AISM (2006) 58: 223–233
DOI 10.1007/s10463-005-0003-y

**Frederic Paik Schoenberg**

# On non-simple marked point processes

Received: 1 September 2003 / Revised: 3 August 2004 / Published online: 4 March 2006


**Abstract** Simple point processes are often characterized by their associated compensators or conditional intensities. Non-simple point processes are not uniquely determined by their conditional intensity and compensator, so instead one may identify with the point process its associated simple point process and corresponding conditional intensity, on an expanded mark space. Some relations between the conditional intensity on the expanded mark space and the ordinary conditional intensity are investigated here, and some classes of separable non-simple processes are presented. Transformations into simple point processes, involving thinning and rescaling, are presented.

**Keywords** Point process · Simplicity · Counting process · Jump process · Conditional intensity · Random time change · Random thinning

## 1 Introduction

The existence and uniqueness of the compensator, for multivariate or marked point processes with at most one point at any time, has long been known and is the basis for using the compensator for modeling these processes (Jacod, 1975). However, relatively little is known about processes which may have simultaneous points. For such non-simple point processes, the non-uniqueness of the compensator is not only an obstacle to modeling, but also in model evaluation: the horizontal rescaling result of Meyer (1971), which is useful for assessing point process models, requires this simplicity assumption, and the same is true for the extensions and alternate proofs of Meyer's theorem, including those involving vertical rescaling

F.P. Schoenberg
Department of Statistics, University of California, Los Angeles,
CA 90095-1554, USA
E-mail: frederic@stat.ucla.edu

(see e.g. Merzbach and Nualart, 1986; Brown and Nair, 1988; Nair, 1990; Schoenberg, 1999).

The failure of the compensator uniquely to characterize non-simple point processes applies even to the case of unmarked point processes on the line. For an elementary example, given a simple temporal Poisson process of unit rate, consider a non-simple point process constructed to have two points at each time at which the Poisson process has a point. The resulting double-point process has a compensator identical to that of a simple Poisson process of rate two.

The requirement of simplicity for the modeling and rescaling of point processes is unfortunate and does not appear to be met for some applications to data. For instance, point processes have been actively used in modeling the occurrences of earthquakes, and recent seismological research on faulting suggests that earthquakes may begin with multiple ruptures at the same or infinitesimally different times (Kagan, 1994); in such cases even if the underlying earthquake process is simple, the recorded observations of such events may not be.

One way to characterize a non-simple point process is described in Chapt. 7.2 of Daley and Vere-Jones (1988). Their method involves identifying with the point process its associated point process on a different space with an expanded mark space. The resulting associated process is simple by construction, and hence may be characterized uniquely by its conditional intensity on the expanded space.

The primary aim of the current paper is to explore some special cases of non-simple point processes, and to investigate the relationships between their ordinary conditional intensities and those of their associated point processes on the expanded space. Introductory definitions and notation are given in Sect. 2, along with the method for characterizing a general non-simple point process given by Daley and Vere-Jones (1988, 2003). Section 3 describes separability criteria for non-simple point processes, defined via conditions on the conditional intensities of the associated point processes on the expanded space, and relations between the conditional intensity of the non-simple process and that of the associated process are derived. Section 4 provides some results on randomly rescaling and thinning non-simple point processes and thus transforming them into simple or separable point processes.

## 2 Preliminaries

In this section we review some basic point process constructs; for further details on point processes and conditional intensities see Papangelou (1972), Jacod (1975), Brémaud (1981), and Daley and Vere-Jones (2003).

A temporal point process $N$ is a $\sigma$-finite random measure on the real-line $\mathbf{R}$ or a portion thereof, taking values in the non-negative integers or infinity. $N(B)$ represents the number of points in a subset $B$ of $\mathbf{R}$. For a temporal *marked* point process (hereafter abbreviated t.m.p.p.), to each point there corresponds a random variable from some measurable mark space $\mathcal{X}$. We consider here the case of temporal marked point processes where the mark space is countable; such a process $N$ may be viewed as a sequence $\{N_i;\ i = 1, 2, \dots\}$ of temporal point processes, where the sum $\sum_i N_i(B)$ is $\sigma$-finite on $\mathbf{R}$ (see e.g. Bremaud, 1981). The results in the subsequent sections can quite trivially be extended to processes where the points lie in more general spaces, provided the definition of a conditional intensity exists (which requires an ordering on the domain as well as measurability of the mark space). For simplicity, we restrict our attention here to the case of a t.m.p.p. with countable mark space.

