HYBRID MARKED POINT PROCESSES:
CHARACTERISATION, EXISTENCE AND UNIQUENESS

Maxime Morariu-Patrichi[*]   Mikko S. Pakkanen[†]

26 August 2019

Abstract

We introduce a class of hybrid marked point processes, which encompasses and extends continuous-time Markov chains and Hawkes processes. While this flexible class amalgamates such existing processes, it also contains novel processes with complex dynamics. These processes are defined implicitly via their intensity and are endowed with a state process that interacts with past-dependent events. The key example we entertain is an extension of a Hawkes process, a state-dependent Hawkes process interacting with its state process. We show the existence and uniqueness of hybrid marked point processes under general assumptions, extending the results of Massoulié (1998) on interacting point processes.

Keywords: marked point processes; Hawkes processes; stochastic intensity; transition probabilities; Poisson embedding; strong existence; strong and weak uniqueness.

2010 Mathematics Subject Classification: 60G55, 60H20, 60K35, 91G99.

1 Introduction

Let $N = (N_1, \ldots, N_d)$ be a d-dimensional counting process, meaning that $N_i(t)$, $i = 1, \ldots, d$, is the number of events of type i that have occurred until time t (Brémaud, 1981; Daley and Vere-Jones, 2003; Last and Brandt, 1995; Sigman, 1995; Jacobsen, 2006). A fundamental concept describing the dynamics of N is the intensity process $\lambda = (\lambda_1, \ldots, \lambda_d)$. Loosely speaking, when it exists, the intensity $\lambda_i(t)$ of N_i at time t is such that

$$\mathbb{E}\left[N_i(t + dt) - N_i(t) \,|\, \mathcal{F}_t^N \right] \approx \lambda_i(t)dt,$$

where $\mathbb{F}^N = (\mathcal{F}_t^N)_{t \geq 0}$ is the natural filtration generated by N. Intuitively, the above equation says that $\lambda_i(t)dt$ is the expected number of events of type i in the infinitesimal time window $(t, t + dt]$, given what has happened so far.

Besides describing counting processes, intensity processes can actually be used to specify them implicitly. A prime example of that is the class of linear Hawkes processes (Hawkes, 1971; Hawkes and Oakes, 1974),

[*]Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK. E-mail: m.morariu-patrichi14@imperial.ac.uk URL: http://www.maximemorariu.com

[†]Corresponding author. Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK and CREATES, Aarhus University, Aarhus, Denmark. E-mail: m.pakkanen@imperial.ac.uk URL: http://www.mikkopakkanen.fi

which are characterised by the intensities

$$\lambda_i(t) = \nu_i + \sum_{j=1}^{d} \int_{[0,t)} k_{ji}(t-s)dN_j(s), \quad t \geq 0, \; i = 1, \ldots, d, \tag{1.1}$$

where each $\nu_i \in \mathbb{R}_{>0}$ is fixed and each $k_{ji} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a non-negative function, usually called a *kernel*, see Laub et al. (2015) for an introduction. This class of processes allows for self-excitation and cross-excitation effects: the arrival at time t of an event of type j increases the intensity $\lambda_i(t+h)$ of events of type i at time $t+h$ by an amount of $k_{ji}(h)$. A Hawkes process is thus a good candidate for a model of interactions between different types of events. Gaining popularity in the last decade, Hawkes processes have indeed found applications in many areas including earthquake modelling (Ogata, 1998; Türkyilmaz et al., 2013; Fox et al., 2016), criminology (Lewis et al., 2012; Mohler, 2013; Loeffler and Flaxman, 2018), social networks analysis (Blundell et al., 2012; Zhou et al., 2013; Farajtabar et al., 2017), neurology (Chornoboy et al., 1988; Chevallier et al., 2015; Gerhard et al., 2017) and finance (Bowsher, 2007; Bacry et al., 2015; Jaisson and Rosenbaum, 2016).

In spite of their success and attractiveness, we notice that these Hawkes-process models account only for the dynamics of events and ignore the state of the underlying system they may influence. For instance, when applied to financial markets, Hawkes processes describe the arrival in time of buy and sell orders (Large, 2007; Bacry et al., 2016; Rambaldi et al., 2017) but capture neither the asset price nor the supply and demand imbalance, which are impacted by the arriving orders. In fact, Hawkes processes can be contrasted with models based on continuous-time Markov chains where the focus is instead on a state process that represents the underlying system (Cont et al., 2010; Cont and de Larrard, 2013; Huang et al., 2015; Huang and Rosenbaum, 2017). However, because of the Markov property inherent to this second group of models, the arrival rates depend only on the current state and, thus, interactions like in general Hawkes processes are not possible. In effect, Hawkes processes and continuous-time Markov chains can have either an *event* viewpoint or a *state* viewpoint, respectively. Note that the dichotomy between Hawkes processes and Markovian models, and the need for something bridging the two, was already suggested by Bacry et al. (2016, p. 1190–1191). In a recent paper, Gonzalez and Schervish (2017) make a similar observation and propose a discrete-time Markov chain to model both the dependence on the most recent event and the current state of the system.

The original idea that started the present work is to address this gap by endowing the counting process N with a state process X that is fully coupled to N in the following manner. On the one hand, we wish to make the intensity dependent on both past events and states by changing (1.1) to

$$\lambda_i(t) = \nu_i + \sum_{j=1}^{d} \int_{[0,t)} k_{ji}(t-s, X_s)dN_j(s), \quad t \geq 0, \; i = 1, \ldots, d,$$

where the kernels now depend on the state process X. On the other hand, we imagine that each event in N prompts a state change according to transition probabilities that depend on the type of the event. Naturally, we call this new process (N, X) a *state-dependent* Hawkes process. The practical relevance and strong potential of this new model is demonstrated in Morariu-Patrichi and Pakkanen (2018), where parametric estimation of the kernels from high-frequency financial data indeed reveals significant state dependence. In fact, this extension of Hawkes processes opens an avenue to novel models of high-frequency data that feature both excitation effects and a feedback loop between events and the state of the underlying system.

The first contribution of this paper is to turn this idea into a class of *hybrid marked point processes* that encompasses and extends continuous-time Markov chains and Hawkes processes. To generalise state-dependent Hawkes processes to hybrid marked point processes, we view the process (N, X) as a single marked point process on a product mark space and allow the intensity of events to be any measurable functional of past events and states. These new hybrid marked point processes are actually defined implicitly via their intensity that takes a specific product form. We prove that the dynamics generated by this product form completely characterise the class of hybrid marked point processes (Theorem 2.13). Offering an *event–state* viewpoint, this new class is well-suited to the joint modelling of events and the time evolution of the state of a system. While this general and flexible class provides a unifying framework for various existing processes, it also contains new processes with complex dynamics, as illustrated by state-dependent Hawkes processes.

The second contribution of this paper is to prove the strong existence and uniqueness of non-explosive state-dependent Hawkes processes and, more generally, hybrid marked point processes. In fact, by dispensing with a Lipschitz condition, we extend the results currently available in the literature for marked point processes defined via their intensity. It is known that a marked point process whose intensity λ is expressed in terms of an intensity functional ψ can be formulated as a solution to a Poisson-driven stochastic differential equation (SDE) (Massoulié, 1998; Brémaud and Massoulié, 1996). However, the existence and uniqueness results available in these works cannot be applied to hybrid marked point processes because their intensity functional may fail to satisfy the Lipschitz condition imposed therein. We show that, under certain integrability or decay conditions, it is enough for ψ to be dominated by either a Hawkes functional or an increasing function of the total number of past events in order to obtain the existence of a strong solution to the Poisson-driven SDE (Theorem 2.17) and, in particular, the existence of hybrid marked point processes (Corollary 2.18). The solution is constructed piece by piece along the time axis in a pathwise manner, taking advantage of the discrete nature of the driving Poisson random measure. A domination argument is then used to show non-explosiveness. In the context of multivariate point processes, a similar construction has already been considered in Çınlar (2011), while a similar domination argument is given in Chevallier (2017). We combine the two in a more general setting (i.e., general mark space, initial conditions and intensity functional). We are also able to obtain strong and weak uniqueness without any specific assumptions (Theorems 2.20 and 2.21) and, in particular, uniqueness of hybrid marked point processes (Corollary 2.22).

The paper is organised as follows. Section 2 introduces the framework that we use and presents the main results. Section 3 proves the results concerning the dynamics of hybrid marked point processes. Section 4 proves the existence and uniqueness results. The Appendix gathers technical results concerning the space $\mathcal{N}^{\#}_{\mathbb{R} \times \mathscr{M}}$ of boundedly finite integer-valued measures and the enumeration representation of marked point processes.

2 Framework and main results

2.1 A framework for point processes

In the following, $\mathcal{U}$ refers to a complete separable metric space and we denote by $\mathcal{B}(\mathcal{U})$ its Borel σ-algebra. We reserve the notation $\mathscr{M}$ for a complete separable metric space that represents the set of marks in the context of marked point processes. For most definitions, we follow closely Daley and Vere-Jones (2008) along with Brémaud (1981). The former reference will be especially used to introduce (marked) point processes while the latter is essential when defining the intensity process.

2.1.1 Spaces of integer-valued measures. Let ξ be a Borel measure on $\mathcal{U}$. We say that ξ is boundedly finite if $\xi(A) < \infty$ for every bounded Borel set $A \in \mathcal{B}(\mathcal{U})$. We denote by $\mathcal{N}_{\mathcal{U}}^\infty$ the space of Borel measures on $\mathcal{U}$ with values in $\mathbb{N} \cup \{\infty\}$. We denote by $\mathcal{N}_{\mathcal{U}}^\#$ the set of all $\xi \in \mathcal{N}_{\mathcal{U}}^\infty$ such that ξ is boundedly finite. We denote by $\mathcal{N}_{\mathbb{R}\times\mathscr{M}}^{\#g}$ the set of all $\xi \in \mathcal{N}_{\mathbb{R}\times\mathscr{M}}^\#$ such that their ground measure $\xi_g(\cdot) := \xi(\cdot \times \mathscr{M})$ satisfies:

(i) $\xi_g \in \mathcal{N}_{\mathbb{R}}^\#$;

(ii) $\xi_g(\{t\}) = 0$ or 1 for all $t \in \mathbb{R}$ (we say that the ground measure is simple).

Observe that $\mathcal{N}_{\mathbb{R}\times\mathscr{M}}^{\#g} \subset \mathcal{N}_{\mathbb{R}\times\mathscr{M}}^\# \subset \mathcal{N}_{\mathbb{R}\times\mathscr{M}}^\infty$. The space $\mathcal{N}_{\mathbb{R}\times\mathscr{M}}^\infty$ corresponds to the realisations of potentially explosive point processes, while the space $\mathcal{N}_{\mathbb{R}\times\mathscr{M}}^\#$ corresponds to the realisations of non-explosive point processes and contains all the realisations of potentially explosive marked point processes. Regarding the space $\mathcal{N}_{\mathbb{R}\times\mathscr{M}}^{\#g}$, each $\xi \in \mathcal{N}_{\mathbb{R}\times\mathscr{M}}^{\#g}$ is a realisation of a non-explosive marked point process. When $\xi(\{(t,m)\}) = 1$ for some $t \in \mathbb{R}$ and $m \in \mathscr{M}$, this should be interpreted as an event happening at time t with characteristics m. The boundedly finite property of the ground measure ensures that, in any finite amount of time, only finitely many events can occur (i.e., the marked point point process is non-explosive). The simpleness constraint on the ground measure means that there cannot be two events at the same time.

The so-called $w^\#$-distance $d^\#$ ("weak-hash") introduced by Daley and Vere-Jones (2003, p. 403), makes $\mathcal{N}_{\mathcal{U}}^\#$ a complete separable metric space, see Theorem A2.6.III in Daley and Vere-Jones (2003, p. 404). The corresponding σ-algebra $\mathcal{B}(\mathcal{N}_{\mathcal{U}}^\#)$ coincides with the one generated by all mappings $\xi \mapsto \xi(A)$, $\xi \in \mathcal{N}_{\mathcal{U}}^\#$, $A \in \mathcal{B}(\mathcal{U})$. Proposition A2.6.II of Daley and Vere-Jones (2003, p. 403) characterises convergence in this topology, called the $w^\#$-topology. These properties of the space $\mathcal{N}_{\mathcal{U}}^\#$ play an important role in this work. Note that Morariu-Patrichi (2018) clarifies the proofs of Proposition A2.6.II and Theorem A2.6.III of Daley and Vere-Jones (2003). Indeed, the original proofs assume a certain function to be monotonic. As this does not seem to actually hold, Morariu-Patrichi (2018) proposes alternative arguments where required. Besides, in the Appendix, we show that $\mathcal{N}_{\mathbb{R}\times\mathscr{M}}^{\#g}$ is indeed a Borel set of $\mathcal{N}_{\mathbb{R}\times\mathscr{M}}^\#$, see Lemma A.1.1.

Finally, for any $u \in \mathcal{U}$, we denote by δ_u the Dirac measure at u.

2.1.2 Non-explosive marked point processes. In the following, the notation $(\Omega, \mathcal{F}, \mathbb{P})$ refers to a probability space. For any σ-algebra $\mathcal{S}$, the trace of $A \in \mathcal{S}$ on $\mathcal{S}$ is defined by $A \cap \mathcal{S} := \{A \cap S : S \in \mathcal{S}\}$.

Definition 2.1 (Non-explosive point process)**.** A *non-explosive point process* on $\mathcal{U}$ is a measurable mapping from $(\Omega, \mathcal{F})$ into $(\mathcal{N}_{\mathcal{U}}^\#, \mathcal{B}(\mathcal{N}_{\mathcal{U}}^\#))$.

Definition 2.2 (Non-explosive marked point process)**.** A *non-explosive marked point process N* on $\mathbb{R} \times \mathscr{M}$ is a non-explosive point process N on $\mathbb{R} \times \mathscr{M}$ such that $N(\omega) \in \mathcal{N}_{\mathbb{R}\times\mathscr{M}}^{\#g}$ for all $\omega \in \Omega$.

Remark 2.3. By applying Lemma 1.6 in Kallenberg (2002, p. 4), we obtain that $\mathcal{B}(\mathcal{N}_{\mathbb{R}\times\mathscr{M}}^{\#g}) = \mathcal{N}_{\mathbb{R}\times\mathscr{M}}^{\#g} \cap \mathcal{B}(\mathcal{N}_{\mathbb{R}\times\mathscr{M}}^\#)$, where $\mathcal{N}_{\mathbb{R}\times\mathscr{M}}^{\#g}$ is also equipped with the $w^\#$-metric $d^\#$. This implies that Definition 2.2 is equivalent to saying that a non-explosive marked point process is a measurable mapping from $(\Omega, \mathcal{F})$ into $(\mathcal{N}_{\mathbb{R}\times\mathscr{M}}^{\#g}, \mathcal{B}(\mathcal{N}_{\mathbb{R}\times\mathscr{M}}^{\#g}))$.

Next, a non-explosive point process induces a probability measure on $\mathcal{N}_{\mathcal{U}}^\#$.

Definition 2.4 (Induced probability)**.** Let N be a non-explosive point process on $\mathcal{U}$. We define the *induced probability measure $\mathcal{P}^N$* on the measurable space $(\mathcal{N}_{\mathcal{U}}^\#, \mathcal{B}(\mathcal{N}_{\mathcal{U}}^\#))$ through the relation

$$\mathcal{P}^N(A) := \mathbb{P}\left(N^{-1}(A)\right), \quad A \in \mathcal{B}(\mathcal{N}_{\mathcal{U}}^\#).$$

2.1.3 Enumeration representation. It is common to define instead marked point processes on $\mathbb{R}_{\geq 0} \times \mathcal{M}$ as a sequence $(T_n, M_n)_{n \in \mathbb{N}}$ of random variables in $(0, \infty] \times \mathcal{M}$ such that $(T_n)_{n \in \mathbb{N}}$ is non-decreasing and $T_n < \infty$ implies $T_n < T_{n+1}$ (Jacod, 1975; Brémaud, 1981). We will call such a sequence an enumeration. Here, T_n is to be interpreted as the time when the nth event occurs while M_n describes the characteristics of that event. Moreover, $T_n < \infty$ with $T_{n+1} = \infty$ means that there are no more events after time T_n. Note that this definition allows for explosion in the sense that $\lim_{n \to \infty} T_n < \infty$ is possible with positive probability. This is why we stress the non-explosive character of marked point processes in Definition 2.2. There is a one-to-one correspondence between non-explosive marked point processes on $\mathbb{R}_{\geq 0} \times \mathcal{M}$ and enumerations such that $\lim_{n \to \infty} T_n = \infty$ a.s. We give a proof of this correspondence in Appendix A.3 for completeness.

2.1.4 Poisson processes. Let ν be a boundedly finite measure on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$. We say that a non-explosive point process N on $\mathcal{U}$ is a Poisson process on $\mathcal{U}$ with parameter measure ν if $N(A_1), \ldots, N(A_n)$ are mutually independent for all disjoint and bounded sets $A_1, \ldots, A_n \in \mathcal{B}(\mathcal{U})$, $n \in \mathbb{N}$, and $N(A)$ follows a Poisson distribution with parameter $\nu(A)$ for all bounded sets $A \in \mathcal{B}(\mathcal{U})$. Their existence can be verified using Theorem 9.2.X in Daley and Vere-Jones (2008, p. 30), see Example 9.2(b) on p. 31 therein.

2.1.5 Pathwise integration. Let N be a non-explosive point process on $\mathcal{U}$. Let $H : \Omega \times \mathcal{U} \to \mathbb{R}_{\geq 0}$ be an $\mathcal{F} \otimes \mathcal{B}(\mathcal{U})$-measurable non-negative mapping. In particular, H, identified henceforth with $(H(\cdot, u))_{u \in \mathcal{U}}$, is an $\mathbb{R}_{\geq 0}$-valued stochastic process on $\mathcal{U}$. One can define the integral of H against N in a pathwise fashion as

$$I(\omega) := \int_{\mathcal{U}} H(\omega, u) N(\omega, du), \quad \omega \in \Omega.$$

Besides, by a monotone class argument, one can check that $\omega \mapsto I(\omega)$ is $\mathcal{F}$-measurable. In the special case where N is actually a non-explosive marked point process on $\mathbb{R}_{\geq 0} \times \mathcal{M}$, the integral can be rewritten as

$$\iint_{\mathbb{R}_{\geq 0} \times \mathcal{M}} H(t, m) N(dt, dm) = \sum_{n \in \mathbb{N}} H(T_n, M_n) \mathbb{1}_{\{T_n < \infty\}}, \quad \text{a.s.},$$

where $(T_n, M_n)_{n \in \mathbb{N}}$ is the enumeration corresponding to N. For any $\xi \in \mathcal{N}^{\#g}_{\mathbb{R} \times \mathcal{M}}$ and any $\tau \in \mathbb{R}$ such that $\xi(\{\tau\} \times \mathcal{M}) > 0$, we abuse the notation and define $\iint_{\{\tau\} \times \mathcal{M}} m \xi(dt, dm)$ as the unique element $m \in \mathcal{M}$ such that $\xi(\{\tau\} \times \{m\}) = 1$.

