_{(1,1)}, \ldots, N_{z_{(1,n)}} \geq k_{(1,n)}, \ldots, N_{z_{(n,1)}} \geq k_{(n,1)}, \ldots, N_{z_{(n,n)}} \geq k_{(n,n)}]$ (where $k_{(1,1)}, \ldots k_{(n,n)} \in \mathbf{N} \cup \{0\}$) forms a determining class for $P \circ N^{-1}$. For such sets, without loss of generality it may be assumed that for any $i, i', j, j' \in \{1, \ldots, n\}$ such that $z_{(i,j)} \leq z_{(i',j')}$, we have $k_{(i,j)} \leq k_{(i',j')}$. Since a labeling of the grid $\{z_{(i,j)}\}_{i,j=1}^n$ always exists (for example, one may enumerate the $z_{(i,j)}$'s for left to right and from bottom to top), by Theorem 6.2.17 there exists a labeling $U = \{u_1; \ldots; u_{n^2}\}$ and an ordered enumeration $\{k_1; \ldots; k_{n^2}\}$ of $\{k_{(i,j)}\}_{i,j=1}^n$ such that

$$[N_{z_{(1,1)}} \geq k_{(1,1)}, \ldots, N_{z_{(1,n)}} \geq k_{(1,n)}, \ldots, N_{z_{(n,1)}} \geq k_{(n,1)}, \ldots, N_{z_{(n,n)}} \geq k_{(n,n)}]$$
$$= [N_{u_1} \geq k_1, \ldots, N_{u_{n^2}} \geq k_{n^2}]$$

and such that, for any $i, j \in \{1, \ldots n\}$ and $l \in \{1, \ldots, n^2\}$, $z_{(i,j)} = u_l$ implies $k_{(i,j)} = k_l$. In particular, for any $l_1, l_2 \in \{1, \ldots, n^2\}$, $l_1 < l_2$ implies $k_1 \leq k_2$. But then, by Lemma

6.2.14,

$$[N_{u_1} \geq k_1, \ldots, N_{u_{n^2}} \geq k_{n^2}] = [\widetilde{N}_{\bigwedge_{1\leq i\leq n^2} D_{u_i}} \geq k_1, \widetilde{N}_{\bigwedge_{2\leq i\leq n^2} D_{u_i}} \geq k_2, \ldots, \widetilde{N}_{D_{u_{n^2}}} \geq k_{n^2}].$$

As

$$P[N_{u_1} \geq k_1, \ldots, N_{u_{n^2}} \geq k_{n^2}] = P[\widetilde{N}_{\bigwedge_{1\leq i\leq n^2} D_{u_i}} \geq k_1, \widetilde{N}_{\bigwedge_{2\leq i\leq n^2} D_{u_i}} \geq k_2, \ldots, \widetilde{N}_{D_{u_{n^2}}} \geq k_{n^2}]$$

is determined by the distribution $P \circ M^{-1} = P \circ \widetilde{N}^{-1}$, it is clear the latter determines $P \circ N^{-1}$ as well. □

6.3 Flow Processes and Flow Compensators

In this section, we first define a flow on $\mathcal{L}$. We then proceed to reveal how a planar point process gives rise to a family of point processes on $\mathbf{R}_+$ which is indexed by all the flows on $\mathcal{L}$; and then, how the distributions of such processes determine that of the planar measure. At last, we provide a sufficient condition on the planar measure for the *compensators* of the flow processes to determine the planar measure's distribution.

Definition 6.3.1 *A flow on $\mathcal{L}$ (a flow for short) is an injective mapping $f : \bar{\mathbf{R}}_+ \longrightarrow \mathcal{L}$ which satisfies the following three conditions:*

i) $f(0) = \{(0,0)\}$ and $f(\infty) = \bar{\mathbf{R}}_+^2$;

ii) for any $s, t \in \bar{\mathbf{R}}_+$, $s \leq t \Longrightarrow f(s) \leq f(t)$;

iii) f is right-continuous, i.e. if $\{t_n\}_{n=1}^{\infty}$ is a decreasing sequence in $\bar{\mathbf{R}}_+$ with infimum t, then $\{f(t_n)\}_{n=1}^{\infty}$ is decreasing in $\mathcal{L}$ and $\bigwedge_{n\in\mathbf{N}} f(t_n) = f(t)$.

For any $L_1, \ldots, L_n \in \mathcal{L}$, a flow f is said to pass through $L_1, \ldots L_n$ if $L_1, \ldots, L_n$ all belong to the range of f.

The following proposition is used in proving Theorem 6.3.5. It also illustrates how common flows are.

Proposition 6.3.2 *For any $L_1, \dots L_n \in \mathcal{L}$ which satisfy $L_1 \leq L_2 \leq \dots \leq L_n$, there exists a flow f passing through $L_1, \dots, L_n$.*

Proof As $f(0) = \{(0,0)\}$ and $f(\infty) = \bar{\mathbf{R}}_+^2$ must hold by definition, one may assume, without loss of generality, that the (distinct) lower layers $L_1 \dots, L_n$ are all different from $\{(0,0)\}$ or $\bar{\mathbf{R}}_+^2$. For any $m \in \mathbf{N}_0 := \mathbf{N} \cup \{0\}$ let $D_m := \{\frac{i}{2^m} : i \in \mathbf{N}_0\}$ and let $D := \bigcup_{m \in \mathbf{N}_0} D_m$ be the set of nonnegative dyadics. As an intermediate step we shall define an injective and increasing map $\tilde{f} : D \longrightarrow \mathcal{L}$ by defining its images $\tilde{f}(D_0)$, $\tilde{f}(D_1)$, $\tilde{f}(D_2), \dots$ successively. First, as regards $\tilde{f}(D_0)$, let

$$\tilde{f}\left(\frac{0}{2^0}\right) = \tilde{f}(0) := \{(0,0)\} =: L_0,$$

$$\tilde{f}\left(\frac{1}{2^0}\right) = \tilde{f}(1) := L_1, \tilde{f}(2) := L_2, \dots, \tilde{f}(n) := L_n$$

and for $j \geq n+1$ define $\tilde{f}(\frac{j}{2^0}) = \tilde{f}(j)$, recursively, as some $L_j \in \mathcal{L}$ which satisfies $L_{j-1} << L_j \neq \bar{\mathbf{R}}_+^2$ (such an L_j exists by Proposition 6.1.4). Now $\tilde{f}(D_0)$ is determined. For any $m \geq 1$ determine the image set $\tilde{f}(D_m)$ recursively as follows. For any $\frac{i}{2^m} \in D_m$, if $\frac{i}{2^m} \in D_{m-1}$ then $\tilde{f}(\frac{i}{2^m})$ is already defined; if $\frac{i}{2^m} \in D_m \setminus D_{m-1}$, then set

$$f\left(\frac{i}{2^m}\right) := L_{\frac{i}{2^m}},$$

where $L_{\frac{i}{2^m}} \in \mathcal{L}$ is such that $L_{\frac{i-1}{2^m}} \leq L_{\frac{i}{2^m}} \leq L_{\frac{i+1}{2^m}}$, $\rho(L_{\frac{i-1}{2^m}}, L_{\frac{i}{2^m}}) < \frac{1}{2^m}$ and $L_{\frac{i-1}{2^m}} \neq L_{\frac{i}{2^m}} \neq L_{\frac{i+1}{2^m}}$ (observe that $\frac{i-1}{2^m}, \frac{i+1}{2^m} \in D_{m-1}$ if $\frac{i}{2^m} \in D_m \setminus D_{m-1}$, and that such an $L_{\frac{i}{2^m}}$ exists by Proposition 6.1.5). Now $\tilde{f} : D \longrightarrow \mathcal{L}$ is completely determined as an increasing and injective map. We exhibit the requisite $f : \bar{\mathbf{R}}_+ \longrightarrow \bar{\mathbf{R}}_+^2$ by defining

$$f(x) := \begin{cases} \{(0,0)\} & \text{if } x = 0 \\ \bar{\mathbf{R}}_+^2 & \text{if } x = \infty \\ \bigwedge\{\tilde{f}(q) : q \in D, q > x\} & \text{if } 0 < x < \infty \end{cases} .$$

Requirements (i) and (ii) of the definition of flow are clearly satisfied by f. This f is also injective as a consequence of the density of D in $\mathbf{R}_+$ and of the antisymmetry of the order $\leq$ on $\mathcal{L}$. Requirement (iii) of the definition is met by the easily verifiable fact that lattice infima of countable subsets of $\mathcal{L}$ are well-defined. Thus f is a flow. Moreover, $f(i) = L_i$ for any $i \in \{1, \dots, n\}$, so that f indeed passes through $L_1, \dots, L_n$. □

We may now define the $\mathbf{R}_+$-indexed point process N^f associated to a given flow f.

Definition 6.3.3 *Let $f : \bar{\mathbf{R}}_+ \longrightarrow \mathcal{L}$ be a flow and $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, (\mathcal{N}))$ a point process on $\mathbf{R}_+^2$. The flow point process N^f on $\mathbf{R}_+$ is defined by writing*

$$N_t^f := \tilde{N}_{f(t)}$$

for all $t \in \mathbf{R}_+$.

Observe that N^f is not a simple point process in general.