We consider the case where the temporal domain is the real half-line $(0, \infty)$. To the random measure $N_i$ there corresponds the right-continuous stochastic process

$$N_i(t) := N_i(0, t] \tag{1}$$

for $t > 0$. To avoid ambiguity, we distinguish between $N_i(t)$ as defined in Eq. (1) and $N_i(\{t\})$, the number of points at exactly time $t$ and mark $x_i$. The collection of processes $\{N_i(t)\}$ is considered adapted to some filtered probability space $(\Omega, \mathcal{F}_t, P)$. A *conditional intensity* $\lambda(t)$ of a temporal point process $N$ is a non-negative, $\mathcal{F}$-predictable process such that $N(t) - \int_0^t \lambda(u)\,\mathrm{d}u$ is an $\mathcal{F}$-martingale; its integral $A(t) = \int_0^t \lambda(u)\,\mathrm{d}u$ defines $N$'s *compensator*, whose general existence and uniqueness are established in Jacod (1975). In the marked setting, $\lambda$ and $A$ are collections $\{\lambda(t, x_i)\}$ and $\{A(t, x_i)\}$ of conditional intensities and compensators, respectively, so that each $N_i(t) - A_i(t)$ is an $\mathcal{F}$-martingale. We assume throughout that the t.m.p.p. $N$ admits a conditional intensity $\lambda$.

A t.m.p.p. is *simple* if with probability one, all its points are unique, i.e., no two points occur at the same time and mark. We say a t.m.p.p. is *completely simple* (or, in the terminology of Daley and Vere-Jones, 2003, has *simple ground process*), if with probability 1, all its points occur at distinct times. For a completely simple t.m.p.p. $N$, the conditional intensity $\lambda$ completely characterizes the finite-dimensional distributions of $N$ (Daley and Vere-Jones, 2003). Hence in modeling $N$ it suffices to prescribe a model for $\lambda$.

The most elementary way to characterize uniquely a non-simple marked point process is via a change in the mark space $\mathcal{X}$, as follows; see also exercise 7.1.6 of Daley and Vere-Jones (1988) or p. 195 of Daley and Vere-Jones (2003).

Let $\mathcal{X}^\cup$ denote the collection of all subsets of $\mathcal{X}$, i.e., $\mathcal{X}^\cup = \cup_{k=1}^{\infty} \mathcal{X}^k$, with $\mathcal{X}^k = \mathcal{X} \times \cdots \times \mathcal{X}$, the collection of all possible combinations of $k$ marks.

Let the function $\phi_i(\underline{x})$ denote the multiplicity of the mark $x_i$ in the vector $\underline{x}$, for $x_i \in \mathcal{X}$ and $\underline{x}$ in $\mathcal{X}^\cup$, and let $\phi(\underline{x}) = \sum_i \phi_i(\underline{x})$. That is, if $\underline{x} = \{x_1^{a_1}, x_2^{a_2}, \ldots\}$, then $\phi_i(\underline{x}) = a_i$ for each $i$, and $\phi(\underline{x}) = \sum_i a_i$.

Define the function $\psi(\underline{x})$ as the number of distinct marks in $\underline{x}$. That is, $\psi(\underline{x}) := \sum_i \mathbf{I}_{\{\phi_i(\underline{x})>0\}}$, with $\mathbf{I}$ the indicator function.

**Definition 2.1** Given a t.m.p.p. $N$ on the mark space $\mathcal{X}$, define the simple t.m.p.p. $N^\cup$ on the expanded mark space $\mathcal{X}^\cup$ as the collection of temporal processes $\{N_{\underline{x}}^\cup\}$, for $\underline{x} \in \mathcal{X}^\cup$, each defined via

$$N_{\underline{x}}^\cup(\{t\}) := \prod_i \mathbf{I}_{\{N_i(\{t\})=\phi_i(\underline{x})\}}, \tag{2}$$

for each $t$. Thus $N_{\underline{x}}^\cup$ has a single point at time $t$ iff. $N$ has $a_i$ points at time $t$ and mark $x_i$ for each $i$, where $\underline{x} = \{x_1^{a_1}, x_2^{a_2}, \ldots\}$.

It is easy to see that $N^\cup$ is a completely simple point process, for any t.m.p.p. $N$. Indeed, it is immediate from Eq. (2) that $N^\cup$ is non-negative and integer-valued. Since for any Borel $B \subset (0, \infty)$, $\sum_{\underline{x}\in\mathcal{X}^\cup} N_{\underline{x}}^\cup(B) \le \sum_i N_i(B)$ and $N_{\underline{x}}^\cup(B) \le$

$N_i(B)$ if $\phi_i(\underline{x}) > 0$, $N^{\cup}$ inherits its $\sigma$-finiteness from $N$. Therefore $N^{\cup}$ is a t.m.p.p., and its complete simplicity follows directly from the construction in Eq. (2). The measurability and uniqueness of the mapping from $N$ to $N^{\cup}$ is established on pp. 208–209 of Daley and Vere-Jones (1988). It thus follows that the conditional intensity $\lambda^{\cup}$ of $N^{\cup}$ (assuming it exists) uniquely characterizes the finite-dimensional distributions of $N^{\cup}$ and hence those of $N$ as well.