2.1.6 Shifts, restrictions, histories and predictability. For all $t \in \mathbb{R}$, define the shift operator $\theta_t : \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{U}} \to \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{U}}$ by $\theta_t \xi(A) := \xi(A + t)$, $A \in \mathcal{B}(\mathbb{R} \times \mathcal{U})$, where $A + t := \{(s + t, u) \in \mathbb{R} \times \mathcal{U} : (s, u) \in A\}$. Then, for any non-explosive point process N on $\mathbb{R} \times \mathcal{U}$, define $\theta_t N$ through $(\theta_t N)(\omega) := \theta_t(N(\omega))$, $\omega \in \Omega$. It will be useful to show that $\theta_t \xi$ is jointly continuous in t and ξ (Lemma A.2.1).

Denote the restriction to the negative real line of any realisation $\xi \in \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{U}}$ by $\xi^{<0}$, which is defined by $\xi^{<0}(A) := \xi(A \cap \mathbb{R}_{<0} \times \mathcal{U})$, $A \in \mathcal{B}(\mathbb{R} \times \mathcal{U})$. We can then define the restriction to the negative real line of any non-explosive point process N on $\mathbb{R} \times \mathcal{U}$ by $N^{<0}(\omega) := (N(\omega))^{<0}$, $\omega \in \Omega$. Similarly, define the notations $\xi^{\leq 0}(A) := \xi(A \cap \mathbb{R}_{\leq 0} \times \mathcal{U})$, $N^{\leq 0}(\omega) := N(\omega)^{\leq 0}$, $\xi^{\geq 0}(A) := \xi(A \cap \mathbb{R}_{\geq 0} \times \mathcal{U})$, $N^{\geq 0}(\omega) := N(\omega)^{\geq 0}$, $\xi^{>0}(A) := \xi(A \cap \mathbb{R}_{>0} \times \mathcal{U})$ and $N^{>0}(\omega) := N(\omega)^{>0}$.

These notations will allow us to refer to the internal history of N. For instance, for any $t \in \mathbb{R}$, $(\theta_t N)^{<0}$ contains the history of the process up to time t, excluding time t. To lighten these notations, we will use the conventions $\theta_t \xi^{<0} := (\theta_t \xi)^{<0}$, $\theta_t \xi^{\leq 0} := (\theta_t \xi)^{\leq 0}$, $\theta_t \xi^{>0} := (\theta_t \xi)^{>0}$ and $\theta_t \xi^{\geq 0} := (\theta_t \xi)^{\geq 0}$. It will be useful

to note that these restriction mappings are measurable (Lemma A.2.2) and that $\theta_t \xi^{<0}$ is left-continuous as a function of $t \in \mathbb{R}$ (Lemma A.2.3).

Let N be a non-explosive point process on $\mathbb{R} \times \mathcal{U}$. We can define the filtration $\mathbb{F}^N = (\mathcal{F}_t^N)_{t \in \mathbb{R}}$ that corresponds to the internal history of N by

$$\mathcal{F}_t^N := \sigma \left\{ N(A \times U) \, : \, A \in \mathcal{B}(\mathbb{R}), \, A \subset (-\infty, t], \, U \in \mathcal{B}(\mathcal{U}) \right\} \quad \text{for all } t \in \mathbb{R}.$$

Using Lemma 1.4 in Kallenberg (2002, p. 4) along with the characterisation of $\mathcal{B}(\mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#})$ given in Theorem A2.6.III in Daley and Vere-Jones (2003, p. 404), one can check that $\mathcal{F}_t^N = \sigma(\theta_t N^{\leq 0})$.

In the following, we call a *history* any filtration that contains the internal history of N, that is any filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ such that $\mathcal{F}_t^N \subset \mathcal{F}_t$, $t \in \mathbb{R}$. Equivalently, one says that N is $\mathbb{F}$-adapted. The notation $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ will always be used to refer to a history. We also need to define the predictable σ-algebra $\mathcal{F}^p$ on $\Omega \times \mathbb{R} \times \mathcal{U}$ corresponding to a history $\mathbb{F}$. The σ-algebra $\mathcal{F}^p$ is the one which is generated by all the sets of the form

$$A \times (s, t] \times U, \quad s, t \in \mathbb{R}, \, s < t, \, U \in \mathcal{B}(\mathcal{U}), \, A \in \mathcal{F}_s.$$

Any mapping $H : \Omega \times \mathbb{R} \times \mathcal{U} \to \mathbb{R}$ that is $\mathcal{F}^p$-measurable is called an $\mathbb{F}$-predictable process. Any mapping $H : \Omega \times \mathbb{R}_{>0} \times \mathcal{U} \to \mathbb{R}$ that is $(\Omega \times \mathbb{R}_{>0} \times \mathcal{U}) \cap \mathcal{F}^p$-measurable is also called an $\mathbb{F}$-predictable process. Given an $\mathbb{F}$-stopping time τ, the strict past $\mathcal{F}_{\tau-}$ is defined as the σ-algebra generated by the all the classes $\{t < \tau\} \cap \mathcal{F}_t$, $t \in \mathbb{R}$.

2.1.7   Intensity process and functional. We equip the mark space $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$ with a reference measure $\mu_{\mathcal{M}}$, allowing us to define the concept of intensity rigorously. Let N be a marked point process on $\mathbb{R} \times \mathcal{M}$ and $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ a history. Let $\lambda : \Omega \times \mathbb{R}_{>0} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$ be a non-negative $\mathbb{F}$-predictable process. We say that λ is the $\mathbb{F}$-intensity of N relative to $\mu_{\mathcal{M}}$ if for every non-negative $\mathbb{F}$-predictable process $H : \Omega \times \mathbb{R}_{>0} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$,

$$\mathbb{E} \left[\iint_{\mathbb{R}_{>0} \times \mathcal{M}} H(t, m) N(dt, dm) \right] = \mathbb{E} \left[\iint_{\mathbb{R}_{>0} \times \mathcal{M}} H(t, m) \lambda(t, m) \mu_{\mathcal{M}}(dm) dt \right]. \tag{2.1}$$

Note that if an intensity exists, it is then unique up to $\mathbb{P}(d\omega) \mu_{\mathcal{M}}(dm) dt$-null sets thanks to the predictability requirement, see Brémaud (1981, Section II.4) and Daley and Vere-Jones (2008, p. 391). In this paper, we will be particularly interested in intensities that are expressed in terms of a functional applied to the point process.

Definition 2.5 (Intensity functional)**.** Let $\psi : \mathcal{M} \times \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\#} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ be a measurable functional. We say that a non-explosive marked point process $N : \Omega \to \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\#g}$ admits ψ as its *intensity functional* if N admits an $\mathbb{F}^N$-intensity $\lambda : \Omega \times \mathbb{R}_{>0} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$ relative to $\mu_{\mathcal{M}}$ such that

$$\lambda(\omega, t, m) = \psi(m \,|\, \theta_t N(\omega)^{<0}), \quad \mathbb{P}(d\omega) dt \mu_{\mathcal{M}}(dm)\text{-a.e.} \tag{2.2}$$

2.1.8   Initial condition. Let $(\Omega_{\leq 0}, \mathcal{F}_{\leq 0}, \mathbb{P}_{\leq 0})$ be a given probability space and $N_{\leq 0}$ be a given marked point process on $\mathbb{R} \times \mathcal{M}$ such that $N_{\leq 0}(\omega_{\leq 0})^{\leq 0} = N_{\leq 0}(\omega_{\leq 0})$ for all $\omega_{\leq 0} \in \Omega_{\leq 0}$ (i.e., there are no events on $\mathbb{R}_{>0}$). We will reserve the notation $N_{\leq 0}$ to refer to an initial condition.

Let $(\Omega_{>0}, \mathcal{F}_{>0}, \mathbb{P}_{>0})$ be another probability space that will correspond to the driving Poisson process in the SDE introduced in Subsection 2.3 below. In the context of strong existence, we will work with

the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which we define as the completion (Kallenberg, 2002, p. 13) of the product probability space given by

$$\Omega := \Omega_{\leq 0} \times \Omega_{> 0}, \quad \tilde{\mathcal{F}} := \mathcal{F}_{\leq 0} \otimes \mathcal{F}_{> 0}, \quad \tilde{\mathbb{P}} := \mathbb{P}_{\leq 0} \times \mathbb{P}_{> 0}. \tag{2.3}$$

Such a structure of the probability space is motivated by the fact that the driving noise and the initial condition are independent.

Definition 2.6 (Strong initial condition). Let $N : \Omega \to \mathcal{N}^{\#g}_{\mathbb{R} \times \mathcal{M}}$ be a non-explosive marked point process on $\mathbb{R} \times \mathcal{M}$. We say that N *satisfies a strong initial condition* $N_{\leq 0}$ if $N(\omega)^{\leq 0} = N_{\leq 0}(\omega_{\leq 0})$ a.s., where $\omega = (\omega_{\leq 0}, \omega_{> 0}) \in \Omega$.

In the context of weak uniqueness, one needs a another concept of initial condition. Let $(\Omega', \mathcal{F}', \mathbb{P}')$ be another probability space potentially different from $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 2.7 (Weak initial condition). Let $N' : \Omega' \to \mathcal{N}^{\#g}_{\mathbb{R} \times \mathcal{M}}$ be a non-explosive marked point process on $\mathbb{R} \times \mathcal{M}$. We say that N' *satisfies a weak initial condition* $N_{\leq 0}$ if the induced probability $\mathcal{P}^{N'^{\leq 0}}$ coincides with $\mathcal{P}^{N_{\leq 0}}$.

2.2 Hybrid marked point processes: an event–state viewpoint

2.2.1 Mark space and state process. Let $(\mathscr{E}, \mathcal{B}(\mathscr{E}), \mu_{\mathscr{E}})$ and $(\mathscr{X}, \mathcal{B}(\mathscr{X}), \mu_{\mathscr{X}})$ be two measure spaces where both $\mathscr{E}$ and $\mathscr{X}$ are complete separable metric spaces and both $\mu_{\mathscr{E}}$ and $\mu_{\mathscr{X}}$ are a boundedly finite Borel measures. Each $e \in \mathscr{E}$ represents a type of event and we call $\mathscr{E}$ the event space. Each $x \in \mathscr{X}$ represents a possible state of a system and we call $\mathscr{X}$ the state space. Motivated by the need to jointly model events and the state of the system (see Introduction), we consider a mark space $(\mathscr{M}, \mathcal{B}(\mathscr{M}), \mu_{\mathscr{M}})$ of the form

$$\mathscr{M} := \mathscr{E} \times \mathscr{X}, \quad \mathcal{B}(\mathscr{M}) = \mathcal{B}(\mathscr{E}) \otimes \mathcal{B}(\mathscr{X}), \quad \mu_{\mathscr{M}} := \mu_{\mathscr{E}} \times \mu_{\mathscr{X}}. \tag{2.4}$$

Such a decomposition of the mark space admits the following interpretation. Let $\xi \in \mathcal{N}^{\#g}_{\mathbb{R} \times \mathcal{M}}$ be a realisation of a marked point process on $\mathbb{R} \times \mathscr{M}$. A point $t \in \mathbb{R}$ and a point $m = (e, x) \in \mathscr{M}$ such that $\xi(\{t, m\}) = 1$ can now be interpreted as an event of type e occurring at time t and moving the state of the system to x. To formalise this viewpoint, we define the state functional and the state process as follows.

Definition 2.8 (State functional and state process). We define the measurable state functional $F : \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M}} \to \mathscr{X}$ by

$$F(\xi) := \begin{cases} \iint_{\{\kappa(\xi)\} \times \mathcal{M}} x\, \xi(dt, de, dx), & \text{if } \xi \in \mathcal{N}^{\#g}_{\mathbb{R} \times \mathcal{M}} \text{ and } \kappa(\xi) > -\infty, \\ x_0, & \text{otherwise,} \end{cases}$$

where $\kappa(\xi) := \inf\{t < 0 : \xi((t, 0) \times \mathscr{M}) = 0\}$ and $x_0 \in \mathscr{X}$ is an arbitrary initial state. Given a non-explosive marked point process N on $\mathscr{M}$, we define the state process $(X_t)_{t \in \mathbb{R}}$ by

$$X_t := F(\theta_t N^{< 0}), \quad t \in \mathbb{R}.$$

Note that $\kappa(\theta_t N^{< 0})$ is the time of the last event up to time t and, thus, X_t is the coordinate $x \in \mathscr{X}$ of the mark $m = (e, x) \in \mathscr{M}$ of the most recent event. As a consequence, we indeed have that a point $t \in \mathbb{R}$

and a point $m = (e, x) \in \mathcal{M}$ such that $N(\{t, m\}) = 1$ can be interpreted as an event of type e occurring at time t and moving the state of the system to x. With this viewpoint, a marked point process on $\mathbb{R} \times \mathcal{M}$ allows to jointly model the evolution of a system with state process $(X_t)_{t \in \mathbb{R}}$ in $\mathcal{X}$ and the arrival in time of the event types $\mathcal{E}$. To check that the state functional F is indeed measurable, one can adapt the proof of Lemma A.3.2 in the Appendix.

2.2.2 Implicit definition through the intensity, implied dynamics and characterisation. We can now introduce the class of hybrid marked point processes, which provides a unified treatment of various types of processes, including continuous-time Markov chains and Hawkes processes, whence the qualifier *hybrid*. Moreover, this class contains new types of processes, such as state-dependent Hawkes processes (Example 2.15 below), which are applicable to the joint modelling of events and systems. This class is specified implicitly through a specific form of the intensity. We were inspired by Cartea et al. (2018, p. 13) who, in the context of a continuous-time Markov chain model, propose a decomposition of the intensity that is similar in spirit. We should also mention the connection to the decomposition of the rate kernel α of a continuous-time Markov chain into a rate function c and transition kernel μ (i.e., $\alpha = \mu c$), see Kallenberg (2002, p. 238-239), even though this is slightly different as c is the total intensity and does not depend on the event variable $e \in \mathcal{E}$.

Definition 2.9 (Hybrid marked point processes). Let $\phi : \mathcal{X} \times \mathcal{E} \times \mathcal{X} \to \mathbb{R}_{\geq 0}$ be a measurable non-negative function such that $\phi(\cdot \mid e, x)$ is a probability density over $(\mathcal{X}, \mathcal{B}(\mathcal{X}), \mu_{\mathcal{X}})$ for all $e \in \mathcal{E}$, $x \in \mathcal{X}$. Let $\eta : \mathcal{E} \times \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M}} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ be a measurable non-negative functional. Define the measurable intensity functional $\psi : \mathcal{M} \times \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M}} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ by $\psi(m \mid \xi) := \phi(x \mid e, F(\xi))\eta(e \mid \xi)$ for all $m = (e, x) \in \mathcal{M}$, $\xi \in \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M}}$. A *hybrid marked point process* with transition function ϕ and event functional η is a non-explosive marked point process $N : \Omega \to \mathcal{N}^{\#g}_{\mathbb{R} \times \mathcal{M}}$ that admits ψ as its intensity functional. In other words, N admits an $\mathbb{F}^N$-intensity $\lambda : \Omega \times \mathbb{R}_{>0} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$ relative to $\mu_{\mathcal{M}}$ that satisfies

$$\lambda(\omega, t, e, x) = \phi(x \mid e, X_t(\omega))\eta(e \mid \theta_t N(\omega)^{<0}), \quad \mathbb{P}(d\omega)dt\mu_{\mathcal{M}}(de, dx)\text{-a.e.} \tag{2.5}$$

To demonstrate the generality and flexibility of hybrid marked point processes, we give three examples of well-known processes that belong to the class. While at first these examples might be understood only at an intuitive level, the reader should become fully convinced of their validity once Theorem 2.13 is introduced.

Example 2.10 (Compound Poisson process). Let $\mathcal{E} = \{0\}$ (i.e., just one type of event), $\mu_{\mathcal{E}} = \delta_0$, $\mathcal{X} = \mathbb{R}$, $\mu_{\mathcal{X}}(dx) = dx$ (i.e., the Lebesgue measure) and suppose that $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ is a probability density function. Consider a hybrid marked point process N with constant event functional $\eta \equiv \nu \in \mathbb{R}_{>0}$ and transition function given by $\phi(x' \mid x) = f(x' - x)$, $x', x \in \mathbb{R}$. Then, the $\mathbb{F}^N$-intensity of N satisfies $\lambda(t, x) = f(x - X_t)\nu$ and the state process $(X_t)_{t \in \mathbb{R}_{\geq 0}}$ is a compound Poisson process with rate ν and jump size distribution $f(x)dx$.

Example 2.11 (Pure jump Markov process). Let $\mathcal{E}$, $\mu_{\mathcal{E}}$, $\mathcal{X}$ and $\mu_{\mathcal{X}}$ be as in Example 2.10. Suppose that a hybrid marked point process N has event functional of the form $\eta(\xi) = c(F(\xi))$, $\xi \in \mathcal{N}^{\#}_{\mathbb{R}^2}$, where c is a positive function. The $\mathbb{F}^N$-intensity of N is then given by $\lambda(t, x) = \phi(x \mid X_t)c(X_t)$ and the state process $(X_t)_{t \in \mathbb{R}_{\geq 0}}$ is a pure jump Markov process with rate function c and transition kernel $\mu(x, B) = \int_B \phi(y \mid x)dy$, $x \in \mathbb{R}$, $B \in \mathcal{B}(\mathbb{R})$.

Example 2.12 (Multivariate Hawkes process). Let $\mathcal{E} = \{1, \ldots, d\}$, $d \in \mathbb{N}$, $\mu_{\mathcal{E}} = \sum_{n=1}^{d} \delta_n$, $\mathcal{X} = \{0\}$ (i.e., only one possible state), $\mu_{\mathcal{X}} = \delta_0$, $\nu = (\nu_1, \ldots, \nu_d) \in \mathbb{R}^d_{>0}$ and $k : \mathbb{R}_{>0} \times \mathcal{E}^2 \to \mathbb{R}_{\geq 0}$. Consider a hybrid

marked point process N with event functional

$$\eta(e\,|\,\xi) = \nu_e + \iint_{(-\infty,0)\times\mathscr{E}} k(-t',e',e)\xi(dt',de'), \quad e = 1,\ldots,d, \,\xi \in \mathcal{N}^{\#}_{\mathbb{R}\times\mathscr{E}},$$

and note that the transition function must satisfy $\phi \equiv 1$. Then N is a multivariate Hawkes process with base rate ν and kernel k.

