Generalized Multivariate Hawkes Processes

Tomasz R. Bielecki
Department of Applied Mathematics
Illinois Institute of Technology
Chicago, IL 60616, USA

Jacek Jakubowski
Institute of Mathematics
University of Warsaw
Warszawa, Poland

Mariusz Niewęgłowski
Faculty of Mathematics and Information Science
Warsaw University of Technology
Warszawa, Poland

April 30, 2020

Abstract

This work contributes to the theory and applications of Hawkes processes. We introduce and examine a new class of Hawkes processes that we call generalized Hawkes processes, and their special subclass – the generalized multivariate Hawkes processes (GMHPs). GMHPs are multivariate marked point processes that add an important feature to the family of the (classical) multivariate Hawkes processes: they allow for explicit modelling of simultaneous occurrence of excitation events coming from different sources, i.e. caused by different coordinates of the multivariate process. We study the issue of existence of a generalized Hawkes process, and we provide a construction of a specific generalized multivariate Hawkes process. We investigate Markovian aspects of GMHPs, and we indicate some plausible important applications of GMHPs.

Keywords: Generalized Hawkes processes, generalized multivariate Hawkes process, Hawkes kernel, multivariate marked point process, random measure, predictable compensator, seismology, epidemiology, finance.
1 Introduction

A very interesting and important class of stochastic processes was introduced by Alan Hawkes in [11, 12]. These processes, called now Hawkes processes, are meant to model self-exciting and mutually-exciting random phenomena that evolve in time. The self-exciting phenomena are modeled as univariate Hawkes processes, and the mutually-exciting phenomena are modeled as multivariate Hawkes processes. Hawkes processes belong to the family of marked
point processes, and, of course, a univariate Hawkes process is just a special case of the multivariate one.

In this paper, which originates from Chapter 11 of [5], we define and study generalized multivariate Hawkes processes (GMHPs). These processes constitute a subclass of the family of the generalized Hawkes processes defined in this paper as well. In addition, we provide a novel construction of a generalized multivariate Hawkes process.

GMHPs are multivariate marked point processes that add an important feature to the family of the (classical) multivariate Hawkes processes: they allow for explicit modelling of simultaneous occurrence of excitation events coming from different sources, i.e. caused by different coordinates of the multivariate process. The importance of this feature is rather intuitive, and it will be illustrated in Section 6. In this regard, GMHPs differ from the multivariate Hawkes processes that were studied in Bremaud and Massouli [6] and Liniger [24].

We need to stress that we limit ourselves here to the case of linear GMHPs, that are counterpart of the linear classical Hawkes processes. That is to say, we do not study here what would be a counterpart of the nonlinear classical Hawkes processes. We refer to e.g. Chapter 1 in [33] for comparison of linear and nonlinear Hawkes processes. We also note that the generalized Hawkes processes introduced here should not be confused with those studied in [32]. In particular, we do not introduce any additional random factors, such as Brownian motions, into the compensators of the multivariate marked point process $N$ showing in the Definition 3.1 below.

We also need to stress that we are not concerned in this study with stationarity and spectral properties of the GMHPs. This is the reason why in the definition of the Hawkes kernel $\kappa$, of the generalized Hawkes process, we use integration over the interval $(0,t)$ rather than integration over $(-\infty,t)$. Please see also Remark 2.2 in this regard.

The paper is organized as follows. In Section 2 we define, prove existence of and provide some discussion of a generalized Hawkes process. Section 3 is devoted to study of the main object of this paper, namely the generalized multivariate Hawkes process. In Section 4 we provide a mathematical construction of and computational pseudo-algorithm for simulation of a generalized multivariate Hawkes process with deterministic kernels $\eta$ and $f$ (cf. (2.4)). Markovian aspects of a generalized multivariate Hawkes process are discussed in Section 5. Section 6 contains a brief description of possible applications of generalized multivariate Hawkes processes in seismology, epidemiology and finance. Finally, in the Appendix, we provide some needed technical results.

In this paper we use various concepts and results from stochastic analysis. For a comprehensive study of these concepts and results we refer to e.g. [16], [22] and [18].

## 2 Generalized Hawkes process

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{X}, \mathcal{A})$ be a Borel space. We take $\partial$ to be a point external to $\mathcal{X}$, and we let $\mathcal{X}^\partial := \mathcal{X} \cup \partial$. On $(\Omega, \mathcal{F}, \mathbb{P})$ we consider a marked point process $N$ with mark space $\mathcal{X}$, that is, a sequence of random elements

\[ N = ((T_n, X_n))_{n \geq 1}, \quad (2.1) \]
where for each $n$:

1. $T_n$ is a random variable with values in $(0, \infty]$,
2. $X_n$ is a random variable with values in $\mathcal{X}^0$,
3. $T_n \leq T_{n+1}$, and if $T_n < +\infty$ then $T_n < T_{n+1}$,
4. $X_n = \partial$ iff $T_n = \infty$.

The explosion time of $N$, say $T_\infty$, is defined as

$$T_\infty := \lim_{n \to \infty} T_n.$$ 

Following the typical techniques used in the theory of Marked Point Processes (MPPs), in particular following Section 1.3 in [22], we associate with the process $N$ an integer-valued random measure on $(\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X})$, also denoted by $N$ and defined as

$$N(dt, dx) := \sum_{n \geq 1} \delta_{(T_n, X_n)}(dt, dx) \mathbb{1}_{\{T_n < \infty\}}, \quad (2.2)$$

so that

$$N((0, t], A) = \sum_{n \geq 1} \mathbb{1}_{\{T_n \leq t\}} \mathbb{1}_{\{X_n \in A\}},$$

where $A \in \mathcal{X}$.

Let $\mathbb{F}^N$ be the natural filtration of $N$, so $\mathbb{F}^N := (\mathcal{F}_t^N, t \geq 0)$, where $\mathcal{F}_t^N$ is the $\mathbb{P}$-completed $\sigma$-field $\sigma(N((s, t] \times A) : 0 \leq s < r \leq t, A \in \mathcal{X}), t \geq 0$. In view of Theorem 2.2.4 in [22] the filtration $\mathbb{F}^N$ satisfies the usual conditions. Moreover, $N$ is $\mathbb{F}^N$-optional, so, using Proposition 4.1.1 in [22] we conclude that $T_n$’s are $\mathbb{F}^N$-stopping times and $X_n$ are $\mathcal{F}_{T_n}$-measurable. In what follows we denote by $\mathcal{P}$ the $\mathbb{F}^N$-predictable $\sigma$-field.

We recall that for a given filtration $\mathbb{F}$ a stochastic process $X : \Omega \times [0, \infty) \to \mathbb{R}$ is said to be $\mathbb{F}$-predictable if it is measurable with respect to the predictable sigma field $\mathcal{P}$ on $\Omega \times [0, \infty)$, which is generated by $\mathbb{F}$-adapted processes whose paths are continuous (equivalently left-continuous, with the left limit at $t = 0$ defined as the value of the path at $t = 0$) functions of time variable. More generally, a function $X : \Omega \times [0, \infty) \times \mathcal{X} \to \mathbb{R}$ is said to be $\mathbb{F}$-predictable function if it is measurable with respect to the predictable sigma field $\mathcal{P} \otimes \mathcal{X}$ on $\Omega \times [0, \infty) \times \mathcal{X}$. The sigma field $\mathcal{P} \otimes \mathcal{X}$ is generated by the sets $A \times \{0\} \times \mathcal{X}$ where $A \in \mathcal{F}_0$ and the sets of the form $B \times \{s, t\} \times D$ where $0 < s \leq t$, $B \in \mathcal{F}_s$ and $D \in \mathcal{X}$.

We now consider a random measure $\nu$ on $(\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X})$ defined as

$$\nu(\omega, dt, dy) := \mathbb{1}_{[0, T_\infty]}(\omega)(t) \kappa(\omega, t, dy) dt, \quad (2.3)$$

where, for $A \in \mathcal{X}$,

$$\kappa(t, A) = \eta(t, A) + \int_{(0, t] \times \mathcal{X}} f(t, s, x, A) N(ds, dx), \quad (2.4)$$

$\eta$ is a finite kernel from $(\Omega \times [0, \infty), \mathcal{P})$ to $(\mathcal{X}, \mathcal{X})$, and $f$ is a kernel from $(\Omega \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{X}, \mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X})$ to $(\mathcal{X}, \mathcal{X})$.

We assume also that $f$ is a kernel satisfying:

\textsuperscript{3}See Appendix A.2 in Last and Brandt [22] for the definition of the kernel.
1. \( f(t, s, x, A) = 0 \) for \( s \geq t \),

2. \( \theta \) defined as

\[
\theta(t, A) := \int_{(0,t) \times \mathcal{X}} f(t, s, x, A)N(ds, dx), \quad t \geq 0, \ A \in \mathcal{X},
\]

is a kernel from \((\Omega \times [0, \infty), \mathcal{P})\) to \((\mathcal{X}, \mathcal{X})\), which is finite for \( t < T_\infty \).

Clearly, we have

\[
\theta(t, A) = \sum_{n: T_n < t} f(t, T_n, X_n, A).
\] (2.5)

Note that \( \kappa(t, \mathcal{X}) \) is finite for any \( t < T_\infty \). We additionally assume that \( \kappa(t, \mathcal{X}) > 0 \) for all \( t \geq 0 \) and that the integral \( \int_{[0,t]} \kappa(s, A)ds \) is finite for any \( A \in \mathcal{X} \) and any \( t < T_\infty \). This last assumption is satisfied under mild boundedness conditions imposed on \( \eta \) and \( f \).

Before we proceed we recall that for a given filtration \( \mathbb{F} \) the random measure \( \nu \) is said to be \( \mathbb{F} \)-compensator of a random measure \( N \) if it is \( \mathbb{F} \)-predictable random measure such that it holds

\[
\mathbb{E} \int_0^\infty \int_{E^\Delta} F(v, x)N(dv, dx) = \mathbb{E} \int_0^\infty \int_{E^\Delta} F(v, x)\nu(dv, dx)
\]

for every non-negative \( \mathbb{F} \)-predictable function \( F : \Omega \times [0, \infty) \times \mathcal{X} \to \mathbb{R} \).

We are ready to state the underlying definition in this paper.

**Definition 2.1.** Let \( N \) be the marked point process introduced in (2.1) with the corresponding random measure \( N \) defined in (2.2). We call \( N \) a generalized Hawkes process on \((\Omega, \mathcal{F}, \mathbb{P})\), if the \((\mathbb{F}^N, \mathbb{P})\)-compensator of \( N \), say \( \nu \), is of the form (2.3). The kernel \( \kappa \) is called the Hawkes kernel for \( N \).

We use this convention here since we are not considering stationarity and spectral properties of the generalized Hawkes processes.

**Remark 2.2.**

(i) Recall that the compensator of a random measure is unique (up to equivalence). Thus, the compensator \( \nu \) of \( N \) is unique. However, the representation (2.3)-(2.4) is not unique, by any means, in general. For any given \( \eta \) and \( f \) in the representation (2.3)-(2.4), one can always find \( \tilde{\eta} \neq \eta \) and \( \tilde{f} \neq f \) such that

\[
\kappa(t, dy) = \tilde{\eta}(t, dy) + \int_{(0,t) \times \mathcal{X}} \tilde{f}(t, s, x, dy)N(ds, dx).
\] (2.6)

(ii) With a slight abuse of terminology we refer to \( \kappa \) as to the Hawkes intensity kernel of \( N \). Accordingly, we refer to the quantity \( \kappa(t, A) \) as to the intensity at time \( t \) of the event
Remark 2.4. Since $\mathbb{F}_0^N$ is a completed trivial $\sigma$-field, then it is a consequence of Theorem 3.6 in [17] that the compensator $\nu$ determines the law of $N$ under $\mathbb{P}$, and, consequently, the Hawkes kernel $\kappa$ determines the law of $N$ under $\mathbb{P}$. □

2.1 Existence of a generalized Hawkes process

We will now demonstrate that for an arbitrary measure $\nu$ of the form (2.3) there exists a Hawkes process having $\nu$ as $\mathbb{F}_N^N$–compensator. Towards this end we will consider the underlying canonical space. Specifically, we take $(\Omega, \mathcal{F})$ to be the canonical space of multivariate marked point processes with marks taking values in $X^\partial$. That is, $\Omega$ consists of elements $\omega = ((t_n, x_n))_{n \geq 1}$, satisfying $(t_n, x_n) \in (0, \infty] \times X^\partial$ and

$$t_n \leq t_{n+1};$$
if $t_n < \infty$, then $t_n < t_{n+1};$
$t_n = \infty$ iff $x_n = \partial.$

The $\sigma$-field $\mathcal{F}$ is defined to be the smallest $\sigma$-field on $\Omega$ such that the mappings $T_n : \Omega \rightarrow ([0, \infty], \mathcal{B}[0, \infty]), X_n : \Omega \rightarrow (X^\partial, \mathcal{X}^\partial)$ defined by

$$T_n(\omega) := t_n, \quad X_n(\omega) := x_n$$

are measurable for every $n$.

Note that the canonical space introduced above agrees with the definition of canonical space considered in [22] (see Remark 2.2.5 therein). On this space we denote by $N$ a sequence of measurable mappings

$$N = ((T_n, X_n))_{n \geq 1}, \quad (2.7)$$

Clearly, these mappings satisfy

1. $T_n \leq T_{n+1}$, and if $T_n < +\infty$ then $T_n < T_{n+1},$

2. $X_n = \partial$ iff $T_n = \infty.$

We call such $N$ a canonical mapping.

The following result provides the existence of a probability measure $\mathbb{P}_\nu$ on $(\Omega, \mathcal{F})$ such that the canonical mapping $N$ becomes a generalized Hawkes process with a given Hawkes kernel $\kappa$, which in a unique way determines the compensator $\nu$.

Theorem 2.5. Consider the canonical space $(\Omega, \mathcal{F})$ and the canonical mapping $N$ given by (2.7). Let measures $N$ and $\nu$ be associated with this canonical mapping through (2.2) and (2.3)–(2.4), respectively. Then, there exists a unique probability measure $\mathbb{P}_\nu$ on $(\Omega, \mathcal{F})$, such that the measure $\nu$ is an $(\mathbb{F}_N, \mathbb{P}_\nu)$–compensator of $N$. So, $N$ is a generalized multivariate Hawkes process on $(\Omega, \mathcal{F}, \mathbb{P}_\nu)$. 

[Note: The document contains mathematical notation and concepts related to Hawkes processes, which are used in the analysis of point processes and their applications in various fields such as finance, neuroscience, and seismology. The text describes the existence of a generalized Hawkes process and provides a framework for understanding the compensator and intensity functions associated with such processes.]
Proof. We will use Theorem 8.2.1 in [22] with \( X = \mathcal{X} \), \( \varphi = \omega \), and with
\[
\bar{\alpha}(\omega, dt) := \nu(\omega, dt, \mathcal{X}) = 1_{[0,T_\infty(\omega)]}(t)\kappa(\omega, t, \mathcal{X})dt,
\]
from which we will conclude the assertion of theorem. Towards this end, we will verify that all assumptions of the said theorem are satisfied in the present case. As already observed, the random measure \( \nu \) is \( F_\infty \)-predictable. Next, let us fix \( \omega \in \Omega \). Given (2.8) we see that \( \bar{\alpha} \) satisfies the following equalities
\[
\bar{\alpha}(\omega, \{0\}) = 0, \quad \bar{\alpha}(\omega, \{t\}) = 0 \leq 1, \quad t \geq 0,
\]
which correspond to conditions (4.2.6) and (4.2.7) in [22], respectively. It remains to show that condition (4.2.8) holds as well, that is
\[
\bar{\alpha}(\omega, [\pi_\infty(\omega), \infty]) = 0,
\]
where
\[
\pi_\infty(\omega) := \inf \{ t \geq 0 : \bar{\alpha}(\omega, (0,t]) = \infty \}.
\]
To see this, we first note that (2.8) implies
\[
\bar{\alpha}(\omega, [T_\infty(\omega), \infty]) = 0.
\]
Thus it suffices to show that \( \pi_\infty(\omega) = T_\infty(\omega) \). By definition of \( \bar{\alpha} \) we can write
\[
\bar{\alpha}(\omega, (0,t]) = \begin{cases} 
\int_0^t \kappa(\omega, s, \mathcal{X})ds, & t < T_\infty(\omega), \\
\int_0^{T_\infty(\omega)} \kappa(\omega, s, \mathcal{X})ds, & t \geq T_\infty(\omega).
\end{cases}
\]
If \( T_\infty(\omega) = \infty \), then we clearly have \( \pi_\infty(\omega) = \infty = T_\infty(\omega) \).

Next, if \( T_\infty(\omega) < \infty \), then \( \lim_{t \uparrow T_\infty(\omega)} \bar{\alpha}(\omega, (0,t]) = a \). We need to consider two cases now: \( a = \infty \) and \( a < \infty \).

If \( a = \infty \), then \( \bar{\alpha}(\omega, (0,t]) = \infty \) for \( t \geq T_\infty(\omega) \), and, \( \bar{\alpha}(\omega, (0,t]) < \infty \) for \( t < T_\infty(\omega) \) in view of our assumptions imposed on \( \kappa \) in the beginning of this section. This implies that \( \pi_\infty(\omega) = T_\infty(\omega) \).

If \( a < \infty \), then \( \bar{\alpha}(\omega, (0,t]) = a < \infty \) for \( t \geq T_\infty(\omega) \), hence \( \pi_\infty(\omega) = \infty \geq T_\infty(\omega) \). Thus, \( \pi_\infty(\omega) \geq T_\infty(\omega) \), which implies that (2.9) holds.

Since \( \omega \) was arbitrary, we conclude that for all \( \omega \in \Omega \) conditions (4.2.6)-(4.2.8) in [22] are satisfied. So, applying Theorem 8.2.1 in [22] with \( \beta = \nu \), we obtain that there exists a unique probability measure \( \mathbb{P}_\nu \) such that \( \nu \) is a \( \mathbb{P}_\nu \)-compensator of \( N \) under \( \mathbb{P}_\nu \). \( \square \)

2.2 Cluster interpretation of the generalized Hawkes processes

The classical Hawkes processes are conveniently interpreted, or represented, in terms of so called clusters. This kind of representation is sometimes called immigration and birth representation. We refer to [14] and [23].

Generalized Hawkes processes also admit cluster representation. The dynamics of cluster centers, or the immigrants, is directed by \( \eta \). Specifically, \( \eta(t, A) \) is the time-\( t \) intensity of
arrivals of immigrants with marks belonging to set $A$. The dynamics of the off-springs is
directed by $f$. Specifically, $f(t, s, x, A)$ represents the time-$t$ intensity of births of offsprings
with marks in set $A$ of either an immigrant with mark $x$ who arrived at time $s$, or of an
offspring with mark $x$ who was born at time $s$.

The cluster interpretation will be exploited in a follow-up work for asymptotic analysis
of generalized Hawkes processes.

3 Generalized multivariate Hawkes process

We now introduce the concept of a generalized multivariate Hawkes process, which is a
particular case of the concept of a generalized Hawkes process.

3.1 Definition

We first construct an appropriate mark space. Specifically, we fix an integer $d \geq 1$ and
we let $(E_i, \mathcal{E}_i)$, $i = 1, \ldots, d$, be some non-empty Borel spaces, and $\Delta$ be a dummy mark,
the meaning of which will be explained below. Very often, in practical modelling, spaces $E_i$ are discrete. The instrumental rationale for considering a discrete mark space is that
in most of the applications of the Hawkes processes that we are familiar with and/or we
can imagine, a discrete mark space is sufficient to account for the intended features of the
modeled phenomenon.