Let $\lambda^{\cup}(t, \mathbf{x})$ denote the conditional intensity of the simple process $N^{\cup}$ on the expanded mark space associated with $N$, and define $\lambda^{\cup}(t) := \sum_{\underline{x}\in\mathcal{X}^{\cup}} \lambda^{\cup}(t, \underline{x})$ as the total ground rate of $N^{\cup}$. In subsequent sections we will make use of the ratio $p^{\cup}(t, \underline{x}) := \lambda^{\cup}(t, \underline{x})/\lambda^{\cup}(t)$, which represents the conditional probability of marks $\underline{x}$ at time $t$ given the history of the process and the fact that at least one point occurs at time $t$. Of use as well is the sum $\bar{\lambda}^{\cup}(t, x_i)$ of the expanded conditional intensities for a particular mark $x_i$, i.e., $\bar{\lambda}^{\cup}(t, x_i) := \sum_{\underline{x}:\phi_i(\underline{x})>0} \lambda^{\cup}(t, \underline{x})$.

Our first result relates the two conditional intensities, $\lambda$ and $\lambda^{\cup}$.

**Theorem 2.1** *Suppose the t.m.p.p. $N$ has conditional intensity $\lambda$ and that its associated process $N^{\cup}$ has conditional intensity $\lambda^{\cup}$. For each $i$, for almost all $t$,*

$$\lambda(t, x_i) = \sum_{\underline{x}\in\mathcal{X}^{\cup}} \phi_i(\underline{x})\lambda^{\cup}(t, \underline{x}). \tag{3}$$

*Proof* Fix $i$. Observe from Eq. (2) that $N_i(t) = \sum_{\underline{x}\in\mathcal{X}^{\cup}} \phi_i(\underline{x}) N^{\cup}_{\underline{x}}(t)$. So

$$N_i(t) - \int_0^t \sum_{\underline{x}\in\mathcal{X}^{\cup}} \phi_i(\underline{x})\lambda^{\cup}(u, \underline{x})\, \mathrm{d}u = \sum_{\underline{x}\in\mathcal{X}^{\cup}} \phi_i(\underline{x}) \left[ N^{\cup}_{\underline{x}}(t) - \int_0^t \lambda^{\cup}(u, \underline{x})\, \mathrm{d}u \right],$$

which is a linear combination of $\mathcal{F}$-martingales and is therefore itself an $\mathcal{F}$-martingale. Hence the sum in Eq. (3) is an $\mathcal{F}$-conditional intensity of $N_i$, and thus coincides almost everywhere with $\lambda(t, x_i)$ by the uniqueness theorem for point process compensators (Jacod, 1975). □

As an alternative, one may consider describing a non-simple t.m.p.p. $N$ in terms of a sequence of point processes whose points are identical to those of $N$ but whose multiplicities are the powers of those of $N$, as defined below.

**Definition 2.2** For $i, j = 1, 2, \ldots,$ let $N_i^{(j)}(B) = \sum_{\underline{x}} \int_B [\phi_i(\underline{x})]^j dN^{\cup}_{\underline{x}}$.

Each $N^{(j)}$ is an $\mathcal{F}$-adapted t.m.p.p. provided the same is true of $N$, and one may consider whether the collection of their conditional intensities $\{\lambda^{(j)}(t, x_i); i, j, = 1, 2, \ldots\}$ uniquely determines the distribution of $N$. The answer in general is no, since $N$ may have simultaneous points at different marks, and $\{\lambda^{(j)}(t, x_i)\}$ do not uniquely determine the likelihood of such occurrences for each combination of marks. However, a separability condition under which the conditional intensities of $N^{(j)}$ do uniquely characterize $N$ is given in Sect. 3.

## 3 Separable conditional intensities

Theorem 2.1 relates the conditional intensity $\lambda^{\cup}$ of the associated simple process $N^{\cup}$ to that of the non-simple point process $N$. Note that the relation between the simple point process $N^{\cup}$ on the expanded mark space to the non-simple point process $N$ on the original mark space is analogous to the relation between a (possibly infinite-dimensional) multivariate distribution to its marginal distributions, where attention is restricted exclusively to non-negative integer-valued distributions.

Certain special cases of non-simple point processes are worthy of closer attention. One natural case to consider is the following.

**Definition 3.1** $N$ has *separable* expanded conditional intensity $\lambda^{\cup}$ if, for all $\underline{x} \in \mathcal{X}^{\cup}$ and almost all $t \in (0, \infty)$,

$$\begin{aligned} \lambda^{\cup}(t, \underline{x}) &= \left(\lambda^{\cup}(t)\right)^{(1-\psi(\underline{x}))} \prod_{\underline{x}'=\{x_i^a\}; a=\phi_i(\underline{x})>0} \lambda^{\cup}(t, \underline{x}') \\ &= \lambda^{\cup}(t) \prod_{\underline{x}'=\{x_i^a\}; a=\phi_i(\underline{x})>0} p^{\cup}(t, \underline{x}'). \end{aligned} \tag{4}$$