Let us next explain the dynamics that the intensity (2.5) implies. If one integrates out the state variable x by using the fact that $\phi(\cdot\,|\,e, X_t)$ is a probability density, one can see that the $\mathbb{F}^N$-intensity of the marked point process $N(\cdot \times \mathscr{X})$ on $\mathbb{R} \times \mathscr{E}$ is exactly $\eta(e\,|\,\theta_t N^{<0})$. In other words, $\eta(e\,|\,\theta_t N^{<0})$ is the intensity of the aggregation of events of type e (irrespectively of how they impact the state process X_t). Then, $\eta(e\,|\,\theta_t N^{<0})$ is distributed in the state space $\mathscr{X}$ according to $\phi(x\,|\,e, X_t)$, specifying the intensity of events with mark (e,x). This suggests that $\phi(\cdot\,|\,e, X_t)$ is the probability density of the next state of the system given that the next event is of type e and that the current state is X_t. This intuition is confirmed by Theorem 2.13, which actually goes further and states that these dynamics characterise hybrid marked point processes. The proof is presented in Subsection 3.2.

Theorem 2.13 (Implied dynamics and characterisation). *Let ϕ and η be as in Definition 2.9. Moreover, suppose that N is a non-explosive marked point process on $\mathbb{R}\times\mathscr{M}$ with an $\mathbb{F}^N$-intensity relative to $\mu_{\mathscr{M}}$, where $\mathscr{M} = \mathscr{E} \times \mathscr{X}$. Then, N is a hybrid marked point process with transition function ϕ and event functional η if and only if the following two statements hold.*

(i) *$N_{\mathscr{E}}(\cdot) := N(\cdot \times \mathscr{X})$ is a non-explosive marked point process on $\mathbb{R} \times \mathscr{E}$ that admits an $\mathbb{F}^N$-intensity $\lambda_{\mathscr{E}} : \Omega \times \mathbb{R}_{>0} \times \mathscr{E} \to \mathbb{R}_{\geq 0}$ relative to $\mu_{\mathscr{E}}$ such that $\lambda_{\mathscr{E}}(\omega, t, e) = \eta(e\,|\,\theta_t N(\omega)^{<0})$ holds $\mathbb{P}(d\omega)dt\mu_{\mathscr{E}}(de)$-a.e.*

(ii) *Let $t \in \mathbb{R}_{\geq 0}$ and define the stopping time $\tau_t := \sup\{u > t : N((t,u) \times \mathscr{M}) = 0\}$ and the random elements $(E, X) := \iint_{\{\tau_t\}\times\mathscr{M}}(e,x)N(du, de, dx)$ such that τ_t is the time of the first event after time t and (E, X) is the corresponding mark. We have that*

$$\mathbb{P}\big(X \in dx\,|\,\sigma(E) \vee \mathcal{F}^N_{\tau_t-}\big)\mathbb{1}_{\{\tau_t<\infty\}} = \phi(x\,|\,E, X_t)\mu_{\mathscr{X}}(dx)\mathbb{1}_{\{\tau_t<\infty\}}, \quad a.s. \tag{2.6}$$

Remark 2.14. As shown in the proof of Theorem 2.13, Equation (2.6) implies that

$$\mathbb{P}\big(X \in dx\,|\,\sigma(E) \vee \mathcal{F}^N_t, \{\tau_t < \infty\}\big) = \phi(x\,|\,E, X_t)\mu_{\mathscr{X}}(dx) \quad \text{a.s.}$$

We now add a fourth example to show that Definition 2.9 contains also new types of processes. This example extends Hawkes processes to what could be called state-dependent Hawkes processes. Together, the four examples demonstrate that hybrid marked point processes provide a common framework to construct and analyse various types of processes.

Example 2.15 (State-dependent Hawkes process). Consider hybrid marked point processes with event functionals η of the form

$$\eta(e\,|\,\xi) = \nu(e) + \iint_{(-\infty,0)\times\mathscr{M}} k(-t',m',e)\xi(dt',dm'), \quad e \in \mathscr{E}, \xi \in \mathcal{N}^{\#}_{\mathbb{R}\times\mathscr{M}}, \tag{2.7}$$

where $\nu : \mathscr{E} \to \mathbb{R}_{\geq 0}$ and $k : \mathbb{R} \times \mathscr{M} \times \mathscr{E} \to \mathbb{R}_{\geq 0}$ are non-negative measurable functions. We show that such functionals are indeed measurable (Proposition A.4.1). By Theorem 2.13, this gives rise to a marked point

process $N_{\mathscr{E}}$ with marks in $\mathscr{E}$ and intensity η that interacts with a state process $(X_t)_{t\in\mathbb{R}}$ on $\mathscr{X}$ with transition probabilities ϕ. On the one hand, events in $N_{\mathscr{E}}$ occur like in a Hawkes process except that now the kernel depends also on the state process. For example, an event of type $e' \in \mathscr{E}$ might precipitate an event of type $e \in \mathscr{E}$ only if it moves the system to some specific state $x_0 \in \mathscr{X}$, i.e., $k(\cdot, e', x', e) \equiv 0$ as soon as $x' \neq x_0$. On the other hand, the occurence of an event in $N_{\mathscr{E}}$ prompts a state change according to the transition probabilities ϕ. Consequently, such a marked point process defines a state-dependent Hawkes process where the state process is fully coupled with the Hawkes process. Viewing $N_{\mathscr{E}}$ and $(X_t)_{t\in\mathbb{R}}$ as one single marked point process N on $\mathscr{E} \times \mathscr{X}$ with intensity $\phi\eta$ will allow us to prove the existence of such dynamics, see Corollary 2.18 and Example 2.19. Finally, we can also consider a state-dependent extension of non-linear Hawkes processes akin to Brémaud and Massoulié (1996), characterised by event functionals η of the form

$$\eta(e \,|\, \xi) = g\left(e, \iint_{(-\infty,0)\times\mathscr{M}} k(-t', m', e)\xi(dt', dm')\right), \quad e \in \mathscr{E}, \xi \in \mathcal{N}^{\#}_{\mathbb{R}\times\mathscr{M}}, \tag{2.8}$$

where $g : \mathscr{E} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ is a non-negative measurable function, while $k : \mathbb{R} \times \mathscr{M} \times \mathscr{E} \to \mathbb{R}$ is measurable, but need not be non-negative, allowing past events to inhibit new events. The measurability of the functional (2.8) follows readily from the proof of Proposition A.4.1.

The subclass of hybrid marked point processes characterised by (2.7) extends the regime-switching model of Vinkovskaya (2014), where the state process triggering the regime switches is not modelled. Besides, since here the events drive the dynamics of the state process, this subclass is different from the Markov-modulated Hawkes processes considered by Cohen and Elliott (2013) or Swishchuk (2017), where the state process is a continuous-time Markov chain that jumps independently of the events. Moreover, the intensity in Cohen and Elliott (2013) depends only on the current state whereas, here and in Swishchuk (2017), it may depend on all past states.

2.3 Existence and uniqueness of hybrid marked point processes

In this subsection, $\mathscr{M}$ is not required to be a product space as in Subsection 2.2 but can be again an arbitrary complete separable metric space.

2.3.1 The existence and uniqueness problem. A hybrid marked point processes (Definition 2.9) is defined implicitly via its intensity process, which, in turn, depends on the history of the hybrid marked point process. Due to the self-referential nature of the definition, it is not clear a priori that such a marked point process exists. More generally, given an initial condition $N_{\leq 0}$ (see Subsection 2.1.8) and a measurable intensity functional $\psi : \mathscr{M} \times \mathcal{N}^{\#}_{\mathbb{R}\times\mathscr{M}} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$, one can ask if there exists a unique non-explosive marked point process N that satisfies the initial condition $N_{\leq 0}$ on $\mathbb{R}_{\leq 0}$ and admits ψ as its intensity functional on $\mathbb{R}_{>0}$.

Massoulié (1998) tackles this question by reformulating the existence problem as a Poisson-driven SDE, extending the works of Brémaud and Massoulié (1996), Grigelionis (1971) and Kerstan (1964). Delattre et al. (2016) also employ this Poisson embedding technique in the context of Hawkes processes on infinite directed graphs. However, in these papers, strong existence and uniqueness is obtained by imposing a Lipschitz condition on the intensity functional ψ. More precisely, it is assumed that there exists a non-negative kernel $\overline{k} : \mathbb{R}_{>0} \times \mathscr{M} \times \mathscr{M} \to \mathbb{R}_{\geq 0}$ such that

$$|\psi(m \,|\, \xi) - \psi(m \,|\, \xi')| \leq \iint_{\mathbb{R}_{<0}\times\mathscr{M}} \overline{k}(-t', m', m)|\xi - \xi'|(dt', dm'), \quad m \in \mathscr{M}, \xi, \xi' \in \mathcal{N}^{\#}_{\mathbb{R}\times\mathscr{M}}. \tag{2.9}$$

Unfortunately, this condition is too restrictive in the context of hybrid marked point processes. A simple, yet natural, example of a hybrid marked point process not satisfying (2.9) is given in Subsection 4.1.1. Hence, our goal is to construct a strong solution to a Poisson-driven SDE without imposing the Lipschitz condition (2.9) on the intensity functional ψ. We will in fact extend the existence result in Massoulié (1998) by imposing only a weaker sublinearity condition on ψ. The idea to define a random measure as a strong solution to an SDE driven by another random measure was also studied by Jacod (1979). Similarly, a Lipschitz condition that does not seem to apply to Hawkes processes and hybrid marked point processes is required (Jacod, 1979, Chapter 14, Section 1).

Let us also briefly review some weak existence and uniqueness results. Jacod (1975) proved that there exists a unique probability measure on the canonical space of marked point processes such that the canonical marked point process admits a given compensator. However, this marked point process may be explosive a priori. Still, we will apply this result in the proof of Theorem 2.21 below to obtain weak uniqueness. A similar approach is followed by Jacobsen (2006, Proposition 4.3.5, Corollary 4.4.4) who, furthermore, gives a domination condition on the intensity functional ensuring that the corresponding marked point process is non-explosive. Proposition 4.15 will be the counterpart of this result in the strong setting. These weak existence results are however limited to intensities with respect to the internal history $\mathbb{F}^N$. The advantage of the strong setting is that the results of Massoulié (1998) and the pathwise construction of this paper also hold when the intensity functional depends additionally on an auxiliary process, meaning that intensities with respect to larger filtrations can be considered. Besides, the Poisson-driven SDE representation of marked point processes considered in the strong setting directly suggests a simulation (thinning) algorithm. In fact, the Poisson embedding lemma (Lemma 4.6 below), which is a stepping stone to the strong setting, was first given in the simulation literature (Lewis and Shedler, 1976; Ogata, 1981).

Finally, there is a third approach to obtain existence, based on a change of measure, see Brémaud (1981, Theorem 11, p. 242) and Sokol and Hansen (2015). While this technique also accommodates filtrations that are larger than the internal history, existence is generally obtained only on finite time intervals.

2.3.2 The Poisson-driven SDE. Let $(\Omega_{>0}, \mathcal{F}_{>0}, \mathbb{P}_{>0})$ be given and let $M_{>0} : \Omega_{>0} \to \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M} \times \mathbb{R}}$ be a Poisson process on $\mathbb{R} \times \mathcal{M} \times \mathbb{R}$ with mean measure $dt \mu_{\mathcal{M}}(dm) dz$. As usual, denote by $(\mathcal{F}_t^{M_{>0}})_{t \in \mathbb{R}}$ the internal history of $M_{>0}$ on $\Omega_{>0}$. In this Subsection, we work under the assumption that the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is the completion of the product probability space defined by (2.3). In particular, $\Omega := \Omega_{\leq 0} \times \Omega_{>0}$, where $\Omega_{\leq 0}$ corresponds to the probability space of an initial condition $N_{\leq 0}$, see Subsection 2.1.8. We extend $M_{>0}$ to a mapping $M : \Omega \to \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M} \times \mathbb{R}}$ by simply setting

$$M(\omega) := M_{>0}(\omega_{>0}), \quad \omega = (\omega_{\leq 0}, \omega_{>0}) \in \Omega. \tag{2.10}$$

Let $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ be the filtration on Ω such that, for all $t \in \mathbb{R}$, $\mathcal{F}_t$ is the $\mathbb{P}$-completion of $\mathcal{F}_t^{N_{\leq 0}} \otimes \mathcal{F}_t^{M_{>0}}$ in $\mathcal{F}$. In particular, the filtration $\mathbb{F}$ is complete (Kallenberg, 2002, p. 123). Similarly to Massoulié (1998), we want to solve the following Poisson-driven SDE.

Definition 2.16 (The Poisson-driven SDE). Let $\psi : \mathcal{M} \times \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M}} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ be a given measurable functional. By a solution to the *Poisson-driven SDE*, we mean an $\mathbb{F}$-adapted non-explosive marked-point

process $N : \Omega \to \mathcal{N}^{\#g}_{\mathbb{R} \times \mathcal{M}}$ that solves

$$\begin{cases} N(dt, dm) = M(dt, dm, (0, \lambda(t, m)]), & t \in \mathbb{R}_{>0}, \text{ a.s.}, \\ \lambda(\omega, t, m) = \psi(m \mid \theta_t N(\omega)^{<0}), & t \in \mathbb{R}_{>0}, m \in \mathcal{M}, \omega \in \Omega, \\ N^{\leq 0}(\omega) = N_{\leq 0}(\omega_{\leq 0}), & \omega = (\omega_{\leq 0}, \omega_{>0}) \in \Omega, \text{a.s.}, \end{cases} \qquad (2.11)$$

where $N_{\leq 0}$ is a given initial condition (see Subsection 2.1.8).

Still, notice that our problem differs slightly as we only search for solutions in the space of non-explosive marked point processes, a smaller space than the one considered in Massoulié (1998).

2.3.3 Assumptions. The following assumptions are only required for the strong existence result (Theorem 2.17 below). We first need to assume that the mark space $\mathcal{M}$ has finite total mass.

Assumption A. *The reference measure $\mu_{\mathcal{M}}$ is finite, i.e., $\mu_{\mathcal{M}}(\mathcal{M}) < \infty$.*

Next, we need to control for both the intensity functional ψ and the initial condition $N_{\leq 0}$. We will prove Theorem 2.17 for two different scenarios. In the first scenario, the intensity is dominated by an increasing function of the total number of past events, while the number of events before time 0 is finite.

Assumption B. *There exists a non-decreasing function $a : \mathbb{N} \cup \{\infty\} \to \mathbb{N} \cup \{\infty\}$ with $a(n) < \infty$ for all $n \in \mathbb{N}$ and $a(\infty) = \infty$ such that:*

(i) $\psi(m \mid \xi) \leq a(\xi((-\infty, 0) \times \mathcal{M}))$, $m \in \mathcal{M}$, $\xi \in \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M}}$;

(ii) $\sum_{n=0}^{\infty} a(n)^{-1} = \infty$.

Assumption C. *The initial condition satisfies $N_{\leq 0}(\omega_{\leq 0}, (-\infty, 0] \times \mathcal{M}) < \infty$ for all $\omega_{\leq 0} \in \Omega_{\leq 0}$.*

In the second scenario, the intensity functional ψ is dominated by a Hawkes functional. Note that this requirement is weaker than the Lipschitz condition (2.9) in Massoulié (1998).

Assumption D. *There exists $\lambda_0 \in \mathbb{R}_{\geq 0}$ and a measurable function $\overline{k} : \mathbb{R}_{>0} \times \mathcal{M} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$ such that:*

(i) $\psi(m \mid \xi) \leq \lambda_0 + \iint_{(-\infty,0) \times \mathcal{M}} \overline{k}(-t', m', m) \xi(dt', dm')$, $m \in \mathcal{M}$, $\xi \in \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M}}$;

(ii) $\rho := \sup_{m \in \mathcal{M}} \iint_{(0,\infty) \times \mathcal{M}} \overline{k}(t', m', m) \mu_{\mathcal{M}}(dm') dt' < 1$;

(iii) $\sup_{m \in \mathcal{M}} k(t', m', m) < \infty$ for all $t' \in \mathbb{R}_{>0}$, $m' \in \mathcal{M}$.

Assumption E. *The initial condition $N_{\leq 0}$ satisfies:*

(i) $\sup_{t>0, \, m \in \mathcal{M}} \mathbb{E}\left[\iint_{(-\infty,0] \times \mathcal{M}} \overline{k}(t - t', m', m) N_{\leq 0}(dt', dm') \right] < \infty$;

(ii) $\tilde{\lambda}_{\leq 0}(\omega_{\leq 0}, t) := \sup_{m \in \mathcal{M}} \iint_{(-\infty,0] \times \mathcal{M}} \overline{k}(t - t', m', m) N_{\leq 0}(\omega_{\leq 0}, dt', dm') < \infty$, $\omega_{\leq 0} \in \Omega_{\leq 0}$, $t \in \mathbb{R}_{>0}$.

Note that Assumptions D.(ii) and E.(i) are needed in order to reuse Theorem 2 in Massoulié (1998). It will allow us to dominate the marked point process N by a Hawkes process with kernel $\overline{k}$.

2.3.4 Existence. We construct a solution to the Poisson-driven SDE in two mains steps. First, by taking advantage of the discrete nature of the driving Poisson process, we construct in a pathwise fashion a mapping $N : \Omega \to \mathcal{N}^\infty_{\mathbb{R} \times \mathcal{M}}$ that solves (2.11) up to each event time, generalising the construction in Çınlar (2011, Chapter 6, p. 302-306) and Lindvall (1988, p. 127). Second, we dominate N by a non-explosive marked point process to show that N is itself non-explosive, generalising the argument in Chevallier (2017, Lemma B.1, p. 30). When working under Assumptions D and E, these two steps must actually be performed concurrently. Then, it turns out that this constructed N admits ψ as its intensity functional on $\mathbb{R}_{>0}$ and, thus, solves the existence problem. The proof of the following theorem, which extends the existence result in Massoulié (1998), is given in Subsection 4.2.

Theorem 2.17 (Strong existence)**.** *Under either Assumptions A, B, C or Assumptions A, D, E, there exists a non-explosive marked point process $N : \Omega \to \mathcal{N}^{\#g}_{\mathbb{R} \times \mathcal{M}}$ that solves the Poisson-driven SDE (Definition 2.16). Any such N satisfies the strong initial condition $N_{\leq 0}$ and admits ψ as its intensity functional on $\mathbb{R}_{>0}$.*

As a corollary, we obtain conditions that ensure the existence of hybrid marked point processes.