We set $E_i^\Delta := E_i \cup \Delta$, and we denote by $\mathcal{E}_i^\Delta$ the sigma algebra on $E_i^\Delta$ generated by $\mathcal{E}_i$. Then, we define a mark space, say $E^\Delta$, as

$$E^\Delta := E_1^\Delta \times E_2^\Delta \times \ldots \times E_d^\Delta \setminus (\Delta, \Delta, \ldots, \Delta).$$

(3.1)

By $\mathcal{E}^\Delta$ we denote a trace sigma algebra of $\otimes_{i=1}^d \mathcal{E}_i^\Delta$ on $E^\Delta$, i.e.

$$\mathcal{E}^\Delta := \{ A \cap E^\Delta : A \in \otimes_{i=1}^d \mathcal{E}_i^\Delta \}.$$\n
Moreover, denoting by $\partial_i$ the point which is external to $E_i^\Delta$, we define $E_i^\partial := E_i^\Delta \cup \{ \partial_i \}$, and we denote by $\mathcal{E}_i^\partial$ the sigma algebra generated by $\mathcal{E}_i$ and $\{ \partial_i \}$. Analogously we define

$$E^\partial := E^\Delta \cup \partial,$$

where $\partial = (\partial^1, \ldots, \partial^d)$ is a point external to $E_1^\Delta \times E_2^\Delta \times \ldots \times E_d^\Delta$ and by $\mathcal{E}^\partial$ we denote the sigma field generated by $\mathcal{E}^\Delta$ and $\{ \partial \}$.

Definition 3.1. A generalized Hawkes process $N = ((T_n, X_n))_{n \geq 1}$ with the mark space $\mathcal{X} = E^\Delta$ given by (3.1), and with $\mathcal{X}^\partial = E^\partial$, is called a generalized multivariate Hawkes process (of dimension $d$).

Note that a necessary condition for generalized Hawkes processes to feature the self-
excitation and mutual-excitation is that $f \neq 0$. We refer to Example 3.9 for interpretation
of the components $\eta$ and $f$ of the kernel $\kappa$ in case of a generalized multivariate Hawkes
process.
We interpret $T_n \in (0, \infty)$ and $X_n \in E^\Delta$ as the event times of $N$ and as the corresponding mark values, respectively. Thus, if $T_n < \infty$ we have\(^2\)

$$X_n = (X^n_i, i = 1, 2, \ldots, d), \quad \text{where} \quad X^n_i \in E^\Delta_i.$$  

Also, we interpret $X^i$ as the marks associated with $i$-th coordinate of $N$ (cf. Definition 3.3). With this interpretation, the equality $X^n_i(\omega) = \Delta$ means that there is no event taking place with regard to the $i$-th coordinate of $N$ at the (general) event time $T_n(\omega)$. In other words, no event occurs with respect to the $i$-th coordinate of $N$ at time $T_n(\omega)$.

**Definition 3.2.** We say that $T_n(\omega)$ is a common event time for a multivariate Hawkes process $N$ if there exist $i$ and $j$, $i \neq j$, such that $X^n_i(\omega) \in E_i$ and $X^n_j(\omega) \in E_j$. We say that process $N$ admits common event times if

$$\mathbb{P}\left(\omega \in \Omega : \exists n \text{ such that } T_n(\omega) \text{ is a common event time} \right) > 0$$

Otherwise we say that process $N$ admits no common event times. \hfill \(\square\)

**Definition 3.2** generalizes that in Bremaud and Massouli [6] and Liniger [24]. In particular, with regard to the concepts of multivariate Hawkes processes studied in Liniger [24], the genuine multivariate Hawkes processes [24] admits no common event times, whereas in the case of pseudo-multivariate Hawkes process [24] all event times are common.

### 3.2 The $i$-th coordinate of a generalized multivariate Hawkes process $N$

We start with

**Definition 3.3.** We define the $i$-th coordinate $N^i$ of $N$ as

$$N^i((0, t], A) := \sum_{n \geq 1} \mathbb{1}_{\{T_n \leq t\}} \mathbb{1}_{\{X_n \in A\}}; \quad (3.2)$$

for $A \in \mathcal{E}_i$ and $t \geq 0$, where

$$A^i = \left( \bigotimes_{j=1}^{i-1} E^\Delta_j \right) \times A \times \left( \bigotimes_{j=i+1}^{d} E^\Delta_j \right). \quad (3.3)$$

Clearly, $N^i$ is a MPP and

$$N^i((0, t], A) = N((0, t], A^i).$$

\(^2\)Note that here $d$ is the number of components in $X_n$, and $n$ is the index of the $n$-th element in the sequence $(X_n)_{n \geq 1}$. 
Indeed, the \(i\)-th coordinate process \(N^i\) can be represented as a sequence \(N^i = (T^i_k, Y^i_k)_{k \geq 1}\), which is related to the sequence \((T_n, X^i_n)_{n \geq 1}\) as follows

\[
(T^i_k, Y^i_k) = \begin{cases} 
(T_{m^i_k}, X^i_{m^i_k}) & \text{if } m^i_k < \infty, \\
(T_{m^i_k + k - \hat{k}^i}, \Delta) & \text{if } m^i_k = \infty \text{ and } T_\infty < \infty, \\
(\infty, \emptyset^i) & \text{if } m^i_k = \infty \text{ and } T_\infty = \infty,
\end{cases}
\]

(3.4)

where \(\hat{k}^i = \max\{n : m^i_n < \infty\}\), with \(m^i\) defined as

\[
m^i_1 = \inf\{n \geq 1 : X^i_n \in E_i\}, \\
m^i_k = \inf\{n > m^i_{k-1} : X^i_n \in E_i\} \quad \text{for } k > 1.
\]

We clearly have

\[
N^i((0, t], A) = \sum_{k \geq 1} \mathbb{1}_{\{T^i_k \leq t\}} \mathbb{1}_{\{Y^i_k \in A\}}, \quad A \in \mathcal{E}_i.
\]

(3.5)

In particular this means that for the \(i\)-th coordinate \(N^i\) the times \(T_n(\omega)\) such that \(X^i_n(\omega) = \Delta\) are disregarded as event times for this coordinate since the events occurring with regard to the entire \(N\) at these times do not affect the \(i\)-th coordinate.

We define the completed filtration \(\mathbb{F}^{N^i} = (\mathcal{F}^{N^i}_t, t \geq 0)\) generated by \(N^i\) in analogy to \(\mathbb{F}^N\); specifically \(\mathcal{F}^{N^i}\) is the \(\mathbb{P}\)-completion of the \(\sigma\)-field \(\sigma(N^i((s, r] \times A) : 0 \leq s < r \leq t, A \in \mathcal{E}_i), t \geq 0\). In view of Theorem 2.2.4 in [22] the filtration \(\mathbb{F}^{N^i}\) satisfies the usual conditions.

We define the explosion time \(T^i_\infty\) of \(N^i\) as

\[
T^i_\infty := \lim_{n \to \infty} T^i_n.
\]

Clearly, \(T^i_\infty \leq T_\infty\).

We conclude this section with providing some more insight into the properties of the measure \(N^i\). Towards this end, we first observe that the measure \(N^i\) is both \(\mathbb{F}^N\)-optional and \(\mathbb{F}^{N^i}\)-optional. Subsequently, we will derive the compensator of \(N^i\) with respect to \(\mathbb{F}^N\) and the compensator of \(N^i\) with respect to \(\mathbb{F}^{N^i}\). The following Proposition 3.4 and Proposition 3.7 come handy in this regard.

**Proposition 3.4.** Let \(N\) be a generalized multivariate Hawkes process with Hawkes kernel \(\kappa\). Then the \((\mathbb{F}^N, \mathbb{P})\)-compensator, say \(\nu^i\), of measure \(N^i\) defined in (3.2) is given as

\[
\nu^i(\omega, dt, dy_i) = \mathbb{1}_{[0, T^\infty_\infty(\omega)]}(t) \kappa^i(\omega, t, dy_i) dt,
\]

(3.6)

where

\[
\kappa^i(t, A) := \kappa(t, A^i), \quad t \geq 0, \quad A \in \mathcal{E}_i,
\]

(3.7)

with \(A^i\) defined in (3.3).

**Proof.** According to Theorems 4.1.11 and 4.1.7 in [22] the \(i\)-th coordinate \(N^i\) admits a unique \(\mathbb{F}^N\)-compensator, say \(\nu^i\), with a property that \(\nu^i([T^\infty_\infty, \infty] \times E_i) = 0\). For every \(n\) and \(A \in \mathcal{E}_i\) the processes \(M^{i,n,A}\) and \(\tilde{M}^{i,n,A}\) given as

\[
M^{i,n,A}_t = N^i((0, t \wedge T_n] \times A) - \int_0^{t \wedge T_n} \mathbb{1}_{[0, T^\infty_\infty](u) \kappa^i(u, A) du, \quad t \geq 0,
\]

hence, the processes \(M^{i,n,A}\) and \(\tilde{M}^{i,n,A}\) are jointly \(\mathbb{F}^N\)-optional with respect to \(\mathbb{P}\). Therefore, there exists a unique \(\mathbb{F}^N\)-optional \(\nu^i\) with the property that

\[
\nu^i(\omega, dt, dy_i) = \mathbb{1}_{[0, T^\infty_\infty(\omega)]}(t) \kappa^i(\omega, t, dy_i) dt,
\]

(3.6)

where

\[
\kappa^i(t, A) := \kappa(t, A^i), \quad t \geq 0, \quad A \in \mathcal{E}_i,
\]

(3.7)
and
\[
\hat{M}^{i,n,A}_t = N^i((0, t \wedge T_n] \times A) - \nu^i((0, t \wedge T_n] \times A), \ t \geq 0,
\]
are \((\mathbb{F}^N, \mathbb{P})\)-martingales. Hence the process
\[
\int_0^{t \wedge T_n} \mathbb{1}_{(0,T_n]}(u) \kappa^i(u, A) du - \nu^i(du, A), \ t \geq 0,
\]
is an \(\mathbb{F}^N\)-predictable martingale. Since it is of integrable variation and null at 
\(t = 0\) (see e.g. Theorem VI.6.3 in [16]). From the above and the fact that 
\(T_{\infty}^i \leq T_{\infty}\) we deduce that
\[
\int_0^{t \wedge T_n} \mathbb{1}_{(0,T_n]}(u) \kappa^i(u, A) du = \nu^i((0, t \wedge T_n] \times A), \ t \geq 0.
\]
This proves the proposition. \(\square\)

**Remark 3.5.** Note that for each \(i\), the function \(\kappa^i\) defined in (3.7) is a measurable kernel from \((\Omega \times \mathbb{R}_+, \mathcal{P} \otimes \mathcal{B}(\mathbb{R}_+))\) to \((\mathcal{E}_i, \mathcal{E}_i)\). It is important to observe that, in general, there is no one-to-one correspondence between the Hawkes kernel \(\kappa\) and all the marginal kernels \(\kappa^i, i = 1, \ldots, d\). We mean by this that may exist another Hawkes kernel, say \(\hat{\kappa}\), such that \(\hat{\kappa} \neq \kappa\) and
\[
\kappa^i(t, A) = \hat{\kappa}(t, A^i), \ t \geq 0, \ A \in \mathcal{E}_i, \ i = 1, \ldots, d.
\]
(3.8)

**Remark 3.6.** As we know from Remark 2.4 the Hawkes kernel \(\kappa\) determines the law of \(N\). However, in view of Remark 3.5, the kernel \(\kappa^i\) may not determine the law of \(N^i\). It remains to be an open problem for now to determine sufficient conditions under which the law of \(N^i\) is determined by \(\kappa^i\). This problem is a special case of a more general problem: what are general sufficient conditions under which characteristics of a semimartingale determine the law of this semimartingale. \(\square\)

The following important result gives the \(\mathbb{F}^N^i\)-compensator of measure \(N^i\).

**Proposition 3.7.** Let \(N\) be a generalized multivariate Hawkes process with Hawkes kernel \(\kappa\). Then the \(\mathbb{F}^N^i\)-compensator of measure \(N^i\), say \(\tilde{\nu}^i\), is given as
\[
\tilde{\nu}^i(\omega, dt, dy_i) = (\nu^i)^{p,F^N^i}(\omega, dt, dy_i),
\]
where \((\nu^i)^{p,F^N^i}\) is the dual predictable projection of \(\nu^i\) on \(\mathbb{F}^N^i\) under \(\mathbb{P}\).

**Proof.** Using Theorems 4.1.9 and 3.4.6 in [22], as well as the uniqueness of the compensator, it is enough to show that for any \(A \in \mathcal{E}_i\) and any \(n \geq 1\) the process \((\nu^i)^{p,F^N^i}((0, t \wedge T^i_n], A)\), where \((\nu^i)^{p,F^N^i}\) is the dual predictable projection of \(\nu^i\) on \(\mathbb{F}^N^i\) under \(\mathbb{P}\), is the \(\mathbb{F}^N^i\)-compensator of the increasing process \(N^i((0, t \wedge T^i_n], A), \ t \geq 0\). This however follows from Theorem 3.3 in [4]. \(\square\)
3.3 Examples

We will provide now some examples of generalized multivariate Hawkes processes.

For \( \omega = (t_n, x_n)_{n \geq 1}, t \geq 0 \) and \( A \in E^\Delta \) we set

\[
N(\omega, (0, t], A) := \sum_{n \geq 1} 1\{t_n \leq t, x_n \in A\}.
\] (3.10)

In all examples below we define the kernel \( \kappa \) of the form (2.4) with \( \eta \) and \( f \) properly chosen, so that we may apply Theorem 2.5 to the effect that there exists a probability measure \( \mathbb{P}_\nu \) on \( (\Omega, \mathcal{F}) \) such that process \( N \) given by (3.10) is a Hawkes process with the Hawkes kernel equal to \( \kappa \). In other words, there exists a probability measure \( \mathbb{P}_\nu \) on \( (\Omega, \mathcal{F}) \) such that \( \nu \) given in (2.3)–(2.4) is the \( \mathbb{P}^N \)-compensator of \( N \) under \( \mathbb{P}_\nu \).

For a Hawkes process \( N \) with a mark space \( E^\Delta \) we introduce the following notation

\[
N_t = N((0, t], E^\Delta), \quad t \geq 0.
\]

Likewise, we denote for \( i = 1, \ldots, d \),

\[
N^i_t = N^i((0, t], E_i), \quad t \geq 0.
\]

**Example 3.8. Classical univariate Hawkes process**

We take \( d = 1 \) and \( E_1 = \{1\} \), so that \( E^\Delta = E_1 = \{1\} \). As usual, and in accordance with (2.2), we identify \( N \) with a point process \( (N_t)_{t \geq 0} \). Now we take

\[
\eta(t, \{1\}) = \lambda(t),
\]

where \( \lambda \) is positive, locally integrable function, and, for \( 0 \leq s \leq t \), we take

\[
f(t, s, 1, \{1\}) = w(t - s)
\]

for some non-negative function \( w \) defined on \( \mathbb{R}_+ \) (recall that \( f(t, s, 1, \{1\}) = 0 \) for \( s \geq t \)). Using these objects we define \( \kappa \) by

\[
\kappa(t, dy) = \bar{\kappa}(t)\delta_{\{1\}}(dy),
\]

where

\[
\bar{\kappa}(t) = \lambda(t) + \int_{(0,t)} w(t - s) dN_s.
\]

In case of the classical univariate Hawkes process sufficient conditions under which the explosion time is almost surely infinite, that is

\[
T_\infty = \infty \quad \mathbb{P}_\nu - \text{a.s.}
\]

are available in terms of the Hawkes kernel. Specifically, sufficient conditions for no-explosion are given in [1]:

\[
\lambda \text{ is locally bounded, and } \int_0^\infty w(u)du < \infty. \quad \square
\]
Example 3.9. Generalized bivariate Hawkes process with common event times

In the case of a generalized bivariate Hawkes process \( N \) we have \( d = 2 \) and the mark space is given as

\[
E^\Delta = E_1^\Delta \times E_2^\Delta \setminus \{ (\Delta, \Delta) \} = \{ (\Delta, y_2), (y_1, \Delta), (y_1, y_2) : y_1 \in E_1, y_2 \in E_2 \}.
\]

Here, in order to define kernel \( \kappa \), we take kernel \( \eta \) in the form

\[
\eta(t, dy) = \eta_1(t, dy_1) \otimes \delta_\Delta(dy_2) + \delta_\Delta(dy_1) \otimes \eta_2(t, dy_2) + \eta_c(t, dy_1, dy_2),
\]

where \( \delta_\Delta \) is a Dirac measure, \( \eta_i \) for \( i = 1, 2 \) are probability kernels, from \((\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))\) to \((E_i, \mathcal{E}_i)\) and \( \eta_c \) is a probability kernel from \((\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))\) to \((E^\Delta, \mathcal{E}^\Delta)\), satisfying

\[
\eta_c(t, E_1 \times \Delta) = \eta_c(t, \Delta \times E_2) = 0.
\]

Kernel \( f \) is given, for \( 0 \leq s \leq t \) and \( x = (x_1, x_2) \), by

\[
f(t, s, x, dy) = \left( w_{1,1}(t, s)g_{1,1}(x_1)\mathbb{1}_{E_1 \times \Delta}(x) + w_{1,2}(t, s)g_{1,2}(x_2)\mathbb{1}_{\Delta \times E_2}(x) \right.
\]
\[
+ w_{1,c}(t, s)g_{1,c}(x)\mathbb{1}_{E_1 \times E_2}(x) \bigg) \phi_1(x, dy_1) \otimes \delta_\Delta(dy_2)
\]
\[
+ \left( w_{2,1}(t, s)g_{2,1}(x_1)\mathbb{1}_{E_1 \times \Delta}(x) + w_{2,2}(t, s)g_{2,2}(x_2)\mathbb{1}_{\Delta \times E_2}(x) \right.
\]
\[
+ w_{2,c}(t, s)g_{2,c}(x)\mathbb{1}_{E_1 \times E_2}(x) \bigg) \delta_\Delta(dy_1) \otimes \phi_2(x, dy_2)
\]
\[
+ \left( w_{c,1}(t, s)g_{c,1}(x_1)\mathbb{1}_{E_1 \times \Delta}(x) + w_{c,2}(t, s)g_{c,2}(x_2)\mathbb{1}_{\Delta \times E_2}(x) \right.
\]
\[
+ w_{c,c}(t, s)g_{c,c}(x_1, x_2)\mathbb{1}_{E_1 \times E_2}(x) \bigg) \phi_c(x, dy_1, dy_2),
\]

where \( \phi_i \) is a probability kernel from \((E^\Delta, \mathcal{E}^\Delta)\) to \((E_i, \mathcal{E}_i)\) for \( i = 1, 2 \) and \( \phi_c \) is a probability kernel from \((E^\Delta, \mathcal{E}^\Delta)\) to \((E^\Delta, \mathcal{E}^\Delta)\) satisfying

\[
\phi_c(x, E_1 \times \Delta) = \phi_c(x, \Delta \times E_2) = 0.
\]

The decay functions \( w_{i,j} \) and the impact functions \( g_{i,j} \), \( i, j = 1, 2, c \), are appropriately regular and deterministic. Moreover, the decay functions are positive and the impact functions are non-negative. In particular, this implies that the kernel \( f \) is deterministic and non-negative.