Note that relation 4 implies that for all $\underline{x} = \{x_1^{a_1}, x_2^{a_2}, \ldots, x_k^{a_k}\}$, for almost all $t$,

$$\lambda^{\cup}(t, \{x_1^{a_1}, x_2^{a_2}, \ldots, x_k^{a_k}\}) = \lambda^{\cup}(t, \{x_1^{a_1}, \ldots, x_{k-1}^{a_{k-1}}\}) \;\; p^{\cup}(t, \{x_k^{a_k}\}).$$

Heuristically, separability means that the likelihood of $N$ simultaneously having $a_1$ overlapping points at mark $x_1$, and $a_2$ overlapping points at mark $x_2$, etc., is proportional to the product of the likelihoods of each of these phenomena occurring individually. Note also that if $\lambda^{\cup}$ is separable then the finite-dimensional distributions of $N$ are completely determined by the collection of processes $\{\lambda^{\cup}(t, \{x^a\})\}$ alone.

**Definition 3.2** $\lambda^{\cup}$ is *completely separable* if it is separable and also, for all $x \in \mathcal{X}$ and integers $a > 0$, for almost all $t$, $\lambda^{\cup}(t, \{x^a\}) = \left(\lambda^{\cup}(t)\right)^{(1-a)} \left(\lambda^{\cup}(t, \{x\})\right)^a = \lambda^{\cup}(t) \left(p^{\cup}(t, \{x\})\right)^a$.

The complete separability of $\lambda^{\cup}$ implies that if $\underline{x} = \{x_1^{a_1}, \ldots, x_k^{a_k}\}$ with $\sum a_k = a$, then $\lambda^{\cup}(t, \underline{x}) = \left(\lambda^{\cup}(t)\right)^{(1-a)} \prod_{i=1}^k \left(\lambda^{\cup}(t, \{x_i\})\right)^{a_i}$, a.e. With complete separability, the addition into $\underline{x}$ of a new term $x_i$ (or equivalently the addition of one to $a_i$ for some $i$) results in the multiplication of $\lambda^{\cup}(t, \underline{x})$ by $p^{\cup}(t, \{x_i\})$, a.e. If $\lambda^{\cup}$ is completely separable, then $\lambda^{\cup}$ is governed (almost everywhere) by the collection of processes $\{\lambda^{\cup}(t, \{x\})\}$ alone.

Daley and Vere-Jones, (2003, p. 195) define point processes with independent marks or with unpredictable marks. These concepts involve the mark distribution's independence with respect to either all other points or all previous points of the process, respectively. The concept of complete separability is in a similar vein, in that the complete separability of $\lambda^{\cup}$ expresses a sort of independence of the marks from one another. For almost all $t$, the case where $\lambda^{\cup}$ is completely separable corresponds to the situation where, given that exactly $m > 0$ points occur at a certain time $t$, these $m$ marks are all independent and identically distributed.

We now relate the separability condition to the processes defined at the end of Sect. 2.

**Theorem 3.1** *Suppose $N$ is a t.m.p.p. with separable expanded conditional intensity $\lambda^{\cup}$, and suppose that a conditional intensity $\lambda^{(j)}(t, x_i)$ as defined following Definition 2.2 exists for any positive integers $i$ and $j$. Then the collection of processes $\{\lambda^{(j)}(t, x_i)\ ;\ i, j = 1, 2, \ldots\}$ uniquely determine the finite-dimensional distributions of $N$.*

*Proof* For each $j = 1, 2, \ldots$, and any $\mathcal{F}$-predictable process $Y(t, x)$ on $(0, \infty) \times \mathcal{X}$,

$$\begin{aligned} &E \sum_i \int Y(t, x_i) \lambda^{(j)}(t, x_i)\, \mathrm{d}t \\ &= E \sum_i \int Y(t, x_i)\, \mathrm{d}N_i^{(j)}(t) \\ &= E \sum_i \sum_{\underline{x}} \int \left(\phi_i(\underline{x})\right)^j Y(t, x_i)\, \mathrm{d}N_{\underline{x}}^{\cup}(t) \\ &= E \sum_i \sum_{\underline{x}} \int \left(\phi_i(\underline{x})\right)^j Y(t, x_i) \lambda^{\cup}(t, \underline{x})\, \mathrm{d}t. \end{aligned}$$

So for $i, j = 1, 2, \ldots$, for almost all $t$,

$$\begin{aligned} \lambda^{(j)}(t, x_i) &= \sum_{\underline{x}} \left(\phi_i(\underline{x})\right)^j \lambda^{\cup}(t, \underline{x}) \\ &= \sum_{\underline{x}} \left(\phi_i(\underline{x})\right)^j \left(\lambda^{\cup}(t)\right)^{(1-\psi(\underline{x}))} \prod_{\underline{x}' = \{x_k^a\}; a = \phi_k(\underline{x}) > 0} \lambda^{\cup}(t, \underline{x}') \\ &= c_i \sum_{a=1}^{\infty} a^j \lambda^{\cup}(t, \{x_i^a\}), \end{aligned}$$

where

$$c_i = 1 + \sum_{\underline{x}: \phi_i(\underline{x}) = 0} \left(\lambda^{\cup}(t)\right)^{(2-\psi(\underline{x}))} \prod_{\underline{x}' = \{x_k^b\}; b = \phi_k(\underline{x}) > 0} \lambda^{\cup}(t, \underline{x}') > 0.$$