Definition 6.3.4 *Given any planar grid $\{z_{(i,j)}\}_{i,j=1}^n$ and labeling $U := \{u_1; \dots; u_{n^2}\}$ of $\{z_{(i,j)}\}_{i,j=1}^n$, a flow $f : \bar{\mathbf{R}}_+ \longrightarrow \mathcal{L}$ is said to exhaust $\{z_{(i,j)}\}_{i,j=1}^n$ along U if f passes through $l_1 := D_{u_1} \wedge \dots \wedge D_{u_{n^2}}$, $l_2 := D_{u_2} \wedge \dots \wedge D_{u_{n^2}}, \dots, l_{n^2} := D_{u_{n^2}} = D_{z_{(1,n)}}$. If the labeling U is not specified, f is just said to exhaust the grid $\{z_{(i,j)}\}_{i,j=1}^n$.*

We shall henceforth denote by $\mathbf{F}$ the family of flows which exhaust non-trivial grids:

$$\mathbf{F} := \{f : \bar{\mathbf{R}}_+ \longrightarrow \mathcal{L} \mid f \text{ is a flow; } f \text{ exhausts a planar grid of size } n > 1.\}$$

This section's first result relates the distribution of a planar measure to those of its associated flow processes. It constitutes the first of this section's two principal results.

Theorem 6.3.5 *Let $N : (\Omega, ,P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ be a point process on $\mathbf{R}_+^2$. The law $P \circ N^{-1}$ uniquely determines, and is uniquely determined by the family $\{P \circ N^{f^{-1}}\}_{f \in \mathbf{F}}$ of flow point processes.*

Proof That $P \circ N^{-1}$ determines $P \circ N^{f^{-1}}$ for any flow f results from the fact that $P \circ N^{-1}$ determines $P \circ \widetilde{N}^{-1}$. Conversely, by Theorem 6.2.18, $P \circ N^{-1}$ is uniquely determined by distributions of random vectors of the form $(\widetilde{N}_{l_1}, \widetilde{N}_{l_2}, \dots, \widetilde{N}_{l_{n^2}})$, where, for some grid $\{z_{(i,j)}\}_{i,j=1}^n$ and labeling $U := \{u_1; \dots; u_{n^2}\}$ of $\{z_{(i,j)}\}_{i,j=1}^n$, $l_1 = D_{u_1} \wedge \dots \wedge D_{u_{n^2}}$, $l_2 = D_{u_2} \wedge \dots \wedge D_{u_{n^2}}, \dots, l_{n^2} = D_{u_{n^2}} = D_{z_{(1,n)}}$ and where, without loss of generality, it may be assumed that $n > 1$. But since $l_1 \leq l_2 \leq \ldots l_{n^2}$, there exists, by Proposition 6.3.2, a flow $f : \bar{\mathbf{R}}_+ \longrightarrow \mathcal{L}$ passing through $l_1, \ldots, l_{n^2}$. Such a flow belongs to $\mathbf{F}$ by definition. Moreover, writing $s_1 := f^{-1}(l_1)$, $\ldots, s_{n^2} := f^{-1}(l_{n^2})$ one immediately realizes that

$$(\widetilde{N}_{l_1}, \widetilde{N}_{l_2}, \ldots, \widetilde{N}_{l_{n^2}}) = (N^f_{s_1}, N^f_{s_2}, \ldots, N^f_{s_{n^2}}).$$

As this random vector's distribution is uniquely determined by that of N^f, it is clear that the family $\{P \circ N^{f^{-1}}\}_{f \in \mathbf{F}}$ determines $P \circ N^{-1}$. □

While it comes as no surprise that the distribution of the planar measure determines the distributions of the flow measures and vice-versa, a less trivial question arises as to whether the distribution of the planar measure is determined by the *compensators* of the flow processes. Let us elaborate: if $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ is a point process on $\mathbf{R}^2_+$ and $f : \bar{\mathbf{R}}_+ \longrightarrow \mathcal{L}$ is a flow, let the filtration $\{\mathcal{F}^f_t\}_{t \geq 0}$ be the internal history of the flow process N^f on $\mathbf{R}_+$, and let $\{\Lambda^f\}_{t \geq 0}$ be the dual predictable projection of N^f with respect to $\{\mathcal{F}^f_t\}_{t \geq 0}$. Retaining the symbol $\mathbf{F}$ to denote the class of all flows exhausting grids of size greater than one, is $P \circ N^{-1}$ determined by the family $\left\{\{\Lambda^f_t\}_{t \geq 0}\right\}_{f \in \mathbf{F}}$?

We know from Jacod ([23]: Theorem 3.4) that if $\{N'\}_{t \geq 0}$ is a *simple* point process on $\mathbf{R}_+$ generating an internal history $\{\mathcal{F}'_t\}$, then the law of N' is uniquely determined by the compensator $\{\Lambda'_t\}_{t \geq 0}$ of N' with respect to $\{\mathcal{F}'_t\}$. So it is reasonable to ask whether for any flow $f : \bar{\mathbf{R}}_+ \longrightarrow \mathcal{L}$, the compensator Λ^f determines the law of N^f (and this would be sufficient, by the previous theorem, to ensure that $P \circ N^{-1}$ be determined by the compensators $\{\Lambda^f_t\}_{t \geq 0}$). The difficulty stems from the fact that N^f is *not* a simple point process in general; as is well known, a point process admitting multiple jumps on a set of nonzero measure cannot be determined by its compensator.

In the following paragraphs we proceed to a condition on the planar point process N sufficient to ensure that, for a certain class of flows f large enough to determine $P \circ N^{-1}$ from the distributions $P \circ N^{f^{-1}}$, the point process N^f admits simple realizations on a set of measure 1. Any N satisfying this condition shall be called *strictly simple.* Strictly simple planar processes have surfaced in connection with the study of "planar compensators" (see, for example, Lemma B.1 of [20]); it is therefore pleasant to acknowledge their usefulness in a new approach to compensators of processes indexed by more general sets then $\mathbf{R}_+$.

Definition 6.3.6 *A planar measure $\mu \in (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ is said to be strictly simple if, for any $x \in \mathbf{R}_+$, $\mu(\{(x,y) : y \in \mathbf{R}_+\}) \leq 1$ and $\mu(\{(y,x) : y \in \mathbf{R}_+\}) \leq 1$. A point process $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ is said to be strictly simple if $N(\omega)$ is strictly simple for P-almost all $\omega \in \Omega$.*

In other words, a point process is N strictly simple if, with probability one, N does not charge any vertical or horizontal line with a measure greater than one. It is now time to grasp the nettle: the determination of $P \circ N^{-1}$ by the class of compensators $\{\Lambda^f\}_{f \in \mathbf{F}}$ will emerge from the following

Theorem 6.3.7 *Let $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ be a strictly simple point process on $\mathbf{R}^2_+$. For any grid $\{z_{(i,j)}\}_{i,j=1}^n$ and labeling $U := \{u_1; \ldots; u_{n^2}\}$ of $\{z_{(i,j)}\}_{i,j=1}^n$, there exists a flow $f_U : \bar{\mathbf{R}}_+ \longrightarrow (\mathcal{L}, \mathcal{B}(\mathcal{L}))$ which exhausts $\{z_{(i,j)}\}_{i,j=1}^n$ along U, and such that the point process N^{f_U} on $\mathbf{R}_+$ is simple.*

Proof Let N, $\{z_{(i,j)}\}_{i,j=1}^n$, and U be as specified. We shall construct f_U so as to satisfy the statement. By definition we must have $f_U(0) := D_{(0,0)} = D_{u_1}$ and $f_U(\infty) := \mathbf{R}^2_+$. Defining $f_U(t)$ for other values of t will require a relabeling of $\{z_{(i,j)}\}_{i,j=1}^n$. For any $k \in \{1, \ldots, n\}$, agree to call x_k the common abscissa of points $z_{(1,k)}, z_{(2,k)}, \ldots z_{(n,k)}$, and to call y_k the common ordinate of points $z_{(n-k+1,1)}, z_{(n-k+1,2)}, \ldots, z_{(n-k+1,n)}$ (thus, $z_{(n,1)} = (0,0) = (x_1, y_1)$, $z_{(1,n)} = (x_n, y_n)$, etc.). Furthermore, for any $i \in \{1, \ldots, n^2\}$ let $\sigma_i, \tau_i \in \{1, \ldots, n\}$ be such that $u_i = (x_{\sigma_i}, y_{\tau_i})$. If $t > n^2 - 1$, simply let

$$f_U(t) := D_{(x_n, y_n + t - (n^2 - 1))}.$$

At last, if there exists $i \in \{1, \ldots, n^2 - 1\}$ such that $i - 1 < t \leq i$, then define a_t and b_t as

$$a_t := \begin{cases} x_{\sigma_i} & \text{if } \sigma_i = n \\ x_{\sigma_i} + [t - (i-1)](x_{\sigma_{i+1}} - x_{\sigma_i}) & \text{otherwise} \end{cases}$$

and

$$b_t := \begin{cases} y_{\tau_i} & \text{if } \tau_i = n \\ y_{\tau_i} + [t - (i-1)](y_{\tau_{i+1}} - x_{\tau_i}) & \text{otherwise} \end{cases}$$

respectively, and put

$$f_U(t) := \begin{cases} D_{u_{i+1}} \wedge \ldots \wedge D_{u_{n^2}} \wedge D_{(x_{\sigma_i}, b_t)} & \text{if } \sigma_i = n \\ D_{u_{i+1}} \wedge \ldots \wedge D_{u_{n^2}} \wedge D_{(a_t, y_{\tau_i})} & \text{if } \tau_i = n \\ D_{u_{i+1}} \wedge \ldots \wedge D_{u_{n^2}} \wedge D_{(x_{\sigma_i}, b_t)} \wedge D_{(a_t, y_{\tau_i})} & \text{if } \sigma_i \vee \tau_i < n \end{cases}.$$

Claim: For any $i \in \{1, \ldots, n^2 - 1\}$,

$$f_U(i) := D_{u_{i+1}} \wedge \ldots \wedge D_{u_{n^2}}.$$

Proof of Claim: We distinguish the mutually exclusive cases $\sigma_i = n$, $\tau_i = n$ and $\sigma_i \vee \tau_i < n$.