In what follows we will need the concept of idiosyncratic group of \( I \) coordinates of a generalized bivariate Hawkes process \( N \). For \( I = \{ 1 \} \) we define

\[
N^{\text{idio},\{1\}}((0, t], A) := N((0, t], A \times \Delta), \quad t \geq 0, \ A \in \mathcal{E}_1
\]

and, likewise, for \( I = \{ 2 \} \) we define

\[
N^{\text{idio},\{2\}}((0, t], A) := N((0, t], \Delta \times A), \quad t \geq 0, \ A \in \mathcal{E}_2.
\]

Finally, for \( I = \{ 1, 2 \} \) we define

\[
N^{\text{idio},\{1,2\}}((0, t], A) := N((0, t], A), \quad t \geq 0, \ A \in \mathcal{E}_1 \otimes \mathcal{E}_2.
\]
Clearly, $N^{idio, i}$ is a MPP. For example, $N^{idio, i}$ is a MPP which records idiosyncratic events occurring with regard to $X^i$; that is, events that only regard to $X^i$, so that $X^j_n = \Delta$ for $j \neq i$ at times $T_n$ at which these events take place. Likewise, $N^{idio, \{1,2\}}$ is a MPP which records idiosyncratic events occurring with regard to $X^1$ and $X^2$ simultaneously. Let us note that

\[
N^1 = N^{idio,\{1\}} + N^{idio,\{1,2\}}, \quad N^2 = N^{idio,\{2\}} + N^{idio,\{1,2\}}.
\]

We will now interpret various terms that appear in the expressions for $\eta$ and $f$ above:

- $\eta_1(t, dy_1) \otimes \delta_\Delta(dy_2)$ represents autonomous portion of the intensity, at time $t$, of marks of the coordinate $N^1$ taking values in the set $dy_1 \subset E_1$ and no marks occurring for $N^2$;

- $\eta_c(t, dy_1, dy_2)$ represents autonomous portion of the intensity, at time $t$, of an event amounting to the marks of both coordinates $N^1$ and $N^2$ taking values in the set $dy_1dy_2 \subset E_1 \times E_2$;

\[
\int_{(0,t) \times E^\Delta} w_{1,1}(t, s)g_{1,1}(x_1)1_{E_1 \times \Delta}(x)\phi_1(x, dy_1) \otimes \delta_\Delta(dy_2)N(ds, dx)
\]

\[
= \int_{(0,t) \times E_1} w_{1,1}(t, s)g_{1,1}(x_1)\phi_1((x_1, \Delta), dy_1) \otimes \delta_\Delta(dy_2)N^{idio,1}(ds, dx_1)
\]

represents idiosyncratic impact of the coordinate $N^1$ alone on the intensity, at time $t$, of marks of the coordinate $N^1$ taking values in the set $dy_1 \subset E_1$ and no marks occurring for $N^2$;

\[
\int_{(0,t) \times E^\Delta} w_{1,2}(t, s)g_{1,2}(x_2)\phi_1((\Delta, x_2), dy_1) \otimes \delta_\Delta(dy_2)N^{idio,2}(ds, dx_2)
\]

represents idiosyncratic impact of the coordinate $N^2$ alone on the intensity, at time $t$, of an event amounting to the marks of coordinate $N^1$ taking value in the set $dy_1 \subset E_1$ and no marks occurring for $N^2$;

\[
\int_{(0,t) \times E^\Delta} w_{1,c}(t, s)g_{1,c}(x)1_{E_1 \times E_2}(x)\phi_1(x, dy_1) \otimes \delta_\Delta(dy_2)N(ds, dx)
\]

represents joint impact of the coordinates $N^1$ and $N^2$ on the intensity, at time $t$, of an event amounting to the marks of coordinate $N^1$ taking value in the set $dy_1 \subset E_1$ and no marks occurring for $N^2$;

\[
\int_{(0,t) \times E_1} w_{c,1}(t, s)g_{c,1}(x_1)\phi_c((x_1, \Delta), dy_1, dy_2)N^{idio,1}(ds, dx_1)
\]

represents idiosyncratic impact of the coordinate $N^1$ alone on the intensity, at time $t$, of an event amounting to the marks of both coordinates $N^1$ and $N^2$ taking values in the set $dy_1dy_2 \subset E_1 \times E_2$;
In particular, the terms contributing to occurrence of common events are \( \eta_c(t, dy_1, dy_2) \) and
\[
\left( g_{c,1}(x_1) \mathbb{1}_{E_1 \times \Delta}(x) + g_{c,2}(x_2) \mathbb{1}_{E_2 \times \Delta}(x) + g_{c,c}(x_1, x_2) \mathbb{1}_{E_1 \times E_2}(x) \right) \phi_c(x, dy_1, dy_2).
\]
Upon integrating \( \kappa(t, dy) \) over \( A_1 \times \{ \Delta, E_2 \} \) we get
\[
\kappa^1(t, A_1) = \kappa(t, A_1 \times \{ \Delta, E_2 \})
= \eta_1(t, A_1) + \eta_c(t, A_1 \times E_2)
+ \int_{(0,t) \times E_1} \left( w_{1,1}(t, s) g_{1,1}(x_1) \phi_1((x_1, \Delta), A_1)
+ w_{c,1}(t, s) g_{c,1}(x_1) \phi_c((x_1, \Delta), A_1 \times E_2) \right) N^{\text{idio},1}(ds, dx_1)
+ \int_{(0,t) \times E_2} \left( w_{1,2}(t, s) g_{1,2}(x_2) \phi_1((\Delta, x_2), A_1)
+ w_{c,2}(t, s) g_{c,2}(x_2) \phi_c((\Delta, x_2), A_1 \times E_2) \right) N^{\text{idio},2}(ds, dx_2)
+ \int_{(0,t) \times E} \left( w_{1,c}(t, s) g_{1,c}(x) \phi_1((x_1, x_2), A_1)
+ w_{c,c}(t, s) g_{c,c}(x_1, x_2) \phi_c((x_1, x_2), A_1 \times E_2) \right) N^{\text{idio},(1,2)}(ds, dx).
\]
To complete this example we note that upon setting \( \eta_c = 0 \) and \( \phi_c = 0 \) we produce a generalized bivariate Hawkes process with no common event times. \( \square \)

4 Mathematical construction of and computational pseudo-algorithm for simulation of a generalized multivariate Hawkes process with deterministic kernels \( \eta \) and \( f \)
Fix an arbitrary \( T > 0 \). In this section we first provide a construction of restriction to \( [0,T] \times E^\Delta \) of a generalized multivariate Hawkes process, with deterministic kernels \( \eta \) and \( f \), via Poisson thinning, that is motivated by a similar construction given in [5]. Then, based on our construction, we present a computational pseudo-algorithm for simulation of a generalized multivariate Hawkes process restricted to \( [0,T] \times E^\Delta \).

We are concerned here with a generalized multivariate Hawkes process admitting the Hawkes kernel of the form
\[
\kappa(t, dy) = \eta(t, dy) + \int_{(0,t) \times E^\Delta} f(t, s, x, dy) N(ds, dx), \tag{4.1}
\]
where \( \eta \) is a deterministic finite kernel from \((\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))\) to \((E^\Delta, \mathcal{E}^\Delta)\) and \( f \) is a deterministic finite kernel from \((\mathbb{R}_+^2 \times E^\Delta, \mathcal{B}(\mathbb{R}_+^2) \otimes \mathcal{E}^\Delta)\) to \((E^\Delta, \mathcal{E}^\Delta)\).

We may, and we do, represent kernels \( \eta, f \) as

\[
\eta(t, dy) = \eta(t, E^\Delta) Q_1(t, dy), \quad \text{where} \quad Q_1(t, dy) = \frac{\eta(t, dy)}{\eta(t, E^\Delta)} \mathbb{1}_{\{\eta(t, E^\Delta) > 0\}} + \delta_\partial(dy) \mathbb{1}_{\{\eta(t, E^\Delta) = 0\}},
\]

and

\[
f(t, s, x, dy) = f(t, s, x, E^\Delta) Q_2(t, s, x, dy), \quad \text{where} \quad Q_2(t, s, x, dy) = \frac{f(t, s, x, dy)}{f(t, s, x, E^\Delta)} \mathbb{1}_{\{f(t, s, x, E^\Delta) > 0\}} + \delta_\partial(dy) \mathbb{1}_{\{f(t, s, x, E^\Delta) = 0\}}.
\]

Note that \( Q_1 \) and \( Q_2 \) are deterministic probability kernels.

Since we are concerned with a restricted Hawkes process we consider a Hawkes kernel \( \kappa_T \) which is a restriction to \([0, T] \) of \( \kappa \) that is

\[
\kappa_T(t, dx) = \mathbb{1}_{[0, T]}(t) \kappa(t, dx).
\] (4.2)

For simplicity of notation we suppress \( T \) in the notation below. So, for example, we will write \( f \) rather than \( f_T := \mathbb{1}_{[0, T]} f \).

We make the following standing assumption:

\[
\sup_{t \in [0, T]} \eta(t, E^\Delta) \leq \hat{\eta} < \infty,
\]

for some constant \( \hat{\eta} > 0 \) and, for \( s \in [0, T] \) and \( x \in E^\Delta \)

\[
\sup_{t \in [s, T]} f(t, s, x, E^\Delta) \leq \hat{f}(s, x) < \infty,
\] (4.3)

for some measurable mapping \( \hat{f} : [0, T] \times E^\Delta \to (0, \infty) \).

### 4.1 Description of the construction

Now we describe a construction of Hawkes process with Hawkes kernel given by (4.2). This construction leads immediately to a pseudo-algorithm, presented in the next section, for simulation of such Hawkes process.

In what follows we will define recursively a sequence of random measures \((N^k)_{k \geq 0}\) that provide building blocks for our Hawkes process.

Towards this end we first let \( \beta \) be the Borel isomorphism between the space \( E^\partial \) and a Borel subset of \( \mathbb{R}^d \cup \partial \), with the convention that \( \beta(\partial) = \partial \).

Our construction will proceed in several steps.

Step 1). Let us consider an array \( \{ (Z_i^{k,j}, (U_i^{k,j}, V_i^{k,j}, W_i^{k,j})_{j=1}^\infty) \}_{k=0, j=1}^\infty \) of independent identically distributed random variables with uniform distribution on \( (0, 1) \). Let \( D : [0, \infty) \times (0, 1) \to \mathbb{N} \) be a measurable function such that

\[
\int_{(0,1]} \mathbb{1}_\{k\}(D(\lambda, u)) du = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, \ldots
\]
where we use the convention that \(0^0 = 1\). Therefore, for a random variable \(U\) uniformly distributed on \((0, 1]\) the random variable \(D(\lambda, U)\) has Poisson distribution with parameter \(\lambda \geq 0\), where we extend the concept of Poisson distribution by allowing \(\lambda = 0\). Moreover let \(G_1 : [0, T] \times (0, 1] \to E^\Delta\) be a measurable function such that
\[
\int_{[0,1]} \mathbb{1}_A(G_1(t, u))du = Q_1(t, A), \quad A \in \mathcal{E}^\Delta,
\]
and \(G_2 : [0, T] \times [0, T] \times E^\Delta \times (0, 1] \to E^\Delta\) be a measurable function such that
\[
\int_{(0,1]} \mathbb{1}_A(G_2(t, s, y, u))du = Q_2(t, s, y, A), \quad A \in \mathcal{E}^\Delta.
\]
Existence of such functions \(G_1\) and \(G_2\) is asserted by Lemma 3.22 in [20].

We use the left open intervals of integration above so to be consistent with the the rest of the construction. The reason that we work with left open intervals in the rest of the construction is that the births of the offsprings occur after the appearance of their parents (e.g., after arrivals of the immigrants), see Section 2.2. This feature is explicitly accounted for in the construction presented here.

Step 2). Using \((Z^{0,1}, (U_i^{0,1}, V_i^{0,1}, W_i^{0,1}))_{i=1}^\infty\) we define a random measure \(N^0\) on \(\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}^\Delta\):
\[
N^0(dt, dx) = \sum_{i=1}^\infty \delta_{(T_0^i, X_0^i)}(dt, dx) \mathbb{1}_{\{i \leq P^0, A_0^i \leq \eta(T_0^i, E^\Delta)\}}, \tag{4.4}
\]
where
\[
P^0 = D(T_0, Z^{0,1}), \quad T_0^i = TU_i^{0,1}, \quad A_0^i = \tilde{\eta}V_i^{0,1}, \quad X_0^i = G_1(T_i^0, W_i^{0,1}).
\]
We note that \(P^0\) is a Poisson random variable with parameter \(T_0\), which is independent of the iid sequence \((T_i^0, X_i^0, A_i^0)_{i=1}^\infty\) of random elements with values in \([0, T] \times [0, \tilde{\eta}] \times E^\Delta\), and that
\[
\mathbb{P}((T_i^0, A_i^0, X_i^0) \in dt \times da \times dx) = \frac{1}{T_0^\tilde{\eta}} \mathbb{1}_{(0, T] \times [0, \tilde{\eta}]}(t, a)Q_1(t, dx)dt da.
\]

Then, we consider a sequence \((S_j^0, Y_j^0)_{j=1}^\infty\) with \(S_j^0 \in (0, T] \cup \{\infty\}\) given as
\[
S_j^0 := \inf\{t \geq 0 : N^0((0, t] \times E^\Delta) \geq j\} = \inf\{t \in [0, T] : N^0((0, t] \times E^\Delta) \geq j\},
\]
and with \(Y_j^0\) constructed as follows:
\[
Y_j^0(\omega) := \begin{cases} X_i^0(\omega), & \text{for } i \text{ such that } S_j^0(\omega) = T_i^0(\omega) < \infty; \\ \partial, & \text{if } S_j^0(\omega) = \infty. \end{cases}
\]
The sequence \((S_j^0, Y_j^0)_{j=1}^\infty\) is well defined because \(N^0\) is a counting measure such that \(N^0(\{t\} \times E^\Delta) \leq 1\) for \(t \geq 0\), so \(N^0(\{S_j^0\} \times E^\Delta) = 1\), provided \(S_j^0 < \infty\), and since \(\mathbb{P}(\exists i \neq k : T_i^0 = T_k^0) = 0\). Moreover, \(Y_j^0\) is a random element. Indeed,
\[
Y_j^0 = \beta^{-1} \left( \sum_{i=1}^\infty \beta(X_i^0) \mathbb{1}_{\{T_i^0 = S_j^0\}} \mathbb{1}_{\{S_j^0 < \infty\}} + \tilde{\mathcal{O}} \mathbb{1}_{\{S_j^0 = \infty\}} \right)
\]
\[
= \beta^{-1} \left( \mathbb{1}_{\{S_j^0 < \infty\}} \int_{E^\Delta} \beta(x)N^0(\{S_j^0\} \times dx) + \tilde{\mathcal{O}} \mathbb{1}_{\{S_j^0 = \infty\}} \right).
\]
Observe that $S_j^0 < S_{j+1}^0$ on $\{S_j^0 < \infty\}$, and that the measure $N^0$ may be identified with the sequence $(S_j^0, Y_j^0)_{j=1}^\infty$. Indeed, defining $\Psi^0 := \text{card}\{j \geq 1 : S_j^0 < \infty\}$, we have $\Psi^0 \leq P^0$, and thus $\mathbb{P}(\Psi^0 < \infty) = 1$. Consequently,

\[
N^0(dt, dx) = \sum_{j=1}^\infty \delta_{(S_j^0, Y_j^0)}(dt, dx)\mathbb{1}_{\{S_j^0 < \infty\}} = \sum_{j=1}^\infty \delta_{(S_j^0, Y_j^0)}(dt, dx)\mathbb{1}_{\{j \leq \Psi^0\}}. \tag{4.5}
\]

The representation (4.5) is more convenient for our needs than the representation (4.4). This is because the sequence $(S_j^0, Y_j^0)_{j=1}^\infty$ is ordered with respect to the first component, so that this sequence is a MPP and thus measure $N^0$ may also be considered as a MPP.

Step 3). Now, we proceed by recurrence. So, for $k \in \mathbb{N}$ suppose that we have constructed a random sequence $(S_j^k, Y_j^k)_{j=1}^\infty$ with the property that $S_j^k \in (0, T)$ if $\{S_j^k < \infty\}$, and $\mathbb{P}(\Psi^k < \infty) = 1$ where $\Psi^k = \text{card}\{j \geq 1 : S_j^k < \infty\}$, and that we have also constructed a random measure $N^k$ on $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}_\Delta$ satisfying

\[
N^k(dt, dx) = \sum_{j=1}^\infty \delta_{(S_j^k, Y_j^k)}(dt, dx)\mathbb{1}_{\{S_j^k < \infty\}} = \sum_{j=1}^\infty \delta_{(S_j^k, Y_j^k)}(dt, dx)\mathbb{1}_{\{j \leq \Psi^k\}}.
\]

Given $N^k$, or equivalently given $(S_j^k, Y_j^k)_{j=1}^\infty$, we will define a sequence of random measures $(N_j^{k+1})_{j=1}^\infty$, which are conditionally independent given $\sigma(N^0, \ldots, N^k)$. Fix $j \in \{1, 2, \ldots\}$. We let the random measure $N_j^{k+1}$ on $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}_\Delta$ be defined by

\[
N_j^{k+1}(dt, dx) = \sum_{i=1}^\infty \delta_{(T_i^{k+1,j}, X_i^{k+1,j})}(dt, dx)\mathbb{1}_{\{S_j^k < T, i \leq P_j^{k+1}, A_i^{k+1,j} \leq f(T_i^{k+1,j}, S_j^k, Y_j^k, \mathcal{E}_\Delta)\}}, \tag{4.6}
\]

where $P_j^{k+1}$, $(T_i^{k+1,j}, A_i^{k+1,j}, X_i^{k+1,j})_{i=1}^\infty$ are random variables defined by transformation of the sequence $Z_i^{k+1,j}, (U_i^{k+1,j}, V_i^{k+1,j}, W_i^{k+1,j})_{i=1}^\infty$ and the pair $(S_j^k, Y_j^k)$ in the following way:

\[
P_j^{k+1} = D((T - S_j^k)\hat{f}(S_j^k, Y_j^k)\mathbb{1}_{\{S_j^k < T\}}, Z_j^{k+1,j}) \tag{4.7}
\]

\[
= D((T - S_j^k)\hat{f}(S_j^k, Y_j^k), Z_j^{k+1,j})\mathbb{1}_{\{S_j^k < T\}}
\]

\[
T_i^{k+1,j} = (S_j^k + (T - S_j^k)U_i^{k+1,j})\mathbb{1}_{\{S_j^k < T\}} + \infty\mathbb{1}_{\{S_j^k \geq T\}},
\]

\[
A_i^{k+1,j} = \hat{f}(S_j^k, Y_j^k)V_i^{k+1,j}\mathbb{1}_{\{S_j^k < T\}},
\]

\[
X_i^{k+1,j} = G_2(T_i^{k+1,j}, S_j^k, Y_j^k, W_i^{k+1,j})\mathbb{1}_{\{S_j^k < T\}} + \partial\mathbb{1}_{\{S_j^k \geq T\}}.
\]

Note that the random variable $P_j^{k+1}$ has $\sigma(N^0, \ldots, N^k)$-conditionally Poisson distribution with parameter $(T - S_j^k)\hat{f}(S_j^k, Y_j^k)\mathbb{1}_{\{S_j^k < \infty\}}$, where $\hat{f}$ given by (4.3). The random elements in the sequence $(T_i^{k+1,j}, A_i^{k+1,j}, X_i^{k+1,j})_{i=1}^\infty$ take values in $(S_j^k, T] \times [0, \hat{f}(S_j^k, Y_j^k)] \times \mathcal{E}_\Delta$ if $S_j^k < T$.

\footnote{Conditional independence between random measures is understood as conditional independence between random elements taking values in the space of probability measures. We refer to Kallenberg [20], Chapter 12, for definition of random elements taking values in the space of $\sigma$-finite measures on a measurable space, and to Chapter 6 therein for definition of conditional independence between random elements.}
otherwise, if \( S_j^k \geq T \), these elements are all constant and equal to \((\infty, 0, \partial)\). Moreover, they are \( \sigma(N^0, \ldots, N^k) \)-conditionally independent random elements, and the \( \sigma(N^0, \ldots, N^k) \)-conditional distribution of \((T_{i+1}^k, A_{i+1}^k, X_{i+1}^k)\) is given by

\[
\mathbb{P}((T_{i+1}^k, A_{i+1}^k, X_{i+1}^k) \in dt \times da \times dx | N^0, \ldots, N^k) = \frac{1}{(T - S_j^k) f(S_j^k, Y_j^k)} \mathbb{I}(S_j^k \leq T) \delta_{(\infty, 0, \partial)}(dt, da, dx) \quad (4.8)
\]

Thus, if \( S_j^k \geq T \), then \( N_j^{k+1} = 0 \). The random measure \( N_j^{k+1} \) can be identified with the random sequence \((S_n^{k+1}, Y_n^{k+1})_{n=1}^\infty\), where

\[
S_n^{k+1}(t, da, dx) \triangleq \inf\{t : N_j^{k+1}((0, t] \times \mathbb{E}) \geq n\}
\]

\[
Y_n^{k+1} := \beta^{-1} \left( \mathbb{I}_{\{S_n^{k+1} < \infty\}} \int_{E^\Delta} \beta(x) N_j^{k+1} \left( \{S_n^{k+1}\} \times dx \right) + \hat{\mathbb{P}} \left( \{S_n^{k+1} \geq \infty\} \right) \right)
\]

Indeed, we have

\[
N_j^{k+1}(dt, dx) = \sum_{i=1}^\infty \mathcal{D}(S_i^{k+1}, Y_i^{k+1}) \mathbb{I}_{\{S_i^{k+1} < \infty\}} \mathcal{D}(dt, dx) = \sum_{i=1}^\infty \mathcal{D}(S_i^{k+1}, Y_i^{k+1}) \mathbb{I}_{\{S_i^{k+1} < \infty\}} \mathcal{D}(dt, dx)
\]

where \( \Psi_i^{k+1} = \text{card}\{i : S_i^{k+1} < \infty\} \) is such that \( \Psi_i^{k+1} \leq P_i^{k+1} < \infty \) with probability 1. Therefore, since the sequence \((S_n^{k+1})_{n=1}^\infty\) is increasing as long as \( S_n^{k+1} < \infty \), the measure \( N_j^{k+1} \) is a MPP. Moreover, if \( S_j^k < T \), then \( T_i^{k+1} \in (S_j^k, T] \) for every \( l \). Hence using the definition of \( S_i^{k+1} \) we conclude that \( S_i^{k+1} \in (S_j^k, T] \), when \( S_j^k \) is finite.