Hence for any $i$ and almost all $t$, the collection $\{\lambda^{(j)}(t, x_i);\ j = 1, 2, \ldots, \}$ uniquely determines $\{\lambda^{\cup}(t, \{x_i^a\});\ a = 1, 2, \ldots\}$ (see e.g. chap. 6 of Berman and Fryer, 1972); thus for almost all $t$, $\{\lambda^{(j)}(t, x_i);\ i, j = 1, 2, \ldots\}$ uniquely determines $\{\lambda^{\cup}(t, \{x_i^a\});\ i, a = 1, 2, \ldots\}$. Since $\lambda^{\cup}$ is separable, this implies that $\{\lambda^j(t, x_i)\}$ determines $\lambda^{\cup}$ a.e. and thus the finite-dimensional distributions of $N$ as well. □

Under the assumption that the expanded conditional intensity $\lambda^{\cup}$ is completely separable, the ordinary conditional intensity $\lambda$ can be written directly in terms of the expanded conditional intensity, as in the following result.

**Theorem 3.2** *Suppose $N$ is a t.m.p.p. with conditional intensity $\lambda$ and completely separable $\lambda^{\cup}$. Then for all $i$ and almost all $t$,*

$$\lambda(t, x_i) = \frac{2\, \lambda^{\cup}(t, \{x_i\})\, \lambda^{\cup}(t)}{\lambda^{\cup}(t) - \lambda^{\cup}(t, \{x_i\})}. \tag{5}$$

*Proof* The following proof makes repeated use of the note following Definition 3.2, as well as the idea that for any $i$, the set $\mathcal{X}^{\cup}$ corresponds one-to-one with the collection $\{\underline{x} : \phi_i(\underline{x}) > 0; \phi(\underline{x}) > 1\}$, where the correspondence is simply the augmentation by one of the multiplicity $\phi_i(\underline{x})$.

Fix $i$. We shall suppress $t$ for brevity, as what follows is true for almost all $t$.

Consider the collection of all $\underline{x}$ such that the multiplicity $\phi_i(\underline{x})$ of the mark $x_i$ is positive. Complete separability implies that the sum of $\lambda^{\cup}_{\underline{x}}$ over all such $\underline{x}$ is given by

$$\begin{aligned}
\bar{\lambda}^{\cup}(x_i) &= \sum_{\underline{x}:\phi_i(\underline{x})>0} \lambda^{\cup}(\underline{x}) \\
&= \lambda^{\cup}(\{x_i\}) + \sum_{\underline{x}:\phi_i(\underline{x})>0;\phi(\underline{x})>1} \lambda^{\cup}(\underline{x}) \\
&= \lambda^{\cup}(\{x_i\}) + \sum_{\underline{x}:\phi(\underline{x})>0} p^{\cup}(\{x_i\})\lambda^{\cup}(\underline{x}) \\
&= \lambda^{\cup}(\{x_i\}) + p^{\cup}(\{x_i\})\lambda^{\cup} \\
&= 2\lambda^{\cup}(\{x_i\}) \qquad (6)
\end{aligned}$$

One may now write $\lambda_i$ in terms of $\lambda^{\cup}$, summing over all possible $\underline{x} \in \mathcal{X}^{\cup}$, as follows:

$$\begin{aligned}
\lambda(x_i) &= \sum_{j=1}^{\infty} \sum_{\underline{x}:\phi_i(\underline{x})=j} j\lambda^{\cup}(\underline{x}) \\
&= \sum_{j=1}^{\infty} j\lambda^{\cup}(\{x_i^j\}) + \sum_{j=1}^{\infty} \sum_{\underline{x}:j=\phi_i(\underline{x})<\phi(\underline{x})} j\lambda^{\cup}(\underline{x}) \\
&= \lambda^{\cup} \sum_{j=1}^{\infty} j\left(p^{\cup}(\{x_i\})\right)^j + \sum_{j=1}^{\infty} \sum_{\underline{x}':\phi_i(\underline{x}')=0} j\left(p^{\cup}(\{x_i\})\right)^j \lambda^{\cup}(\underline{x}') \qquad (7) \\
&= \left[\sum_{j=1}^{\infty} j\left(p^{\cup}(\{x_i\})\right)^j\right]\left[\lambda^{\cup} + (\lambda^{\cup} - \bar{\lambda}^{\cup}(x_i))\right] \qquad (8) \\
&= \frac{p^{\cup}(\{x_i\})}{[1 - p^{\cup}(\{x_i\})]^2}\left[2\lambda^{\cup} - 2\lambda^{\cup}(\{x_i\})\right] \qquad (9) \\
&= \frac{2\lambda^{\cup}(\{x_i\})\lambda^{\cup}}{\lambda^{\cup} - \lambda^{\cup}(\{x_i\})},
\end{aligned}$$