Corollary 2.18 (Existence of hybrid marked point processes)**.** *Suppose that Assumption A holds and $\|\phi\|_\infty < \infty$. Moreover, suppose that either Assumptions B and C or Assumptions D and E hold with $\psi(m \,|\, \xi)$ replaced by $\eta(e \,|\, \xi)$, where the dominating kernel $\overline{k}$ is now a function $\overline{k} : \mathbb{R}_{>0} \times \mathcal{M} \times \mathcal{E} \to \mathbb{R}_{\geq 0}$, and with the constraint $\rho < \|\phi\|^{-1}_\infty$. Then, there exists a hybrid marked point process $N : \Omega \to \mathcal{N}^{\#g}_{\mathbb{R} \times \mathcal{M}}$ with transition function ϕ and event functional η that satisfies the strong initial condition $N_{\leq 0}$.*

Example 2.19 (State-dependent Hawkes processes, continuation of Example 2.15)**.** When the transition function ϕ is bounded, the above corollary establishes the existence of state-dependent Hawkes processes, as characterised by (2.7), where the kernel k is either bounded without an integrability constraint or unbounded (up to constraint (iii) of Assumption D) with an integrability constraint. Likewise, the existence of state-dependent non-linear Hawkes processes (2.8) follows provided the function g therein is such that for some $\bar{\lambda}_0 \in \mathbb{R}_{>0}$ we have $\sup_{e \in \mathcal{E}} g(e, x) \leq \bar{\lambda}_0 + \max\{x, 0\}$ for all $x \in \mathbb{R}$. In this case, the boundedness and integrability assumptions above apply to the positive part of the kernel k.

2.3.5 Uniqueness. As Massoulié (1998) considers point processes on $\mathbb{R} \times \mathcal{M}$ that are not necessarily non-explosive marked point processes, he uses the Lipschitz condition (2.9) to obtain strong uniqueness in a space of regular point processes. Here, since we restrict ourselves to non-explosive marked point processes, the enumeration representation allows us to prove strong uniqueness more easily without any specific assumptions. The proof is deferred until Subsection 4.3.

Theorem 2.20 (Strong uniqueness)**.** *Let $N : \Omega \to \mathcal{N}^{\#g}_{\mathbb{R} \times \mathcal{M}}$ and $N' : \Omega \to \mathcal{N}^{\#g}_{\mathbb{R} \times \mathcal{M}}$ be two non-explosive marked point processes solving the Poisson-Driven SDE (Definition 2.16). Then $N = N'$ a.s.*

By applying Theorem 3.4 in Jacod (1975), we can also obtain weak uniqueness. Alternatively, we could also have applied Theorem 14.2.IV in Daley and Vere-Jones (2008, p. 381). The idea is that the intensity and the conditional distributions $\mathbb{P}((T_{n+1}, M_{n+1}) \in \cdot \,|\, \mathcal{F}^N_{T_n})$ uniquely determine each other, see also Last and Brandt (1995) and Jacobsen (2006, Theorem 4.3.2, p. 54). Another approach, as suggested by Massoulié (1998), could be to use the fact that any marked point process with an intensity functional can be represented as the strong solution to a Poisson-driven SDE like in Definition 2.16, see Jacod (1979, Theorem 14.56, p. 472), and use the strong uniqueness result. We prove the following result in Subsection 4.3.

Theorem 2.21 (Weak uniqueness). *Let N_1 and N_2 be two non-explosive marked point processes (possibly on distinct probability spaces) that admit the same intensity functional ψ on $\mathbb{R}_{>0}$. Assume also that both N_1 and N_2 satisfy the weak initial condition $N_{\leq 0}$. Then, we have that $\mathcal{P}^{N_1} = \mathcal{P}^{N_2}$, i.e., the induced probabilities measures on $\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#}$ coincide.*

As a corollary, we obtain the weak uniqueness of hybrid marked point processes.

Corollary 2.22 (Uniqueness of hybrid marked point processes). *All hybrid marked point processes with transition function ϕ and event functional η that satisfy the weak initial condition $N_{\leq 0}$ induce the same probability measure on $\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#}$.*

Remark 2.23. Note that weak uniqueness might not hold for a general history $\mathbb{F}$. Given an $\mathbb{F}$-predictable process λ, there could be two marked point processes N and N' that both admit λ as their $\mathbb{F}$-intensity, but such that $\mathcal{P}^N \neq \mathcal{P}^{N'}$, see Proposition 9.54.(ii) in Kallenberg (2017) for such an example. The fact the we restrict ourselves to the natural filtration $\mathbb{F}^N$ is crucial here.

3 Dynamics of hybrid marked point processes

In this section, we prove Theorem 2.13, which characterises the dynamics of hybrid marked point processes.

3.1 Preliminaries

We first present a lemma that helps us reuse some results in the literature that require a specific form for the filtration. It simply says that the information up to time u is equal to the information up to time t to which we add the information between time t and u, where $t < u$.

Lemma 3.1. *Let N be a non-explosive point process on $\mathbb{R} \times \mathcal{U}$. Let $t, u \in \mathbb{R}$ such that $u > t$. Then, we have that $\mathcal{F}_u^N = \mathcal{F}_t^N \vee \mathcal{F}_u^{\theta_t N^{>0}}$.*

Proof. Note that

$$\mathcal{F}_u^{\theta_t N^{>0}} = \sigma \left\{ N(A \times U) \,:\, A \in \mathcal{B}(\mathbb{R}), \, A \subset (t, u], \, U \in \mathcal{B}(\mathcal{U}) \right\} \quad \text{for all } u > t.$$

Then, clearly $\mathcal{F}_u^{\theta_t N^{>0}} \subset \mathcal{F}_u^N$. Also, $\mathcal{F}_t^N \subset \mathcal{F}_u^N$ and, thus $\mathcal{F}_t^N \vee \mathcal{F}_u^{\theta_t N^{>0}} \subset \mathcal{F}_u^N$. On the other hand, let $A \in \mathcal{B}(\mathbb{R})$ be such that $A \subset (-\infty, u]$ and let $U \in \mathcal{B}(\mathcal{U})$. We have that

$$N(A \times U) = N(A \cap (-\infty, t]) \times U) + N(A \cap (t, u] \times U).$$

The first term is $\mathcal{F}_t^N$-measurable while the second term is $\mathcal{F}_u^{\theta_t N^{>0}}$-measurable. Hence, $N(A \times U)$ is $\mathcal{F}_t^N \vee \mathcal{F}_u^{\theta_t N^{>0}}$ measurable. Since, by definition, $\mathcal{F}_u^N$ is the smallest σ-algebra that makes all the $N(A \times U)$ measurable, this implies that $\mathcal{F}_u^N \subset \mathcal{F}_t^N \vee \mathcal{F}_u^{\theta_t N^{>0}}$, which concludes the proof. $\qquad\square$

As defined in Subsection 2.1.7, an intensity process has to always be finite. We verify that, if one finds a potentially infinite process that satisfies the definition of the intensity, then one can take a finite version of this process and identify it with the intensity.

Lemma 3.2. *Let N be a non-explosive marked point process on $\mathbb{R} \times \mathscr{M}$ and let $\tilde{\lambda} : \Omega \times \mathbb{R}_{>0} \times \mathscr{M} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ be an $\mathbb{F}$-predictable process that satisfies (2.1) for all non-negative $\mathbb{F}$-predictable processes H. Then N admits an $\mathbb{F}$-intensity $\lambda : \Omega \times \mathbb{R}_{>0} \times \mathscr{M} \mapsto \mathbb{R}_{\geq 0}$ relative to $\mu_{\mathscr{M}}$ such that $\lambda(\omega, t, m) = \tilde{\lambda}(\omega, t, m)$ holds $\mathbb{P}(d\omega)\mu_{\mathscr{M}}(dm)dt$-a.e.*

Proof. Since the marked point process N is non-explosive, using similar arguments as in Lemma L2 of Brémaud (1981, p. 24), one can show that, for all bounded sets $A \in \mathcal{B}(\mathbb{R}_{>0})$,

$$\iint_{A \times \mathscr{M}} \tilde{\lambda}(t, m)\mu_{\mathscr{M}}(dm)dt < \infty, \quad \text{a.s.,}$$

which implies that $\tilde{\lambda}(\omega, t, m) < \infty$ holds $\mathbb{P}(d\omega)dt\mu_{\mathscr{M}}(dm)$-a.e. By a composition argument (see the beginning of the proof of Lemma 4.6), since $\tilde{\lambda}$ is $\mathbb{F}$-predictable, we have that $(\omega, t, m) \mapsto \mathbb{1}_{\{\tilde{\lambda}(\omega, t, m) < \infty\}}$ is also $\mathbb{F}$-predictable. It is then easy to check that $\lambda(\omega, t, m) := \mathbb{1}_{\{\tilde{\lambda}(\omega, t, m) < \infty\}}\tilde{\lambda}(\omega, t, m)$ is the $\mathbb{F}$-intensity of N where we use the convention $0 \times \infty = 0$. □

The next lemma says that by integrating the intensity against the state variable x, we obtain the intensity of the marked point process that tracks the event types, ignoring the state process.

Lemma 3.3. *Let N be a marked point process on $\mathbb{R} \times \mathscr{M}$ with $\mathbb{F}$-intensity λ relative to $\mu_{\mathscr{M}}$. Then, $N_{\mathscr{E}}(\cdot) := N(\cdot \times \mathscr{X})$ is a non-explosive marked point process on $\mathbb{R} \times \mathscr{E}$ with $\mathbb{F}$-intensity $\lambda_{\mathscr{E}} : \Omega \times \mathbb{R}_{>0} \times \mathscr{E} \to \mathbb{R}_{\geq 0}$ relative to $\mu_{\mathscr{E}}$ such that $\lambda_{\mathscr{E}}(\omega, t, e) = \int_{\mathscr{X}} \lambda(\omega, t, e, x)\mu_{\mathscr{X}}(dx)$ holds $\mathbb{P}(d\omega)dt\mu_{\mathscr{E}}(de)$-a.e.*

Proof. Let $H : \Omega \times \mathbb{R}_{>0} \times \mathscr{E} \to \mathbb{R}_{\geq 0}$ be an $\mathbb{F}$-predictable non-negative process. Then, by applying the definition of $N_{\mathscr{E}}$ and using Tonelli's theorem, we obtain that

$$\mathbb{E}\left[\iint_{\mathbb{R}_{>0} \times \mathscr{E}} H(t, e)N_{\mathscr{E}}(dt, de)\right] = \mathbb{E}\left[\iiint_{\mathbb{R}_{>0} \times \mathscr{E} \times \mathscr{X}} H(t, e)N(dt, de, dx)\right]$$

$$= \mathbb{E}\left[\iiint_{\mathbb{R}_{>0} \times \mathscr{E} \times \mathscr{X}} H(t, e)\lambda(t, e, x)\mu_{\mathscr{X}}(dx)\mu_{\mathscr{E}}(de)dt\right]$$

$$= \mathbb{E}\left[\iint_{\mathbb{R}_{>0} \times \mathscr{E}} H(t, e)\left(\int_{\mathscr{X}} \lambda(t, e, x)\mu_{\mathscr{X}}(dx)\right)\mu_{\mathscr{E}}(de)dt\right].$$

The process $\int_{\mathscr{X}} \lambda(t, e, x)\mu_{\mathscr{X}}(dx)$, $t \in \mathbb{R}_{>0}$, $e \in \mathscr{E}$, is $\mathbb{F}$-predictable, see for example Lemma 25.23 in Kallenberg (2002, p. 503) and we conclude using Lemma 3.2. □

We now check that an intensity functional applied to the history of a point process defines a predictable process.

Lemma 3.4. *Let $\psi : \mathcal{U} \times \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{U}} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ be a measurable functional and N be a non-explosive point process on $\mathbb{R} \times \mathcal{U}$ that is $\mathbb{F}$-adapted. Then, the process $\lambda : \Omega \times \mathbb{R} \times \mathcal{U} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ defined by $\lambda(\omega, t, u) = \psi(u \,|\, \theta_t N(\omega)^{<0})$, $\omega \in \Omega, t \in \mathbb{R}, u \in \mathcal{U}$, is $\mathbb{F}$-predictable.*

Proof. By Lemma A.2.3, $\theta_t N(\omega)^{<0}$ is left-continuous in t and, by assumption, the process $(\theta_t N^{<0})_{t \in \mathbb{R}}$ is $\mathbb{F}$-adapted. As a consequence, the mapping $(\omega, t) \mapsto \theta_t N(\omega)^{<0}$ is $\mathbb{F}$-predictable, see for example Lemmas 25.1 and 1.10 in Kallenberg (2002, p. 491, p. 6). We then obtain that λ is $\mathbb{F}$-predictable by viewing it as the composition $(\omega, t, u) \mapsto (u, \theta_t N(\omega)^{<0}) \mapsto \psi(u \,|\, \theta_t N(\omega)^{<0})$ and using the measurability of ψ. □

The next lemma essentially says that if two predictable processes coincide at all event times of a marked point process, then they coincide everywhere under positive intensity. A less general variant of this result and its proof are suggested in Brémaud (1981, Theorem T12, p. 31).

Lemma 3.5. *Let $N_{\mathscr{E}}$ be a non-explosive marked point process on $\mathbb{R} \times \mathscr{E}$ with $\mathbb{F}$-intensity $\lambda_{\mathscr{E}}$ relative to $\mu_{\mathscr{E}}$. Let $H_1 : \Omega \times \mathbb{R}_{>0} \times \mathscr{M} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ and $H_2 : \Omega \times \mathbb{R}_{>0} \times \mathscr{M} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ be two non-negative $\mathbb{F}$-predictable processes. Then, $H_1 = H_2$ holds $\mathbb{P}(d\omega) N_{\mathscr{E}}(\omega, dt, de) \mu_{\mathscr{X}}(dx)$-a.e. if and only if $H_1 = H_2$ holds $\mathbb{P}(d\omega) \lambda_{\mathscr{E}}(\omega, t, e) dt \mu_{\mathscr{M}}(de, dx)$-a.e.*

Proof. By a composition argument, since H_1 and H_2 are $\mathbb{F}$-predictable, we have that the function $(\omega, t, m) \mapsto \mathbb{1}_{\{H_1(\omega,t,m) \neq H_1(\omega,t,m)\}}$ is $\mathbb{F}$-predictable (see the beginning of the proof of Lemma 4.6). By Lemma 25.23 in Kallenberg (2002, p. 503), we also have that the process $\int_{\mathscr{X}} \mathbb{1}_{\{H_1(\cdot,\cdot,\cdot,x) \neq H_1(\cdot,\cdot,\cdot,x)\}} \mu_{\mathscr{X}}(dx)$ is $\mathbb{F}$-predictable. Using the definition of the intensity and Tonelli's theorem, we obtain that

$$\int_{\Omega} \int_{\mathbb{R}_{>0} \times \mathscr{M}} \mathbb{1}_{\{H_1(\omega,t,m) \neq H_1(\omega,t,m)\}} \mu_{\mathscr{X}}(dx) N_{\mathscr{E}}(\omega, dt, de) \mathbb{P}(d\omega)$$

$$= \int_{\Omega} \int_{\mathbb{R}_{>0} \times \mathscr{E}} \left(\int_{\mathscr{X}} \mathbb{1}_{\{H_1(\omega,t,m) \neq H_1(\omega,t,m)\}} \mu_{\mathscr{X}}(dx) \right) N_{\mathscr{E}}(\omega, dt, de) \mathbb{P}(d\omega)$$

$$= \int_{\Omega} \int_{\mathbb{R}_{>0} \times \mathscr{M}} \mathbb{1}_{\{H_1(\omega,t,m) \neq H_1(\omega,t,m)\}} \mu_{\mathscr{X}}(dx) \lambda_{\mathscr{E}}(\omega, t, e) \mu_{\mathscr{E}}(de) dt \mathbb{P}(d\omega),$$

from which the assertion follows. $\square$

Finally, we show that the link between joint densities and conditional densities still holds when we pre-condition on a sub-σ-algebra. Since $\mathscr{M}$ is a complete separable metric space and, in particular, Borel, random elements in $\mathscr{M}$ always have regular conditional distributions (Kallenberg, 2002, p. 106, Theorem A1.2, p. 561).

Lemma 3.6. *Let (E, X) be a random element in $\mathscr{M}$. Let $\mathcal{G}$ be a sub-σ-algebra, i.e., $\mathcal{G} \subset \mathcal{F}$, and let $A \in \mathcal{G}$ such that $\mathbb{P}(A) > 0$. Moreover, let $f : \Omega \times \mathscr{M} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ be a non-negative measurable function that is $\mathcal{G} \otimes \mathcal{B}(\mathscr{M})$-measurable. If we have*

$$\mathbb{P}\left(E \in de, X \in dx \,|\, \mathcal{G}\right) \mathbb{1}_A = f(e, x) \mu_{\mathscr{M}}(de, dx) \mathbb{1}_A, \quad a.s., \tag{3.1}$$

then

$$\mathbb{P}\left(X \in dx \,|\, \sigma(E) \vee \mathcal{G}\right) \mathbb{1}_A = \frac{f(E, x)}{\int_{\mathscr{X}} f(E, x') \mu_{\mathscr{X}}(dx')} \mu_{\mathscr{X}}(dx) \mathbb{1}_A, \quad a.s.$$

Proof. Let $B \in \mathcal{B}(\mathscr{X})$, $G \in \mathcal{G}$ and $H \in \sigma(E)$. On the one hand,

$$\mathbb{E}\left[\mathbb{1}_G \mathbb{1}_H \mathbb{1}_{\{X \in B\}} \mathbb{1}_A\right] = \mathbb{E}\left[\mathbb{1}_G h(E) \mathbb{1}_{\{X \in B\}} \mathbb{1}_A\right] = \mathbb{E}\left[\mathbb{1}_G \mathbb{E}\left[h(E) \mathbb{1}_{\{X \in B\}} \mathbb{1}_A \,|\, \mathcal{G}\right]\right]$$

$$= \mathbb{E}\left[\mathbb{1}_G \mathbb{1}_A \int_{\mathscr{E}} \int_B h(e) f(e, x) \mu_{\mathscr{X}}(dx) \mu_{\mathscr{E}}(de)\right], \tag{3.2}$$

where we successively used Lemma 1.13 in Kallenberg (2002, p. 7) to write $\mathbb{1}_H = h(E)$ using a measurable function $h : \mathscr{E} \to \{0, 1\}$, the Tower property, the disintegration theorem in Kallenberg (2002, Theorem 6.4, p. 108) with the regular conditional distribution of (3.1) and, finally, the product form of $\mu_{\mathscr{M}}$. Note that,

here, the disintegration theorem is applied to the probability measure $\mathbb{P}(\cdot \cap A)/\mathbb{P}(A)$ on the measurable space $(A, A \cap \mathcal{F})$. On the other hand, observe that (2.4) and (3.1) imply that

$$\mathbb{P}\left(E \in de \mid \mathcal{G}\right)\mathbb{1}_A = \int_{\mathscr{X}} f(e, x)\mu_{\mathscr{X}}(dx)\mu_{\mathscr{E}}(de)\mathbb{1}_A, \quad \text{a.s.}$$

Then, using similar arguments,

$$\mathbb{E}\left[\mathbb{1}_G\mathbb{1}_H \int_B \frac{f(E, x)}{\int_{\mathscr{X}} f(E, x')\mu_{\mathscr{X}}(dx')}\mu_{\mathscr{X}}(dx)\mathbb{1}_A\right] = \mathbb{E}\left[\mathbb{1}_G h(E) \int_B \frac{f(E, x)}{\int_{\mathscr{X}} f(E, x')\mu_{\mathscr{X}}(dx')}\mu_{\mathscr{X}}(dx)\mathbb{1}_A\right]$$

$$= \mathbb{E}\left[\mathbb{1}_G \mathbb{E}\left[h(E) \int_B \frac{f(E, x)}{\int_{\mathscr{X}} f(E, x')\mu_{\mathscr{X}}(dx')}\mu_{\mathscr{X}}(dx)\mathbb{1}_A \,\Big|\, \mathcal{G}\right]\right]$$

$$= \mathbb{E}\left[\mathbb{1}_G\mathbb{1}_A \int_{\mathscr{E}}\left(h(e) \int_B \frac{f(e, x)}{\int_{\mathscr{X}} f(e, x')\mu_{\mathscr{X}}(dx')}\mu_{\mathscr{X}}(dx)\right) \int_{\mathscr{X}} f(e, x')\mu_{\mathscr{X}}(dx')\mu_{\mathscr{E}}(de)\right].$$

Tonelli's theorem and (3.2) then imply that

$$\mathbb{E}\left[\mathbb{1}_G\mathbb{1}_H\mathbb{1}_{\{X \in B\}}\mathbb{1}_A\right] = \mathbb{E}\left[\mathbb{1}_G\mathbb{1}_H \int_B \frac{f(E, x)}{\int_{\mathscr{X}} f(E, x')\mu_{\mathscr{X}}(dx')}\mu_{\mathscr{X}}(dx)\mathbb{1}_A\right]. \tag{3.3}$$

Using a monotone class argument, we show below that (3.3) can be extended to

$$\mathbb{E}\left[\mathbb{1}_F\mathbb{1}_{\{X \in B\}}\mathbb{1}_A\right] = \mathbb{E}\left[\mathbb{1}_F \int_B \frac{f(E, x)}{\int_{\mathscr{X}} f(E, x')\mu_{\mathscr{X}}(dx')}\mu_{\mathscr{X}}(dx)\mathbb{1}_A\right] \tag{3.4}$$

for all $F \in \sigma(E) \vee \mathcal{G}$, which means exactly that

$$\mathbb{P}\left(X \in B \mid \sigma(E) \vee \mathcal{G}\right)\mathbb{1}_A = \int_B \frac{f(E, x)}{\int_{\mathscr{X}} f(E, x')\mu_{\mathscr{X}}(dx')}\mu_{\mathscr{X}}(dx)\mathbb{1}_A, \quad \text{a.s.},$$

as asserted.