Next, we define the random measure \( N^{k+1} \) on \( \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}_\Delta \):

\[
N^{k+1}(dt, dx) := \sum_{j=1}^\infty N_j^{k+1}(dt, dx) = \sum_{j=1}^\infty N_j^{k+1}(dt, dx) \mathbb{I}_{\{S_j^k < T\}}, \quad (4.9)
\]

Similarly as above we observe that the random measure \( N^{k+1} \) can be identified with the random sequence \((S_n^{k+1}, Y_n^{k+1})_{n=1}^\infty\), where

\[
S_n^{k+1} := \inf\{t : N^{k+1}((0, t] \times \mathbb{E}) \geq n\},
\]

\[
Y_n^{k+1} := \beta^{-1} \left( \mathbb{I}_{\{S_n^{k+1} < \infty\}} \int_{E^\Delta} \beta(x) N^{k+1} \left( \{S_n^{k+1}\} \times dx \right) + \hat{\mathbb{P}} \left( \{S_n^{k+1} \geq \infty\} \right) \right)
\]

Indeed, we have

\[
\Psi^{k+1} := \text{card}\{i : S_i^{k+1} < \infty\} = \sum_{j=1}^{\Psi_j^{k+1}} \psi_j^{k+1},
\]

and so \( \mathbb{P}(\Psi^{k+1} < \infty) = 1 \). Moreover, we observe that \( N^{k+1}((S_n^{k+1}) \times \mathbb{E}) = 1 \), provided that \( S_n^{k+1} < \infty \) and hence \( Y_n^{k+1}(\omega) = Y_i^{k+1}(\omega) \) for all \( i \) such that \( S_n^{k+1}(\omega) = S_i^{k+1}(\omega) < \infty \). Thus

\[
N^{k+1}(dt, dx) = \sum_{i=1}^\infty \mathcal{D}(S_i^{k+1}, Y_i^{k+1}) \mathbb{I}_{\{S_i^{k+1} < \infty\}} = \sum_{i=1}^\infty \mathcal{D}(S_i^{k+1}, Y_i^{k+1}) \mathbb{I}_{\{i \leq \Psi^{k+1}\}}.
\]
Since the sequence \((S_n^{k+1}, Y_n^{k+1})_{n=1}^\infty\) forms a MPP, then \(N^{k+1}\) may be considered as a MPP.

Recall that if \(S_k^j < T\), then \(S^{k+1}, j \in (S_k^j, T]\) for every \(l\), which implies that \(S^{k+1} \in [0, T]\) if \(S^{k+1} < \infty\).

Step 4). Define a sequence of random measures \(H^k, k \geq 1\), on \(\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}^\Delta\) in terms of the previously constructed marked point processes \((N^j)_{j \geq 0}\) by

\[
H^k(dt, dx) = \sum_{m=0}^{k} N^m(dt, dx), \quad k \geq 1.
\]

Step 5). Repeat Step 3 and Step 4 infinitely many times to obtain limiting random measure \(H^\infty\) on \(\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}^\Delta\) given by

\[
H^\infty(dt, dx) = \sum_{m=0}^{\infty} N^m(dt, dx).
\]

Remark 4.1. It is important to note that all random measures introduced in the construction above do not charge any set \(F \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}^\Delta\) such that \(F \subset (T, \infty] \times E^\Delta\). So, for example, for any such set we have \(H^\infty(F) = 0\).

### 4.2 Justification of the construction

Now we will justify that the construction given in Steps 1–5 above delivers a generalized multivariate Hawkes process with the Hawkes kernel given in (4.1). Towards this end let us introduce the following filtrations:

\[
\mathbb{H}^k = \{H^k_t\}_{t \in [0, \infty)}, \quad \text{where} \quad H^k_t := \mathcal{F}^{N^0}_t \vee \ldots \vee \mathcal{F}^{N^k}_t,
\]

\[
\mathbb{H}^\infty = \{H^\infty_t\}_{t \in [0, \infty)}, \quad \text{where} \quad H^\infty_t := \bigvee_{k \geq 0} \mathcal{F}^{N^k}_t,
\]

\[
\hat{\mathbb{H}}^{k+1} = \{\hat{H}^{k+1}_t\}_{t \in [0, \infty)}, \quad \text{where} \quad \hat{H}^{k+1}_t := H^\infty_t \vee \mathcal{F}^{N^{k+1}}_t.
\]

Our first aim is to compute \(\mathbb{H}^\infty\)-compensator of the limiting random measure \(H^\infty\) given in (4.11). We begin with following key result,

**Proposition 4.2.** i) The marked point process \(N^0\) is an \(\mathbb{H}^0\)-doubly stochastic marked Poisson process. The random measure \(\nu^0\) given by

\[
\nu^0((s, t] \times D) = \int_s^t \eta(v, D) dv, \quad 0 \leq s \leq t, \quad D \in \mathcal{E}^\Delta,
\]

is the \(\mathcal{H}^0\)-intensity kernel of \(N^0\). Moreover, \(\nu^0\) is the \(\mathbb{H}^0\)-compensator of \(N^0\).
ii) For each $j$ the marked point process $N_j^{k+1}$ is an $\mathbb{H}_k^{k+1}$-doubly stochastic marked Poisson process. The random measure $\nu_j^{k+1}$ given by

$$\nu_j^{k+1}(s, t] \times D) = \int_s^t f(v, S_j^k, Y_j^k, D) \mathbbm{1}_{(s_t, \infty)}(v)dv, \quad 0 \leq s \leq t, \; D \in \mathcal{E}_{\Delta}, \quad (4.13)$$

is the $\mathbb{F}_0^k$-intensity kernel of $N_j^{k+1}$. Moreover, $\nu^0$ is the $\mathbb{H}_k^{k+1}$-compensator of $N_j^{k+1}$.

Proof. i). Note that from Lemma 7.2 by taking

$$\mathcal{G} = \mathcal{H}_0^k, \; \mathcal{Y} = \{1\}, \; Y = 1, \; \ell(y) = 0, \; g = \eta,$n
it follows that $N^0$ is $\mathcal{H}_0^k$-conditional Poisson random measure with intensity measure $\nu^0$ given by (4.12). Now, the assertion follows from the point i) of Proposition 7.3.

ii). We first note that from Lemma 7.2 by taking

$$\mathcal{G} = \mathcal{H}_k^k, \; \mathcal{Y} = [0, T] \times \mathcal{E}_{\Delta}, \; Y = (S_j^k, Y_j^k), \; \ell(s, y) = s, \; g = f,$n
it follows that $N_j^{k+1}$ defined by (4.6) is $\mathcal{H}_k^\infty$-conditionally Poisson random measure with intensity measure $\nu_j^{k+1}$ given by (4.13).

To complete the proof, in view of assertion ii) of Proposition 7.3, it suffices show that the marked point processes $(N_j^{k+1})_{j \geq 1}$ are $\mathcal{H}_0^{k+1}$-conditionally independent. Since $\mathcal{H}_0^{k+1} = \mathcal{H}_\infty^k = \sigma(N^0, \ldots, N^k)$ it suffices to verify that $(N_j^{k+1})_{j \geq 1}$ are conditionally independent given $\sigma(N^0, \ldots, N^k)$. For this we first note that for each $j$ the random measure $N_j^{k+1}$ is defined by (4.6), so it is constructed from the pair $(S_j^k, Y_j^k)$, which is $\sigma(N^0, \ldots, N^k)$-measurable and from the family

$$I_j := (Z_j^{k+1}, U_i^{k+1,j}, V_i^{k+1,j}, W_i^{k+1,j})_{i=1}^\infty.$$

Now, using the fact that $I_1, I_2, \ldots$ are independent between themselves and also independent from $\sigma(N^0, \ldots, N^k)$, we conclude that $N_1^{k+1}, N_2^{k+1}, \ldots$ are $\sigma(N^0, \ldots, N^k)$-conditionally independent. So we see that $N_j^{k+1}$ is a $\mathcal{H}_0^{k+1}$-conditional Poisson random measure for any $j \geq 1$, and that $(N_j^{k+1})_{j \geq 1}$ are $\mathcal{H}_0^{k+1}$-conditionally independent random measures. Thus we may use Proposition 7.3 to conclude that $N_j^{k+1}$ is an $\mathbb{H}_k^{k+1}$-doubly stochastic marked Poisson process whose $\mathcal{H}_0^{k+1}$-intensity kernel is $\nu_j^{k+1}$ given by (4.13).

From Proposition 4.2 and from its proof we conclude that the random measure $N^{k+1}$ given by (4.9) is a sum of $\mathcal{H}_\infty^k$-conditionally independent $\mathbb{H}_k^{k+1}$-doubly stochastic marked Poisson processes. We will prove now that $N^{k+1}$ is also an $\mathbb{H}_k^{k+1}$-doubly stochastic marked Poisson process whose intensity kernel is simply the sum of intensity kernels of $N_j^{k+1}, j \geq 0$.

**Proposition 4.3.** The marked point process $N^{k+1}$ is an $\mathbb{H}_k^{k+1}$-doubly stochastic marked Poisson process with intensity kernel $\nu^{k+1}$ given by

$$\nu^{k+1}(s, t] \times D) = \sum_{j=1}^\infty \nu_j^{k+1}(s, t] \times D) = \int_s^t \int_{(0,v) \times \mathcal{E}_{\Delta}} f(v, u, y, D) N^k(du, dy) dv, \quad (4.14)$$

for $0 \leq s \leq t, \; D \in \mathcal{E}_{\Delta}$. Moreover, the intensity kernel $\nu^{k+1}$ of $N^{k+1}$ is the $\mathbb{H}_k^{k+1}$-compensator of $N^{k+1}$.
Proof. To prove the first assertion, in view of Proposition 6.1.4 in [22], it suffices to show that \( \nu^{k+1} \) is the \( \hat{\mathbb{H}}^{k+1} \)-compensator of \( N^{k+1} \). Indeed this compensating property implies that

\[
\mathbb{E}(N^{k+1}((s, t] \times B)|\hat{\mathcal{H}}_0^{k+1}) = \mathbb{E}(\nu^{k+1}((s, t] \times B)|\hat{\mathcal{H}}_0^{k+1}) = \nu^{k+1}((s, t] \times B),
\]

where the last equality follows from \( \hat{\mathcal{H}}_0^{k+1} \)-measurability of \( \nu^{k+1} \). So, if \( \nu^{k+1} \) is \( \hat{\mathbb{H}}^{k+1} \)-compensator of \( N^{k+1} \) then it is \( \hat{\mathcal{H}}_0^{k+1} \)-intensity kernel and, thus, Theorem 6.1.4 in [22] implies the first assertion. Therefore it remains to show that \( \nu^{k+1} \) is \( \hat{\mathbb{H}}^{k+1} \)-compensator of \( N^{k+1} \).

Towards this end we first note that from Proposition 4.2 it follows that \( \hat{\mathcal{H}}_0^{k+1} \) is the \( \hat{\mathbb{H}}^{k+1} \)-compensator of \( N^{k+1} \). Indeed this compensating property implies that \( \nu^{k+1} \) is \( \hat{\mathbb{H}}^{k+1} \)-predictable function \( F : \Omega \times [0, T] \times \mathcal{E}^{\Delta} \to \mathbb{R} \) it holds

\[
\mathbb{E}
\left(
\int_0^\infty \int_{E^{\Delta}} F(u, y) N_j^{k+1}(du, dy)
\right) = \mathbb{E}
\left(
\int_0^\infty \int_{E^{\Delta}} F(u, y) \nu_j^{k+1}(du, dy)
\right),
\]

This implies that

\[
\lim_{m \to \infty} \mathbb{E}
\left(
\int_0^\infty \int_{E^{\Delta}} F(u, y) \left( \sum_{j=1}^m \nu_j^{k+1} \right)(du, dy)
\right) = \lim_{m \to \infty} \mathbb{E}
\left(
\int_0^\infty \int_{E^{\Delta}} F(u, y) \left( \sum_{j=1}^m \nu_j^{k+1} \right)(du, dy)
\right).
\]

Since, for every \( A \in \mathcal{B}([0, \infty)) \otimes \mathcal{E}^{\Delta} \),

\[
N_j^{k+1}(\omega, A) = \lim_{m \to \infty} \sum_{j=1}^m \nu_j^{k+1}(\omega, A), \quad \text{and} \quad \nu^{k+1}(\omega, A) = \lim_{m \to \infty} \sum_{j=1}^m \nu_j^{k+1}(\omega, A)
\]

almost surely, using Lemma 1.3 for

\[
\mu_j(dt, dy, d\omega) = N_j^{k+1}(\omega, dt, dy) \mathbb{P}(d\omega), \quad \mu(dt, dy, d\omega) = N^{k+1}(\omega, dt, dy) \mathbb{P}(d\omega)
\]

and once again for

\[
\mu_j(dt, dy, d\omega) = \nu_j^{k+1}(\omega, dt, dy) \mathbb{P}(d\omega), \quad \mu(dt, dy, d\omega) = \nu^{k+1}(\omega, dt, dy) \mathbb{P}(d\omega)
\]

we see that

\[
\mathbb{E}
\left(
\int_0^\infty \int_{E^{\Delta}} F(u, y) N_j^{k+1}(du, dy)
\right) = \mathbb{E}
\left(
\int_0^\infty \int_{E^{\Delta}} F(u, y) \nu^{k+1}(du, dy)
\right).
\]

Now, since for any \( 0 \leq s \leq t, D \in \mathcal{E}^{\Delta} \)

\[
\nu^{k+1}((s, t] \times D) = \sum_{j=1}^\infty \nu_j^{k+1}((s, t] \times D) = \sum_{j=1}^\infty \int_s^t f(v, S_j^k, Y_j^k, D) \mathbb{I}_{(s_j, \infty)}(v) dv
\]

\[
= \int_s^t \sum_{j=1}^\infty f(v, S_j^k, Y_j^k, D) \mathbb{I}_{(s_j, \infty)}(v) dv,
\]
and since
\[
\int_{(0,v)\times E^k} f(v, u, y, D) N^k(du, dy) = \sum_{j=1}^{\infty} f(v, S_j^k, Y_j^k, D) \mathbb{1}_{(S_j^k, \infty)}(v),
\]
we obtain that (4.14) holds. This concludes the proof of the first assertion.

Now we will prove that the \( \hat{\nu}^{k+1} \)-intensity kernel \( \nu^{k+1} \) of \( N^{k+1} \) is the \( \hat{\mathbb{H}}^{k+1} \)-compensator of \( N^{k+1} \). For this, we first observe that from Theorem 6.1.4 in [22] it follows that the intensity kernel of \( N^{k+1} \) is the \( \hat{\mathbb{H}}^{k+1} \)-compensator of \( N^{k+1} \). So, for an arbitrary non-negative \( \hat{\mathbb{H}}^{k+1} \)-predictable function \( F : \Omega \times [0, T] \times E^k \to \mathbb{R} \) it holds
\[
\mathbb{E}\left( \int_0^\infty \int_{E^k} F(u, y) N^{k+1}(du, dy) \right) = \mathbb{E}\left( \int_0^\infty \int_{E^k} F(u, y) \nu^{k+1}(du, dy) \right).
\]
(4.16)
Since \( \hat{\mathbb{H}}^{k+1} \supset \mathbb{H}^{k+1} \), the \( \mathbb{H}^{k+1} \)-predictable functions are also \( \hat{\mathbb{H}}^{k+1} \)-predictable. So (4.16) holds for an arbitrary non-negative \( \mathbb{H}^{k+1} \)-predictable function \( F \). From (4.14) we see that for an arbitrary \( D \) the process \( \nu^{k+1}(((0, t] \times D)) \) is \( \mathbb{H}^{k} \)-adapted and continuous. Hence it is \( \mathbb{H}^{k} \)-predictable and thus also \( \hat{\mathbb{H}}^{k+1} \)-predictable (since \( \mathbb{H}^{k+1} \supset \mathbb{H}^{k} \)). So, \( \nu^{k+1} \) is an \( \hat{\mathbb{H}}^{k+1} \)-predictable random measure such that (4.16) holds for arbitrary non-negative \( \hat{\mathbb{H}}^{k+1} \)-predictable function \( F \). This means that \( \nu^{k+1} \) is the \( \hat{\mathbb{H}}^{k+1} \)-compensator of \( N^{k+1} \).

In order to proceed we will need the following auxiliary result.

**Lemma 4.4.** Let \( F \) and \( G \) be filtrations in \((\Omega, \mathcal{F}, \mathbb{P})\). Then \( F \) is \( \mathbb{P} \)-immersed in \( F \vee G \) if and only if for every \( t \geq 0 \) and every bounded \( \mathcal{G}_t \)-measurable random variable \( \eta \) we have
\[
\mathbb{E}(\eta | \mathcal{F}_t) = \mathbb{E}(\eta | \mathcal{F}_\infty).
\]
(4.17)

**Proof.** The necessity follows from Proposition 5.9.1.1 in [19]. To prove sufficiency it is enough to show, again by Proposition 5.9.1.1 in [19], that for every \( t \geq 0 \) and every bounded \( \mathcal{F}_\infty \)-measurable random variable \( \xi \) it holds that
\[
\mathbb{E}(\xi | \mathcal{F}_t \vee \mathcal{G}_t) = \mathbb{E}(\xi | \mathcal{F}_t).
\]
(4.18)
Fix \( \xi \), we need to show that
\[
\mathbb{E}(\xi 1_A) = \mathbb{E}(\mathbb{E}(\xi | \mathcal{F}_t) 1_A),
\]
(4.18)
for every \( A \in \mathcal{F}_t \vee \mathcal{G}_t \). Towards this end let us consider a family \( \mathcal{U} \) of sets defined as
\[
\mathcal{U} = \{ A : A = B \cap C, B \in \mathcal{F}_t, C \in \mathcal{G}_t \}.
\]
Note that \( \mathcal{U} \) is a \( \pi \)-system of sets which generates \( \mathcal{F}_t \vee \mathcal{G}_t \). Observe that family of all sets for which (4.18) holds is a \( \lambda \)-system. Thus, by the Sierpinski’s Monotone Class Theorem (cf. Theorem 1.1 in [20]), which is also known as the Dynkin’s \( \pi - \lambda \) Theorem, it suffices to prove (4.18) for the sets from \( \mathcal{U} \), which we will do now so to complete the proof.

For \( A \in \mathcal{U} \), we have
\[
\mathbb{E}(\xi 1_A) = \mathbb{E}(\xi 1_{B \cap C}) = \mathbb{E}(\mathbb{E}(\xi 1_{B \cap C} | \mathcal{F}_\infty)) = \mathbb{E}(\xi 1_B \mathbb{E}(1_C | \mathcal{F}_t)) = \mathbb{E}(\xi 1_C \mathbb{E}(1_B | \mathcal{F}_t)) = \mathbb{E}(\xi 1_B \mathbb{E}(1_C | \mathcal{F}_t)) = \mathbb{E}(\xi 1_C \mathbb{E}(1_B | \mathcal{F}_t)) = \mathbb{E}(\mathbb{E}(\xi | \mathcal{F}_t) 1_{B \cap C} | \mathcal{F}_t) = \mathbb{E}(\mathbb{E}(\xi | \mathcal{F}_t) 1_{B \cap C}) = \mathbb{E}(\mathbb{E}(\xi | \mathcal{F}_t) 1_A),
\]
where the fourth equality follows from (4.17) for \( \eta = 1_C \).