where the observation that $\sum_{\underline{x}':\phi_i(\underline{x}')=0} \lambda^{\cup}(\underline{x}') = \lambda^{\cup} - \bar{\lambda}^{\cup}(x_i)$ is used to go from Eq. (7) to (8), and the relation $\bar{\lambda}^{\cup}(x_i) = 2\lambda^{\cup}(\{x_i\})$ of Eq. (6) is used to go from Eq. (8) to (9), while the definition $p^{\cup}(\{x_i\}) = \lambda^{\cup}(\{x_i\})/\lambda^{\cup}$ and some elementary cancellations are used for the last step. □

We now consider the case where $N$ may have multiple points simultaneously, but where no two such points may occur at the same mark. We previously defined such a process as simple, but to distinguish it from a completely simple process, in what follows we will use the term singular instead.

**Definition 3.3** $\lambda^{\cup}$ is *singular* if for all $\underline{x}$ such that $\phi_i(\underline{x}) > 1$ for any $i$, $\lambda^{\cup}(t, \underline{x}) = 0$ a.e.

It is obvious that, assuming separability and given the same values of the expanded conditional intensity $\lambda^{\cup}(t, \{x_i\})$, the ordinary conditional intensity $\lambda(t, x_i)$ will generally be smaller for a singular t.m.p.p. than in the non-singular case, since singularity prohibits the existence of multiple points at mark $x_i$. Interestingly, for a singular t.m.p.p. with separable expanded conditional intensity, the form of the ordinary conditional intensity of points of mark $x_i$ is very similar to that of processes with completely separable expanded conditional intensities, except that the difference between the two terms in the denominator of Eq. (5) is replaced by their sum. This is established in the next result.

**Theorem 3.3** *Suppose $N$ is a singular t.m.p.p. with conditional intensity $\lambda$ and separable $\lambda^{\cup}$. Then for all $i$, for almost all $t$,*

$$\lambda(t, x_i) = \frac{2\lambda^{\cup}(t, \{x_i\})\lambda^{\cup}(t)}{\lambda^{\cup}(t) + \lambda^{\cup}(t, \{x_i\})}. \tag{10}$$

*Proof* Again, we suppress $t$; the following is true for almost all $t$.

$$\begin{aligned}\lambda(x_i) &= \lambda^{\cup}(\{x_i\}) + \sum_{\underline{x}:\phi_i(\underline{x})=1;\phi(\underline{x})>1} \lambda^{\cup}(\underline{x}) \\ &= \lambda^{\cup}(\{x_i\}) + p^{\cup}(\{x_i\}) \sum_{\underline{x}:\phi_i(\underline{x})=0} \lambda^{\cup}(\underline{x}) \\ &= \lambda^{\cup}(\{x_i\}) + p^{\cup}(\{x_i\}) \left(\lambda^{\cup} - \lambda(x_i)\right).\end{aligned}$$

So $\lambda(x_i)\left(1 + \frac{\lambda^{\cup}(\{x_i\})}{\lambda^{\cup}}\right) = \lambda^{\cup}(\{x_i\}) + \lambda^{\cup} p^{\cup}(\{x_i\}) = 2\lambda^{\cup}(\{x_i\})$, which establishes Eq. (10). □

We now consider the antithesis of singularity, i.e. the case where $N$ may have multiple points at a given time, but such multiple points must all occur at the same mark.

**Theorem 3.4** *Suppose $N$ is a t.m.p.p. with conditional intensity $\lambda$ and that $\lambda^{\cup}(t, \underline{x}) = 0$ a.e. for all $\underline{x}$ such that $\phi(\underline{x}) > \phi_i(\underline{x})$ for all $i$. Suppose also that for almost all $t$, $\lambda^{\cup}(\{x_i^{a_i}\}) = (\lambda^{\cup})^{1-a_i}\left(\lambda^{\cup}(\{x_i\})\right)^{a_i}$. Then for each $i$, for almost all $t$,*

$$\lambda(t, x_i) = \frac{\lambda^{\cup}(t, \{x_i\})\left(\lambda^{\cup}(t)\right)^2}{(\lambda^{\cup}(t) - \lambda^{\cup}(t, \{x_i\}))^2}.$$

*Proof* Again we suppress $t$, and the following is true a.e. Under the stated conditions, one need only consider elements of $\mathcal{X}^{\cup}$ of the form $\underline{x} = \{x_i^k\}$. Hence

$$\begin{aligned}\lambda(x_i) &= \lambda^{\cup}(\{x_i\}) + 2\lambda^{\cup}(\{x_i^2\}) + 3\lambda^{\cup}(\{x_i^3\}) + \cdots \\ &= \lambda^{\cup} \sum_{k=1}^{\infty} k\left(p^{\cup}(\{x_i\})\right)^k \\ &= \frac{\lambda^{\cup} p^{\cup}(\{x_i\})}{(1 - p^{\cup}(\{x_i\}))^2} \\ &= \frac{\lambda^{\cup}(\{x_i\})(\lambda^{\cup})^2}{(\lambda^{\cup} - \lambda^{\cup}(\{x_i\}))^2}.\end{aligned}$$

□

## 4 Transformations

One way to transform a non-simple point process into a simple one is via Definition 2.1, i.e., by expanding the mark space. Alternatively, one may randomly transform the process to obtain a Poisson process or a process with completely separable expanded conditional intensity, as in the following three results.