If $\sigma_i = n$, then $\tau_i < n$ because $u_i \neq u_{n^2}$. This implies $(x_{\sigma_i}, b_i) = (x_{\sigma_i}, y_{\tau_i+1})$. Suppose $(x_{\sigma_i}, y_{\tau_i+1}) \notin \{u_{i+1} \ldots, u_{n^2}\}$. Then there exists $i^* \in \{1, \ldots, n\}$ such that $(x_{\sigma_i}, y_{\tau_i+1}) = u_{i^*} = (x_{\sigma_{i^*}}, y_{\tau_{i^*}})$. If $i^* = i$, then $(x_{\sigma_i}, y_{\tau_i+1}) = u_i = (x_{\sigma_i}, y_{\tau_i})$, which is a contradiction. If $i^* < i$, then $u_i \not\leq u_{i^*}$ because U is a labeling; but this contradicts $u_i = (x_{\sigma_i}, y_{\tau_i}) \leq (x_{\sigma_i}, y_{\tau_i+1}) = u_{i^*}$. Thus, $(x_{\sigma_i}, y_{\tau_i+1}) \in \{u_{i+1} \ldots, u_{n^2}\}$ if $\sigma_i = n$; in particular, $f_U(i) = D_{u_{i+1}} \wedge \ldots \wedge D_{u_{n^2}} \wedge D_{(x_{\sigma_i}, y_{\tau_i+1})} = D_{u_{i+1}} \wedge \ldots \wedge D_{u_{n^2}}$ in such a case.

If $\tau_i = n$, the reasoning is similar.

If $\sigma_i \vee \tau_i < n$, then $(x_{\sigma_i}, b_i) = (x_{\sigma_i}, y_{\tau_i+1})$ and $(a_i, y_{\tau_i}) = (x_{\sigma_i+1}, y_\tau)$, in which case we show $(x_{\sigma_i}, y_{\tau_i+1}), (x_{\sigma_i+1}, y_\tau) \in \{u_{i+1}, \ldots, u_{n^2}\}$ as was done previously. It will follow that $f_U(i) = D_{u_{i+1}} \wedge \ldots \wedge D_{u_{n^2}}$. //

As the previous claim's statement remains trivially true for $i = 0$, f_U satisfies the requirement that it pass through $D_{(0,0)} = D_{u_1} \wedge \ldots \wedge D_{u_{n^2}}$, $D_{u_2} \wedge \ldots \wedge D_{u_{n^2}}, \ldots$, $D_{u_{n^2}}$. It may also be checked by inspection that f_U is both continuous and injective. As f_U is now a *bona fide* flow, there only remains to show the $\mathbf{R}_+$-indexed point process N^{f_U} is simple. It can be assumed, without changing the essence of the argument (but alleviating its notation significantly!), that $N(\omega)$ is a strictly simple planar measure for *all* $\omega \in \Omega$, rather than just for all ω outside of a P-null set. With this in mind, let us adopt the notation that for any $t > 0$ and $x, y \in \mathbf{R}_+$, $N^{f_U}_{t-} = \lim_{s\uparrow\uparrow t} N^{f^U}_s := \lim_{s\to t, s<t} N^{f^U}_s$, $\Delta N^{f_U}_t := N^{f_U}_t - N^{f_U}_{t-}$, $N_{(x,t-)} = \lim_{s\uparrow\uparrow t} N_{(x,s)} := \lim_{s\to t, s<t} N_{(x,s)}$ and $N_{(t-,y)} = \lim_{s\uparrow\uparrow t} N_{(s,y)} := \lim_{s\to t, s<t} N_{(s,y)}$. Since N^{f_U} does not charge $(0,0)$, it suffices to check that $\Delta N^{f_U}_t \le 1$ for all $t > 0$. For $t > n^2 - 1$ we have

$$
\begin{aligned}
\Delta N^{f_U}_t &= N^{f_U}_t - N^{f_U}_{t-} \\
&= \widetilde{N}_{f_U(t)} - \lim_{s\uparrow\uparrow t} \widetilde{N}_{f_U(s)} \\
&= \widetilde{N}_{D_{(x_n, y_n + t - (n^2-1))}} - \lim_{s\uparrow\uparrow t} \widetilde{N}_{D_{(x_n, y_n + s - (n^2-1))}} \\
&= N_{(x_n, y_n + t - (n^2-1))} - N_{(x_n, y_n + t - (n^2-1) -)} \\
&\le 1
\end{aligned}
$$

the inequality owing to our assumption that N be everywhere strictly simple. If there exists an $i \in \{1, \ldots, n^2 - 1\}$ such that $i - 1 < t \le i$, we face the usual and mutually exclusive cases $\sigma_i = n$, $\tau_i = n$ and $\sigma_i \vee \tau_i < n$. We only address the case $\sigma_i \vee \tau_i < n$, the other two cases being simpler and handled similarly. In this event,

$$
\begin{aligned}
\Delta N^{f_U}_t &= \widetilde{N}_{f_U(t)} - \lim_{s\uparrow\uparrow t} \widetilde{N}_{f_U(s)} \\
&= N_{u_{i+1}} \wedge \ldots \wedge N_{u_{n^2}} \wedge N_{(x_{\sigma_i}, b_t)} \wedge N_{(a_t, y_{\tau_i})} \\
&\quad - N_{u_{i+1}} \wedge \ldots \wedge N_{u_{n^2}} \wedge N_{(x_{\sigma_i}, b_t -)} \wedge N_{(a_t -, y_{\tau_i})}. \qquad (7)
\end{aligned}
$$

By virtue of n's strict simplicity, $N_{(x_{\sigma_i}, b_t)} - N_{(x_{\sigma_i}, b_t -)} \le 1$ and $N_{(a_t, y_{\tau_i})} - N_{(a_t -, y_{\tau_i})} \le 1$. The right-hand-side of equation (1) must therefore be bounded by unity. □

The chapter's main result is now within hand's reach:

Corollary 6.3.8 *If $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ is a strictly simple point process on $\mathbf{R}_+^2$, then the distribution $P \circ N^{-1}$ is uniquely determined by the family $\{\Lambda^f\}_{f \in \mathbf{F}}$ where, for every $f \in \mathbf{F}$, $\{\Lambda_t^f\}_{t \geq 0}$ is the compensator of the point process N^f on $\mathbf{R}_+$.*

Proof By Theorem 6.2.18, $P \circ N^{-1}$ is determined by the distributions of the random vectors $(\widetilde{N}_{l_1}, \ldots, \widetilde{N}_{l_{n^2}})$ where, for some grid $\{z_{(i,j)}\}_{i,j=1}^n$ and labeling $U := \{u_1, \ldots, u_{n^2}\}$ of $\{z_{(i,j)}\}_{i,j=1}^n$, $l_k = D_{u_k} \wedge \ldots \wedge D_{u_{n^2}}$ for $k = 1, 2, \ldots n^2$. Let f_U be the flow exhausting $\{z_{(i,j)}\}_{i,j=1}^n$ along U garanteed by Theorem 6.3.7. $(\widetilde{N}_{l_1}, \ldots, \widetilde{N}_{l_{n^2}}) = (N_0^{f_U}, \ldots, N_{n^2-1}^{f_U})$ by definition of N^{f_U}. But since N^{f_U} is simple, Theorem 3.4 of [23] ensures the distribution of N^{f_U} - and hence, that of $(N_0^{f_U}, \ldots, N_{n^2-1}^{f_U})$ - is uniquely determined by the compensator $\{\Lambda_t^{f_U}\}_{t \geq 0}$. As $f \in \mathbf{F}$, the theorem ensues. □

The next section dicusses the "past information" carried by the family $\{\Lambda^f : f \text{ a flow}\}$ of $\mathbf{R}_+$-indexed compensators, as opposed to the "past information" carried by the $\mathbf{R}_+^2$-indexed compensators researchers have hitherto considered. It also addresses the question of how variously defined "compensators" have - or lack - the power to characterize the distribution of a planar point process.

6.4 Of Compensators and Filtrations

Let $(\Omega, \mathcal{F}, P)$ be a probability space. For any $\mathbf{R}_+$-indexed filtration $\{\mathcal{F}_t\}_{t \geq 0}$ of sub-σ-fields of $\mathcal{F}$, there is no ambiguity in what one means when referring to the "past" at any point $t \in \mathbf{R}_+$: the past is the σ-field $\mathcal{F}_t$ itself, which represents all the information available up to time t: $\mathcal{F}_t = \bigvee_{s \leq t} \mathcal{F}_s$.

The situation is not as simple in the planar context. If $\{\mathcal{F}_z\}_{z \in \mathbf{R}_+^2}$ is a $\mathbf{R}_+^2$-indexed filtration (i.e. a $\mathbf{R}_+^2$-indexed collection of sub-σ-fields of $\mathcal{F}$ such that $\mathcal{F}_{z_1} \subseteq \mathcal{F}_{z_2}$ whenever $z \leq z'$ coordinatewise), there are many possible interpretation of the "past" at a given point $z \in \mathbf{R}_+^2$. To our knowledge, the literature identifies two types of "pasts" at point z. The first is the so-called "strict past", which corresponds to $\mathcal{F}_z$ itself, and therefore embraces only the information available at points z', where $z' \leq z$: $\mathcal{F}_z = \bigvee_{z' \leq z} \mathcal{F}_{z'}$ (see Figure 1). The second type of "past" heretofore encountered is the so-called "wide past", which, at point z, covers all the information comprising the strict pasts of points not strictly larger than z coordinatewise. More precisely, the "wide past" at z is represented by the σ-field $\mathcal{F}^*_z$ obtained via the relation $\mathcal{F}^*_z := \bigvee_{\{z' : z \not\ll z'\}} \mathcal{F}_{z'}$ where, for any $z = (x, y)$, $z' = (x', y') \in \mathbf{R}^2_+$, $z << z'$ if and only if $x < x'$ and $y < y'$ (see Figure 2).