\( \square \)
We are now ready to demonstrate the following proposition.

**Proposition 4.5.** Filtration $\mathbb{H}^k$ is $\mathbb{P}$-immersed in $\mathbb{H}^{k+1}$.

**Proof.** Since $\mathbb{H}^{k+1} = \mathbb{H}^k \vee \mathbb{F}^{N^{k+1}}$ we use Lemma 4.4 to prove immersion of $\mathbb{H}^k$ in $\mathbb{H}^k \vee \mathbb{F}^{N^{k+1}}$. It suffices to show that
\[ \mathbb{P}(A|\mathbb{H}_\infty^k) = \mathbb{P}(A|\mathbb{H}_u^k), \]
for every $u \geq 0$ and every $A \in \mathcal{U}$, where
\[ \mathcal{U} = \left\{ A : A = \bigcap_{i=1}^n \{ N^{k+1}((s_i, t_i] \times D_i) = l_i \}, D_1, \ldots, D_n \text{ are disjoint sets, and } 0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \ldots \leq s_n < t_n \leq u, n \in \mathbb{N} \right\}. \]

Indeed, if (4.19) holds for $A \in \mathcal{U}$, then since $\mathcal{U}$ is a $\pi$-system which generates $\mathcal{F}^{N^{k+1}}_u$ the monotone class theorem implies that (4.17) holds. It remains to show (4.19) for $A \in \mathcal{U}$. Using Proposition 4.3 and invoking (7.11) we have
\[ \mathbb{P}\left( \bigcap_{i=1}^n \{ N^{k+1}((s_i, t_i] \times D_i) = l_i \}|\widehat{\mathbb{H}}_0^{k+1} \right) = \prod_{i=1}^n e^{-\nu^{k+1}((s_i, t_i] \times D_i)} \frac{(\nu^{k+1}((s_i, t_i] \times D_i))^{l_i+1}}{l_i!}. \]

Since $\widehat{\mathbb{H}}_0^{k+1} = \mathbb{H}_\infty^k$ and $\nu^{k+1}((s_i, t_i] \times D)$ is $\mathbb{H}_n^k$-measurable we infer that the right hand side of (4.20) is $\mathbb{H}_n^k$-measurable and hence also $\mathbb{H}_n^k$-measurable for arbitrary $u \geq t_n$. Consequently by taking conditional expectations with respect to $\mathbb{H}_u^k$ for $u \geq t_n$ we conclude that (4.19) holds for $A \in \mathcal{U}$. The proof is complete. \qed

We will determine now the compensators for $H^0 := N^0$ and for $H^k$ given by (4.10) for $k \geq 1$.

**Proposition 4.6.** The $\mathbb{H}_0^0$-compensator of $H^0$, is given by
\[ \vartheta^0((s, t] \times D) = \int_s^t \eta(v, D)dv, \quad 0 \leq s \leq t, \ D \in \mathcal{E}^\Delta, \]
where the kernel $\eta$ appears in (4.1).

The $\mathbb{H}_k^k$-compensator of $H^k$, for $k \geq 1$, is given by
\[ \vartheta^k((s, t] \times D) = \int_s^t \left( \eta(v, D) + \int_{(0, v] \times \mathcal{E}^\Delta} f(v, u, y, D)H^{k-1}(du, dy) \right)dv, \quad 0 \leq s \leq t, \ D \in \mathcal{E}^\Delta. \]

**Proof.** The proof goes by induction.

Since $H^0 = N^0$, then the form of $\mathbb{H}_0^0$-compensator of $H^0$ follows from assertion i) of Proposition 4.2 and from Proposition 6.1.4 [22].
Suppose now that $H_k$-compensator of $H^k$ is given by (4.21). This means that for every $D \in \mathcal{E}^\Delta$ the process
\[ M^k_t(D) = (H^k - \vartheta^k)((0, t] \times D), \quad t \geq 0, \] 
(4.22)
is an $H_k$-local martingale. Proposition 4.5 implies that $M^k(D)$ is an $H_{k+1}$-local martingale. We know from Proposition 4.3 that
\[ L^{k+1}_t(D) = (N^{k+1} - \nu^{k+1})((0, t] \times D), \quad t \geq 0, \]
is an $H_{k+1}$-local martingale. Thus $M^k(D) + L^{k+1}_t(D)$ is an $H_{k+1}$-local martingale. This $H_{k+1}$-local martingale can be written in the form
\[ M^k_t(D) + L^{k+1}_t(D) = (H^k + N^{k+1} - (\vartheta^k + \nu^{k+1}))(0, t] \times D, \]
where the second equality follows from $H_{k+1} = H^k + N^{k+1}$.

Note that the random measure $\vartheta^k + \nu^{k+1}$ is $H_{k+1}$-predictable so it is the $H_{k+1}$-compensator of $H^k$. To complete the proof it suffices to show that $\vartheta^k + \nu^{k+1} = \vartheta^{k+1}$. By the induction hypothesis on $\vartheta^k$ and by (4.14) we have
\[ (\vartheta^k + \nu^{k+1})((s, t] \times D) = \int_s^t \left( \eta(v, D) + \int_{(0,v) \times E^\Delta} f(v, u, y)(H^{k-1} + N^k)(du, dy) \right) dv \]
\[ = \vartheta^{k+1}((s, t] \times D). \]
The proof is complete. \qed

Before we conclude our construction of a generalized multivariate Hawkes process, we derive the following result.

**Proposition 4.7.** The $H^\infty$-compensator of $H^\infty$ is given by
\[ \vartheta^\infty((s, t] \times D) = \int_s^t \left( \eta(v, D) + \int_{(0,v) \times E^\Delta} f(v, u, y)H^\infty(du, dy) \right) dv. \]
(4.23)

**Proof.** Proposition 4.5 and Proposition 4.6 imply that for every $k \geq 1$, the $H^\infty$-compensator of $H^k$ is given by (4.21). Thus, we see that for any $k \geq 1$ and for an arbitrary non-negative $H^\infty$-predictable function $F: \Omega \times [0, T] \times E^\Delta \to \mathbb{R}$ it holds
\[ \mathbb{E}\left( \int_0^\infty \int_{E^\Delta} F(v, y)H^k(du, dy) \right) = \mathbb{E}\left( \int_0^\infty \int_{E^\Delta} F(v, y)\vartheta^k(du, dy) \right). \]
Using Lemma 7.5 in an analogous way as in the proof of Proposition 4.3 we obtain
\[ \mathbb{E}\left( \int_0^\infty \int_{E^\Delta} F(u, y)H^\infty(du, dy) \right) = \mathbb{E}\left( \int_0^\infty \int_{E^\Delta} F(u, y)\vartheta^\infty(du, dy) \right). \]
(4.24)
This completes the proof. \qed
We are now ready to conclude our construction of a generalized multivariate Hawkes process. Let $T_\infty$ be the first accumulation time of $H^\infty$. Then we have the following

**Theorem 4.8.** The process $N := \mathbb{1}_{0,T_\infty}[H^\infty]$ is an $\mathbb{E}^N$-Hawkes process with the Hawkes kernel $\kappa(t, dx) = \eta(t, dx) + \int_{(0,t) \times E^\Delta} f(t, u, y, dx)N(du, dy)$.

**Proof.** Let us define a sequence $(T_n, X_n)_{n \geq 1}$ by

$$T_n = \inf \{ t > 0 : H^\infty((0, t] \times E^\Delta) \geq n \},$$

$$X_n = \beta^{-1}\left(\mathbb{1}_{T_n < \infty} \int_{E^\Delta} \beta(x) H^\infty(\{T_n\} \times dx) + \hat{\partial}_{\{T_n = \infty\}}\right),$$

and the random measure

$$N(dt, dx) = \sum_{n > 0} \delta_{(T_n, X_n)}(dt, dx)\mathbb{1}_{\{T_n < \infty\}}.$$

Then

$$N(dt, dx) = H^\infty(dt, dx)|_{0,T_\infty[\times E^\Delta}.$$

Consequently, such restriction of $H^\infty$ to $[0, T_\infty[ \times E^\Delta$ is a marked point process. Moreover, since $[0, T_\infty[ \times E^\Delta$ is an $\mathbb{H}^\infty$-predictable set, we have for arbitrary non negative $\mathbb{H}^\infty$-predictable function $F : \Omega \times \mathbb{R}_+ \times E^\Delta \to \mathbb{R}$

$$\mathbb{E}\left(\int_0^\infty \int_{E^\Delta} F(u, x)N(du, dx)\right) = \mathbb{E}\left(\int_0^\infty \int_{E^\Delta} F(u, x)\mathbb{1}_{0,T_\infty[\times E^\Delta} H^\infty(du, dx)\right),$$

$$= \mathbb{E}\left(\int_0^\infty \int_{E^\Delta} F(u, x)\mathbb{1}_{0,T_\infty[\times E^\Delta} \vartheta^\infty(du, dx)\right), \tag{4.25}$$

where $\vartheta^\infty$ is given in (4.23).

So the compensator of the restriction of $H^\infty$ to $[0, T_\infty[ \times E^\Delta$ is the restriction to $[0, T_\infty[ \times E^\Delta$ of compensator of $H^\infty$. Now we will prove that

$$\vartheta^\infty(dt, dx)|_{0,T_\infty[\times E^\Delta} = \mathbb{1}_{(0,T_\infty(\omega))}(t)\kappa(t, dx)dt,$$

where

$$\kappa(t, dx) = \eta(t, dx) + \int_{(0,t) \times E^\Delta} f(t, u, y, dx)N(du, dy).$$

Towards this end note that for arbitrary $0 \leq s < t \leq T$ and $D \in E^\Delta$ we have

$$\vartheta^\infty|_{0,T_\infty[\times E^\Delta}((s, t] \times D) = \vartheta^\infty((s, t] \cap (0,T_\infty(\omega)) \times D)$$

$$= \int_s^t \mathbb{1}_{(0,T_\infty(\omega))}(v)\left(\eta(v, D) + \int_{(0,v) \times E^\Delta} f(v, u, y, D)H^\infty(du, dy)\right)dv.$$
The second term above can be written as

\[
1_{(0, T_\infty(\omega))}(v) \int_{(0,v) \times E^\Delta} f(v, u, y, D) H^\infty(du, dy)
\]

\[
= \int_{(0,T] \times E^\Delta} 1_{(0,T)}(u) 1_{(0,T_\infty(\omega))}(v) f(v, u, y, D) H^\infty(du, dy)
\]

\[
= \int_{(0,T] \times E^\Delta} 1_{u<v<T_\infty(\omega)}(u) f(v, u, y, D) 1_{(0,T_\infty(\omega))}(v) H^\infty(du, dy)
\]

\[
= \int_{(0,T_\infty(\omega))}(v) \int_{(0,v) \times E^\Delta} f(v, u, y, D) N(du, dy).
\]

Hence

\[
\vartheta^\infty|_{[0,T_\infty] \times E^\Delta}((s,t] \times D)
\]

\[
= \int_s^t 1_{(0,T_\infty(\omega))}(v) \left[ \eta(v, D) + \int_{(0,v) \times E^\Delta} f(v, u, y, D) N(du, dy) \right] dv.
\]

\[
= \int_s^t 1_{(0,T_\infty(\omega))}(v) \kappa(v, D) dv.
\]

This and (4.25) imply that \( \vartheta^\infty|_{[0,T_\infty] \times E^\Delta} \) is an \( \mathbb{F}^N \)-predictable random measure such that for arbitrary non-negative \( \mathbb{F}^N \)-predictable function \( F : \Omega \times \mathbb{R}_+ \times E^\Delta \to \mathbb{R} \) we have

\[
\mathbb{E}\left( \int_0^\infty \int_{E^\Delta} F(u, x) N(du, dx) \right) = \mathbb{E}\left( \int_0^\infty \int_{E^\Delta} F(u, x) 1_{[0,T_\infty]}(u) \kappa(u, dx) du \right).
\]

Thus \( N \) is a \( \mathbb{F}^N \)-Hawkes process (restricted to \([0, T] \times E^\Delta\)) with the Hawkes kernel \( \kappa \).

### 4.3 The pseudo-algorithm

In the description of the pseudo-algorithm below we use the objects \( \eta, f, \hat{\eta}, \hat{f}, G_1, G_2 \) that underly the construction of our Hawkes process given in Section

The steps of the pseudo-algorithm are based on the steps presented in our construction of a generalized multivariate Hawkes process with deterministic kernels \( \eta \) and \( f \), and they are:

**Step 0.** Choose a positive integer \( K \), set \( C^0 = \emptyset \).

**Step 1.** Generate a realization, say \( p \), of a Poisson random variable with parameter \( T\hat{\eta} \).

**Step 2.** If \( p = 0 \), then go to Step 3. Else, if \( p > 0 \), then for \( i = 1, \ldots, p \):

- Generate realizations \( u \) and \( v \) of independent random variables uniformly distributed on \([0,1]\). Set \( t = Tu, a = \hat{\eta} \).
– If \( a \leq \eta(t, E^\Delta) \), then generate a realization \( w \) of random variable uniformly distributed on \([0, 1]\), compute \( x = G_1(t, w) \) and include \((t, x)\) into the cluster \( C^0 \).

Step 3. Set \( N = C^0, C^{\text{prev}} = C^0, k = 0 \).

Step 4. While \( C^{\text{prev}} \neq \emptyset \) and \( k \leq K \):

– Set \( C^{\text{new}} = \emptyset \).

– For every \((s, y) \in C^{\text{prev}}\):

  * generate a realization \( p \) of Poisson random variable with parameter \((T - s)\hat{f}(s, y)\).
  * for \( i = 1, \ldots, p \):
    
    ◦ Generate realizations \( u \) and \( v \) of independent random variables uniformly distributed on \([0, 1]\). Set \( t = s + (T - s)u, a = \hat{f}(s, y)v \).
    
    ◦ If \( a \leq f(t, s, y, E^\Delta) \), then generate a realization \( w \) of random variable uniformly distributed on \([0, 1]\), compute \( x = G_2(t, s, y, w) \) and include \((t, x)\) into the cluster \( C^{\text{new}} \).

– Set \( N = N \cup C^{\text{new}}, C^{\text{prev}} = C^{\text{new}} \).

– Set \( k = k + 1 \).

Step 5. Return \( N \).

4.3.1 Numerical examples via simulation

The pseudo-algorithm presented above is implemented here in two cases. In the first case, presented in Example 4.9, we implemented the algorithm for a generalized bivariate Hawkes process with \( E_1 = E_2 = \{1\} \). In the second case, presented in Example 4.10, we set \( E_1 = E_2 = \mathbb{R} \).

We used Python to run the simulations and to plot graphs.

Example 4.9. Bivariate point Hawkes process

Here we implement our pseudo-algorithm for a bivariate point Hawkes process, that is the generalized bivariate Hawkes process with \( E_1 = E_2 = \{1\} \), and hence with

\[ E^\Delta = \{(1, \Delta), (\Delta, 1), (1, 1)\} \]

Moreover, we let

\[ \eta(t, dy) := \eta_1(t)\delta_{(1, \Delta)}(dy) + \eta_2(t)\delta_{(\Delta, 1)}(dy) + \eta_c(t)\delta_{(1, 1)}(dy), \]

where

\[ \eta_i(t) := \alpha_i + (\eta_i(0) - \alpha_i)e^{-\beta_i t}, \quad i \in \{1, 2, c\}, \]

and \( \alpha_i, \eta_i(0), \beta_i \) are non-negative constants. We assume that, for \( 0 \leq s \leq t \), the kernel \( f \) is given as in (3.11) with the decay functions \( w_{i,j} \) in the exponential form:

\[ w_{i,j}(t, s) = e^{-\beta_i(t-s)}, \quad i, j \in \{1, 2, c\}, \]
with constant non-negative impact functions:
\[
\begin{align*}
&g_{1,1}(x_1) = \vartheta_{1,1}, \quad g_{1,2}(x_2) = \vartheta_{1,2}, \quad g_{1,c}(x) = \vartheta_{1,c}, \\
&g_{2,1}(x_1) = \vartheta_{2,1}, \quad g_{2,2}(x_2) = \vartheta_{2,2}, \quad g_{2,c}(x) = \vartheta_{2,c}, \\
&g_{c,1}(x_1) = \vartheta_{c,1}, \quad g_{c,2}(x_2) = \vartheta_{c,2}, \quad g_{c,c}(x) = \vartheta_{c,c},
\end{align*}
\]

and with Dirac kernels:
\[
\phi_1(x, dy_1) = \delta_1(dy_1), \quad \phi_2(x, dy_2) = \delta_1(dy_2), \quad \phi_c(x, dy_1, dy_2) = \delta_{(1,1)}(dy_1, dy_2).
\]

Thus, the kernel \( f \) is of the form
\[
f(t, s, x, dy) = e^{-\beta_1(t-s)} \left( \vartheta_{1,1} \mathbb{1}_{\{1\} \times \Delta}(x) + \vartheta_{1,2} \mathbb{1}_{\Delta \times \{1\}}(x) + \vartheta_{1,c} \mathbb{1}_{\{1\} \times \{1\}}(x) \right) \delta_{(1,\Delta)}(dy)
+ e^{-\beta_2(t-s)} \left( \vartheta_{2,1} \mathbb{1}_{\{1\} \times \Delta}(x) + \vartheta_{2,2} \mathbb{1}_{\Delta \times \{1\}}(x) + \vartheta_{2,c} \mathbb{1}_{\{1\} \times \{1\}}(x) \right) \delta_{(\Delta,1)}(dy)
+ e^{-\beta_c(t-s)} \left( \vartheta_{c,1} \mathbb{1}_{\{1\} \times \Delta}(x) + \vartheta_{c,2} \mathbb{1}_{\Delta \times \{1\}}(x) + \vartheta_{c,c} \mathbb{1}_{\{1\} \times \{1\}}(x) \right) \delta_{(1,1)}(dy).
\]

The coordinates of \( N \) (cf. \((3.2)\)) reduce here to counting (point) processes, so that
\[
N^1_t = N^1((0, t], \{1\}) = N((0, t], \{1\} \times \{1, \Delta\}),
\]

and
\[
N^2_t = N^2((0, t], \{1\}) = N((0, t], \{1, \Delta\} \times \{1\}).
\]

Moreover, \( N^{\text{idio},\{1,2\}}_t \) – the MPP of idiosyncratic group of \{1,2\} coordinates – reduces here to the process counting the number of occurrences of the common events:
\[
N^c_t = N^{\text{idio},\{1,2\}}((0, t], \{(1, 1)\}) = N((0, t], \{(1, 1)\}).
\]

We take the following values of parameters:

| \( i \) | \( \eta_i(0) \) | \( \alpha_i \) | \( \beta_i \) | \( \vartheta_{i,j} \) | \( j \) |
|---|---|---|---|---|---|
| 1 | 0.5 | 0.5 | 2.5 | | 1 | 0.5 | 0.25 | 0.25 |
| 2 | 0.5 | 0.5 | 2.5 | | 2 | 0.25 | 0.5 | 0.25 |
| c | 0.25 | 0.25 | 5.0 | | c | 0.25 | 0.25 | 0.25 |

Simulated sample paths of \( N \) corresponding to the above setting are presented in Figure 1 and Figure 2.