To prove (3.4), define the functions

$$\mu_1 : \quad \sigma(E) \vee \mathcal{G} \to [0, 1]$$
$$F \mapsto \mu_1(F) := \mathbb{E}\left[\mathbb{1}_F\mathbb{1}_{\{X \in B\}}\mathbb{1}_A\right],$$
$$\mu_2 : \quad \sigma(E) \vee \mathcal{G} \to [0, 1]$$
$$F \mapsto \mu_2(F) := \mathbb{E}\left[\mathbb{1}_F \int_B \frac{f(E, x)}{\int_{\mathscr{X}} f(E, x')\mu_{\mathscr{X}}(dx')}\mu_{\mathscr{X}}(dx)\mathbb{1}_A\right].$$

One can check that μ_1 and μ_2 are bounded measures on $(\Omega, \sigma(E) \vee \mathcal{G})$ (to swap an expectation with an infinite sum, use the monotone convergence theorem, see for example Theorem 1.19 in Kallenberg (2002, p. 11)). Define also the class $\mathcal{C} := \{G \cap H \,:\, G \in \mathcal{G}, \, H \in \sigma(E)\}$. Equation (3.3) means that $\mu_1(C) = \mu_2(C)$ for all $C \in \mathcal{C}$. Moreover, $\mathcal{C}$ is a π-system such that $\Omega \in \mathcal{C}$. Also, note that $\sigma(E) \cup \mathcal{G} \subset \mathcal{C} \subset \sigma(E) \vee \mathcal{G}$ and, thus, $\sigma(\mathcal{C}) = \sigma(E) \vee \mathcal{G}$. As a consequence, we can apply Lemma 1.17 in Kallenberg (2002, p. 9) to conclude that $\mu_1(F) = \mu_2(F)$ for all $F \in \sigma(E) \vee \mathcal{G}$, meaning that (3.4) holds. $\qquad\square$

3.2 Implied dynamics and characterisation

Proof of Theorem 2.13. Recall that we denote by λ the $\mathbb{F}^N$-intensity of N relative to $\mu_{\mathscr{M}}$ and by $\lambda_{\mathscr{E}}$ the $\mathbb{F}^N$-intensity of $N_{\mathscr{E}} := N(\cdot \times \mathscr{X})$ relative to $\mu_{\mathscr{E}}$.

Necessity. Assume that N is a hybrid marked point process with transition function ϕ and event functional η. We first observe that statement (i) holds simply by applying Lemma 3.3 and using the fact that $\phi(\cdot \,|\, e, x)$ is a probability density for all $e \in \mathscr{E}$ and $x \in \mathscr{X}$.

Next, we show that statement (ii) holds. This is clearly true when $\mathbb{P}(\tau_t < \infty) = 0$ and, thus, we assume that $\mathbb{P}(\tau_t < \infty) > 0$. By applying Theorem T6 in Brémaud (1981, p. 236), we obtain that, for all $M \in \mathcal{B}(\mathscr{M})$,

$$\mathbb{P}\left((E, X) \in M \,|\, \mathcal{F}^N_{\tau_t -}\right) \mathbb{1}_{\{\tau_t < \infty\}} = \frac{\int_M \lambda(\tau_t, m) \mu_{\mathscr{M}}(dm)}{\int_{\mathscr{M}} \lambda(\tau_t, m') \mu_{\mathscr{M}}(dm')} \mathbb{1}_{\{\tau_t < \infty\}}, \quad \text{a.s.}$$

This is allowed since Lemma 3.1 tells us that the filtration $\mathbb{F}^N$ is within the framework of this result. Hence, we have identified the unique regular conditional distribution of (E, X) given $\mathcal{F}^N_{\tau_t -}$ on the measurable space $(\{\tau_t < \infty\}, \{\tau_t < \infty\} \cap \mathcal{F})$ equipped with the measure $\mathbb{P}(\cdot \cap \{\tau_t < \infty\})/\mathbb{P}(\{\tau_t < \infty\})$ (Kallenberg, 2002, Theorem 6.3, p. 107). Besides, observe that the mapping $(\omega, m) \mapsto \lambda(\omega, \tau_t(\omega), m) \mathbb{1}_{\{\tau_t(\omega) < \infty\}}$ is $\mathcal{F}^N_{\tau_t -} \otimes \mathcal{B}(\mathscr{M})$-measurable, see for example Lemma 25.3 in Kallenberg (2002, p. 492). Using Lemma 1.26 in Kallenberg (2002, p. 14), we obtain that the function f defined by

$$f(\omega, e, x) = \frac{\lambda(\omega, \tau_t(\omega), e, x)}{\int_{\mathscr{M}} \lambda(\omega, \tau_t(\omega), m') \mu_{\mathscr{M}}(dm')} \mathbb{1}_{\{\tau_t(\omega) < \infty\}}, \quad \omega \in \Omega, e \in \mathscr{E}, x \in \mathscr{X},$$

is $\mathcal{F}^N_{\tau_t -} \otimes \mathcal{B}(\mathscr{M})$-measurable. We can then apply Lemma 3.6 with $\mathcal{G} = \mathcal{F}^N_{\tau_t -}$ and $A = \{\tau_t < \infty\}$. This yields that

$$\mathbb{P}\left(X \in dx \,|\, \sigma(E) \vee \mathcal{F}^N_{\tau_t -}\right) \mathbb{1}_{\{\tau_t < \infty\}} = \frac{\lambda(\tau_t, E, x)}{\int_{\mathscr{X}} \lambda(\tau_t, E, x') \mu_{\mathscr{X}}(dx')} \mu_{\mathscr{X}}(dx) \mathbb{1}_{\{\tau_t < \infty\}}, \quad \text{a.s.} \tag{3.5}$$

By viewing the term $\phi(x \,|\, e, X_t)$ as a measurable function φ applied to $(e, x, \theta_t N^{<0})$ where $\varphi(x, e \,|\, \xi) = \phi(x \,|\, e, F(\xi))$, and using the measurability of the state functional F and the transition function ϕ, we obtain by Lemma 3.4 that $\phi(x \,|\, e, \theta_t N^{<0})$, $t \in \mathbb{R}$, $e \in \mathscr{E}$, $m \in \mathscr{M}$, is $\mathbb{F}^N$-predictable. Similarly, note that $\eta(e \,|\, \theta_t N^{<0})$, $t \in \mathbb{R}$, $e \in \mathscr{E}$, is also $\mathbb{F}^N$-predictable (this will be useful when proving sufficiency). Besides, thanks to the assumption on λ,

$$\frac{\lambda(\omega, t, e, x)}{\int_{\mathscr{X}} \lambda(\omega, t, e, x') \mu_{\mathscr{X}}(dx')} = \phi(x \,|\, e, X_t(\omega)), \quad \mathbb{P}(d\omega) dt \mu_{\mathscr{M}}(de, dx)\text{-a.e.}$$

Hence, using Lemma 3.5, (3.5) becomes

$$\mathbb{P}\left(X \in dx \,|\, \sigma(E) \vee \mathcal{F}^N_{\tau_t -}\right) \mathbb{1}_{\{\tau_t < \infty\}} = \phi(x \,|\, E, X_{\tau_t}) \mu_{\mathscr{X}}(dx) \mathbb{1}_{\{\tau_t < \infty\}}, \quad \text{a.s.}$$

To obtain (2.6), it remains to notice that $X_{\tau_t} = X_{t+}$ on $\{\tau_t < \infty\}$ since there is no event on the time interval (t, τ_t) by definition of τ_t. Also, since the ground point process $N(\cdot \times \mathscr{M})$ admits an $\mathbb{F}^N$-intensity, we have that $N(\{t\} \times \mathscr{M}) = 0$ a.s., implying that $X_{t+} = X_t$ a.s. To show the statement in Remark 2.14, simply use (2.6) and the tower property to obtain that

$$\mathbb{E}\left[\mathbb{1}_F \mathbb{1}_{\{\tau_t < \infty\}} \mathbb{1}_{X \in B}\right] = \mathbb{E}\left[\mathbb{1}_F \mathbb{1}_{\{\tau_t < \infty\}} \mathbb{E}\left[\mathbb{1}_{X \in B} \,|\, \sigma(E) \vee \mathcal{F}^N_{\tau_t -}\right]\right] = \mathbb{E}\left[\mathbb{1}_F \mathbb{1}_{\{\tau_t < \infty\}} \int_B \phi(x \,|\, E, X_t) \mu_{\mathscr{X}}(dx)\right]$$

for all $F \in \sigma(E) \vee \mathcal{F}^N_t$, $B \in \mathcal{B}(\mathscr{X})$ and observe that $\int_B \phi(x \,|\, E, X_t) \mu_{\mathscr{X}}(dx)$ is $\sigma(E) \vee \mathcal{F}^N_t$-measurable.

Sufficiency. Assume that N is a non-explosive marked point process on $\mathbb{R} \times \mathscr{M}$ such that it admits an $\mathbb{F}^N$-intensity relative to $\mu_{\mathscr{M}}$ and such that statements (i) and (ii) hold. We want to show that $\lambda(\omega, t, e, x) =$

$\phi(x \mid e, X_t(\omega))\eta(e \mid \theta_t N(\omega)^{<0})$ holds $\mathbb{P}(d\omega)\mu_{\mathscr{M}}(de, dx)dt$-a.e. For all $t \in \mathbb{R}_{\geq 0}$, by using statement (ii), (3.5), Lemmas 3.3 and 3.5, and statement (i), we obtain that

$$
\begin{aligned}
\phi(x \mid E, X_{\tau_t})\mu_{\mathscr{X}}(dx)\mathbb{1}_{\{\tau_t < \infty\}} &= \mathbb{P}\left(X \in dx \mid \sigma(E) \vee \mathcal{F}_{\tau_t -}^N\right)\mathbb{1}_{\{\tau_t < \infty\}} \\
&= \frac{\lambda(\tau_t, E, x)}{\int_{\mathscr{X}} \lambda(\tau_t, E, x')\mu_{\mathscr{X}}(dx')}\mu_{\mathscr{X}}(dx)\mathbb{1}_{\{\tau_t < \infty\}} \\
&= \frac{\lambda(\tau_t, E, x)}{\lambda_{\mathscr{E}}(\tau_t, E)}\mu_{\mathscr{X}}(dx)\mathbb{1}_{\{\tau_t < \infty\}} \\
&= \frac{\lambda(\tau_t, E, x)}{\eta(E \mid \theta_{\tau_t} N^{<0})}\mu_{\mathscr{X}}(dx)\mathbb{1}_{\{\tau_t < \infty\}}, \quad \text{a.s.}
\end{aligned}
$$

This means that, for all $t \in \mathbb{R}_{\geq 0}$, we have that

$$
\lambda(\tau_t, E, x)\mathbb{1}_{\{\tau_t < \infty\}} = \phi(x \mid E, X_{\tau_t})\eta(E \mid \theta_{\tau_t} N^{<0})\mathbb{1}_{\{\tau_t < \infty\}}, \quad \mu_{\mathscr{X}}(dx)\text{-a.e., a.s.}
$$

This holds a.s. simultaneously for all $t \in \mathbb{Q} \cap \mathbb{R}_{\geq 0}$, whence, using that the number of events in N is countable and finite in any bounded time interval,

$$
\lambda(\omega, t, e, x) = \phi(x \mid e, X_t(\omega))\eta(e \mid \theta_t N(\omega)^{<0}), \quad \mathbb{P}(d\omega)N_{\mathscr{E}}(\omega, dt, de)\mu_{\mathscr{X}}(dx)\text{-a.e.}
$$

By Lemma 3.5, the above equality then implies that

$$
\lambda(\omega, t, e, x) = \phi(x \mid e, X_t(\omega))\eta(e \mid \theta_t N(\omega)^{<0}), \quad \mathbb{P}(d\omega)\lambda_{\mathscr{E}}(\omega, t, e)dt\mu_{\mathscr{M}}(de, dx)\text{-a.e.}
$$

By noticing that, on $\lambda_{\mathscr{E}}(\omega, t, e) = 0$, we have that $\eta(e \mid \theta_t N(\omega)^{<0}) = 0$ holds $\mathbb{P}(d\omega)dt\mu_{\mathscr{E}}(de)$-a.e. and that $\lambda(\omega, t, e, x) = 0$ holds $\mathbb{P}(d\omega)dt\mu_{\mathscr{E}}(de)\mu_{\mathscr{X}}(dx)$-a.e. (using again Lemma 3.3), we conclude that the above equation actually holds $\mathbb{P}(d\omega)dt\mu_{\mathscr{M}}(de, dx)$-a.e. $\qquad\square$

4 Existence and uniqueness

In this section, we prove the strong existence result (Theorem 2.17) by means of a Poisson embedding lemma given below. Subsequently, we also prove the strong and weak uniqueness results (Theorems 2.20 and 2.21).

4.1 Preliminaries

4.1.1 Example violating the Lipschitz condition. We give here an example of a hybrid marked point process that does not satisfy the Lipschitz condition (2.9), implying that the existence and uniqueness results in Massoulié (1998) do not apply.

Example 4.1. Set $\mathscr{E} = \{0, 1\}$ and $\mathscr{X} = \{0, 1\}$ with $\mu_{\mathscr{E}} = \delta_0 + \delta_1$ and $\mu_{\mathscr{X}} = \delta_0 + \delta_1$. Consider an intensity functional ψ that corresponds to a hybrid marked point process with transition function ϕ and event functional η (see Definition 2.9). Take η to be a Hawkes functional of the form

$$
\eta(e \mid \xi) = \nu + \iint_{\mathbb{R}_{<0} \times \mathscr{M}} k(-t', x', e)\xi(dt', de', dx'),
$$

19

where $\nu \in \mathbb{R}_{>0}$ and $k : \mathbb{R}_{>0} \times \mathscr{X} \times \mathscr{E} \to \mathbb{R}_{>0}$ is continuous in time and strictly positive. Let $t_0 \in \mathbb{R}_{<0}$ and choose $\xi_0, \xi_1 \in \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}$ such that ξ_0 and ξ_1 coincide on $(-\infty, t_0]$ (i.e., $\theta_{t_0} \xi_0^{\leq 0} = \theta_{t_0} \xi_1^{\leq 0}$) but $F(\xi_0) = 0$ and $F(\xi_1) = 1$ (thus, ξ_0 and ξ_1 do not coincide on $(t_0, 0)$). Assume also that $\phi(0 \,|\, 0, 1) > \phi(0 \,|\, 0, 0)$, $\eta(0 \,|\, \xi_1) < \infty$, and $\iint_{(t_0,0) \times \mathscr{M}} k(-t', x', 0) \xi_1(dt', de', dx') > \iint_{(t_0,0) \times \mathscr{M}} k(-t', x', 0) \xi_0(dt', de', dx')$. Then, following some computations that are left to the reader,

$$|\psi(0, 0 \,|\, \xi_1) - \psi(0, 0 \,|\, \xi_0)| \geq (\phi(0 \,|\, 0, 1) - \phi(0 \,|\, 0, 0)) \iint_{(-\infty, t_0] \times \mathscr{M}} k(-t, x, 0) \xi_0(dt, dx).$$

Next, consider any non-negative kernel $\overline{k} : \mathbb{R}_{>0} \times \mathscr{M} \times \mathscr{M} \to \mathbb{R}_{\geq 0}$. We have that

$$\iint_{\mathbb{R}_{<0} \times \mathscr{M}} \overline{k}(-t, m, 0, 0) |\xi_1 - \xi_0|(dt, dm) = \iint_{(t_0,0) \times \mathscr{M}} \overline{k}(-t, m, 0, 0) |\xi_1 - \xi_0|(dt, dm).$$

We can now add as many points as necessary to ξ_0 and ξ_1 on $(-\infty, t_0]$ to guarantee that

$$|\psi(0, 0 \,|\, \xi_1) - \psi(0, 0 \,|\, \xi_0)| > \iint_{\mathbb{R}_{<0} \times \mathscr{M}} \overline{k}(-t, m, 0, 0) |\xi_1 - \xi_0|(dt, dm).$$

Consequently, the intensity functional ψ does not satisfy the Lipschitz condition (2.9).