**Theorem 4.1** *Suppose $N$ is a t.m.p.p. with expanded conditional intensity $\lambda^{\cup}$, such that for each $\underline{x} \in \mathcal{X}^{\cup}$, $\int_0^{\infty} \lambda^{\cup}(u, \underline{x})du = \infty$. Then the time transformation which moves each point of $N^{\cup}$ from $(t, \underline{x})$ to $\left(\int_0^t \lambda^{\cup}(u, \underline{x})du, \underline{x}\right)$ results in a sequence $\{\tilde{N}_{\underline{x}}; \underline{x} \in \mathcal{X}^{\cup}\}$ of independent Poisson processes of unit rate.*

*Proof* Since $N^{\cup}$ is completely simple, the result follows from application of the random time change theorem of Meyer (1971). □

Our final two results involve randomly thinning a point process, where the thinning depends on a uniformly distributed random variable that is independent of the point process. Hence we suppose there exists a white noise process $U_t$ on $(\Omega, \mathcal{F}_t, P)$ with $\{U_t; t \geq 0\}$ independent of $N$, and where the $U_t$ are i.i.d. uniformly distributed on $(0, 1)$.

**Theorem 4.2** *Suppose $N$ has expanded conditional intensity $\lambda^{\cup}$ and that for each $i = 1, 2, \ldots$, for almost all $t$, $\bar{\lambda}^{\cup}(t, x_i) > 0$. Let $b(t, x)$ be any strictly positive predictable process on $(0, \infty) \times \mathcal{X}$, independent of $\{U_t\}$, and such that for almost all $t$, $\sum_{x \in \mathcal{X}} c(t, x) < 1$, where $c(t, x_i) := b(t, x_i)/\bar{\lambda}^{\cup}(t, x_i)$. Let $c(t, x_0) = 0$, and consider the transformation $N \to \tilde{N}$ where $\tilde{N}$ has a single point at $(t, x_i)$ provided $N^{\cup}$ has a point at $(t, \underline{x})$ with $\phi_i(\underline{x}) > 0$ and $\sum_{j=0}^{i-1} c(t, x_j) \leq U_t < \sum_{j=0}^{i} c(t, x_j)$. Then $\tilde{N}$ is a completely simple $\mathcal{F}$-adapted marked point process with conditional intensity $\tilde{\lambda}(t, x_i) = b(t, x_i)$.*

*Proof* It is clear that $\tilde{N}$ is an $\mathcal{F}$-adapted point process, and $\tilde{N}$ is completely simple by construction since for any $t$, $\sum_{j=0}^{i-1} c(t, x_j) \leq U_t < \sum_{j=0}^{i} c(t, x_j)$ can be true for at most one $i$. For any $\mathcal{F}$-predictable process $Y(t, x)$ on $(0, \infty) \times \mathcal{X}$,

$$
\begin{aligned}
E \sum_i \int Y(t, x_i)\, \mathrm{d}\tilde{N} &= E \sum_i \sum_{\underline{x}:\phi_i(\underline{x})>0} c(t, x_i) \int Y(t, x_i)\, \mathrm{d}N^{\cup}_{\underline{x}} \\
&= E \sum_i \sum_{\underline{x}:\phi_i(\underline{x})>0} c(t, x_i) \int Y(t, x_i) \lambda^{\cup}(t, \underline{x})\, \mathrm{d}t \\
&= E \sum_i \int Y(t, x_i) c(t, x_i) \sum_{\underline{x}:\phi_i(\underline{x})>0} \lambda^{\cup}(t, \underline{x})\, \mathrm{d}t.
\end{aligned}
$$

Therefore a version of the conditional intensity of $\tilde{N}$ is given by

$$
\begin{aligned}
\tilde{\lambda}(t, x_i) &= c(t, x_i) \sum_{\underline{x}:\phi_i(\underline{x})>0} \lambda^{\cup}(t, \underline{x}) \\
&= \frac{b(t, x_i)}{\bar{\lambda}^{\cup}(t, x_i)} \bar{\lambda}^{\cup}(t, x_i) \\
&= b(t, x_i).
\end{aligned}
$$

□

The previous two results involve rescaling or thinning $N$ in order to form a completely simple point process. We now turn our attention to the problem of transforming $N$ instead into a t.m.p.p. with completely separable expanded conditional intensity.