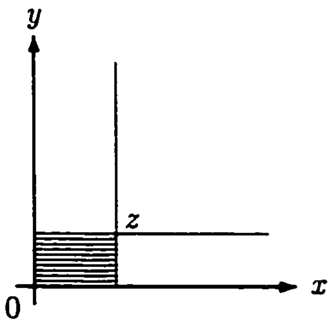

Figure 1

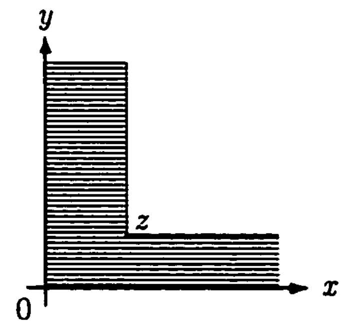

Figure 2

The two filtrations $\{\mathcal{F}_z\}_{z \in \mathbf{R}^2_+}$ and $\{\mathcal{F}^*_z\}_{z \in \mathbf{R}^2_+}$ give rise to two different $\mathbf{R}^2_+$-indexed compensators for planar point processes. If $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{L}, \mathcal{B}(\mathcal{L}))$ is a point process on $\mathbf{R}^2_+$ adapted to the filtration $\{\mathcal{F}_z\}_{z \in \mathbf{R}^2_+}$, then:

- a *compensator* of N is an $\mathbf{R}^2_+$-indexed increasing process $\Lambda := \{\Lambda_z\}_{z \in \mathbf{R}^2_+}$ which is adapted to $\{\mathcal{F}_z\}_{z \in \mathbf{R}^2_+}$ and such that, for any $z, z' \in \mathbf{R}^2_+$ with $z << z'$,

$$\mathbf{E}\left[N((z, z'])|\mathcal{F}_z\right] = \mathbf{E}\left[\Lambda((z, z'])|\mathcal{F}_z\right];$$

- a **-compensator* of N is a $\mathbf{R}^2_+$ indexed-increasing process $M := \{M_z\}_{z \in \mathbf{R}^2_+}$ which is adapted to $\{\mathcal{F}^*_z\}_{z \in \mathbf{R}^2_+}$ and such that, for any $z, z' \in \mathbf{R}^2_+$ with $z << z'$,

$$\mathbf{E}\left[N((z, z'])|\mathcal{F}^*_z\right] = \mathbf{E}\left[M((z, z'])|\mathcal{F}^*_z\right].$$

Remember that a simple point process on $\mathbf{R}_+$ which does not charge 0 and admits at most one accumulation point per realization is guaranteed a compensator with respect to its internal history; this compensator is unique up to indistinguishability and characterizes the distribution of the point process ([23]: Theorems 2.1 and 3.4). The essential features of such a compensator are therefore its existence, its uniqueness and its power to characterize the process' distribution. As we have stated in the chapter introduction, however, no planar compensator - whether it is the *compensator per se* or the **-compensator* - replicates these three features in full generality. The *compensator* may not exist if "F4" is not assumed, and is known not to characterize the point process distribution ([21]: (2), pg. 240); on the other hand, the **-compensator* is only known to characterize the point process distribution when the latter is Poisson or Cox ([22]: Theorem 5.3.1).

Let us therefore appraise the present chapter's material in view of the limitations inherent to compensators defined as $\mathbf{R}_+^2$-indexed increasing processes. In our jargon, every planar point process $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ gives rise to the family $\{N^f : f \text{ a flow}\}$ of non-simple point processes on $\mathbf{R}_+$, the internal histories of which, in turn, produce the family $\{\Lambda^f : f \text{ a flow}\}$ of corresponding compensators. Combining the statements of Theorem 6.3.5 and Corollary 6.3.8, we obtain this characterization:

Theorem 6.4.1 *Let $N : (\Omega, \mathcal{F}, P) \longrightarrow (\mathcal{N}, \mathcal{B}(\mathcal{N}))$ be a planar point process. The distribution $P \circ N^{-1}$ determines the family*

$$\left\{\Lambda^f : f \text{ a flow; } \Lambda^f \text{ the compensator of } N^f\right\}.$$

Furthermore, if N is strictly simple, then the family

$$\left\{\Lambda^f : f \text{ a flow; } \Lambda^f \text{ the compensator of } N^f\right\}$$

determines $P \circ N^{-1}$.

Since the family $\{\Lambda^f : f \text{ a flow}\}$ always exists as a result of the Doob-Meyer decomposition theorem for $\mathbf{R}_+$-indexed processes, we may claim that, in essence, this family acts as analogue to the compensator of simple point processes on $\mathbf{R}_+$. Indeed, the three principal features of the $\mathbf{R}_+$-indexed compensator, namely its existence, its uniqueness and its power to characterize the distribution of the simple point process, are all replicated by the current scheme. This suggests that, for purposes of fixing a certain planar point process distribution via a "compensator", focusing on the family $\{\Lambda^f : f \text{ a flow}\}$ of linear compensators may yield more satisfactory results than attempting to determine the distribution via the $\mathbf{R}_+^2$-indexed dual predictable projection, if it exists. Of course, we are still ignorant of conditions required for a given family $\{A^f : f \text{ a flow}\}$ of increasing processes on the line to be the family of flow compensators associated to a given distribution. Futher investigation will be required to determine any such conditions.

We would like to close this section by addressing some of the questions an enquiring reader may have about the kind of "past" the family $\{\Lambda^f : f \text{ a flow}\}$ generates. For any flow $f : \bar{\mathbf{R}}_+ \longrightarrow (\mathcal{L}, \mathcal{B}(\mathcal{L}))$, let $\{\mathcal{F}_t^f\}_{t\geq 0}$ denote the internal history of the point process N^f. If $\{\mathcal{F}_z\}_{z\in\mathbf{R}_+^2}$ denotes the internal history of the planar point process N, we would like to compare, for any $z \in \mathbf{R}_+^2$, the σ-fields $\mathcal{F}_z$ and $\mathcal{F}_z^*$ to the σ-field $\mathcal{G}_z := \sigma(\mathcal{F}_t^f : f \text{ a flow passing through } D_z, t = f^{-1}(D_z))$. It appears that $\mathcal{G}_z$ properly contains $\mathcal{F}_z$ but is strictly included in $\mathcal{F}_z^*$. In the following illustration (see Figure 3), z is an arbitrary point of the plane, whereas points marked with an "x" and labeled A, B, C and D respectively, are the atoms of a given realization μ of a planar point process N. The lines labeled L_1, L_2, L_3 and L_4 represent the lines L_1^μ, L_2^μ, L_3^μ and L_4^μ respectively (we omitted the superscript so as not to clutter the figure).

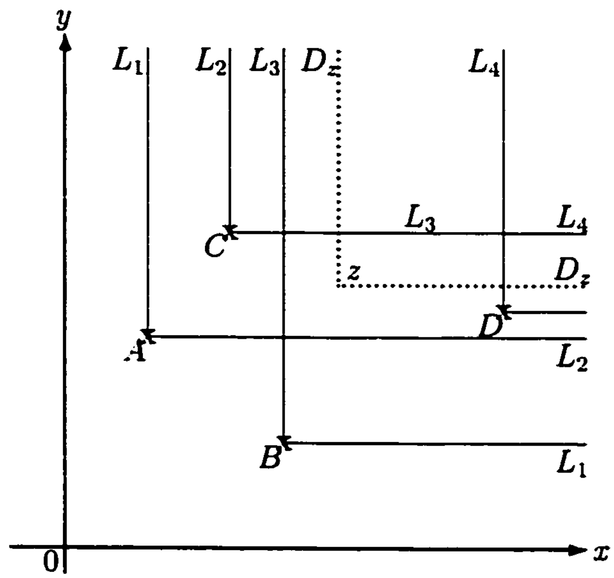

Figure 3

As can be verified, points A and B are captured by all three σ-fields. Point D is only captured by $\mathcal{F}_z^*$: the "strict past" of z does not capture D because point D is not smaller than z and, moreover, since the smallest of the L_i^μ's on which point D lies is L_3^μ and $L_3^\mu \not\leq D_z$, point D cannot be captured by $\mathcal{G}_z$ either. Finally, point C is captured by both $\mathcal{F}_z^*$ and $\mathcal{G}_z$, but obviously not by $\mathcal{F}_z$. That C is captured by $\mathcal{G}_z$ is due to the fact that point C is an exposed point of the line L_2^μ and $L_2^\mu \leq D_z$: gleaning the information produced by all flows passing through the current line L_3^μ and D_z, it is possible to trace back the location of point C.

Part IV

Appendices

Appendix A

Stopping Times and Predictability

We recall elements of the classical theory of point processes on $\mathbf{R}_+ = [0, \infty)$ which pertain to filtrations, predictability, martingales and the Doob-Meyer decomposition.

The symbol $\mathcal{B}(\mathbf{R}_+)$ designates the σ-field of Borel subsets of $\mathbf{R}_+$. The Alexandroff compactification $\bar{\mathbf{R}}_+ = [0, \infty]$ of $\mathbf{R}_+$ also generates a Borel σ-field which we denote by $\mathcal{B}(\bar{\mathbf{R}}_+)$. We observe that $\mathcal{B}(\bar{\mathbf{R}}_+) \cap \mathbf{R}_+ = \mathcal{B}(\mathbf{R}_+)$ and that $\{\infty\} \in \mathcal{B}(\bar{\mathbf{R}}_+)$.