4.1.2 Integration with respect to Poisson processes.

We first clarify briefly the link between point processes and random measures. A random measure M on a measurable space $(S, \mathcal{S})$ is a mapping $M : \Omega \times \mathcal{S} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ such that $M(\omega, \cdot)$ is a measure on $(S, \mathcal{S})$ for all $\omega \in \Omega$ and $M(\cdot, A)$ is a random variable for all $A \in \mathcal{S}$, see Kallenberg (2002, p. 106) and Çınlar (2011, Chapter 6, p. 243). Note that the concepts of internal history and adaptedness of Subsection 2.1.6 can be directly extended to random measures. Not surprisingly, point processes are exactly the boundedly finite integer-valued random measures.

Proposition 4.2. *Let N be a random measure on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$ such that $N(\omega, \cdot) \in \mathcal{N}_{\mathcal{U}}^{\#}$ for all $\omega \in \Omega$. Then N is a non-explosive point process on $\mathcal{U}$. In return, any non-explosive point process N on $\mathcal{U}$ is a random measure on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$ such that $N(\omega, \cdot) \in \mathcal{N}_{\mathcal{U}}^{\#}$ for all $\omega \in \Omega$.*

Proof. See Proposition 9.1.VIII in Daley and Vere-Jones (2008, p. 8). $\square$

One can show that Poisson processes are Poisson random measures in the sense of Çınlar (2011, Chapter 6, p. 249). This enables us to apply an important result on integration with respect to Poisson random measures. Before stating the result, we need to define what it means for a Poisson process to be Poisson relative to a filtration.

Definition 4.3. Let N be a Poisson process on $\mathbb{R} \times \mathcal{U}$ and $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ be a filtration. We say that N is Poisson *relative* to $\mathbb{F}$ if for all $t \in \mathbb{R}$, the point process $\theta_t N^{\leq 0}$ is $\mathcal{F}_t$-measurable and $\sigma(\theta_t N^{>0})$ is independent of $\mathcal{F}_t$.

Trivially, a Poisson process N is always Poisson relative to its internal history $\mathbb{F}^N$. The next result plays a crucial role in the Poisson embedding technique, which is later used to construct marked point processes with given intensities.

Theorem 4.4. *Let N be a Poisson process on $\mathbb{R} \times \mathcal{U}$ with parameter measure ν. Let $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{R}}$ be a filtration and suppose that N is Poisson relative to $\mathbb{F}$. Then, for every non-negative $\mathbb{F}$-predictable process $H : \Omega \times \mathbb{R} \times \mathcal{U} \to \mathbb{R}_{\geq 0}$, we have that*

$$\mathbb{E}\left[\iint_{\mathbb{R} \times \mathcal{U}} H(t, u) N(dt, du)\right] = \mathbb{E}\left[\iint_{\mathbb{R} \times \mathcal{U}} H(t, u) \nu(dt, du)\right].$$

Proof. See Theorem 6.2 in Çınlar (2011, Chapter 6, p. 299). $\qquad\square$

4.1.3 Driving Poisson process. We prove that the mapping $M : \Omega \to \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M} \times \mathbb{R}}$ defined by (2.10) is still a Poisson process.

Lemma 4.5. *The mapping $M : \Omega \to \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M} \times \mathbb{R}}$ is a Poisson process on $\mathbb{R} \times \mathcal{M} \times \mathbb{R}$ with parameter measure $dt\mu_{\mathcal{M}}(dm)dz$. Moreover, M is Poisson relative to $\mathbb{F}$.*

Proof. By composition, using the measurability of $M_{>0}$, it is easy to check that M is a measurable mapping and, thus, it is a non-explosive point process. To show that M is a Poisson process with parameter measure $dt\mu_{\mathcal{M}}(dm)dz$, it is enough notice that, for any $n \in \mathbb{N}$, for every family of bounded sets $(A_i)_{i \in \{1,\dots,n\}}$, for all $k_1, \dots, k_n \in \mathbb{N}$,

$$\mathbb{P}(M(A_i) = k_i,\, i = 1, \dots, n) = \mathbb{P}_{>0}(M_{>0}(A_i) = k_i,\, i = 1, \dots, n),$$

and use the fact that $M_{>0}$ is a Poisson process with parameter measure $dt\mu_{\mathcal{M}}(dm)dz$. To show that $\theta_t M^{\leq 0}$ is $\mathcal{F}_t$-measurable for any $t \in \mathbb{R}$, use the fact that $\theta_t M^{\leq 0}_{>0}$ is $\mathcal{F}^{M_{>0}}_t$-measurable (since a Poisson process is always Poisson relative to its internal history) along with a composition argument. Similarly, one can show that $\sigma(\theta_t M^{>0}) \subset \{\varnothing, \Omega_{\leq 0}\} \otimes \sigma(\theta_t M^{>0}_{>0})$ and, thus, to show that $\sigma(\theta_t M^{>0})$ is independent of $\mathcal{F}_t$, it is enough to show that $\{\varnothing, \Omega_{\leq 0}\} \otimes \sigma(\theta_t M^{>0}_{>0})$ is independent of $\mathcal{F}_t$. For this, let $A_{\leq 0} \in \{\varnothing, \Omega_{\leq 0}\}$, $A_{>0} \in \sigma(\theta_t M^{>0}_{>0})$, $B_{\leq 0} \in \mathcal{F}^{N_{\leq 0}}_t$ and $B_{>0} \in \mathcal{F}^{M_{>0}}_t$. Then, using the fact that $M_{>0}$ is Poisson relative to $\mathcal{F}^{M_{>0}}$, we have that

$$\begin{aligned}
\mathbb{P}(A_{\leq 0} \times A_{>0} \cap B_{\leq 0} \times B_{>0}) &= \mathbb{P}(A_{\leq 0} \cap B_{\leq 0} \times A_{>0} \cap B_{>0}) = \mathbb{P}_{\leq 0}(A_{\leq 0} \cap B_{\leq 0})\mathbb{P}_{>0}(A_{>0} \cap B_{>0}) \\
&= \mathbb{P}_{\leq 0}(A_{\leq 0})\mathbb{P}_{\leq 0}(B_{\leq 0})\mathbb{P}_{>0}(A_{>0})\mathbb{P}_{>0}(B_{>0}) \\
&= \mathbb{P}(A_{\leq 0} \times A_{>0})\mathbb{P}(B_{\leq 0} \times B_{>0}).
\end{aligned}$$

This shows that two π-systems generating $\{\varnothing, \Omega_{\leq 0}\} \otimes \sigma(\theta_t M^{>0}_{>0})$ and $\mathcal{F}^{N_{\leq 0}}_t \otimes \mathcal{F}^{M_{>0}}_t$, respectively, are independent. We conclude using Lemma 3.6 in Kallenberg (2002, p. 50) that $\{\varnothing, \Omega_{\leq 0}\} \otimes \sigma(\theta_t M^{>0}_{>0})$ and $\mathcal{F}^{N_{\leq 0}}_t \otimes \mathcal{F}^{M_{>0}}_t$ are independent. We can then verify that $\{\varnothing, \Omega_{\leq 0}\} \otimes \sigma(\theta_t M^{>0}_{>0})$ remains independent of the completion of $\mathcal{F}^{N_{\leq 0}}_t \otimes \mathcal{F}^{M_{>0}}_t$, which by definition is $\mathcal{F}_t$. Indeed, remember that $\mathcal{F}_t := \sigma(\mathcal{C})$ with $\mathcal{C} := (\mathcal{F}^{N_{\leq 0}}_t \otimes \mathcal{F}^{M_{>0}}_t) \cup \mathcal{A}$ and where $\mathcal{A}$ denotes the class of all subsets of $\mathbb{P}$-null sets in $\mathcal{F}$. It then suffices to notice that $\mathcal{C}$ is a π-system and that $\{\varnothing, \Omega_{\leq 0}\} \otimes \sigma(\theta_t M^{>0}_{>0})$ remains independent of $\mathcal{C}$. $\qquad\square$

4.1.4 Poisson-embedding lemma. We are now able to show the following key lemma which demonstrates how the extra-dimension of the Poisson process M allows us to generate a marked point process with a given intensity.

Lemma 4.6 (Poisson embedding). *Let $\lambda : \Omega \times \mathbb{R}_{>0} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$ be an $\mathbb{F}$-predictable process. Then, the mapping*

$$\begin{aligned}
N : \quad &\Omega \times \mathcal{B}(\mathbb{R}_{>0} \times \mathcal{M}) \to \mathbb{R}_{\geq 0} \cup \{\infty\} \\
&(\omega, A) \mapsto N(\omega, A) := \iint_A \int_{(0, \lambda(\omega, t, m)]} M(\omega, dt, dm, dz) \quad (4.1)
\end{aligned}$$

is an $\mathbb{F}$-adapted integer-valued random measure on $\mathbb{R}_{>0} \times \mathcal{M}$. Moreover, for every non-negative $\mathbb{F}$-predictable process $H : \Omega \times \mathbb{R}_{>0} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$, we have that

$$\mathbb{E}\left[\iint_{\mathbb{R}_{>0} \times \mathcal{M}} H(t, m) N(dt, dm)\right] = \mathbb{E}\left[\iint_{\mathbb{R}_{>0} \times \mathcal{M}} H(t, m)\lambda(t, m)\mu_{\mathcal{M}}(dm)dt\right].$$

Proof. First, let $A \in \mathcal{B}(\mathbb{R}_{>0} \times \mathcal{M})$ and consider the following composition

$$(\omega, t, m, z) \mapsto (\lambda(\omega, t, m), z) \mapsto \mathbb{1}_{(0, \lambda(\omega, t, m)]}(z)$$

to notice that $\mathbb{1}_{(0, \lambda(\omega, t, m)]}(z)$ is $\mathbb{F}$-predictable by means of Lemma 1.7 and Lemma 1.8 in Kallenberg (2002, p. 5). Then, the product $\mathbb{1}_A(t, m)\mathbb{1}_{(0, \lambda(\omega, t, m)]}(z)$ of two $\mathbb{F}$-predictable processes is also $\mathbb{F}$-predictable by Lemma 1.12 in Kallenberg (2002, p. 7). This ensures that the integral

$$\iint_A \int_{(0, \lambda(\omega, t, m)]} M(\omega, dt, dm, dz) = \iiint_{\mathbb{R}_{>0} \times \mathcal{M} \times \mathbb{R}} \mathbb{1}_A(t, m)\mathbb{1}_{(0, \lambda(\omega, t, m)]}(z)M(\omega, dt, dm, dz)$$

is well defined for all $\omega \in \Omega$ and that $N(\cdot, A)$ is a random variable (see Subsection 2.1.5).

Second, let $\omega \in \Omega$. For any finite family of disjoint sets $A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R}_{>0} \times \mathcal{M})$, $n \in \mathbb{N}$, we clearly have that $N(\omega, \bigcup_{i \leq n} A_i) = \sum_{i \leq n} N(\omega, A_i)$, which means that $N(\omega, \cdot)$ is finitely additive. To prove that $N(\omega, \cdot)$ is countably additive, invoke finite additivity and apply the monotone convergence theorem. These first two steps show that N is indeed a random measure.

Third, to show that N is $\mathbb{F}$-adapted, first consider processes $\lambda : \Omega \times \mathbb{R}_{>0} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$ of the form $\lambda(\omega, t, m) = \mathbb{1}_F(\omega)\mathbb{1}_{(s, u]}(t)\mathbb{1}_C(m)$ where $F \in \mathcal{F}_s$, $s, u \in \mathbb{R}_{>0}$, $s < u$, $C \in \mathcal{B}(\mathcal{M})$. For any $t \in \mathbb{R}_{>0}$, any $A \in \mathcal{B}(\mathbb{R}_{>0})$ such that $A \subset (0, t]$ and any $B \in \mathcal{B}(\mathcal{M})$, we obtain that

$$N(\omega, A \times B) = \mathbb{1}_F(\omega)M(\omega, A \cap (s, u] \times B \cap C \times (0, 1]),$$

which is $\mathcal{F}_t$-measurable since M is $\mathbb{F}$-adapted by Lemma 4.5. Hence, N is $\mathbb{F}$-adapted. To extend this result to any $\mathbb{F}$-predictable process λ, one can use a monotone class argument like in the proof of Proposition A.4.1 for example.

Fourth, let $\omega \in \Omega$. By the definition of N and by the linearity of the integral, for all simple non-negative functions f on $\mathbb{R}_{>0} \times \mathcal{M}$, we have that

$$\iint_{\mathbb{R}_{>0} \times \mathcal{M}} f(t, m)N(\omega, dt, dm) = \iiint_{\mathbb{R}_{>0} \times \mathcal{M} \times \mathbb{R}} f(t, m)\mathbb{1}_{(0, \lambda(\omega, t, m)]}(z)M(\omega, dt, dm, dz).$$

Then, by Lemma 1.11 in Kallenberg (2002, p. 7) and the monotone convergence theorem, we have that the above equality holds for any $\mathcal{B}(\mathbb{R}_{>0} \times \mathcal{M})$-measurable non-negative function f. In particular, we have that, for all $\omega \in \Omega$,

$$\iint_{\mathbb{R}_{>0} \times \mathcal{M}} H(\omega, t, m)N(\omega, dt, dm) = \iiint_{\mathbb{R}_{>0} \times \mathcal{M} \times \mathbb{R}} H(\omega, t, m)\mathbb{1}_{(0, \lambda(\omega, t, m)]}(z)M(\omega, dt, dm, dz).$$

Fifth, using Lemma 4.5 and Theorem 4.4, we deduce that

$$\mathbb{E}\left[\iint_{\mathbb{R}_{>0} \times \mathcal{M}} H(t, m)N(dt, dm)\right] = \mathbb{E}\left[\iiint_{\mathbb{R}_{>0} \times \mathcal{M} \times \mathbb{R}} H(t, m)\mathbb{1}_{(0, \lambda(t, m)]}(z)M(dt, dm, dz)\right]$$

$$= \mathbb{E}\left[\iiint_{\mathbb{R}_{>0} \times \mathcal{M} \times \mathbb{R}} H(t, m)\mathbb{1}_{(0, \lambda(t, m)]}(z)dt\mu_{\mathcal{M}}(dm)dz\right]$$

$$= \mathbb{E}\left[\iint_{\mathbb{R}_{>0} \times \mathcal{M}} H(t, m)\lambda(t, m)\mu_{\mathcal{M}}(dm)dt\right]. \qquad \square$$

Remark 4.7. Similar results are given by Brémaud and Massoulié (1996, Lemma 3, p. 1571), Massoulié (1998, Lemma 1, p. 3) and Torrisi (2016, Lemma 2.1, p. 4). They refer to Lewis and Shedler (1976)

and Ogata (1981) for proofs. The fifth part of our proof follows Daley and Vere-Jones (2008, Proposition 14.7.I, p. 427), but we could not find the first four parts anywhere. For the special case $\mathscr{M} = \{0\}$ (i.e., for univariate point processes), a similar proof is given by Çınlar (2011, Theorem 6.11, p. 303) while an alternative proof is given by Chevallier et al. (2015, Theorem B.11). Besides, our version of this lemma does not impose any local integrability condition on λ and, thus, does not say if the obtained random measure N is boundedly finite. Finally, note that (4.1) can be rewritten using the compact notation of Massoulié (1998) as $N(dt, dm) = M(dt, dm, [0, \lambda(t, m)])$, $t \in \mathbb{R}_{>0}$.

We can now prove the final statement in Thoerem 2.17, which we restate here as a corollary.

Corollary 4.8. *Let* $N : \Omega \to \mathcal{N}^{\#g}_{\mathbb{R} \times \mathscr{M}}$ *be a solution to the Poisson-driven SDE (Definition 2.16) under either Assumptions A, B, C, or Assumptions A, D, E. Then, N admits ψ as its intensity functional on $\mathbb{R}_{>0}$.*

Proof. Let $G \in \mathcal{F}$ be the almost sure event that (2.11) holds. Consider the following modifications of N and λ, where λ is defined as in (2.11):

$$\tilde{N}(\omega) := N(\omega)\mathbb{1}_G(\omega), \quad \omega \in \Omega, \quad \text{and} \quad \tilde{\lambda}(\omega, t, m) := \lambda(\omega, t, m)\mathbb{1}_G(\omega), \quad \omega \in \Omega, t \in \mathbb{R}_{>0}, m \in \mathscr{M}.$$

Then, $\tilde{N}$ and $\tilde{\lambda}$ satisfy (4.1) and, using either Assumptions B.(i) and C or Assumptions D.(i) and E.(ii), one can check that $\tilde{\lambda}(\omega, t, m) < \infty$ for all $\omega \in \Omega$, $t \in \mathbb{R}_{>0}$, $m \in \mathscr{M}$. Moreover, by Lemma 3.4, λ is $\mathbb{F}^N$-predictable, and, thus, $\mathbb{F}$-predictable as N is $\mathbb{F}$-adapted. Since the filtration $\mathbb{F}$ is complete, this implies that $\tilde{\lambda}$ is also $\mathbb{F}$-predictable. Now, consider any non-negative $\mathbb{F}^N$-predictable process $H : \Omega \times \mathbb{R}_{>0} \times \mathscr{M} \to \mathbb{R}_{\geq 0}$ and apply Lemma 4.6 to obtain

$$\mathbb{E}\left[\iint_{\mathbb{R}_{>0} \times \mathscr{M}} H(t, m)N(dt, dm)\right] = \mathbb{E}\left[\iint_{\mathbb{R}_{>0} \times \mathscr{M}} H(t, m)\tilde{N}(dt, dm)\right]$$
$$= \mathbb{E}\left[\iint_{\mathbb{R}_{>0} \times \mathscr{M}} H(t, m)\tilde{\lambda}(t, m)\mu_{\mathscr{M}}(dm)dt\right]$$
$$= \mathbb{E}\left[\iint_{\mathbb{R}_{>0} \times \mathscr{M}} H(t, m)\lambda(t, m)\mu_{\mathscr{M}}(dm)dt\right].$$

We conclude that N admits ψ as its intensity functional using Lemma 3.2. □

Given a non-explosive point process N on $\mathbb{R} \times \mathscr{M}$ that solves (2.11) or is defined through a Poisson embedding as in Lemma 4.6, one can ask when N is in fact a non-explosive marked point process. To this end, it is useful to define the following random measures induced by the driving Poisson process M:

$$L_n(\omega, \cdot) := M(\omega, \cdot \times \mathscr{M} \times (0, n]), \quad \omega \in \Omega, n \in \mathbb{N}.$$

We are then able to find the following sufficient condition on λ.