**Theorem 4.3** *Suppose a t.m.p.p. $N$ has expanded conditional intensity $\lambda^{\cup}$ and that a strictly positive predictable process $b(t, x)$ can be found that is independent of $\{U_t\}$ and so that for all $\underline{x}$ and almost all $t$,*

$$\left(\lambda^{\cup}(t)\right)^{1-\phi(\underline{x})} \prod_i b(t, x_i)^{\phi_i(\underline{x})} \le \lambda^{\cup}(t, \underline{x}).$$

*Let $\tilde{N}$ be a thinned version of $N$ so that $\tilde{N}^{\cup}$ has a point at $(t, \underline{x})$ whenever $N^{\cup}$ has a point at $(t, \underline{x})$, provided $U_t < \left(\lambda^{\cup}(t, \underline{x})\right)^{-1} \left(\lambda^{\cup}(t)\right)^{1-\phi(\underline{x})} \prod_i b(t, x_i)^{\phi_i(\underline{x})}$. Then $\tilde{N}$ is an $\mathcal{F}$-adapted marked point process with completely separable expanded conditional intensity $\tilde{\lambda}^{\cup}$ such that $\tilde{\lambda}^{\cup}(t, \{x_i\}) = b(t, x_i)$.*

*Proof* That $\tilde{N}$ is an $\mathcal{F}$-adapted point process is clear as it inherits the necessary properties directly from $N$, and $\tilde{N}^{\cup}$ is completely simple since the same is true of $N^{\cup}$. Since $N^{\cup}$ has conditional intensity $\lambda^{\cup}$, for any $\mathcal{F}$-predictable process $Y(t, \underline{x})$ on $(0, \infty) \times \mathcal{X}^{\cup}$,

$$\begin{aligned} E \sum_{\underline{x}} \int Y(t, \underline{x})\, \mathrm{d}\tilde{N}^{\cup} &= E \sum_{\underline{x}} \int Y(t, \underline{x}) \left(\lambda^{\cup}(t, \underline{x})\right)^{-1} \left(\lambda^{\cup}(t)\right)^{1-\phi(\underline{x})} \\ &\quad \times \prod_i b(t, x_i)^{\phi_i(\underline{x})}\, \mathrm{d}N^{\cup} \\ &= E \sum_{\underline{x}} \int Y(t, \underline{x}) \left(\lambda^{\cup}(t, \underline{x})\right)^{-1} \left(\lambda^{\cup}(t)\right)^{1-\phi(\underline{x})} \\ &\quad \times \prod_i b(t, x_i)^{\phi_i(\underline{x})} \lambda^{\cup}(t, \underline{x})\, \mathrm{d}t. \end{aligned}$$

Hence $\left(\lambda^{\cup}(t)\right)^{1-\phi(\underline{x})} \prod_i b(t, x_i)^{\phi_i(\underline{x})}$ is a conditional intensity for $\tilde{N}^{\cup}$, which in view of Definition 3.2 establishes the desired result. □

**Acknowledgements** The author thanks David Vere-Jones for numerous stimulating conversations. Thanks also to Jean Jacod for useful conversations on this subject and also for his invitation to Université Paris VI where much of this research was conducted. Further thanks to the associate editor and two anonymous referees for comments which led to substantial improvements. This material is based upon work supported by the National Science Foundation under Grant No. 0306526. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## References

Berman, G., Fryer, K.D. (1972). *Introduction to combinatorics*. New York: Academic
Brémaud, P. (1981). Point processes and queues: martingale dynamics. Berlin Heidelberg New York:vadjust Springer

Brown, T., Nair, M. (1988). Poisson approximations for time-changed point processes, *Stochastic Processes and Their Applications*. *29*, 247–256.
Daley, D.J., Vere-Jones, D. (1988). *An introduction to the theory of point processes*. Berlin Heidelberg New York: Springer
Daley, D.J., Vere-Jones, D. (2003). *An introduction to the theory of point processes*, (2nd ed.). Berlin Heidelberg New York: Springer
Jacod, J. (1975). Multivariate point processes: predictable projections, Radon-Nikodym derivatives, representation of martingales, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, *31*, 235–253.
Kagan, Y.Y. (1994). Observational evidence for earthquakes as a nonlinear dynamic process, *Physica D*, *77*, 160–192.
Merzbach, E., Nualart, D. (1986). A characterization of the spatial Poisson process and changing time, *Annals of Probability*, *14*, 1380–1390.
Meyer, P. (1971). Démonstration simplifée d'un théorème de Knight, *Lecture Notes in Mathematics*, *191*, 191–195.
Nair, M. (1990). Random space change for multiparameter point processes, *Annals of Probability*, *18*, 1222–1231.
Papangelou, F. (1972). Integrability of expected increments of point processes and a related random change of scale, *Transactions of the American Mathematical Society*, *165*, 483–506.
Schoenberg, F. (1999). Transforming spatial point processes into Poisson processes, *Stochastic Processes and Their Applications*, *81*, 155–164.