We fix a probability space $(\Omega, \mathcal{F}, P)$.

A.1 Filtration and Stopping Times

Definition A.1.1 *A filtration on $\mathbf{R}_+$ (a filtration for short) is an $\mathbf{R}_+$-indexed collection $\{\mathcal{F}_t\}_{t\geq 0}$ of sub-σ-field of $\mathcal{F}$ such that, for any $s, t \in \mathbf{R}_+$, $s \leq t$ implies $\mathcal{F}_s \subseteq \mathcal{F}_t$. Furthermore, if $\{\mathcal{F}_t\}_{t\geq 0}$ satisfies the property that*

$$\mathcal{F}_t = \bigcap_{u>t} \mathcal{F}_u,$$

then $\{\mathcal{F}_t\}_{t\geq 0}$ is said to be a right-continuous filtration (right-continuous for short). By convention one writes $\mathcal{F}_\infty := \mathcal{F}$ and $\mathcal{F}_{\infty-} := \bigvee_{t\in\mathbf{R}_+} \mathcal{F}_t$.

For the sequel we suppose $(\Omega, \mathcal{F}, P)$ is endowed with the right-continuous filtration $\{\mathcal{F}_t\}_{t\geq 0}$.

Definition A.1.2 *A random variable $T : (\Omega, \mathcal{F}, P) \longrightarrow (\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+))$ is called a stopping time with respect to $\{\mathcal{F}_t\}_{t\geq 0}$ (a stopping time if the filtration is unambiguous) if, for any $t \in \mathbf{R}_+$,*

$$\{T \leq t\} \in \mathcal{F}_t.$$

If T is a stopping time, the class

$$\mathcal{F}_T := \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t \ \forall\ t \in \mathbf{R}_+\}$$

is called the σ-field of events that occur by time T, and

$$\mathcal{F}_{T-} := \mathcal{F}_0 \vee \sigma\left(A \cap \{t < T\} : t \in \mathbf{R}_+ \text{ and } A \in \mathcal{F}_t\right)$$

is called the σ-field of events strictly prior to T.

One may check that $\mathcal{F}_T = \mathcal{F}_t$ and $\mathcal{F}_{T-} = \mathcal{F}_{t-} := \bigvee_{s<t} \mathcal{F}_s$ is unambiguous when $T \equiv t \in \bar{\mathbf{R}}_+$ is constant.

Remark A.1.3 *If $T : (\Omega, \mathcal{F}, P) \longrightarrow (\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+))$ is a stopping time, then $\sigma(T) \subseteq \mathcal{F}_{T-} \subseteq \mathcal{F}_T$.*

Definition A.1.4 *Let $T : (\Omega, \mathcal{F}, P) \longrightarrow (\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+))$ be a stopping time and let $A \in \mathcal{F}_T$. We define $T_A : (\Omega, \mathcal{F}, P) \longrightarrow (\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+))$ by writing*

$$T_A(\omega) := \begin{cases} T(\omega) & \text{if } \omega \in A \\ +\infty & \text{otherwise} \end{cases}.$$

T_A is a stopping time.

The following facts are classical (see, for example [12]).

Proposition A.1.5 *Let $t \in \mathbf{R}_+$; let S and T be stopping times, and let $\{T_n\}_{n\in\mathbf{N}}$ be a sequence of stopping times.*

i) $\bigwedge_{n\in\mathbf{N}} T_n$ *and* $\bigvee_{n\in\mathbf{N}} T_n$ *are stopping times;*

ii) $T+t$ *is a stopping time;*

iii) for any $A \in \mathcal{F}_S$, $A \cap \{S \le t\} \in \mathcal{F}_T$ *(in particular,* $\mathcal{F}_S \subseteq \mathcal{F}_T$ *if* $S \le T$*);*

iv) events $\{S \le T\}, \{S < T\}$ *and* $\{S = T\}$ *belong to both* $\mathcal{F}_S$ *and* $\mathcal{F}_T$.

A.2 Stochastic Intervals and Predictability

We now focus our attention on certain classes of sets of $\mathcal{F} \otimes \mathcal{B}(\bar{\mathbf{R}}_+)$.

Definition A.2.1 *Let* $X = (X_t : t \in \mathbf{R}_+)$ *be a stochastic process.* X *is said to be measurable if the map* $(\omega, t) \mapsto X_t(\omega)$*, defined on* $\Omega \times \mathbf{R}_+$*, is measurable with respect the the* σ*-field* $\mathcal{F} \otimes \mathcal{B}(\mathbf{R}_+)$.

Definition A.2.2 *Let* S *and* T *be stopping times. We define the stochastic intervals* $]S,T]$, $[S,T]$, $]S,T[$ *and* $[S,T[$ *via the relations*

$$]S,T] := \{(\omega,t) \in \Omega \times \mathbf{R}_+ : S(\omega) < t \le T(\omega)\},$$

$$[S,T] := \{(\omega,t) \in \Omega \times \mathbf{R}_+ : S(\omega) \le t \le T(\omega)\},$$

$$]S,T[:= \{(\omega,t) \in \Omega \times \mathbf{R}_+ : S(\omega) < t < T(\omega)\},$$

and

$$[S,T[:= \{(\omega,t) \in \Omega \times \mathbf{R}_+ : S(\omega) \le t < T(\omega)\}$$

respectively. The graph $[T]$ *of* T *is defined as* $[T] := [T,T]$.

It can be shown by an approximation argument that if S and T are stopping times, then the stochastic intervals $]S,T]$, $[S,T]$, $]S,T[$ and $[S,T[$ all belong to $\mathcal{F} \otimes \mathcal{B}(\bar{\mathbf{R}}_+)$. More can be said:

Definition A.2.3 *Let* $X = (X_t : t \in \mathbf{R}_+)$ *be a measurable stochastic process.* X *is said to be adapted (to* $\{\mathcal{F}_t\}_{t\in\mathbf{R}_+}$*) if, for any* $t \in \mathbf{R}_+$, X_t *is measurable with respect to* $\mathcal{F}_t$.

Remark A.2.4 *Let S and T be stopping times. The stochastic processes $\mathbf{1}_{]S,T]}$, $\mathbf{1}_{[S,T]}$, $\mathbf{1}_{]S,T[}$ and $\mathbf{1}_{[S,T[}$ are adapted.*

Stochastic intervals are used to define the optional and predictable sub-σ-fields of $\mathcal{F}$.

Definition A.2.5 *The optional σ-field $\mathcal{O} \subseteq \mathcal{F} \otimes \mathcal{B}(\bar{\mathbf{R}}_+)$ is defined via the relation*

$$\mathcal{O} := \sigma\left([0, T[: T \text{ a stopping time}\right);$$

the predictable σ-field $\mathcal{P} \subseteq \mathcal{F} \otimes \mathcal{B}(\bar{\mathbf{R}}_+)$ is defined by writing

$$\mathcal{P} := \sigma\left([0_A, T] : T \text{ a stopping time}, A \in \mathcal{F}_0\right).$$

Processes and sets measurable with respect to $\mathcal{O}$ are said to be optional; processes and sets measurable with respect to $\mathcal{P}$ are said to be predictable.

If we agree to define a finite, measurable process $X = (X_t : t \in \mathbf{R}_+)$ as left-continuous (respectively, right-continuous, continuous) if, for every $\omega \in \Omega$, the path $t \mapsto X_t(\omega)$ is left-continuous (respectively, right-continuous, continuous), then we obtain the following characterizations of optional and predictable processes (see [26]: Definition I.1.20 and Remark I.1.26; [32]: Theorems 3.4.15 and 3.4.17 and Corollary 3.4.16).

Proposition A.2.6 *i) The optional σ-field is generated by all adapted right-continuous processes;*

ii) the predictable σ-field is generated by all adapted continuous processes;

iii) the predictable σ-field is generated by all adapted left-continuous processes.

Corollary A.2.7 *All predictable processes are optional.*

A straightforward approximation argument which uses adapted left-continuous (respectively, right-continuous) generators of $\mathcal{P}$ (respectively, $\mathcal{O}$) produces other characterizations of $\mathcal{P}$ (respectively, $\mathcal{O}$):

Proposition A.2.8 *i)*
$$\begin{aligned}\mathcal{P} &= \sigma\left(B \times \{0\}, A \times (s, t] : B \in \mathcal{F}_0, s < t \text{ and } A \in \mathcal{F}_s\right) \\ &= \sigma\left(B \times \{0\},]s_A, t] : B \in \mathcal{F}_0, s < t \text{ and } A \in \mathcal{F}_s\right);\end{aligned}$$

ii) $\mathcal{O} = \sigma\left(A \times [s,t] : s < t \text{ and } A \in \mathcal{F}_s\right) = \sigma\left([\![s_A, t]\!] : s < t \text{ and } A \in \mathcal{F}_s\right)$.

It is immediate for the definition that if T is a stopping time, then the set $[\![0,T]\!]$ is predictable (this is equivalent to saying that $]\!]0,T]\!]$ is predictable because $[\![0,T]\!] = [\![0]\!] \cup]\!]0,T]\!]$ and $[\![0]\!] = \Omega \times \{0\}$ is predictable). Likewise, the set $[\![0,T[\![$ is optional, but not predictable in general.

Definition A.2.9 *Let T be a stopping time. T is said to be predictable if the stochastic interval $[\![0,T[\![$ is predictable.*

Predictable stopping times are characterized by the fact that they are *announceable* (see [26]: Theorem I.2.15):

Theorem A.2.10 *Let $T : (\Omega, \mathcal{F}, P) \longrightarrow (\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+))$ be a stopping time. These two statements are equivalent:*

i) T is predictable;

ii) there exist a sequence $\{T_n\}_{n\in\mathbf{N}}$ of stopping times such that, for any $n \in \mathbf{N}$, $T_n \leq T_{n+1} < T$ P-almost surely on $\{T > 0\}$, and such that $\lim_{n\in\mathbf{N}} T_n = T$ almost surely.