**Example 4.10. Bivariate Hawkes process**

Here we apply our pseudo-algorithm to Example 3.9 with \( d = 2 \) and \( E_1 = E_2 = \mathbb{R} \). We let:
\[
\begin{align*}
\eta_1(t, dy_1) &= \alpha_1 \varphi_{\mu_1,\sigma_1}(y_1)dy_1, \\
\eta_2(t, dy_2) &= \alpha_2 \varphi_{\mu_2,\sigma_2}(y_2)dy_2, \\
\eta_1(t, dy_1) &= \alpha_c \varphi_{\mu_c,\sigma_c}(y_1)\varphi_{\mu_c,\sigma_c}(y_2)dy_1dy_2.
\end{align*}
\]
where $\alpha_i \geq 0, i \in \{1, 2, c\}$, $\varphi_{\mu, \sigma}$ is the one dimensional Gaussian density function with mean $\mu$ and variance $\sigma^2$, and:

$$w_{1,i}(t, s) = w_{2,i}(t, s) = w_{c,i}(t, s) = e^{-\beta_i(t-s)}, \quad i = 1, 2, c.$$  

Moreover, we set:

$$g_{1,1}(x_1) = g_{1,1}, \quad g_{2,1}(x_1) = 0, \quad g_{c,1}(x_1) = g_{c,1},$$

$$g_{1,2}(x_2) = 0, \quad g_{2,2}(x_2) = g_{2,2}, \quad g_{c,2}(x_2) = g_{c,2},$$

$$g_{1,c}(x) = 0, \quad g_{2,c}(x) = 0, \quad g_{c,c}(x) = g_{c,c};$$

and we take

$$\phi_1(x, dy_1) = 1_{E_1 \times \Delta(x)} \varphi_{a_{1,1}, \sigma_1}(y_1) dy_1 + 1_{\Delta \times E_2(x)} \varphi_{0, \sigma_1}(y_1) dy_1,$$

$$\phi_2(x, dy_2) = 1_{\Delta \times E_2(x)} \varphi_{a_{2,2}, \sigma_2}(y_2) dy_2 + 1_{E_1 \times \Delta(x)} \varphi_{0, \sigma_1}(y_1) dy_2,$$

$$\phi_c(x, dy_1, dy_2) = 1_{E_1 \times \Delta(x)} \varphi_{a_{c,1}, \sigma_c}(y_1) \varphi_{0, \sigma_c}(y_2) + 1_{\Delta \times E_2(x)} \varphi_{0, \sigma_c}(y_1) \varphi_{a_{c,2}, \sigma_c}(y_2) + 1_{E_1 \times E_2(x)} \varphi_{a_{c,1}, \sigma_c}(y_1) \varphi_{a_{c,2}, \sigma_c}(y_2);$$

**Figure 1:** Bar plot of 10 paths of a bivariate point Hawkes process. Red bars represent common events, black bars represents idiosyncratic events.
Figure 2: Plot of a single path of counting processes associated with 2-variate Hawkes process.

with the following values of the parameters:

\[
\begin{array}{cccccc}
  \alpha_i & \mu_i & \sigma_i & \beta_i & \alpha_i & g_{i,i} & g_{c,i} \\
  1 & 0.4 & 2 & 0.16331 & 0.41175 & 0.9 & 0.3 & 0.1 \\
  2 & 0.4 & -2 & 0.16331 & 0.41175 & 0.9 & 0.3 & 0.1 \\
  c & 0.2 & 0 & 0.16331 & 0.81175 & 1.1 & 0.4 & 0.4 \\
\end{array}
\]

A simulated sample path is presented on Figure 3.

5 Markovian aspects of a generalized bivariate Hawkes process

An important class of Hawkes processes considered in the literature is the one of Hawkes processes for which the Hawkes kernel is given in terms of exponential decay functions. See, e.g., [7], [25], [33]. One interesting and useful aspect of such processes is that they can be extended to Markov processes, a feature that we term the Markovian aspects of a generalized bivariate Hawkes process.

To simplify the presentation, we will discuss Markovian aspects of generalized bivariate Hawkes processes specified in Example 4.9. Using this specification we end up with the
Figure 3: Plot of a simulated path of the bivariate Hawkes process specified in Example 4.10.

Hawkes kernel $\kappa$ of the form:

$$\kappa(t, dy) = \lambda_1^t \delta_{(1, \Delta)}(dy) + \lambda_2^t \delta_{(\Delta, 1)}(dy) + \lambda_c^t \delta_{(1, 1)}(dy),$$

where, for $i = 1, 2, c$, we have $\lambda_i^0 := \eta_i(0)$ and

$$\lambda_i^t = \alpha_i + (\lambda_i^0 - \alpha_i)e^{-\beta_i t}$$

$$+ \int_{(0,t) \times E \Delta} e^{-\beta_i (t-u)} \left( \vartheta_{i,1} 1_{\{1\} \times \Delta}(x) + \vartheta_{i,2} 1_{\Delta \times \{1\}}(x) + \vartheta_{i,c} 1_{\{1\} \times \{1, \Delta\}}(x) \right) N(du, dx).$$

We now refer to canonical space as in Section 2.1 and to the random measure $\nu$ corresponding to $\kappa$ as in (2.3). So, using Theorem 2.5 we see that there exists a unique probability $P_\nu$ such that the canonical process $N$ given as in (3.10) is a generalized multivariate Hawkes process with Hawkes kernel $\kappa$.

The coordinates of $N$ (cf. (3.2)) reduce here to counting (point) processes

$$N_t^1 = N^1((0, t], \{1\}) = N((0, t], \{1\} \times \{1, \Delta\}),$$

and

$$N_t^2 = N^2((0, t], \{1\}) = N((0, t], \{1, \Delta\} \times \{1\}).$$
It is straightforward to verify (upon appropriate integration of the kernel $\kappa$ i.e. over $\{1\} \times \{1, \Delta\}$ for $N^1$ and $\{1, \Delta\} \times \{1\}$ for $N^2$) that the $\mathbb{R}^N$–intensity of process $N^i$, say $\tilde{\lambda}^i$, is given as
\begin{equation}
\tilde{\lambda}^i_t = \lambda^i_t + \lambda^c_i, \quad t \geq 0,
\end{equation}
for $i = 1, 2$. Let

$$
\tilde{N}^c_i = [N^1, N^2]_u, \quad \tilde{N}^1_i = N^1_u - \tilde{N}^c_i, \quad \tilde{N}^2_i = N^2_u - \tilde{N}^c_i,
$$

where $[N^1, N^2]$ is the square bracket of $N^1$, $N^2$. Then, for $i = 1, 2, c$, the equality (5.2) can be written as

$$
\lambda^i_t = \alpha_i + (\lambda^i_0 - \alpha_i)e^{-\beta_i t} + \int_{(0,t)} e^{-\beta_i (t-u)} \left( \vartheta_{i,1} d\tilde{N}^1_u + \vartheta_{i,2} d\tilde{N}^2_u + \vartheta_{i,c} d\tilde{N}^c_u \right)
$$

for $t \geq 0$. This follows from the fact that $[N^1, N^2]$ counts common jumps of $N^1$ and $N^2$, so for $i = 1, 2$ the process $\tilde{N}^i$ is counts the idiosyncratic jumps of $N^i$, that is the jumps that do not occur simultaneously with the jumps of $N^j$, $j \neq i$. In particular, expression (5.4) allows us to give the interpretation of the parameters $\vartheta_{i,j}$, $i, j \in \{1, 2, c\}$, namely the parameter $\vartheta_{i,j}$ describes the impact of the jump of the process $\tilde{N}^j$ on the intensity of $\tilde{N}^i$.

Now, let us consider a bivariate counting process $\tilde{N} := (N^1, N^2)$. Note that we may, and we do, identify process $\tilde{N}$ with our bivariate generalized Hawkes process $N$:

$$
T_0 = 0, \quad T_n = \inf \{ t > T_{n-1} : \Delta \tilde{N}_t \neq (0,0) \},
$$

and for $i = 1, 2$

$$
X^i_n = \begin{cases} 1 & \text{if } \Delta \tilde{N}^i_n = 1, \\
\Delta & \text{if } \Delta \tilde{N}^i_n = 0. \end{cases}
$$

Also, note that we may, and we do, identify the process $\tilde{N}$ with a random measure $\mu^{\tilde{N}}$ on $\mathbb{R}_+ \times \tilde{E}$, where $\tilde{E} = \{(1,0), (0,1), (1,1)\}$, given by

$$
\mu^{\tilde{N}}(dt, dy) = \sum_{n \geq 0} \delta_{(T_n, Y^i_n)}(dt, dy) \mathbb{1}_{\{T_n < \infty\}},
$$

where $Y^i_n = \mathbb{1}_{\{X^i_n = 1\}}$. Using (5.1) we see that $\mathbb{R}^{\tilde{N}}$–compensator of $\mu^{\tilde{N}}$ is given by

$$
\tilde{\nu}(dt, dy) = \mathbb{1}_{[0, T_{\infty}]} \tilde{\kappa}(t, dy) dt,
$$

where

$$
\tilde{\kappa}(t, dy) = \lambda^1 t \delta_{(1,0)}(dy) + \lambda^2 t \delta_{(0,1)}(dy) + \lambda^c t \delta_{(1,1)}(dy).
$$

Thus, we may slightly abuse terminology and call $\tilde{N}$ a generalized bivariate Hawkes process.
Theorem 5.1. Let $N$ be a Hawkes process defined as above. Then
i) The process $\tilde{N} = (N^1_t, N^2_t)_{t \geq 0}$ is not a Markov process.
ii) The process
$$Z = (\lambda^1_t, \lambda^2_t, \lambda^c_t, N^1_t, N^2_t)_{t \geq 0}$$
is a Markov process with the strong generator $A$ acting on $C^\infty_c(\mathbb{R}_+^5)$ given by

\begin{equation}
Av(\lambda^1, \lambda^2, \lambda^c, n^1, n^2)
= \beta_1(\alpha_1 - \lambda^1) \frac{\partial}{\partial \lambda^1} v(\lambda^1, \lambda^2, \lambda^c, n^1, n^2) + \beta_2(\alpha_2 - \lambda^2) \frac{\partial}{\partial \lambda^2} v(\lambda^1, \lambda^2, \lambda^c, n^1, n^2)
+ \beta_c(\alpha_c - \lambda^c) \frac{\partial}{\partial \lambda^c} v(\lambda^1, \lambda^2, \lambda^c, n^1, n^2)
+ (v(\lambda^1 + \vartheta_{1,1}, \lambda^2 + \vartheta_{2,1}, \lambda^c + \vartheta_{c,1}, n^1 + 1, n^2) - v(\lambda^1, \lambda^2, \lambda^c, n^1, n^2)) \lambda^1
+ (v(\lambda^1 + \vartheta_{1,2}, \lambda^2 + \vartheta_{2,2}, \lambda^c + \vartheta_{c,2}, n^1, n^2 + 1) - v(\lambda^1, \lambda^2, \lambda^c, n^1, n^2)) \lambda^2
+ (v(\lambda^1 + \vartheta_{1,c}, \lambda^2 + \vartheta_{2,c}, \lambda^c + \vartheta_{c,c}, n^1 + 1, n^2 + 1) - v(\lambda^1, \lambda^2, \lambda^c, n^1, n^2)) \lambda^c.
\end{equation}

Proof. i) From (5.4) and (5.6) we see that for any $t > 0$ the quantity $\tilde{N}(dt, dy)$ given in (5.5) depends on the entire path of $\tilde{N}$ until time $t$. Thus, by Theorem 4 in [15], the process $\tilde{N}$ is not a Markov process.

ii) First note that (5.4) can be written as
$$\lambda^i_t - \alpha_i = e^{-\beta_t} \left( \lambda^i_0 - \alpha_i + \int_{(0,t)} e^{\beta_u} \left( \vartheta_{i,1} d\tilde{N}^1_u + \vartheta_{i,2} d\tilde{N}^2_u + \vartheta_{i,c} d\tilde{N}^c_u \right) \right).$$

Hence using stochastic integration by parts one can show that $\lambda^i$ can be represented as
$$\lambda^i_t = \lambda^i_0 + \int_0^t \beta_i(\alpha_i - \lambda^i_u) du + \int_{(0,t)} \left( \vartheta_{i,1} d\tilde{N}^1_u + \vartheta_{i,2} d\tilde{N}^2_u + \vartheta_{i,c} d\tilde{N}^c_u \right).$$

This and (5.6) implies that the process $Z$ is an $\mathbb{F}^Z$–semimartingale with characteristics (with respect to cut-off function $h(x) = x 1_{|x| < 1}$)
$$B_t = \int_0^t b_u du, \quad C_t = 0_{4 \times 4},$$

$$\nu(dt, dy_1, dy_2, dy_c, dz_1, dz_2) = \nu_t(dy_1, dy_2, dy_c, dz_1, dz_2) dt,$$

where
$$b_t := (\beta_1(\alpha_1 - \lambda^1_{t-}), \beta_2(\alpha_2 - \lambda^2_{t-}), \beta_c(\alpha_c - \lambda^c_{t-}), 0, 0)'$$

and
\begin{equation}
\nu_t(dy_1, dy_2, dy_c, dz_1, dz_2)
:= \lambda^1_{-u} \delta(\vartheta_{1,1}, \vartheta_{1,2}, \vartheta_{c,1}, 1, 0)(dy_1, dy_2, dy_c, dz_1, dz_2)
+ \lambda^2_{-u} \delta(\vartheta_{1,2}, \vartheta_{2,2}, \vartheta_{c,2}, 0, 1)(dy_1, dy_2, dy_c, dz_1, dz_2)
+ \lambda^c_{-u} \delta(\vartheta_{1,c}, \vartheta_{2,c}, \vartheta_{c,c}, 1, 1)(dy_1, dy_2, dy_c, dz_1, dz_2).
\end{equation}
This, by Theorem II.2.42 in [18], implies that for any function \( v \in C_b^2(\mathbb{R}^5) \) the process \( M^v \) given as
\[
M^v_t = v(Z_t) - \int_0^t Av(Z_u)du = v(\lambda^1_t, \lambda^2_t, \lambda^c_t, N^1_t, N^2_t) - \int_0^t Av(\lambda^1_u, \lambda^2_u, \lambda^c_u, N^1_u, N^2_u)du, \quad t \geq 0,
\]
is an \( \mathbb{F}^Z \)-local martingale. Hence, for any \( v \in C_c^\infty(\mathbb{R}^5) \) the process defined above is a martingale under \( \mathbb{P} \), since \( v \) and \( Av \) are bounded, which follows from the fact that \( v \in C_c^\infty(\mathbb{R}^5) \) has compact support, and thus the local martingale \( M^v \) is a martingale for such \( v \). Consequently, the process \( Z \) solves martingale problem for \( (A, \rho) \), where \( \rho \) is the deterministic initial distribution of \( Z \), that is \( \rho(dz) = \delta_{Z_0}(dz) \).

We will now verify that \( Z \) is a Markov process with generator \( A \) given in (5.7) using Theorem 4.4.1 in [10].

For this, we first observe that parameters determining \( A, \) i.e.
\[
\mathcal{I} = \{1, \ldots, 5\}, \quad \mathcal{J} = \emptyset, \quad a = 0, \quad \alpha = 0, \quad c = 0, \quad \gamma = 0, \quad m = 0,
\]
\[
b = (\alpha_1 \beta_1, \alpha_2 \beta_2, \alpha_3 \beta_3, 0, 0)' , \quad \beta = \text{diag}(-\beta_1, -\beta_2, -\beta_3, 0, 0),
\]
\[
\mu_1 = \delta(\vartheta_{1,1}, \vartheta_{1,2}, \vartheta_{1,1}, 0, 0), \quad \mu_2 = \delta(\vartheta_{1,2}, \vartheta_{2,2}, \vartheta_{1,1}, 0, 1),
\]
\[
\mu_3 = \delta(\vartheta_{1,c}, \vartheta_{2,c}, \vartheta_{1,1}, 1, 1), \quad \mu_4 = \mu_5 = 0,
\]
are admissible in the sense of Definition 2.6 in [8].

Thus, invoking Theorem 2.7 in [8] we conclude that there exists a unique regular affine semigroup \( (P_t)_{t \geq 0} \) with infinitesimal generator \( A \) given by (5.7). Hence, there exists a unique regular affine process with generator \( A \) and with transition function \( P \) defined by \( (P_t)_{t \geq 0} \). Since \( A \) is a generator of regular affine process it satisfies the Hille-Yosida conditions (cf. Theorem 1.2.6 in [10]) relative to the Banach space \( B(\mathbb{R}^5) \) of real valued, bounded and measurable functions on \( \mathbb{R}^5 \). Moreover, from Corollary 1.1.6 in [10] it follows that \( A \) is a closed operator. Now, using Theorem 4.4.1 in [10] we obtain that \( Z \) is a Markov process with generator \( A \). Moreover, \( P \) is the transition function of \( Z \).

Let us note that using analogous argument as in the proof of Theorem 5.1 we can prove that the process \( Y^1 := (\lambda^1_t + \lambda^c_t, N^1_t)_{t \geq 0} \) is a Markov process in filtration \( \mathbb{F}^Z \) provided that parameters of \( \lambda^k, k \in \{1, c\} \), satisfy
\[
\vartheta_{1,2} = \vartheta_{c,2} = 0, \quad \beta_1 = \beta_c, \quad \vartheta_{1,c} + \vartheta_{c,c} = \vartheta_{1,1} + \vartheta_{c,1}.
\]
Analogous statement is valid for \( Y^2 := (\lambda^2 + \lambda^c, N^2) \).

### 6 Applications

The are numerous potential applications of the generalized multivariate Hawkes processes. Here we present a brief description of possible applications in seismology, in epidemiology and in finance.
6.1 Seismology

In the Introduction to [27] the author writes:

"Lists of earthquakes are published regularly by the seismological services of most countries in which earthquakes occur with frequency. These lists supply at least the epicenter of each shock, focal depth, origin time and instrumental magnitude.

Such records from a self-contained seismic region reveal time series of extremely complex structure. Large fluctuations in the numbers of shocks per time unit, complicated sequences of shocks related to each other, dependence on activity in other seismic regions, fluctuations of seismicity on a larger time scale, and changes in the detection level of shocks, all appear to be characteristic features of such records. In this manuscript the origin times are mainly considered to be modeled by point processes, with other elements being largely ignored, except that the ETAS model and its extensions use data of magnitudes and epicenters."

In particular, the dependence on (simultaneous) seismic activity in other seismic regions has been ignored in the classical univariate ETAS\textsuperscript{[6]} model, and in all other models that we are aware of.

The ETAS model is a univariate self-exciting point process, in which the shock intensity at time $t$, corresponding to a specific seismic location, is designed as (cf. Equation (17) in [27])

$$\lambda(t|H_t) = \mu + \sum_{t_m < t} \frac{K_m}{(t - t_m + c)^p}. \tag{6.1}$$

In the above formula, $H_t$ stands for the history of after-shocks at the given location, $\mu$ represents the background occurrence rate of seismic activity at his location, $t_m$s are the times of occurrences of all after-shocks that took place prior to time $t$ at the specific seismic location, and

$$K_m = K_0 e^{\alpha(M_m - M_0)},$$

where $M_m$ is the magnitude of the shock occurring at time $t_m$, and $M_0$ is the cut-off magnitude of the data set; we refer to [27] for details. As said above, dependence between (simultaneous) seismic activity in different seismic regions has been ignored in the classical univariate ETAS model.

Below we suggest a possible method to construct a generalized multivariate Hawkes process that may offer a good way of modeling of joint seismic activities at various locations, accounting for dependencies between seismic activities at different locations and for consistencies with local data.

We will now briefly describe this construction that leads to a plausible model, which we name the \textit{multivariate generalized ETAS model}. Towards this end we consider a GMHP $N$ (cf. Definition 3.1), where the index $i = 1, \ldots, d$ represents the $i$-th seismic location, and where the set $E_i = M_i := \{m_1, m_2, \ldots, m_{n_i}\}$ of marks is a discrete set whose elements represent possible magnitudes of seismic shocks with epicenter at location $i$. In the corresponding Hawkes kernel $\kappa$ the measure $\eta(t, dy)$ represents the time-$t$ background distribution of shocks’ across all seismic regions, and the measure $f(t, s, dy, x)$ represents the feedback effect.