Lemma 4.9 (Simple ground measure). *Let* $\lambda : \Omega \times \mathbb{R}_{>0} \times \mathscr{M} \to \mathbb{R}_{\geq 0}$ *be an $\mathbb{F}$-predictable process and let N be the $\mathbb{F}$-adapted integer-valued random measure on $\mathbb{R}_{>0} \times \mathscr{M}$ defined by (4.1). Then, if Assumption A holds and if $\sup_{m \in \mathscr{M}} \lambda(t, m) < \infty$ for all $t \in \mathbb{R}_{>0}$, a.s., we have that $N(\{t\} \times \mathscr{M}) \leq 1$ for all $t \in \mathbb{R}_{>0}$, a.s.*

Proof. Each L_n is a Poisson random measure on $\mathbb{R}$ in the sense of Çınlar (2011, Chapter 6, p. 249) with boundedly finite parameter measure $n\mu_{\mathscr{M}}(\mathscr{M})dt$. Applying Theorem 2.17 in Çınlar (2011, Chapter 6, p. 256) for each $n \in \mathbb{N}$, there exists a set $B \in \mathcal{F}$ such that $\mathbb{P}(B) = 1$ and such that, for all $\omega \in B$ and $n \in \mathbb{N}$,

$L_n(\omega) \in \mathcal{N}_{\mathbb{R}}^{\#g}$ (i.e., L_n, $n \in \mathbb{N}$, are simultaneously simple). Next, let A be the almost sure event that $\sup_{m \in \mathcal{M}} \lambda(t, m) < \infty$ for all $t \in \mathbb{R}_{>0}$. Fix $\omega \in A \cap B$ and use the assumption on λ to find that

$$N(\omega, \{t\} \times \mathcal{M}) = \int_{\{t\}} \int_{\mathcal{M}} \int_{(0, \lambda(\omega, s, m)]} M(\omega, ds, dm, dz)$$

$$\leq M\left(\omega, \{t\} \times \mathcal{M} \times \left(0, \sup_{m \in \mathcal{M}} \lambda(\omega, t, m)\right]\right)$$

$$\leq M\left(\omega, \{t\} \times \mathcal{M} \times (0, p(\omega, t)]\right) = L_{p(\omega, t)}(\omega, \{t\}) \leq 1,$$

where $p(\omega, t) \in \mathbb{N}$ is such that $\sup_{m \in \mathcal{M}} \lambda(\omega, t, m) \leq p(\omega, t)$. $\qquad\square$

4.2 Strong existence: pathwise construction via Poisson embedding

4.2.1 Existence under Assumptions A, B, C. We begin by proposing a construction of a candidate solution $N : \Omega \to \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\infty}$ to (2.11). We proceed in a pathwise fashion. Under Assumption A, using the definition of a Poisson process, it is not difficult to see that, given $n \in \mathbb{N}$, $L_n \in \mathcal{N}_{\mathbb{R}}^{\#}$ a.s. This implies that

$$F_1 := \{\omega \in \Omega \mid L_n(\omega) \in \mathcal{N}_{\mathbb{R}}^{\#}, n \in \mathbb{N}\} \in \mathcal{F}$$

is an almost sure event, which plays a key role in our pathwise construction.

Algorithm 4.10. Construct the mapping $N : \Omega \to \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\infty}$ as follows. For all $\omega = (\omega_{\leq 0}, \omega_{>0}) \in F_1$, initialise $N_0(\omega) := N_{\leq 0}(\omega_{\leq 0})$, $T_0(\omega) := 0$, $\mathbb{M}_0(\omega) := \varnothing$, and $\lambda_0(\omega, t, m) := \psi(m \mid \theta_t N_0(\omega)^{<0})$ for all $t \in \mathbb{R}_{>0}$, $m \in \mathcal{M}$. Define recursively the sequences $(N_n)_{n \in \mathbb{N}}$, $(T_n)_{n \in \mathbb{N}}$, $(\mathbb{M}_n)_{n \in \mathbb{N}}$, and $(\lambda_n)_{n \in \mathbb{N}}$ as follows. For all $n \in \mathbb{N}$,

- if $T_n(\omega) < \infty$, then

$$T_{n+1}(\omega) := \sup\left\{u > T_n(\omega) : \iint_{(T_n(\omega), u) \times \mathcal{M}} \int_{(0, \lambda_n(\omega, t, m)]} M(\omega, dt, dm, dz) = 0\right\}; \qquad (4.2)$$

  – if $T_{n+1}(\omega) < \infty$, then

$$\mathbb{M}_{n+1}(\omega) := \{m \in \mathcal{M} : M(\omega, \{T_{n+1}(\omega)\} \times \{m\} \times (0, \lambda_n(\omega, T_{n+1}(\omega), m)]) > 0\};$$

$$N_{n+1}(\omega) := \sum_{i=1}^{n+1} \sum_{m \in \mathbb{M}_i(\omega)} M(\omega, \{T_i(\omega)\} \times \{m\} \times (0, \lambda_{i-1}(\omega, T_i(\omega), m)]) \delta_{(T_i(\omega), m)}; \qquad (4.3)$$

$$\lambda_{n+1}(\omega, t, m) := \psi(m \mid \theta_t N_{n+1}(\omega)^{<0}), \quad t \in \mathbb{R}_{>0}, m \in \mathcal{M}; \qquad (4.4)$$

  – if $T_{n+1}(\omega) = \infty$, then

$$\mathbb{M}_{n+1}(\omega) := \varnothing;$$

$$N_{n+1}(\omega) := N_n(\omega);$$

$$\lambda_{n+1}(\omega, t, m) := \lambda_n(\omega, t, m), \quad t \in \mathbb{R}_{>0}, m \in \mathcal{M};$$

- if $T_n(\omega) = \infty$, then

$$T_{n+1}(\omega) := \infty;$$

$$\mathbb{M}_{n+1}(\omega) := \varnothing;$$

$$N_{n+1}(\omega) := N_n(\omega);$$

$$\lambda_{n+1}(\omega, t, m) := \lambda_n(\omega, t, m), \quad t \in \mathbb{R}_{>0}, m \in \mathcal{M}.$$

For all $\omega = (\omega_{\leq 0}, \omega_{>0}) \in \Omega \setminus F_1$, set $N_n(\omega) := N_{\leq 0}(\omega_{\leq 0})$, $T_n(\omega) := \infty$, $\mathbb{M}_n(\omega) := \varnothing$, $\lambda_n(\omega, t, m) := 0$, $t \in \mathbb{R}_{>0}$, $m \in \mathscr{M}$, for all $n \in \mathbb{N}$. Then, for all $\omega \in \Omega$, for all $n \in \mathbb{N}$, define $N(\omega)$ on $(-\infty, T_{n+1}(\omega))$ by $\theta_{T_{n+1}(\omega)} N(\omega)^{<0} := \theta_{T_{n+1}(\omega)} N_n(\omega)^{<0}$. Define also the explosion time $T_\infty(\omega) := \lim_{n \to \infty} T_n(\omega)$. If $T_\infty(\omega) < \infty$, extend $N(\omega)$ to $[T_\infty(\omega), \infty)$ by $\theta_{T_\infty(\omega)} N(\omega)^{\geq 0} := 0$. This is equivalent to defining $N(\omega)$ as

$$N(\omega) := \lim_{n \to \infty} N_n(\omega) = \sum_{n=1}^{\infty} \sum_{m \in \mathbb{M}_n(\omega)} M(\omega, \{T_n(\omega)\} \times \{m\} \times (0, \lambda_{n-1}(\omega, T_n(\omega), m)]) \delta_{(T_n(\omega), m)} \mathbb{1}_{\{T_n(\omega) < \infty\}}.$$

Algorithm 4.10 would be ill-defined if the set in (4.2) were empty. This would mean that there are infinitely many events just after the time T_n. The following proposition shows that this actually never happens and, thus, ensures that Algorithm 4.10 is well-defined. We also need to prove that the set $\mathbb{M}_n$ is finite and that $N_n(\omega) \in \mathcal{N}^{\#}_{\mathbb{R} \times \mathscr{M}}$ for all $n \in \mathbb{N}$, because otherwise $\lambda_n(\omega, t, m)$ might be ill-defined (ψ is a functional on $\mathscr{M} \times \mathcal{N}^{\#}_{\mathbb{R} \times \mathscr{M}}$).

Proposition 4.11. *In Algorithm 4.10, under Assumptions A, B.(i) and C, we have that, for every $\omega \in F_1$,* $\mathrm{card}(\mathbb{M}_n(\omega)) < \infty$, $N_n(\omega) \in \mathcal{N}^{\#}_{\mathbb{R} \times \mathscr{M}}$, $\|\lambda_i\|(\omega) := \sup_{t>0, m \in \mathscr{M}} \lambda_i(\omega, t, m) < \infty$, *for all $n \in \mathbb{N}$, and*

$$\left\{ u > T_n(\omega) : \iint_{(T_n(\omega), u) \times \mathscr{M}} \int_{(0, \lambda_n(\omega, t, m)]} M(\omega, dt, dm, dz) = 0 \right\} \neq \varnothing \quad \text{for all } n \in \mathbb{N} \text{ s.t. } T_n(\omega) < \infty.$$

Hence, Algorithm 4.10 is well-defined under these assumptions.

Proof. We show the desired result by induction. Take any $\omega = (\omega_{\leq 0}, \omega_{>0}) \in F_1$. Let $n \in \mathbb{N}$ and for all $i \in \mathbb{N}$ such that $i < n$ and $T_i(\omega) < \infty$, assume that

$$\left\{ u > T_i(\omega) : \iint_{(T_i(\omega), u) \times \mathscr{M}} \int_{(0, \lambda_i(\omega, t, m)]} M(\omega, dt, dm, dz) = 0 \right\} \neq \varnothing. \tag{4.5}$$

For all $i \in \mathbb{N}$ such that $i \leq n$, assume that $N_i(\omega) \in \mathcal{N}^{\#}_{\mathbb{R} \times \mathscr{M}}$ and that $\|\lambda_i\|(\omega) := \sup_{t>0, m \in \mathscr{M}} \lambda_i(\omega, t, m) < \infty$. If $T_n(\omega) = \infty$, then, by construction, this is also true for $n + 1$.

Now, assume that $T_n(\omega) < \infty$. We first show that (4.5) holds also for $i = n$. Take any $\varepsilon > 0$. We have that

$$\iint_{(T_n(\omega), T_n(\omega)+\varepsilon) \times \mathscr{M}} \int_{(0, \lambda_n(\omega, t, m)]} M(\omega, dt, dm, dz) \leq M(\omega, (T_n(\omega), T_n(\omega) + \varepsilon) \times \mathscr{M} \times (0, \|\lambda_n\|(\omega)])$$

$$\leq L_{p_n(\omega)}(\omega, (T_n(\omega), T_n(\omega) + \varepsilon)) =: U_n(\omega, \varepsilon) < \infty,$$

where $p_n(\omega) \in \mathbb{N}$ is such that $\|\lambda_n\|(\omega) \leq p_n(\omega)$ and we used the fact that $L_{p_n(\omega)}(\omega) \in \mathcal{N}^{\#}_{\mathbb{R}}$. If $U_n(\omega, \varepsilon) = 0$, then clearly (4.5) is satisfied for $i = n$. If not, $M(\omega)$ has a finite number of points in $(T_n(\omega), T_n(\omega) + \varepsilon) \times \mathscr{M} \times (0, \|\lambda_n\|(\omega)]$ and there exists $0 < \varepsilon' < \varepsilon$ such that $U_n(\omega, \varepsilon') = 0$, in which case (4.5) is again satisfied for $i = n$. Note that the integral in (4.5) is well-defined since $\lambda_i(\omega, \cdot, \cdot)$ is a measurable function on $\mathbb{R}_{>0} \times \mathscr{M}$ for all $\omega \in \Omega$. To see this, consider the composition $(t, m) \mapsto (m, \theta_t N_i(\omega)^{<0}) \mapsto \psi(m \,|\, \theta_t N_i(\omega)^{<0})$ and use Lemma A.2.3, the measurability of ψ and Lemma 1.8 in Kallenberg (2002).

Second, we show that $\mathrm{card}(\mathbb{M}_{n+1}) < \infty$ and $N_{n+1}(\omega) \in \mathcal{N}^{\#}_{\mathbb{R} \times \mathscr{M}}$. If $T_{n+1}(\omega) = \infty$, then this is immediate. If not, using again that $L_{p_n(\omega)}(\omega) \in \mathcal{N}^{\#}_{\mathbb{R}}$,

$$\sum_{m \in \mathbb{M}_{n+1}} M(\omega, \{T_{n+1}(\omega)\} \times \{m\} \times (0, \lambda_n(\omega, T_{n+1}(\omega), m)]) \leq M(\omega, \{T_{n+1}(\omega)\} \times \mathscr{M} \times (0, \|\lambda_n\|(\omega)])$$

$$\leq L_{p_n(\omega)}(\omega, \{T_{n+1}(\omega)\}) < \infty,$$

which implies that the set $\mathbb{M}_{n+1}(\omega)$ is finite and, in view of (4.3), that $N_{n+1}(\omega) \in \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#}$. Note that this also proves that $N_{n+1}(\omega, (0, T_{n+1}(\omega)) \times \mathscr{M}) < \infty$.

Third, we show that $\|\lambda_{n+1}\|(\omega) < \infty$. If $T_{n+1}(\omega) = \infty$, then this is immediate. If not, by (4.4) and using Assumptions B.(i) and (4.3), we have that for all $t > 0$, $m = (x, e) \in \mathscr{M}$,

$$\begin{aligned}
\lambda_{n+1}(\omega, t, m) &\leq a\left(N_{n+1}(\omega, (-\infty, t) \times \mathscr{M})\right) \\
&= a\left(N_{n+1}(\omega, (-\infty, 0] \times \mathscr{M}) + N_{n+1}(\omega, (0, t) \times \mathscr{M})\right) \\
&\leq a\left(N_{\leq 0}(\omega_{\leq 0}, (-\infty, 0] \times \mathscr{M}) + N_{n+1}(\omega, (0, T_{n+1}(\omega)) \times \mathscr{M})\right).
\end{aligned} \tag{4.6}$$

Since $N_{n+1}(\omega, (0, T_{n+1}(\omega)) \times \mathscr{M}) < \infty$ and, by Assumption C, $N_{\leq 0}(\omega_{\leq 0}, (-\infty, 0] \times \mathscr{M}) < \infty$, this implies that $\|\lambda_{n+1}\|(\omega) < \infty$.

Regarding the basis of this induction, it is immediate that $N_0(\omega) = N_{\leq 0}(\omega_{\leq 0}) \in \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#}$. To see that $\|\lambda_0\|(\omega) < \infty$, simply set $n = -1$ in (4.6). $\qquad\square$

We show that the constructed mapping $N : \Omega \to \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\infty}$ satisfies indeed (2.11) up to each event time.

Proposition 4.12. *Under Assumptions A, B.(i) and C, the mapping $N : \Omega \to \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\infty}$ given by Algorithm 4.10 is such that $N(\omega)$ solves (2.11) on $(-\infty, T_n(\omega))$ for all $n \in \mathbb{N}$, for all $\omega \in F_1$.*

Proof. Define the process $\lambda(\omega, t, m) := \psi(m \,|\, \theta_t N(\omega)^{<0})$, $\omega \in \Omega, t \in (0, T_\infty(\omega))$, $m \in \mathscr{M}$. Take any $\omega = (\omega_{\leq 0}, \omega_{>0}) \in F_1$. By construction, $N(\omega)^{\leq 0} = N_0(\omega) = N_{\leq 0}(\omega_{\leq 0})$ and, thus, N satisfies the strong initial condition $N_{\leq 0}$. Take any $n \in \mathbb{N}$ such that $T_{n+1}(\omega) < \infty$ in Algorithm 4.10 and consider the time interval $(T_n(\omega), T_{n+1}(\omega)]$. By construction, we have that

$$N_{n+1}(\omega, dt, dm) = M(\omega, dt, dm, (0, \lambda_n(\omega, t, m)]) \quad \text{for all } t \in (T_n(\omega), T_{n+1}(\omega)]. \tag{4.7}$$

But, by definition, on $(-\infty, T_{n+1}(\omega)]$, $N(\omega) = N_{n+1}(\omega)$ and thus, for all $t \in (0, T_{n+1}(\omega)]$, $m \in \mathscr{M}$,

$$\lambda(\omega, t, m) = \psi(m \,|\, \theta_t N(\omega)^{<0}) = \psi(m \,|\, \theta_t N_{n+1}(\omega)^{<0}) = \psi(m \,|\, \theta_t N_n(\omega)^{<0}) = \lambda_n(\omega, t, m),$$

by the definition (4.4) of λ_n, since $N_{n+1}(\omega)$ and $N_n(\omega)$ can only differ by a mass at time $T_{n+1}(\omega)$. Consequently, (4.7) can be rewritten on $(T_n(\omega), T_{n+1}(\omega)]$ as

$$N(\omega, dt, dm) = M(\omega, dt, dm, (0, \lambda(\omega, t, m)]). \tag{4.8}$$

This shows that the constructed $N(\omega)$ solves (2.11) on $(-\infty, T_n(\omega)]$ for all $n \in \mathbb{N}$ such that $T_n(\omega) < \infty$. Now, if there is $n \in \mathbb{N}$ such that $T_n(\omega) < \infty$ and $T_{n+1}(\omega) = \infty$, then clearly the constructed $N(\omega)$ is null on $(T_n(\omega), \infty)$ and by similar arguments, (4.8) holds on $(T_n(\omega), \infty)$. This now allows us to conclude that $N(\omega)$ solves (2.11) on $(-\infty, T_n(\omega))$ for all $n \in \mathbb{N}$ in both cases $T_n(\omega) < \infty$ and $T_n(\omega) = \infty$. $\qquad\square$

It will also be crucial for the strong existence proof to show that, for all $n \in \mathbb{N}$, N_n is adapted to the filtration $\mathbb{F}$ and λ_n is $\mathbb{F}$-predictable.