If T is a predictable stopping time and $A \in \mathcal{F}_T$, then the stopping time T_A may or may not be predictable. However, we have this result ([26]: Proposition I.2.10):

Proposition A.2.11 *Let $T : (\Omega, \mathcal{F}, P) \longrightarrow (\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+))$ be a predictable stopping stopping time. If $A \in \mathcal{F}_{T-}$, then the stopping time T_A is predictable.*

For any stopping time T and any constant $t > 0$, the stopping time $T+t$ is predictable. As $[\![0_A, T]\!] = \bigcap_{n\in\mathbf{N}} [\![0_A, T + \frac{1}{n}[\![$ for any $A \in \mathcal{F}_0$, an ultimate characterization of the predictable σ-field is obtained:

Proposition A.2.12

$$\begin{aligned}\mathcal{P} &= \sigma\left([\![S,T[\![: S \text{ and } T \text{ are predictable stopping times}\right)\\ &= \sigma\left([\![0,T[\![: T \text{ is a predictable stopping time}\right).\end{aligned}$$

Appendix B

Martingales and Increasing Processes

We hereby define martingales and increasing processes so as to state the Doob-Meyer decomposition for "submartingales of class (D)".

We fix a probability space $(\Omega, \mathcal{F}, P)$, which we endow with a right-continous filtration $\{\mathcal{F}_t\}_{t\geq 0}$.

B.1 Martingales

If $f : \mathbf{R}_+ \longrightarrow \mathbf{R}$ is a function, agree to call f *càdlàg* if, for every $t \in (0, \infty)$, $f(t) = \lim_{s \to t+} f(s)$ and $\lim_{s \to t-} f(s)$ exists.

Definition B.1.1 *Let $X = (X_t : t \in \mathbf{R}_+)$ be a finite-valued, measurable process such that, for P-almost all $\omega \in \Omega$, the path $t \mapsto X_t(\omega)$ is càdlàg, and such that X_t is integrable for all $t \in \mathbf{R}_+$. We say X is*

i) a submartingale if $X_s \leq \mathbf{E}(X_t|\mathcal{F}_s)$ for all $s, t \in \mathbf{R}_+$ with $s \leq t$;

ii) s supermartingale if $X_s \geq \mathbf{E}(X_t|\mathcal{F}_s)$ for all $s, t \in \mathbf{R}_+$ with $s \leq t$;

iii) a martingale if $X_s = \mathbf{E}(X_t|\mathcal{F}_s)$ for all $s, t \in \mathbf{R}_+$ with $s \leq t$.

Doob's Optional Sampling Theorem states that, provided a supermartingale is *closed* by a terminal random variable, the defining inequalities extend to indices which are stopping times (see [26]: Theorem I.1.39):

Theorem B.1.2 *Let X be a supermartingale such that there exists an integrable random variable Y with $X_t \geq \mathbf{E}(Y|\mathcal{F}_t)$ for all $t \in \mathbf{R}_+$. If S and T are two stopping times, then the variables $X_S \mathbf{1}_{\{S<\infty\}}$ and $X_T \mathbf{1}_{\{T<\infty\}}$ are integrable and satisfy $X_S \geq \mathbf{E}(X_T|\mathcal{F}_S)$ on $\{S \leq T < \infty\}$.*

We close this section by defining *local* martingales, submartingales and supermartingales.

Definition B.1.3 *A stochastic process X is said to be a local martingale (respectively local submartingale, local supermartingale) if there exist an increasing sequence $\{T_n\}_{n\in\mathbf{N}}$ of stopping times such that $\sup_{n\in\mathbf{N}} T_n = \infty$ almost surely and such that for every $n \in \mathbf{N}$, the process*

$$X^{T_n} := (X_{t\wedge T_n} : t \in \mathbf{R}_+)$$

is a martingale (respectively submartingale, supermartingale).

B.2 Increasing Processes

Definition B.2.1 *A measurable stochastic process A is said to be increasing if A adapted and for every $\omega \in \Omega$, the path $t \mapsto A_t(\omega)$ is non-decreasing and càdlàg with $A_0(\omega) = 0$.*

It is thus clear that if A is an increasing process, then every $\omega \in \Omega$ determines a path $t \mapsto A_t(\omega)$ which is the distribution function of a measure on the Borel subsets if $\mathbf{R}_+$. The "differential" dA therefore makes sense when employed in a Stieltjes integral.

Notation B.2.2 *Throughout this thesis, if A is an increasing process and B is a Borel subset of $\mathbf{R}_+$, for any $\omega \in \Omega$ the symbol $A(\omega, B)$ denotes the quantity $\int_B dA(\omega)$ and the symbol $A(B)$ denotes the corresponding random variable.*

We recall the definition of indistinguishability prior to introducing the result which establishes a correspondence between σ-finite measures on $\mathcal{F} \otimes \mathcal{B}(\mathbf{R}_+)$ and "locally integrable increasing processes".

Definition B.2.3 *Let $X = (X_t : t \in \mathbf{R}_+)$ and $Y = (Y_t : t \in \mathbf{R}_+)$ be stochastic processes. X and Y are said to be indistinguishable if the set*

$$\{\omega \in \Omega : \exists\, t \in \mathbf{R}_+ \text{ such that } X_t(\omega) \neq Y_t(\omega)\}$$

is P-null.

When a process possessing a certain characteristic is unique up to indistinguishability, we also say that it is "essentially unique".

The following result can be found in Dellacherie ([12]: IV.T41).

Theorem B.2.4 *i) Let $A = (A_t : t \in \mathbf{R}_+)$ be an increasing process such that $\mathbf{E}(A_t) < \infty$ for all $t \in \mathbf{R}_+$. There exists a unique σ-finite measure μ_A on $\mathcal{F} \otimes \mathcal{B}(\mathbf{R}_+)$ such that, for any nonnegative measurable process $X = (X_t : t \in \mathbf{R}_+)$,*

$$\int X d\mu_A = \int_\Omega \int_{(0,\infty)} X_t(\omega) dA_t(\omega) dP(\omega) = \mathbf{E}\left[\int X dA\right].$$

i) Let μ be a σ-finite measure on $\mathcal{F} \otimes \mathcal{B}(\mathbf{R}_+)$. There exists an essentially unique increasing process $A = (A_t : t \in \mathbf{R}_+)$ such that $\mathbf{E}(A_t) < \infty$ for all $t \in \mathbf{R}_+$ and such that

$$\int X d\mu = \mathbf{E}\left[\int X dA\right]$$

for all nonnegative measurable processes $X = (X_t : t \in \mathbf{R}_+)$.

B.3 Doob-Meyer Decomposition

This is where martingale theory and increasing processes meet. We shall invoke the Doob-Meyer decomposition result in the subsequent appendix.

Let us first define what is meant for a process to be "of class (D)".

Definition B.3.1 *Let $X = (X_t : t \in \mathbf{R}_+)$ be a finite-valued, measurable stochastic process. X is said to be of class (D) if the collection of random variables $\{X_T : T \text{ a finite-valued stopping time}\}$ is uniformly integrable.*

The Doob-Meyer decomposition for submartingales may be stated as follows (see [26]: Theorem I.3.15):

Theorem B.3.2 *Let X be a submartingale of class (D). There exists an essentially unique predictable increasing process A such that the process*

$$X - A = (X_t - A_t : t \in \mathbf{R}_+)$$

is a martingale of class (D).

A special case of Theorem B.3.2, stated in slightly different form, arises when X is a "locally integrable" increasing process. In this case, X is a submartingale which is not necessarily of class (D), but which nevertheless admits a Doob-Meyer decomposition. According to Dellacherie ([12]: V.T28), the following holds.

Theorem B.3.3 *Let A be an increasing process such that $\mathbf{E}(A_t) < \infty$ for all $t \in \mathbf{R}_+$. There exists an essentially unique predictable increasing process A^p such that, for any nonnegative, predictable process X, we have*

$$\mathbf{E} \int X dA = \mathbf{E} \int X dA^p.$$

Let us give a name to this process A^p.

Definition B.3.4 *Let A and A^p be as in Theorem B.3.3. A^p is called the dual predictable projection, or compensator, of A.*

Appendix C

Point Processes and their Compensators

C.1 Defining a Point Process Compensator

A simple, unmarked point process on $\mathbf{R}_+$ is an $\bar{\mathbf{R}}_+$-valued stochastic process $N = (N_t : t \in \mathbf{R}_+)$ constructed as follows: we let

$$T_1, T_2, \ldots : (\Omega, \mathcal{F}, P) \longrightarrow \left(\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+)\right)$$

be such that

$$T_0 := 0 < T_1 < T_2 < \ldots$$

and for any $\omega \in \Omega, t \in \mathbf{R}_+$ we let

$$N_t(\omega) := \sum_{n \in \mathbf{N}} \mathbf{1}_{\{T_n \leq t\}}(\omega).$$

Observe that $N_t(\omega) < \infty$ for all $t \in \mathbf{R}_+$ if and only if $T_n(\omega) \to \infty$ as $n \to \infty$; N thus is not an increasing process in general as the trajectories $t \mapsto N_t(\omega)$, $\omega \in \Omega$ may not necessarily be the graphs of $\mathbf{R}_+$-valued functions. Speaking of the "compensator of N" does therefore not make sense in general, according to Definition B.3.4. However, it is easy to see that N induces a measure μ_N on $\mathcal{F} \otimes \mathbf{R}_+$ via the relation

$$\mu_N \left(A \times (s, t]\right) := \mathbf{E}\left[\sum_{n \in \mathbf{N}} \mathbf{1}_{A \cap \{s < T_n \leq t\}}\right]$$

for all $s, t \in \mathbf{R}_+$ with $s < t$ and all $A \in \mathcal{F}$ (the expectation may be infinite). Additionally, μ_N thus defined is σ-finite because μ_N takes value at most 1 at the sets

$$\{(\omega, t) \in \Omega \times \mathbf{R}_+ : T_{n-1}(\omega) < t \leq T_n(\omega)\},\ n \in \mathbf{N},$$

and equals 0 at the set

$$\{(\omega, t) \in \Omega \times \mathbf{R}_+ : \sup_{n \in \mathbf{N}} T_n(\omega) \leq t < \infty\}.$$