\footnotesize{\textsuperscript{6}The Epidemic-type Aftershock-sequences Model}
For the purpose of illustration, let $d = 2$. Suppose that local seismic data are collected for each location to the effect of producing local kernels of the form

$$
\kappa^i(t, \{y_i\}) = \chi^i(t, \{y_i\}) + \int_{(0,t) \times E_1} h_{i,1}(t, s, x_1, \{y_i\}) \mathcal{N}^{idio,1}(ds, dx_1)
+ \int_{(0,t) \times E_2} h_{i,2}(t, s, x_2, \{y_i\}) \mathcal{N}^{idio,2}(ds, dx_2)
+ \int_{(0,t) \times E_1 \times E_2} h_{i,c}(t, s, x, \{y_i\}) \mathcal{N}(ds, dx), \quad i = 1, 2.
$$

In particular, the quantity $\lambda^i(t) := \kappa^i(t, E_i) = \sum_{y_i \in \mathfrak{M}_i} \kappa^i(t, y_i)$ represents the time-$t$ intensity of seismic activity at the $i$-th location.

In order to produce an ETAS type model, we postulate that

$$
\sum_{y_i \in \mathfrak{M}_i} h_{i,j}(t, s, x_j, \{y_i\}) = K_{i,j,0} e^{\alpha_{i,j,0}(x_j-x_{j,0})} (t - s + c)^{p_{i,j}},
$$

for $j = 1, 2$ and

$$
\sum_{y_i \in \mathfrak{M}_i} h_{i,c}(t, s, x, \{y_i\}) = K_{i,c,0} e^{\alpha_{i,c,0}[(x_1-x_{1,0})+(x_2-x_{2,0})]} (t - s + c)^{p_{i,c}}.
$$

Thus,

$$
\lambda^i(t) = \sum_{y_i \in \mathfrak{M}_i} \left( \chi^i(t, \{y_i\}) + \sum_{j=1}^2 \sum_{t_{j,m} < t} K_{i,j,0} e^{\alpha_{i,j}(X_{j,t_{j,m}}-x_{j,0})} (t - s + c)^{p_{i,j}}
+ \sum_{t_{c,m} < t} K_{i,c,0} e^{\alpha_{i,c}[(X_{1,t_{c,m}}-x_{1,0})+(X_{2,t_{c,m}}-x_{2,0})]} (t - s + c)^{p_{i,c}} \right),
$$

where

- $t_{j,m}$ are the times of occurrences of after-shocks that took place prior to time $t$ only at the $i$-th seismic location, and $X_{j,t_{j,m}}$ is the magnitude of the aftershock at location $i$ that took place at time $t_{j,m}$;

- $t_{c,m}$ are the times of occurrences of after-shocks that took place prior to time $t$ both seismic locations, and $X_{j,t_{c,m}}$ is the magnitude of the aftershock at location $i$ that took place at time $t_{c,m}$.

The classical univariate ETAS model has been extended in [26] to the (classical) univariate space-time ETAS model (see also Section 5 in [27]). It is important to note that our generalized multivariate Hawkes process may also be used as an useful generalization of the space-time extension of the multivariate generalized ETAS model. In order to see this, let
us consider the model (2.1) in [26] with $g$ as in Section 2.1 in [26], that is (in the original notation of [26], which should not be confused with our notation)

$$
\lambda(t, x, y|H_t) = \mu(x, y) + \int_0^t \int_0^\infty g(t-s, x-\xi, y-\eta; M) N(ds, d\xi, d\eta, dM).
$$

(6.4)

Then, coming back to our generalized multivariate Hawkes process, let the seismic location $i = 1, 2$ be identified with a point in the plane with coordinates $(a_i, b_i) \in \mathbb{R}^2$. Next, let the set of marks $E_i$ be given as

$$
E_i := D_i \times M_i,
$$

(6.5)

where $D_i = [a_i - a_i', b_i - b_i'] \times [a_i + a_i'', b_i + b_i'']$ for some positive numbers $a_i', a_i'', b_i', b_i''$. This will lead to a space-time generalized multivariate Hawkes process that will be studied elsewhere.

### 6.2 Epidemiology

It was already observed by Hawkes in [12] that Hawkes processes may find applications in epidemiology for modeling spread of epidemic diseases accounting for various types of cases, such as children or adults, that can be taken as marks. This insight has been validated over the years in numerous studies. We refer for example to [30, 28, 21] and the references therein.

It is important to account for the temporal and spatial aspects in the modeling of spread and intensity of epidemic and pandemic diseases, such as COVID-19. We believe that the variant of the generalized multivariate Hawkes process that we described at the end of Section 6.1 may offer a valuable tool in this regard. This will be investigated in a follow-up work.

### 6.3 Finance

Hawkes processes have found important applications in finance over the past two decades. We refer to [13] for a relevant survey. Here, we briefly discuss a possible application in finance of the generalized multivariate Hawkes processes.

In a series of papers [2, 3, 1] introduced a multidimensional model for stock prices driven by (multivariate) Hawkes processes. The model for stock prices is formulated in [2] via a marked point process $N = (T_n, Z_n)_{n \geq 1}$, where $Z_n$ is a random variable taking values in $\{1, \ldots, 2d\}$, and the compensator $\nu$ of $N$ has the form (it is assumed that $T_\infty = \lim_{n \to \infty} T_n = \infty$)

$$
\nu(dt, dy) = \sum_{i=1}^{2d} \delta_i(dy)\lambda_i(t)dt,
$$

where

$$
\lambda_i(t) = \mu_i + \sum_{j=1}^{2d} \int_{(0, t]} \phi_{i,j}(t-s)N(ds \times \{j\}), \quad t \geq 0,
$$

with $\mu_i \in \mathbb{R}_+$ and functions $\phi_{i,j}$ from $\mathbb{R}_+$ to $\mathbb{R}_+$. Let us define the processes $N^i$, $i = 1, \ldots, 2d$, by

$$
N^i((0, t]) = \sum_{n \geq 1} \mathbb{1}_{\{T_n \leq t\} \cap \{Z_n = i\}}, \quad t \geq 0.
$$
Note that the above implies that $N^1, \ldots, N^{2d}$ have no common jumps and the $\mathbb{P}^N$-intensity of $N^i$ is given by $\lambda_i$ and can be written in the form

$$
\lambda_i(t) = \mu_i + \sum_{j=1}^{2d} \int_{(0,t)} \phi_{i,j}(t-s) N^j(ds), \quad t \geq 0.
$$

In [2] it is assumed that a $d$-dimensional vector of assets prices $S = (S^1, \ldots, S^d)$ is based on $N$ via representation

$$
S^i_t = N^{2i-1}((0,t]) - N^{2i}((0,t]), \quad t \geq 0, \; i = 1, \ldots, d.
$$

The obvious interpretation is that $N^{2i-1}$ corresponds to an upward jump of the $i$-th asset whereas $N^{2i}$ corresponds to a downward jump of $i$-th asset. Bacry et.al. [2] showed that within such framework some stylised facts about high frequency data, such as microstructure noise and the Epps effect, are reproduced.

Using the GMHPs we can easily generalize their model in several directions. In particular, a model of stock price movements driven by a generalized multivariate Hawkes process $N$ allows for common jumps in upward and/or downward direction. This can be done by setting the multivariate mark space of $N$ to be

$$
E^\Delta = \{e = (e_1, \ldots, e_{2d}) : e_i \in \{1, \Delta\} \} \setminus \{(\Delta, \ldots, \Delta)\},
$$

and the $\mathbb{P}^N$-compensator of $N$ to be

$$
\nu(dt, dy) = \mathbb{1}_{[0,t]}(t) \sum_{e \in E^\Delta} \delta_e(dy) \lambda_e(t) dt,
$$

where

$$
\lambda_e(t) = \mu_e + \int_{E^\Delta \times (0,t]} \phi_{e,x}(t-s) N(ds \times dx), \quad e \in E^\Delta, \quad t \geq 0,
$$

and where $\mu_e \in \mathbb{R}_+$ and $\phi_{e,x}$ is a function from $\mathbb{R}_+$ to $\mathbb{R}_+$.

Including possibility of embedding co-jumps of the prices of various stocks in the book in the common excitation mechanism, may turn out to be important in modeling the book evolution in general, and in pricing basket options in particular.

7 Appendix

In this appendix we provide some auxiliary concepts and results that are needed in the rest of the paper.

7.1 Conditional Poisson random measure: definition and specific construction

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{X}, \mathcal{X})$ be a Borel space. For a given sigma field $\mathcal{G} \subseteq \mathcal{F}$, we define a $\mathcal{G}$-conditionally Poisson random measure on $(\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X})$ as follows:
Definition 7.1. Let $\nu$ be a $\sigma$-finite random measure on $(\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X})$. A random measure $N$ on $(\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X})$ is a $\mathcal{G}$-conditionally Poisson random measure with intensity measure $\nu$ if the following two properties are satisfied:

1. For every $C \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X}$ such that $\nu(C) < \infty$, we have
   $$\mathbb{P}(N(C) = k|\mathcal{G}) = e^{-\nu(C)}(\nu(C))^k/k!.$$  

2. For arbitrary $n = 1, 2, \ldots$, and arbitrary disjoint sets $C_1, \ldots, C_n$ from $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X}$, such that $\nu(C_m) < \infty$, $m = 1, \ldots, n$, the random variables
   $$N(C_1), \ldots, N(C_n)$$  
   are $\mathcal{G}$-conditionally independent.

Clearly $\nu$ is $\mathcal{G}$-measurable. Note that if $\mathcal{G}$ is trivial $\sigma$-field (or if $N$ is independent of $\mathcal{G}$), then $N$ is a Poisson random measure (see Chapter 4.19 in [29]), which sometimes referred to as the Poisson process on $\mathbb{R}_+ \times \mathcal{X}$ (see e.g. [20]). In this case $\nu$ is a deterministic $\sigma$-finite measure. For $\mathcal{G} = \sigma(\nu)$, the $\sigma(\nu)$-conditional Poisson random measure is also known in the literature as Cox process directed by $\nu$ (see [20]).

Now we will provide a construction of a $\mathcal{G}$-conditional Poisson random measure with the intensity measure given in terms of a specific kernel $g$. In fact, the measure constructed below is supported on sets from $\mathcal{B}((0, T]) \otimes \mathcal{X}$, in the sense that for any set $C$ that has an empty intersection with $(0, T] \times \mathcal{X}$ the value of the measure is 0 almost surely.

We begin by letting $g(t, y, dx)$ be a finite kernel from $(\mathbb{R}_+ \times \mathcal{Y}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{Y})$ to $(\mathcal{X}, \mathcal{X})$, where $(\mathcal{Y}, \mathcal{Y})$ and $(\mathcal{X}, \mathcal{X})$ are Borel spaces, satisfying

$$g(t, y, \mathcal{X}) = 0 \quad \text{for} \quad t > T. \quad (7.1)$$

Next, let $\partial$ be an element external to $\mathcal{X}$, and define kernel $g^\partial$ from $(\mathbb{R}_+ \times \mathcal{Y}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{Y})$ to $(\mathcal{X}^\partial, \mathcal{X}^\partial)$ as

$$g^\partial(t, y, dx) = \lambda(t, y)\gamma(t, y, dx),$$

where

$$\lambda(t, y) = g(t, y, \mathcal{X}), \quad \gamma(t, y, dx) = \frac{g(t, y, dx)}{g(t, y, \mathcal{X})} \mathbb{1}_{g(t, y, \mathcal{X}) > 0} + \delta_\partial(dx) \mathbb{1}_{g(t, y, \mathcal{X}) = 0}.$$  

Suppose that

$$\sup_{t \in [\ell(y), T]} \lambda(t, y) \leq \hat{\lambda}(y) < \infty, \quad \gamma(t, y, A) = \int_{(0, 1]} \mathbb{1}_A(\Gamma(t, y, u))du, \quad A \in \mathcal{X},$$

for some measurable mappings $\ell : \mathcal{Y} \to [0, T) \cup \{\infty\}$, $\hat{\lambda} : \mathcal{Y} \to (0, \infty)$ and $\Gamma : \mathbb{R}_+ \times \mathcal{Y} \times (0, 1] \to \mathcal{X}$. Existence of such mapping $\Gamma$ is asserted by Lemma 3.22 in [20]. In addition, let $D : [0, \infty) \times (0, 1] \to \mathbb{N}$ be as in Step 1 of our construction done in Section 4.1.
Next, take $Y$ to be a $(\mathcal{Y}, \mathcal{Y})$-valued random element, which is $\mathcal{G}$-measurable, and let $Z$ and $(U_m, V_m, W_m)_{m=1}^{\infty}$ be independent random variables uniformly distributed on $(0, 1]$ and independent of $\mathcal{G}$. We now define a random measure $N$ on $(\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X})$ as

\begin{equation}
N(dt, dx) = \sum_{m=1}^{\infty} \delta_{(T_m, X_m)}(dt, dx) \mathbb{1}_{\{\ell(Y) < T, t \leq P, A_m \leq \lambda(T_m, Y)\}} \tag{7.2}
\end{equation}

where $P, (T_m, A_m, X_m)_{m=1}^{\infty}$ are random variables defined by transformation of the sequence $Z, (U_m, V_m, W_m)_{m=1}^{\infty}$ and the random element $Y$ in the following way:

\begin{align*}
P &= D((T - \ell(Y))\hat{\lambda}(Y) \mathbb{1}_{\{\ell(Y) < T\}}, Z) \\
&= D((T - \ell(Y))\hat{\lambda}(Y), Z) \mathbb{1}_{\{\ell(Y) < T\}}, \\
T_m &= (\ell(Y) + (T - \ell(Y))U_m) \mathbb{1}_{\{\ell(Y) < T\}} + \infty \mathbb{1}_{\{\ell(Y) \geq T\}}, \\
A_m &= \hat{\lambda}(Y)V_m \mathbb{1}_{\{\ell(Y) < T\}}, \\
X_m &= \begin{cases} 
\Gamma(T_m, Y, W_m), & \text{if } \ell(Y) < T, \\
\partial, & \text{if } \ell(Y) \geq T.
\end{cases}
\end{align*}

Using the above set-up we see that, for each $m = 1, 2, \ldots$,

\begin{align*}
\mathbb{P}((T_m, A_m, X_m) \in dt \times da \times dx | \mathcal{G}) &= \mathbb{1}_{\{\ell(Y) < T\}} \frac{1}{(T - \ell(Y))\hat{\lambda}(Y)} \mathbb{1}_{(\ell(Y), T) \times (0, \hat{\lambda}(Y))} (t, a) \gamma(t, Y, dx) dt da \\
&\quad + \mathbb{1}_{\{\ell(Y) \geq T\}} \delta_{(\infty, 0, \partial)}(dt, da, dx), \tag{7.4}
\end{align*}

where $\delta_{(\infty, 0, \partial)}$ is a Dirac measure.

Note that even though the random elements $X_m, m = 1, 2, \ldots$, may take value $\partial$, the measure $N$ given in (7.2) is a random measure on $(\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X})$ having support belonging to $\mathcal{B}([0, T]) \otimes \mathcal{X}$.

Given the above, we now have the following result.

**Lemma 7.2.** The random measure $N$ defined by (7.2) is a $\mathcal{G}$-conditionally Poisson random measure with intensity measure $\nu$ given by

\begin{equation}
\nu(C) = \int_C g(v, Y, dx) \mathbb{1}_{(\ell(Y), \infty)}(v) dv, \quad C \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X}. \tag{7.5}
\end{equation}

**Proof.** To prove the result we consider $N((s, t] \times B)$ for fixed $0 \leq s \leq t, B \in \mathcal{X}$. We have

\begin{align*}
N((s, t] \times B) &= \sum_{m=1}^{\infty} \delta_{(T_m, X_m)}((s, t] \times B) \mathbb{1}_{\{\ell(Y) < T, t \leq P, A_m \leq \lambda(T_m, Y)\}} \\
&= \sum_{m=1}^{P} \mathbb{1}_{\{\ell(Y) < T, s \leq T_m \leq t, X_m \in B, A_m \leq \lambda(T_m, Y)\}}.
\end{align*}
First we will prove that, conditionally on $G$, the random variable $N((s,t] \times B)$ has the Poisson distribution with mean $\nu((s,t] \times B)$. Towards this end we observe that $P$ has, conditionally on $G$, the Poisson distribution with mean $(T - \ell(Y))\hat{\lambda}(Y)\mathbb{1}_{\{\ell(Y) < T\}}$ (see (7.3)), so

$$\mathbb{P}(P = k|G) = e^{-\ell(Y)T}\hat{\lambda}(Y)\mathbb{1}_{\{\ell(Y) < T\}} \frac{((T - \ell(Y))\hat{\lambda}(Y)\mathbb{1}_{\{\ell(Y) < T\}})^k}{k!}, \quad k = 0, 1, \ldots .$$

Moreover, we conclude from (7.4) that for $m = 1, 2, \ldots ,$

$$\mathbb{P}(\ell(Y) < T, s < T_m \leq t, X_m \in B, A_m \leq \lambda(T_m, Y)|G)$$

$$= \mathbb{1}_{\{\ell(Y) < T\}} \int_s^t \left( \int_B \left( \int_0^{\lambda(u, Y)} \frac{1}{(T - \ell(Y))\hat{\lambda}(Y)\mathbb{1}_{\{\ell(Y) < T\}}} (u, a) da \right) \gamma(u, Y, dx) \right) du$$

$$= \mathbb{1}_{\{\ell(Y) < T\}} \int_s^t \frac{1}{(T - \ell(Y))\hat{\lambda}(Y)} \int_s^t \mathbb{1}_{\{\ell(Y) < T\}}(u)\lambda(u, Y)\gamma(u, Y, B) du$$

$$= \mathbb{1}_{\{\ell(Y) < T\}} \int_s^t \frac{1}{(T - \ell(Y))\hat{\lambda}(Y)} \int_s^t \mathbb{1}_{\{\ell(Y) < T\}}(u)g(u, Y, B) du =: p(Y), \quad (7.6)$$

where the last equality follows from (7.3). Note that for $u \in \mathbb{R}$ and $m = 1, 2, \ldots ,$ we have $^7$

$$\mathbb{E}(e^{iu\mathbb{1}_{\{\ell(Y) < T\}}(s < T_m \leq t, X_m \in B, A_m \leq \lambda(T_m, Y))}|G) = (1 - p(Y)) + p(Y)e^{iu}.$$

This and the $G$-conditional independence of $P$ and $(T_m, A_m, X_m)_{m=1}^\infty$ imply that

$$\mathbb{E}(e^{iuN((s,t] \times B)}|G) = e^{(e^{iu} - 1)p(Y)(T - \ell(Y))\hat{\lambda}(Y)\mathbb{1}_{\{\ell(Y) < T\}}} = e^{(e^{iu} - 1)\int_s^t \mathbb{1}_{\{\ell(Y) < T\}}(u)\gamma(u, Y, B) du}$$

$$= e^{(e^{iu} - 1)\nu((s,t] \times B)}.$$

Thus, the random variable $N((s,t] \times B)$ has the $G$-conditional Poisson distribution with mean equal to $\nu((s,t] \times B)$.

Using standard monotone class arguments we obtain that for arbitrary $C \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X}$ random variable $N(C)$ has, conditionally on $G$, the Poisson distribution with mean $\nu(C)$.

Next, we will show that for $0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \ldots \leq s_n < t_n$ and for sets $B_1, \ldots , B_n \in \mathcal{X}$ the random variables

$$N((s_1, t_1] \times B_1), \ldots , N((s_n, t_n] \times B_n) \quad (7.7)$$

are conditionally independent given $G$. Towards this end let us define

$$S_r((s,t] \times B) := \sum_{m=1}^r I_m((s,t] \times B),$$

for $r \in \mathbb{N}, 0 \leq s < t, B \in \mathcal{X}$, where

$$I_m((s,t] \times B) := \mathbb{1}_{\{s < T_m \leq t, X_m \in B, A_m \leq \lambda(T_m, Y)\}}.$$

$^7$In the ensuing two formulæ $i = \sqrt{-1}$. 