Theorem B.3.3 suggests, by its formulation, a way of defining a compensator for N. In this instance, supposing the process $N = (N_t : t \in \mathbf{R}_+)$ is adapted some right-continuous filtration $\{\mathcal{F}_t\}_{t \geq 0}$, we would define a compensator of N as a predictable, $\bar{\mathbf{R}}_+$ process $\Lambda = (\Lambda_t : t \in \mathbf{R}_+)$ such that, for P-almost all $\omega \in \Omega$,

i) $\Lambda_0(\omega) = 0$,

ii) $\Lambda_s(\omega) \leq \Lambda_t(\omega)$ for all $s, t \in \mathbf{R}_+$ with $s \leq t$ and

ii) $\Lambda_t(\omega) = \lim_{u \to t+} \Lambda_u(\omega)$ for all $t \in \mathbf{R}_+$,

and such that, for all nonnegative, predictable processes X,

$$\mathbf{E} \int X_t dN_t = \mathbf{E} \int X_t d\Lambda_t$$

(for the right-hand side to make sense, it suffices to consider the stopping time $T_\infty^\Lambda := \inf\{t \in \mathbf{R}_+ : \Lambda_t = \infty\}$ and set $\int Y_t d\Lambda_t \equiv 0$ for $Y :=]T_\infty^\Lambda, \infty[$). It so happens that if there exists a sub-σ-field $\mathcal{G}$ of $\mathcal{F}$ such that, for any $t \in \mathbf{R}_+$, $\mathcal{F}_t = \mathcal{G} \vee \sigma(N_s : s \leq t)$, then such process Λ exists and is unique up to indistinguishability (see [23]: Theorem 2.1, as well as the remark on same page guaranteeing the right-continuity of $\{\mathcal{F}_t\}_{t \geq 0}$ in this case). In such a situation we shall refer to the process Λ as the compensator of the point process N.

In Chapter 3, we argue that the compensator still exists when the hypothesis

$$0 < T_1 < T_2 < \ldots$$

is relaxed to

$$0 < T_1 \leq T_2 \leq \ldots,$$

provided that for any $m \in \mathbf{N}$ and any $\omega \in \Omega$, $T_m(\omega) < \sup_{n\in\mathbf{N}} T_n(\omega)$ holds. In fact, the result is generalized to compensators of "marked point processes" which Chapter 3 generically refers to as "the random measure μ".

C.2 A Finite Compensator for a Finite Point Process

Chapters 2, 4 and 5 assume that if N is a simple, unmarked point process on $\mathbf{R}_+$ such that $N_t(\omega) < \infty$ for all $\omega \in \Omega$ and $t \in \mathbf{R}_+$ (i.e. for all $\omega \in \Omega$, $T_n(\omega) \to \infty$ as $n \to \infty$), then, supposing the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfies $\mathcal{F}_t = \sigma(N_s : s \leq t)$ for all $t \in \mathbf{R}_+$, the compensator Λ of N can be chosen so that $\Lambda_t(\omega) < \infty$ for all $\omega \in \Omega$ and $t \in \mathbf{R}_+$. Since $\mathbf{E}(N_t)$ may not necessarily be finite for all $t \in \mathbf{R}_+$, Theorem B.3.3 cannot be used to infer this. Let us first prove a preliminary result:

Proposition C.2.1 *Let N be a simple, unmarked point process on $\mathbf{R}_+$ such that $N_t(\omega) < \infty$ for all $\omega \in \Omega$ and all $t \in \mathbf{R}_+$. Suppose the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfies $\mathcal{F}_t = \sigma(N_s : s \leq t)$, and let Λ be a version of N's compensator. We have that*

$$P\left[\omega \in \Omega : \Lambda_t(\omega) = \infty \text{ for some } t \in \mathbf{R}_+\right] = 0.$$

Proof Let us first prove $\{\omega \in \Omega : \Lambda_t(\omega) = \infty \text{ for some } t \in \mathbf{R}_+\}$ is a measurable set. As before, define the $\bar{\mathbf{R}}_+$-valued random variable T_∞^Λ be writing

$$T_\infty^\Lambda := \inf\{t \in \mathbf{R}_+ : \Lambda_t = \infty\}.$$

T_∞^Λ is a stopping time, so that

$$[\omega \in \Omega : \Lambda_t(\omega) = \infty \text{ for some } t \in \mathbf{R}_+] = \{T_\infty^\Lambda < \infty\} \in \mathcal{F}.$$

Suppose $P\{T_\infty^\Lambda < \infty\} = \eta$ for some $\eta > 0$. Since $\lim_{n\to\infty} T_n = \infty$ be hypothesis, we have

$$\{T_\infty^\Lambda < \infty\} = \bigcup_{n\in\mathbf{N}} \{T_\infty^\Lambda < T_n\},$$

so that there exists a $n^* \in \mathbf{N}$ and a $\delta > 0$ such that $P\{T_\infty^\Lambda < T_{n^*}\} = \delta$. Observe that $\{T_\infty^\Lambda < T_{n^*}\}$ is the projection onto Ω of the stochastic interval $]T_\infty^\Lambda, T_{n^*}]$, which is a predictable set. Therefore, by the Predictable Section Theorem ([12]: IV.T10), there exists a stopping time T such that $T_\infty^\Lambda < T \leq T_{n^*}$ almost surely on $\{T < \infty\}$, and such that

$$P\{T < \infty\} \geq P\{T_\infty^\Lambda < T_{n^*}\} - \frac{\delta}{2} \geq \frac{\delta}{2}.$$

Since $\Lambda_T = \infty$ almost surely on $\{T < \infty\}$ and $P\{T < \infty\} > 0$, it must follow that $\mathbf{E}\left[\Lambda_T \mathbf{1}_{\{T<\infty\}}\right] = \infty$. However, computations reveal

$$\begin{aligned}
\mathbf{E}\left[\Lambda_T \mathbf{1}_{\{T<\infty\}}\right] &= \mathbf{E}\left[\int \mathbf{1}_{\{T<\infty\}} \mathbf{1}_{]0,T]} d\Lambda\right] \\
&= \mathbf{E}\left[\int \mathbf{1}_{\{T<\infty\}} \mathbf{1}_{]0,T]} \mathbf{1}_{]0,T_{n^*}]} d\Lambda\right] \\
&\leq \mathbf{E}\left[\int \mathbf{1}_{]0,T_{n^*}]} d\Lambda\right] \\
&= \mathbf{E}\left[\int \mathbf{1}_{]0,T_{n^*}]} dN\right] \qquad \text{(because }]0, T_{n^*}] \in \mathcal{P}) \\
&\leq n^*,
\end{aligned}$$

which is a contradiction. We conclude $P\{T_\infty^\Lambda < \infty\} = 0$. □

The next proposition will establish the existence of a finite compensator for the finite, simple, unmarked point process.

Proposition C.2.2 *Let N, Λ and $\{\mathcal{F}_t\}_{t\geq 0}$ be as in the statement of Proposition C.2.1. There exists a version $\widetilde{\Lambda}$ of N's compensator which is such that $\widetilde{\Lambda}_t(\omega) < \infty$ for all $\omega \in \Omega$ and all $t \in \mathbf{R}_+$.*

Proof Define $T_\infty^\Lambda : (\Omega, \mathcal{F}, P) \longrightarrow (\bar{\mathbf{R}}_+, \mathcal{B}(\bar{\mathbf{R}}_+))$ as in the proof of Proposition C.2.1. Because Λ is predictable, the stopping time T_∞^Λ is predictable as well ([26]: Proposition I.2.13), so that stochastic intervals $]0, T_\infty^\Lambda[$ and $[T_\infty^\Lambda, \infty[$ are predictable. For any $\omega \in \Omega, t \in \mathbf{R}_+$, let

$$\widetilde{\Lambda}_t(\omega) := \Lambda_t(\omega) \cdot \mathbf{1}_{]0,T_\infty^\Lambda[}(\omega, t).$$

Since $\Lambda < \infty$ everywhere on $]0, T_\infty^\Lambda[$, $\widetilde{\Lambda}$ is always finite. Since $P\{T_\infty^\Lambda < \infty\} = 0$, it is also clear that the paths $t \mapsto \widetilde{\Lambda}_t$ are almost everywhere càdlàg and non-decreasing. □

C.3 The Poisson and Cox Processes

For the sake of completeness, we hereby define some special point processes mentioned in this thesis.

Definition C.3.1 *Let N be a simple, unmarked point process on $\mathbf{R}_+$.*

- *N is called a Poisson process with intensity f if there exists a Borel function $f : \mathbf{R}_+ \longrightarrow \mathbf{R}_+$ such that $\int_0^t f(x)dx < \infty$ for all $t \in \mathbf{R}_+$ and such that, for any $t_1, \ldots, t_n \in \mathbf{R}_+$ with $0 < t_1 < \ldots < t_n$, random variables $N_{t_1} - N_0$, $N_{t_2} - N_{t_1}, \ldots, N_{t_n} - N_{t_{n-1}}$ are independent and Poisson distributed with parameters $\int_0^{t_1} f(x)dx, \int_{t_1}^{t_2} f(x)dx, \ldots, \int_{t_{n-1}}^{t_n} f(x)dx$ respectively;*
- *N is called a Standard Poisson Process if it is a Poisson process with intensity 1;*
- *N is called a Cox (or doubly stochastic) process if there exists an increasing process $\Lambda = (\Lambda_t : t \in \mathbf{R}_+)$ such that, for any $t_1, \ldots, t_n \in \mathbf{R}_+$ with $0 < t_1 \ldots < t_n$, random variables $N_{t_1} - N_0, N_{t_2} - N_{t_1}, \ldots, N_{t_n} - N_{t_{n-1}}$, when conditioned on the event $\left[\Lambda_{t_1} - \Lambda_0 = x_1, \Lambda_{t_2} - \Lambda_{t_1} = x_2, \ldots, \Lambda_{t_n} - \Lambda_{t_{n-1}} = x_n\right]$, are independently Poisson distributed with respective parameters $x_1, x_2, \ldots, x_n$.*

Appendix D

Lusin Space Compactification

A Lusin topological space is a Borel subset of a compact metric space. We prove that an element can be added to a Lusin space in a way which preserves the Lusin property.

Proposition D.0.2 *Let E be a Lusin topological space. It is possible to find $\Delta \notin E$ such that $\bar{E} := E \cup \{\Delta\}$ can be topologized as a Lusin space.*

Proof Let X be a compact metric space such that E is a Borel subset of X. If $E \neq X$, simply choose $\Delta \in X \setminus E$; $\bar{E} := E \cup \{\Delta\}$ will then be a Borel subset of the compact metric space X. If $E = X$, let $\bar{E} := E \cup \{\Delta\} =: \bar{X}$ be the Alexandroff compactification of X. Since X is compact - hence locally compact - and Hausdorff, $\bar{E} = \bar{X}$ is Hausdorff ([31]; p. 150). By [31] (p. 125), to show that $\bar{X}$ is metrizable, it suffices to ascertain that $\bar{X}$ is T_1 and regular, and that its topology admits a countable basis. That $\bar{X}$ is T_1 (i.e. its singletons are closed sets) follows from the fact that $\bar{X}$ is Hausdorff. To show that $\bar{X}$ is regular (i.e. every point admits a system of closed neighborhoods), observe that, since X is compact, Δ must be an isolated point of $\bar{X}$ ([31]; p. 150). From the fact that X itself is a metric space follows that evey point of $\bar{X}$, including Δ, admits a system of closed neighborhoods. Thus, $\bar{X}$ is regular. Finally, $\bar{X}$'s topology admits a countable basis because X is Lusin (and therefore, homeomorphic to a Borel subset of $[0,1]^{\mathbf{N}}$ by [13], pp. 48-49) and Δ is an isolated point of $\bar{X}$. □

Bibliography

[1] Ash, R. B. (1972). *Real Analysis and Probability.* Academic Press, New York.

[2] Bakry, D. (1979). Sur la régularité des trajectoires des martingales à deux indices. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **50**, 149-157.

[3] Bakry, D. (1981). Théorèmes de section et de projection pour les processus à deux indices. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **55**, 55-71.

[4] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[5] Billingsley, P. (1995). *Probability and Measure.* Wiley, New York.

[6] Brennan, M. D. (1979). Planar semimartingales. *Journal of Multivariate Analysis* **9**, 465-486.

[7] Brémaud, P. (1981). *Point Processes and Queues, Martingale Dynamics.* Springer-Verlag, New York.

[8] Brown, T. (1978). A martingale approach to the Poisson convergence of simple point processes. *Annals of Probability*, Vol. 6, No. 4, 615-628.

[9] Burton, R. and Waymire, E. (1985). Scaling limits for associated random measures. *Annals of Probability* **13**, 1267-1278.

[10] Cox. J. T. and Grimmett, G. (1984). Central limit theorems for associated random variables and the percolation model. *Annals of Probability* **12**, 514-528.

[11] Dabrowski, A. R. and Jakubowski, A. (1994). Stable limits for associated random variables. Annals of Probability **22**, 1-16.

[12] Dellacherie, C. (1972). *Capacités et processus stochastiques.* Springer-Verlag, New York.

[13] Dellacherie, C. and Meyer, P.-A. (1978). *Probabilities and Potential*, North-Holland, New York.

[14] Doléans, C. and Meyer, P.-A. (1979). Un petit théorème de projection pour processus à deux indices, pp. 204-215, *Lecture Notes in Mathematics* **721**, Springer, New York.

[15] Dozzi, M. (1981). On the decomposition and integration of two-parameter stochastic processes, pp. 162-171, *Lecture Notes in Mathematics* **863**, Springer, New York.

[16] Esary, J. D. and Proschan, F. (1970). A reliability bound for systems of maintained, interdependent components. *Journal of the American Statistical Association* —bf 65, 329-338.

[17] Esary, J. D., Proschan, F. and Walkup, D. W. (1967). Association of random variables, with applications. *Annals of Mathematical Statistics* **38**, 1466-1474.

[18] Grandell, J. (1977). Point processes and random measures. *Advances in Applied Probability* **9**, 502-526.

[19] Gushchin, A. A. (1982). On the general theory of random fields on the plane. *Russian Mathematical Surveys* **37**:6, 55-80.

[20] Ivanoff, B. G. (1985). Poisson Convergence for Point Processes on the Plane. *Journal of the Australian Mathematical Society* **39**, 253-269.

[21] Ivanoff, B. G. and Merzbach, E. (1990). Characterization of Compensators for Point Processes on the Plane. *Stochastics and Stochastics Reports* **29**, 395-405.

[22] Ivanoff, B. G. and Merzbach, E. (2000). *Set-Indexed Martingales.* Chapman and Hall/CRC, Boca Raton.

[23] Jacod, J. (1975). Multivariate Point Processes: Predictable Projection, Radon-Nikodym Derivatives, Representation of Martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **31**, 235-253.

[24] Jacod, J. (1979). *Calcul Stochastique et Problèmes de Martingales*, Springer-Verlag, New York.

[25] Jacod. J. (1987). Sur la convergence des processus ponctuels. *Probability Theory and Related Fields* **76**, 573-586.

[26] Jacod, J. and Shiryaev, A. N. (1987). *Limit Theorems for Stochastic Processes.* Springer-Verlag, Berlin.

[27] Kabanov, Y., Liptser, R. S. and Shiryaev, A. N. (1981). Some limit theorems for simple point processes (martingale approach). *Stochastics and Stochastics Reports* **3**, 203-216.

[28] Kallenberg, O. (1983). *Random Measures.* Academic Press, London.

[29] Kamae, T., Krengel, U. and O'Brien, G. L. (1977). Stochastic inequalities on partially ordered spaces. *Annals of Probability*, Vol. 5, No. 6, 899-912.

[30] Karr, A. F. (1991). *Point processes and their statistical inference*, second edition. Marcel Dekker, Inc., New York.

[31] Kelley, John L. (1955). *General Topology.* Van Nostrand, Princeton.

[32] Kopp, P. E. (1984). *Martingales and stochastic integrals.* Cambridge University Press, Cambridge.

[33] Kuratowski, K. and Mostovski, A. (1976). *Set Theory.* PWN, Warsaw.

[34] Kurtz, T. G. (1980). The optional sampling theorem for martingales indexed by directed sets. *Annals of Probability* **8**, 675-681.

[35] Kurtz, T. G. (1982). Representation and approximation of counting processes. *Lecture Notes in Control and Information Sciences* **42**, pp. 177-199. Springer-Verlag, New York.

[36] Kwieciński, A. and Szekli, R. (1991). Compensator conditions for stochastic ordering of point processes. *Journal of Applied Probability* **28**, 751-761.

[37] Kwieciński, A. and Szekli, R. (1996). Some monotonicity and dependence properties of self-exciting point processes. *Annals of Applied Probability*, Vol. 6, No. 4, 1211-1231.

[38] Last, G. and Szekli, R. (1998). Stochastic comparison of repairable systems by coupling.

[39] Lindqvist, B. H. (1988). Association of Probability Measures on Partially Ordered Spaces. *Journal of Multivariate Analysis* **26**, 111-132.

[40] Matheron, G. (1975). *Random Sets and Integral Geometry.* John Wiley and Sons, Toronto.

[41] McDonald, D. (1994). *Elements of Applied Probability.* Quality Advice, Ottawa.

[42] Mazziotto, G. and Merzbach, E. (1988). Point Processes Indexed by Directed Sets. *Stochastic Processes and their Applications* **30**, 105-119.

[43] Merzbach, E. and Zakai, M. (1980). Predictable and Dual Predictable Projections of Two-Parameter Stochastic Processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **53**, 263-269.

[44] Nachbin, L. (1965). *Topology and Order.* Van Nostrand, Princeton.

[45] Newman, C. M. and Wright, A. L. (1981). An invariance principle for certain dependent sequences. *Annals of Probability* **9**, 671-675.

[46] Plante, M. (1999). A note on self-exciting point processes. *Technical Report Series of the Laboratory for Research in Statistics and Probability (University of Ottawa - Carleton University)*, No. 334.

[47] Rolski, T. and Szekli, R. (1991). Stochastic ordering and thinning of point processes. *Stochastic Processes and their Applications* **37**, 299-312.

[48] Strassen, V. (1965). The existence of probability measures with given marginals. *Annals of Mathematical Statistics* **36**, 423-439.

[49] Whitt, W. (1980). Uniform conditional stochastic order. *Journal of Applied Probability* **17**, 112-123.