April 30, 2020

Generalized multivariate Hawkes processes 42 of 49
Note that the random variable \( N((s, t] \times B) \) can be represented as
\[
N((s, t] \times B) = S_P((s, t] \times B).
\]
Using this representation we obtain that
\[
J := \mathbb{P}(N((s_1, t_1] \times B_1) = l_1, \ldots, N((s_n, t_n] \times B_n) = l_n | \mathcal{G})
= \sum_{r=0}^{\infty} \mathbb{P}
\left( \bigcap_{j=1}^{n} S_r((s_j, t_j] \times B_j) = l_j, P = r | \mathcal{G} \right)
= \sum_{r=l}^{\infty} \mathbb{P}
\left( \bigcap_{j=1}^{n} S_r((s_j, t_j] \times B_j) = l_j, S_r\left( \mathbb{R}_+ \times \mathcal{X} \setminus \bigcup_{j=1}^{n}(s_j, t_j] \times B_j \right) = r - l | \mathcal{G} \right) \mathbb{P}(P = r | \mathcal{G}),
\]
where \( l = \sum_{j=1}^{n} l_j \). Now, from (7.6), we see that the random vector
\[
\left( S_r((s_1, t_1] \times B_1), \ldots, S_r((s_n, t_n] \times B_n), S_r\left( \mathbb{R}_+ \times \mathcal{X} \setminus \bigcup_{j=1}^{n}(s_j, t_j] \times B_j \right) \right)
\]
has, conditionally on \( \mathcal{G} \), the multinomial distribution with parameters \( p_1, \ldots, p_{n+1} \) given by:
\[
p_j = p_j(Y) := \mathbb{P}(\ell(Y) < T, s_j < T_1 \leq t_j, X_1 \in B_j, A_1 \leq \lambda(T_1, Y) | \mathcal{G}),
\]
for \( j = 1, \ldots, n \), and
\[
p_{n+1} = 1 - p_1 - \ldots - p_n.
\]
Hence, using the fact that \( l = \sum_{j=1}^{n} l_j \), we deduce that
\[
J = \sum_{r=l}^{\infty} \frac{r!}{l_1! \ldots l_n!(r-l)!} p_1^{l_1} \ldots p_n^{l_n} p_{n+1}^{r-l} \mathbb{P}(P = r | \mathcal{G})
= \frac{1}{l_1! \ldots l_n!} p_1^{l_1} \ldots p_n^{l_n} \sum_{r=0}^{\infty} \frac{(r+l)!}{r!} p_{n+1}^{r-l} \mathbb{P}(P = r + l | \mathcal{G})
= \frac{1}{l_1! \ldots l_n!} p_1^{l_1} \ldots p_n^{l_n} \sum_{r=0}^{\infty} \frac{(r+l)!}{r!} p_{n+1}^{r-l} e^{-\lambda(T_1, Y)} \mathcal{A}(Y) 1_{\{\ell(Y) < T\}} \left( (T - \ell(Y)) (T - \lambda(Y)) 1_{\{\ell(Y) < T\}} \right)^{r+l}
= \prod_{j=1}^{n} \frac{p_j(T - \ell(Y)) \mathcal{A}(Y) 1_{\{\ell(Y) < T\}}^{l_j}}{l_j!} e^{(p_{n+1} - 1) (T - \ell(Y)) (T - \lambda(Y)) 1_{\{\ell(Y) < T\}}}
= \prod_{j=1}^{n} \frac{p_j(T - \ell(Y)) \mathcal{A}(Y) 1_{\{\ell(Y) < T\}}^{l_j}}{l_j!} e^{-p_j(T - \ell(Y)) \mathcal{A}(Y) 1_{\{\ell(Y) < T\}}}
= \prod_{j=1}^{n} \mathbb{P}(N((s_j, t_j] \times B_j) = l_j | \mathcal{G}),
\]
where the last equality follows from the fact that \( N((s_j, t_j] \times B_j) \) has the \( \mathcal{G} \)-conditional Poisson distribution with mean equal to \( \nu((s_j, t_j] \times B_j) = p_j(T - \ell(Y)) \mathcal{A}(Y) 1_{\{\ell(Y) < T\}} \), which
is a consequence of (7.5), (7.6) and (7.8). Using standard use the monotone class arguments we conclude from (7.7) that for arbitrary disjoint sets \( C_1, \ldots, C_n \in \mathcal{B}(R_+ \otimes \mathcal{X}) \) that random variables \( N(C_1), \ldots, N(C_n) \) are \( \mathcal{G} \)-conditionally independent. The proof is now complete.

\[ \square \]

### 7.2 Relation between conditional Poisson random measures and doubly stochastic marked Poisson processes

We begin by recalling (cf. Chapter 6 in [22]) the concept of a doubly stochastic marked Poisson process. For this, we consider a filtration \( \mathbb{F} \) on \( (\Omega, \mathcal{F}, \mathbb{P}) \). A marked point process \( N \) on \( (\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X}) \) is an \( \mathbb{F} \)-doubly stochastic marked Poisson process if there exist an \( \mathcal{F}_0 \)-measurable random measure \( \nu \) on \( (\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X}) \) such that

\[
\mathbb{P}(N((s, t] \times B) = k | \mathcal{F}_s) = e^{\nu((s, t] \times B)} \frac{(\nu((s, t] \times B))^k}{k!}, \quad 0 \leq s < t, \ B \in \mathcal{X}. \tag{7.9}
\]

Thus, for \( 0 \leq s < t, \ B \in \mathcal{X} \) we have

\[
\nu((s, t] \times B) = \mathbb{E}(N((s, t] \times B) | \mathcal{F}_0). \tag{7.10}
\]

Hence, by analogy with the concept of the intensity of a Poisson random measure, the measure \( \nu \) is called the \( \mathcal{F}_0 \)-intensity kernel of \( N \) (see Chapter 6 in [22]).

Let now \( \tilde{N} \) be marked point process on \( (\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X}) \), such that its \( \mathbb{F} \)-compensator \( \tilde{\nu} \) is the \( \mathcal{F}_0 \)-intensity kernel in a sense that the property analogous to (7.10) holds,

\[
\tilde{\nu}((s, t] \times B) = \mathbb{E}(\tilde{N}((s, t] \times B) | \mathcal{F}_0), \quad 0 \leq s < t, \ B \in \mathcal{X}.
\]

Then, one can show (see Theorem 6.1.4 in [22]) that \( \tilde{N} \) is an \( \mathbb{F} \)-doubly stochastic marked Poisson process, i.e. the analog of (7.9) holds with \( \tilde{N} \) and \( \tilde{\nu} \). The opposite statement is true as well (see Theorem 6.1.4 in [22]): if \( \tilde{N} \) is an \( \mathbb{F} \)-doubly stochastic marked Poisson process, then the \( \mathbb{F} \)-compensator \( \tilde{\nu} \) of \( \tilde{N} \) is an \( \mathcal{F}_0 \)-intensity kernel of \( \tilde{N} \).

Conditional Poisson random measures on \( (\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X}) \) are closely related to \( \mathbb{F} \)-doubly stochastic marked Poisson processes. It can be shown that if \( N \) is an \( \mathbb{F} \)-doubly stochastic marked Poisson process with intensity kernel \( \nu \), then \( N \) considered as a random measure is an \( \mathcal{F}_0 \)-conditionally Poisson random measure with intensity kernel \( \nu \).

This implies that for sets \( B_1, \ldots, B_n \in \mathcal{X} \) and for \( 0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \ldots \leq s_n < t_n \leq t, \ n \in \mathbb{N} \), we have

\[
\mathbb{P}\left( \bigcap_{i=1}^{n} \{N((s_i, t_i] \times B_i) = l_i\} | \mathcal{F}_0 \right) = \prod_{i=1}^{n} e^{\nu((s_i, t_i] \times B_i)} \frac{(\nu((s_i, t_i] \times B_i))^{l_i}}{l_i!} \tag{7.11}
\]

\[
= \prod_{i=1}^{n} \mathbb{P}\{N((s_i, t_i] \times B_i) = l_i\} | \mathcal{F}_0\).
\]

The next result, in a sense, complements our discussion of conditional Poisson random measures and doubly stochastic marked Poisson processes.
Proposition 7.3. i) Let $M$ be a marked point process on $(\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X})$, which is a $\mathcal{G}$-conditional Poisson random measure with intensity measure $\nu$, and let $\hat{\mathcal{F}}^M_t$ be a filtration defined by family of $\sigma$-fields
\[
\hat{\mathcal{F}}^M_t = \mathcal{G} \vee \mathcal{F}^M_t, \quad t \geq 0.
\]
Then $M$ is an $\hat{\mathcal{F}}^M$-doubly stochastic marked Poisson process with $\hat{\mathcal{F}}^M$-intensity kernel $\nu$ being also $\hat{\mathcal{F}}^M$-compensator of $M$.

ii) Let $N = (N_j)_{j \geq 1}$ be a family of marked point processes on $(\mathbb{R}_+ \times \mathcal{X}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{X})$, which are $\mathcal{G}$-conditional Poisson random measures (each $N_j$ with intensity measure $\nu_j$), and let $\hat{\mathcal{F}}^N_t$ be a filtration defined by the family of $\sigma$-fields
\[
\hat{\mathcal{F}}^N_t = \mathcal{G} \vee \bigvee_{k \geq 1} \mathcal{F}^{N_k}_t, \quad t \geq 0.
\]
Suppose that $(N_j)_{j \geq 1}$ are $\mathcal{G}$-conditionally independent. Then each $N_j$ is an $\hat{\mathcal{F}}^N$-doubly stochastic marked Poisson process with $\hat{\mathcal{F}}^N$-intensity kernel $\nu_j$ being also $\hat{\mathcal{F}}^N$-compensator of $N_j$.

In the proof of Proposition 7.3 we will use the following elementary result, whose derivation is omitted:

Lemma 7.4. Let $\mathcal{G}$ be a sigma field and let $A \in \mathcal{G}$. Then for arbitrary measurable sets $B$ and $C$ which are conditionally independent given $\mathcal{G}$ we have
\[
\mathbb{E}(1_B 1_{A \cap C}) = \mathbb{E}(\mathbb{P}(B|\mathcal{G}) 1_{A \cap C}).
\]

Proof. (of Proposition 7.3) We will prove ii), the proof of i) is similar in spirit to the proof of ii) and in fact a bit simpler. Fix arbitrary $j \geq 1$. By assumption $N_j$ is a $\mathcal{G}$-conditional Poisson random measure, so we have for fixed $0 \leq s < t$ and $D \in \mathcal{X}$
\[
\mathbb{P}(N_j((s,t] \times D) = k|\mathcal{G}) = e^{-\nu_j((s,t] \times D)} \frac{(\nu_j((s,t] \times D))^k}{k!},
\]
(7.12)
where $\nu_j$ is $\mathcal{G} = \hat{\mathcal{F}}^N_0$-measurable random measure. In view of the definition of $\hat{\mathcal{F}}^N$-doubly stochastic marked Poisson process, of the above formula and of Proposition 6.1.4 in [22] it suffices to show that for arbitrary set $F \in \hat{\mathcal{F}}^N_0$ it holds
\[
\mathbb{E}(1_{\{N_j((s,t] \times D) = k\}} 1_F) = \mathbb{E}(\mathbb{P}(N_j((s,t] \times D) = k|\mathcal{G}) 1_F).
\]
(7.13)
Indeed (7.13) and (7.12) imply
\[
\mathbb{P}(N_j((s,t] \times D) = k|\hat{\mathcal{F}}^N_s) = e^{-\nu_j((s,t] \times D)} \frac{(\nu_j((s,t] \times D))^k}{k!}.
\]
for $\hat{\mathcal{F}}^N_0$-measurable random measure $\nu_j$. So that $N_j$ is a $\hat{\mathcal{F}}^N$-doubly stochastic marked Poisson process with $\hat{\mathcal{F}}^N$-intensity kernel $\nu_j$. Then Proposition 6.1.4 in [22] implies that $\nu_j$ is $\hat{\mathcal{F}}^N$-compensator of $N_j$. 
To prove (7.13) we will use the Monotone Class Theorem. First note that sets \( F \) for which (7.13) holds constitute \( \lambda \)-system. Thus it suffices to show the above equality for a \( \pi \)-system of sets which generates \( \hat{\mathcal{F}}_s^N \). Towards this end consider family of sets:

\[
\mathcal{A}_s := \left\{ A \cap C : A \in \mathcal{G}, C = \cap_{i=1}^n \cap_{i=1}^{p_r} \{ N_{m_r}((s^r_i, t^r_i] \times D^r_i) = k^r_i \}, \quad 0 \leq s^r_1 < t^r_1 \leq \ldots \leq s^r_{p_r} < t^r_{p_r} \leq s, \quad D^r_1, \ldots, D^r_{p_r} \in \mathcal{X}, \quad k^r_1, \ldots, k^r_{p_r} \in \mathbb{N}, \quad 0 \leq p_1 \leq \ldots \leq p_r, \quad 0 \leq m_1 \leq \ldots \leq m_r \right\},
\]

Clearly, \( \mathcal{A}_s \) is a \( \pi \)-system and \( \sigma(\mathcal{A}_s) = \hat{\mathcal{F}}_s^N \). Let us take \( F \in \mathcal{A}_s \), so \( F = A \cap C \), and let \( (s, t) \times D \) be disjoint with sets \((s^r_i, t^r_i] \times D^r_i\) which define \( C \). This and \( \mathcal{G}\)-conditional independence of \( \{N_j\}_{j \geq 1} \) imply that events \( \{N_j((s, t) \times D) = k\} \) and \( C \) are conditionally independent given \( \mathcal{G} \). Hence, by applying Lemma 7.4, we obtain that (7.13) holds for \( F \in \mathcal{A}_s \). Then, invoking the Monotone Class Theorem, we conclude that (7.13) holds for sets \( F \in \hat{\mathcal{F}}_s^N \). The proof is complete.

### 7.3 Additional Technical Result

**Lemma 7.5.** Let \((\mu^k)_{k=1}^{\infty}\) be a sequence of measures. Let \( \mu \) be a mapping \( \mu : \mathcal{X} \to [0, \infty] \) defined by

\[
\mu(A) = \lim_{n \to \infty} \sum_{k=1}^{n} \mu_k(A).
\]

Then \( \mu \) is a measure. Moreover for any measurable non negative function \( F : \mathcal{X} \to \mathbb{R}_+ \) we have

\[
\int_{\mathcal{X}} F d\mu = \lim_{n \to \infty} \sum_{k=1}^{n} \int_{\mathcal{X}} F d\mu_k
\]

**Proof.** The first part follows from the Nikodym convergence theorem (see e.g. Theorem 7.48 in Swartz [31]).

To prove the second assertion it suffices to consider simple step functions only, i.e. functions \( F \) of the form

\[
F(x) := \sum_{i=1}^{n} a_i 1_{A_i}(x), \quad a_i \in \mathbb{R}_+, A_i \in \mathcal{X}.
\]

For such \( F \) it holds

\[
\int_{\mathcal{X}} F d\mu = \sum_{i=1}^{n} a_i \mu(A_i) = \sum_{i=1}^{n} a_i \sum_{k=1}^{\infty} \mu_k(A_i) = \sum_{k=1}^{\infty} \sum_{i=1}^{n} a_i \mu_k(A_i) = \sum_{k=1}^{\infty} \int_{\mathcal{X}} F d\mu_k.
\]

Using usual approximation technique and the monotone convergence theorem we finish the proof. \( \square \)
References

[1] E. Bacry, S. Delattre, C. Hoffmann, and J.F. Muzy. Some limit theorems for Hawkes processes and application to financial statistics. Stochastic Processes and their Applications, 123(7):2475–2499, 2013.

[2] E. Bacry, S. Delattre, M. Hoffmann, and J. F. Muzy. Modelling microstructure noise with mutually exciting point processes. Quant. Finance, 13(1):65–77, 2013.

[3] E. Bacry and J.F. Muzy. Hawkes model for price and trades high-frequency dynamics. Quant. Finance, 14(7):1147–1166, 2014.

[4] T.R. Bielecki, J. Jakubowski, M. Jeanblanc, and M. Niewegowski. Semimartingales and shrinkage of filtration. Submitted, 2019. https://arxiv.org/pdf/1803.03700.pdf.

[5] T.R. Bielecki, J. Jakubowski, and M. Niewegowski. Structured Dependence between Stochastic Processes. Cambridge University Press, forthcoming, 2020.

[6] P. Brémaud and L. Massoulié. Stability of nonlinear Hawkes processes. Ann. Probab., 24(3):1563–1588, 1996.

[7] E. Çınlar. Probability and stochastics, volume 261 of Graduate Texts in Mathematics. Springer, New York, 2011.

[8] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. Ann. Appl. Probab., 13(3):984–1053, 2003.

[9] P. Embrechts, T. Liniger, and L. Lin. Multivariate Hawkes processes: an application to financial data. J. Appl. Probab., 48A(New frontiers in applied probability: a Festschrift for Søren Asmussen):367–378, 2011.

[10] S.N. Ethier and T.G. Kurtz. Markov processes: Characterization and convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.

[11] A.G. Hawkes. Point Spectra of Some Mutually Exciting Point Processes. Journal of the Royal Statistical Society. Series B (Methodological), 33(3):438–443, 1971.

[12] A.G. Hawkes. Spectra of Some Self-Exciting and Mutually Exciting Point Processes. Biometrika, 58(1):83–90, 1971.

[13] A.G. Hawkes. Hawkes processes and their applications to finance: a review. Quantitative Finance, 18(2):193–198, dec 2017.

[14] A.G. Hawkes and D. Oakes. A cluster process representation of a self-exciting process. J. Appl. Probab., 11:493–503, 1974.

[15] S.-W. He and J.-G. Wang. Two results on jump processes. In Séminaire de Probabilités XVIII 1982/83, pages 256–267. Springer, 1984.
[16] S.-W. He, J.-G. Wang, and Ji.-A. Yan. Semimartingale Theory and Stochastic Calculus. Kexue Chubanshe (Science Press), Beijing, 1992.

[17] J. Jacod. Multivariate point processes: predictable projection, Radon-Nikodým derivatives, representation of martingales. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 31:235–253, 1974/75.

[18] J. Jacod and A.N. Shiryaev. Limit theorems for stochastic processes, volume 288 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, second edition, 2003.

[19] M. Jeanblanc, M. Yor, and M. Chesney. Mathematical methods for financial markets. Springer Finance. Springer-Verlag London, Ltd., London, 2009.

[20] O. Kallenberg. Foundations of modern probability. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[21] J.D. Kelly, J. Park, R.J. Harrigan, and et al. Real-time predictions of the 2018 – 2019 Ebola virus disease outbreak in the Democratic Republic of the Congo using Hawkes point process models. Epidemics, 28(100354), 2019.

[22] G. Last and A. Brandt. Marked point processes on the real line: The dynamic approach. Probability and its Applications (New York). Springer-Verlag, New York, 1995.

[23] P.J. Laub, T. Taimre, and P.K. Pollett. Hawkes processes. https://arxiv.org/abs/1507.02822.

[24] T.J. Liniger. Multivariate hawkes processes. PhD thesis, ETH Zurich, 2009.

[25] D. Oakes. The Markovian self-exciting process. J. Appl. Probability, 12:69–77, 1975.

[26] Y. Ogata. Space-time Point-process Models for Earthquake Occurrences. Ann. Inst. Math. Statist., 50:379–402, 1998.

[27] Y. Ogata. Seismicity Analysis through Point-process Modeling: A Review. Pure appl. geophys., 155:471–507, 1999.

[28] M.A. Rizoiu, S. Mishra, Q. Kong, M. Carman, and L. Xie. SIR-Hawkes: Linking epidemic models and Hawkes processes to model diffusions in finite populations. In WWW ’18: Proceedings of the 2018 World Wide Web Conference, pages 419–428, 2018.

[29] K.-i. Sato. Lévy processes and infinitely divisible distributions, volume 68 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2013. Translated from the 1990 Japanese original, Revised edition of the 1999 English translation.

[30] F.P. Schoenberg, M. Hoffmann, and R.J. Harrigan. A recursive point process model for infectious diseases. Annals of the Institute of Statistical Mathematics, 71:1271–1287, 2019.
[31] C. Swartz. Multiplier convergent series. World Scientific, 2009.

[32] A. Vacarescu. Filtering and parameter estimation for partially observed generalized Hawkes processes. PhD thesis, Stanford University, 2011.

[33] L. Zhu. Nonlinear Hawkes Processes. PhD thesis, New York University, May 2013.