Proposition 4.13. *In Algorithm 4.10, for all $n \in \mathbb{N}$, λ_n is $\mathbb{F}$-predictable, N_n is an $\mathbb{F}$-adapted non-explosive point process and T_n is an $\mathbb{F}$-stopping time.*

Proof. We proceed by induction. Regarding the basis, as the filtration $\mathbb{F}$ is complete, clearly N_0 is $\mathbb{F}$-adapted and T_0 is an $\mathbb{F}$-stopping time. Now assume that N_n is $\mathbb{F}$-adapted and T_n is an $\mathbb{F}$-stopping time for some $n \in \mathbb{N}$. First, observe that this implies that λ_n is $\mathbb{F}$-predictable by simply using the identity $\lambda_n(\omega, t, m) = \psi(m \,|\, \theta_t N_n(\omega)^{<0}) \mathbb{1}_{F_1}(\omega)$ and invoking Lemma 3.4, the fact that $\mathcal{F}_t^{N_n} \subset \mathcal{F}_t$, $t \in \mathbb{R}$, and the assumption that $\mathbb{F}$ is complete. Second, let $t \in \mathbb{R}$ and notice that

$$\{T_{n+1} \leq t\} = \left\{ \iiint_{\mathbb{R} \times \mathscr{M} \times \mathbb{R}} \mathbb{1}_{(T_n, t]}(s) \mathbb{1}_{(0, \lambda_n(s,m)]}(z) M(ds, dm, dz) > 0 \right\} \cap \{T_n < \infty\} \cap F_1. \tag{4.9}$$

Because T_n is an $\mathbb{F}$-stopping time, we have that $(\mathbb{1}_{(T_n, t]}(s))_{s \in \mathbb{R}}$ is $\mathbb{F}$-adapted and left-continuous, implying that it is $\mathbb{F}$-predictable, see for example Lemma 25.1 in Kallenberg (2002, p. 491). Adapting the arguments of the third part of the proof of Lemma 4.6, we deduce that the first event on the right-hand side of (4.9) belongs to $\mathcal{F}_t$ and so T_{n+1} is an $\mathbb{F}$-stopping time. Third, using Proposition 4.12 and looking at Algorithm 4.10, notice that N_{n+1} satisfies

$$\begin{cases} N_{n+1}(\omega, dt, dm) = M(\omega, dt, dm, (0, \lambda_n(\omega, t, m) \mathbb{1}_{\{t \leq T_{n+1}(\omega)\}}]), & \omega \in \Omega, t \in \mathbb{R}_{>0}, \\ N^{\leq 0}(\omega) = N_{\leq 0}(\omega_{\leq 0}), & \omega = (\omega_{\leq 0}, \omega_{>0}) \in \Omega, \end{cases}$$

where $\lambda_n(t, m) \mathbb{1}_{\{t \leq T_{n+1}\}}$ is $\mathbb{F}$-predictable as a product of $\mathbb{F}$-predictable processes, note that $\mathbb{1}_{\{t \leq T_{n+1}\}}$ is $\mathbb{F}$-adapted and left-continuous since T_{n+1} is an $\mathbb{F}$-stopping time. Now, applying Lemma 4.6, it follows that N_{n+1} is indeed $\mathbb{F}$-adapted. $\square$

We are now in a position to prove Theorem 2.17 under Assumption A, B and C for the following intensity functional:

$$\psi' : \mathscr{M} \times \mathcal{N}^{\#}_{\mathbb{R} \times \mathscr{M}} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$$
$$(m, \xi) \mapsto \psi'(m \,|\, \xi) := a\left(\xi((-\infty, 0) \times \mathscr{M}) \right).$$

Still, note that the first step of the following proof remains true for general intensity functionals ψ that satisfy Assumption B and will be reused in other parts of the proof of Theorem 2.17.

Proof of Theorem 2.17, Part 1. Let $N : \Omega \to \mathcal{N}^{\infty}_{\mathbb{R} \times \mathscr{M}}$ be given by Algorithm 4.10 under Assumptions A, B and C, which is well-defined by Proposition 4.11, and consider here the special case $\psi = \psi'$. We will prove that N admits a version that solves the Poisson-driven SDE. We proceed in four steps.

First, notice that for all $\omega \in \Omega$, $t < T_\infty(\omega)$, there exists $n \in \mathbb{N}$ such that $\theta_t N(\omega)^{<0} = \theta_t N_n(\omega)^{<0}$, which implies by Proposition 4.15 that the process

$$\lambda(\omega, t, m) := \psi(m \,|\, \theta_t N(\omega)^{<0}) \mathbb{1}_{F_1}(\omega) \mathbb{1}_{\{t < T_\infty(\omega)\}}, \quad \omega \in \Omega, t \in \mathbb{R}_{>0}, m \in \mathscr{M},$$

is well-defined and finite, and that, for all $\omega \in \Omega$, $t \in \mathbb{R}_{>0}$, $m \in \mathscr{M}$,

$$\lambda(\omega, t, m) = \lim_{n \to \infty} \psi(m \,|\, \theta_t N_n(\omega)^{<0}) \mathbb{1}_{F_1}(\omega) \mathbb{1}_{\{t < T_\infty(\omega)\}} = \lim_{n \to \infty} \lambda_n(\omega, t, m) \mathbb{1}_{\{t < T_\infty(\omega)\}}.$$

By Proposition 4.12, and because of the way we constructed N, we have that N and λ satisfy (4.1). By Proposition 4.13, for all $n \in \mathbb{N}$, λ_n is $\mathbb{F}$-predictable and T_n is an $\mathbb{F}$-stopping time. Since $T_\infty = \lim_{n \to \infty} T_n$, we have that T_∞ is an $\mathbb{F}$-predictable time, which implies by Lemma 25.3.(ii) in Kallenberg (2002, p. 492) that $\mathbb{1}_{\{t < T_\infty\}}$ is $\mathbb{F}$-predictable. As λ is a limit of $\mathbb{F}$-predictable processes, we have that λ is also $\mathbb{F}$-predictable

by Lemma 1.9 in Kallenberg (2002, p. 6). Consequently, we can apply Lemma 4.6 to obtain that N is an $\mathbb{F}$-adapted integer-valued random measure. The main goal of the next steps is to show that $T_\infty = \infty$ a.s.

Second, following Proposition 4.11, we can see that $\sup_{m \in \mathcal{M}} \lambda(\omega, t, m) < \infty$ for all $t \in \mathbb{R}_{>0}$, $\omega \in \Omega$. Hence, by Lemma 4.9, there exists and almost sure event $G \in \mathcal{F}$ on which $N(\{t\} \times \mathcal{M}) \leq 1$ for all $t \in \mathbb{R}$. Let $\tilde{N}$, $(\tilde{N}_n)_{n \in \mathbb{N}}$, $(\tilde{T}_n)_{n \in \mathbb{N}}$ and $\tilde{T}_\infty$ coincide with N, $(N_n)_{n \in \mathbb{N}}$, $(T_n)_{n \in \mathbb{N}}$ and T_∞ on G. Outside G, set $\tilde{N} := 0$, $\tilde{N}_n := 0$, $\tilde{T}_n := \infty$, for all $n \in \mathbb{N}$, and $\tilde{T}_\infty := \infty$. Define the random measures on $\mathbb{R}$

$$\tilde{N}_{\mathcal{M}}(\cdot) := \tilde{N}(\cdot \times \mathcal{M}), \quad \tilde{N}_{\mathcal{M},n}(\cdot) := \tilde{N}_n(\cdot \times \mathcal{M}), \quad n \in \mathbb{N},$$

and define the process

$$\tilde{\lambda}(\omega, t) := \lim_{n \to \infty} a(\tilde{N}_{\mathcal{M},n}(\omega, (-\infty, t))) \mathbb{1}_{\{t < \tilde{T}_\infty(\omega)\}} = a(\tilde{N}_{\mathcal{M}}(\omega, (-\infty, t))) \mathbb{1}_{\{t < \tilde{T}_\infty(\omega)\}}, \quad \omega \in \Omega, t \in \mathbb{R}_{>0}.$$

Since $\{\tilde{T}_n \leq t\} = \{\tilde{N}_{\mathcal{M}}((0, t]) \geq n\}$, $\tilde{T}_n$ is in fact an $\mathbb{F}^{\tilde{N}_{\mathcal{M}}}$-stopping time and, thus, reusing the argument in the first step, we have that $\mathbb{1}_{\{t < \tilde{T}_\infty\}}$ is $\mathbb{F}^{\tilde{N}_{\mathcal{M}}}$-predictable. Moreover, by Lemma 3.4, we have that $(a(\tilde{N}_{\mathcal{M},n}((-\infty, t))))_{t>0}$ is $\mathbb{F}^{\tilde{N}_{\mathcal{M},n}}$-predictable and, thus, $\mathbb{F}^{\tilde{N}_{\mathcal{M}}}$-predictable. Hence, using again Lemma 1.9 in Kallenberg (2002, p. 6), we have that $\tilde{\lambda}$ is also $\mathbb{F}^{\tilde{N}_{\mathcal{M}}}$-predictable. Next, because $\tilde{N} = N$ a.s. and $\tilde{\lambda}(t) = \lambda(t, m)$ for all $t \in \mathbb{R}_{>0}$, $m \in \mathcal{M}$, a.s., and because Lemma 4.6 applies to N and λ, we have that, for any non-negative $\mathbb{F}^{\tilde{N}_{\mathcal{M}}}$-predictable process $H : \Omega \times \mathbb{R}_{>0} \to \mathbb{R}_{\geq 0}$,

$$\mathbb{E}\left[\int_{\mathbb{R}_{>0}} H(t)\tilde{N}_{\mathcal{M}}(dt)\right] = \mathbb{E}\left[\iint_{\mathbb{R}_{>0} \times \mathcal{M}} H(t)N(dt, dm)\right]$$

$$= \mathbb{E}\left[\iint_{\mathbb{R}_{>0} \times \mathcal{M}} H(t)\lambda(t, m)\mu_{\mathcal{M}}(dm)dt\right]$$

$$= \mathbb{E}\left[\int_{\mathbb{R}_{>0}} H(t)\tilde{\lambda}(t)\mu_{\mathcal{M}}(\mathcal{M})dt\right].$$

Consequently, $\tilde{N}_{\mathcal{M}}$, or equivalently $(\tilde{T}_n)_{n \in \mathbb{N}}$, defines a simple point process on $\mathbb{R}_{>0}$ with $\mathbb{F}^{\tilde{N}_{\mathcal{M}}}$-predictable projection $(\mu_{\mathcal{M}}(\mathcal{M}) \int_0^t \tilde{\lambda}(s)ds)_{t>0}$ in the sense of Jacod (1975).

Third, by Lemma 3.1, $\mathcal{F}_t^{\tilde{N}_{\mathcal{M}}} = \mathcal{F}_0^{\tilde{N}_{\mathcal{M}}} \vee \mathcal{F}_t^{\tilde{N}_{\mathcal{M}}^{>0}}$, and, thus, Assumption A.1 of Jacod (1975) holds, see also the proof of Theorem 2.21 and Remark 4.17. Then, by Proposition 3.1 in Jacod (1975), we have that, conditional on $N_{\leq 0}((-\infty, 0]) = n_0 \in \mathbb{N}$, $S_n := \tilde{T}_{n+1} - \tilde{T}_n$, $n \in \mathbb{N}$, follows an exponential distribution with parameter $a(n + n_0)\mu_{\mathcal{M}}(\mathcal{M})$ and the $(S_n)_{n \in \mathbb{N}}$ are independent. Thanks to Assumption B.(ii), by Example 3.1.4 in Jacobsen (2006, p.20), we deduce that, conditional on $N_{\leq 0}((-\infty, 0]) = n_0$, $\tilde{T}_\infty = \lim_{n \to \infty} \tilde{T}_n = \infty$ a.s., see also Proposition 12.19 in Kallenberg (2002, p.240). Consequently, it holds that $\tilde{T}_\infty = \infty$ a.s. unconditionally.

Fourth, following these first three steps, we have proved that there exists a version of N such that $N \in \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\#g}$, i.e., this version of N is a non-explosive marked point process, see Proposition 4.2, and such that N solves the Poisson-driven SDE. We conclude by Corollary 4.8. $\square$

To prove Theorem 2.17 under Assumptions A, B and C in the general case, we will use a solution to the special case $\psi = \psi'$ to show that the constructed mapping $N : \Omega \to \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\infty}$ actually takes values in $\mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\#g}$. First, we need to define what we mean for a marked point process N to be dominated by another marked point process $\overline{N}$.

Definition 4.14. Let $\xi, \overline{\xi} \in \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\infty}$. We say that ξ is *dominated* by $\overline{\xi}$ and write $\xi \prec \overline{\xi}$ if, for all $A \in \mathcal{B}(\mathbb{R} \times \mathcal{M})$, $\xi(A) \leq \overline{\xi}(A)$. Let $T \in \mathbb{R}$. We say that ξ is dominated by $\overline{\xi}$ on $(-\infty, T]$ if $\theta_T \xi^{\leq 0} \prec \theta_T \overline{\xi}^{\leq 0}$. Consider two mappings $N : \Omega \to \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\infty}$ and $\overline{N} : \Omega \to \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\infty}$. We say that N is *dominated* by $\overline{N}$ if $N \prec \overline{N}$ a.s.

When $N \prec \overline{N}$ a.s., one could also say that N is a thinning of $\overline{N}$. Indeed, notice that $\xi \prec \overline{\xi}$ implies that all the atoms of ξ are also atoms of $\overline{\xi}$.

We will now show that the constructed mapping $N : \Omega \to \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\infty}$ is dominated by any solution to the special case $\psi = \psi'$.

Proposition 4.15. *Let $N' : \Omega \to \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\#g}$ be a solution to the Poisson-driven SDE with intensity functional ψ'. Then, under Assumptions A, B.(i) and C, the mapping $N : \Omega \to \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\infty}$ obtained from Algorithm 4.10 satisfies $N \prec N'$ a.s.*

Proof. Fix $\omega = (\omega_{\leq 0}, \omega_{>0}) \in A \cap F_1$, where $A \in \mathcal{F}$ is the almost sure event that N' solves (2.11), where ψ is replaced by ψ'. Clearly, we have that $N(\omega) \prec \overline{N}(\omega)$ on $(-\infty, 0]$. Now take any $n \in \mathbb{N}$ such that $T_n(\omega) < \infty$ and assume that $N(\omega) \prec N'(\omega)$ on $(-\infty, T_n(\omega)]$. If $T_{n+1}(\omega) = \infty$, then $N(\omega)$ is null on (T_n, ∞) and we have $N(\omega) \prec N'(\omega)$. If $T_{n+1}(\omega) < \infty$, we have that for all $t \in (T_n(\omega), T_{n+1}(\omega)]$, $m \in \mathcal{M}$,

$$\lambda(\omega, t, m) = \psi(m \,|\, \theta_t N(\omega)^{<0})$$
$$\text{(by construction)} \quad = \psi(m \,|\, \theta_t N_n(\omega)^{<0})$$
$$\text{(by Assumption B.(i))} \quad \leq \psi'(m \,|\, \theta_t N_n(\omega)^{<0})$$
$$\text{(by the definition of } \psi', \text{ Assumption B and since } N_n(\omega) \prec N'(\omega)) \quad \leq \psi'(m \,|\, \theta_t N'(\omega)^{<0}) =: \lambda'(\omega, t, m).$$

By Proposition 4.12 for $N(\omega)$ and by assumption for $N'(\omega)$, we have that $N(\omega)$ and $N'(\omega)$ both satisfy (2.11) on $(-\infty, T_{n+1}(\omega)]$, where λ is replaced by λ' for $N'(\omega)$. Consequently, we must have $N(\omega) \prec N'(\omega)$ on $(-\infty, T_{n+1}(\omega)]$. As, by construction, $N(\omega)$ has mass on $\mathbb{R}_{>0}$ only at the times $T_1(\omega) < T_2(\omega) < \ldots < \infty$, we have shown that $N(\omega) \prec N'(\omega)$. This implies that $N \prec N'$ a.s. $\qquad\square$

This allows us to conclude the proof of Theorem 2.17 under Assumptions A, B and C.

Proof of Theorem 2.17, Part 2. Repeat the first step of Part 1. Then, by Proposition 4.15, we deduce that $N \in \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\#g}$ and $T_\infty = \infty$ a.s. We then conclude by repeating the fourth step of Part 1. $\qquad\square$

4.2.2 Existence under Assumptions A, D, E. To prove Theorem 2.17 under Assumptions A, D and E, we will also use Algorithm 4.10 to construct a candidate solution, but the almost sure event F_1 needs to be replaced by another almost sure event F_2 that guarantees that the algorithm is well-defined under these new assumptions. Whereas under Assumptions B and C, we were able to first construct the candidate solution and then dominate it by a solution to the special case $\psi = \psi'$, here we will dominate the candidate solution while constructing it. The dominating non-explosive marked point process is nothing else than a solution to the Poisson-driven SDE with the Hawkes intensity functional

$$\overline{\psi} : \mathcal{M} \times \mathcal{N}_{\mathbb{R} \times \mathcal{M}}^{\#} \to \mathbb{R}_{>0} \cup \{\infty\}$$
$$(m, \xi) \mapsto \overline{\psi}(m \,|\, \xi) := \lambda_0 + \iint_{(-\infty, 0) \times \mathcal{M}} \overline{k}(-t', m', m) \xi(dt', dm'),$$

where λ_0 and $\overline{k}$ are as in Assumption D. Indeed, by applying the results of Massoulié (1998) and Lemma 4.9, we can prove Theorem 2.17 under Assumptions A, D and E for the special case $\psi = \overline{\psi}$.

Proof of Theorem 2.17, Part 3. Clearly, $\overline{\psi}$ satisfies the Lipschitz condition (2.9) with the kernel $\overline{k}$. Under Assumptions A, D.(ii) and E.(i), by Theorem 2 of Massoulié (1998), we know that there exists a non-explosive point process $\overline{N} : \Omega \to \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M}}$ that solves (2.11), where ψ is replaced by $\overline{\psi}$, and such that $\overline{N}(\cdot \times \mathcal{M}) \in \mathcal{N}^{\#}_{\mathbb{R}}$. Moreover, applying Assumptions D.(iii) and E.(ii), we obtain that

$$\overline{\lambda}(\omega, t, m) := \overline{\psi}(m \mid \theta_t \overline{N}(\omega)^{<0}) = \lambda_0 + \iint_{(-\infty, t) \times \mathcal{M}} \overline{k}(t - t', m', m) \overline{N}(\omega, dt', dm')$$

$$\leq \lambda_0 + \tilde{\lambda}_{\leq 0}(\omega_{\leq 0}, t) + \iint_{(0,t) \times \mathcal{M}} \sup_{m'' \in \mathcal{M}} \overline{k}(t - t', m', m'') \overline{N}(\omega, dt', dm')$$

$$< \infty, \quad \omega \in \Omega, t \in \mathbb{R}_{>0}, m \in \mathcal{M},$$

which proves that $\sup_{m \in \mathcal{M}} \overline{\lambda}(t, m) < \infty$, for all $t \in \mathbb{R}_{>0}$, a.s. Hence, by Lemma 4.9, we conclude that $\overline{N}$ admits a version such that $\overline{N}(\omega) \in \mathcal{N}^{\#g}_{\mathbb{R} \times \mathcal{M}}$, $\omega \in \Omega$, meaning that this version solves the Poisson-driven SDE. Conclude by Corollary 4.8. $\qquad\square$

From now on, denote by $\overline{N}$ a solution to the Poisson-driven SDE in the special case $\psi = \overline{\psi}$ and by $F_2 \in \mathcal{F}$ the almost sure event that $\overline{N}$ solves (2.11), where ψ is replaced by $\overline{\psi}$. The following statement is the analogue of Proposition 4.11 and ensures that Algorithm 4.10 is well-defined under this different set of assumptions.

Proposition 4.16. *In Algorithm 4.10, where F_1 is replaced by F_2, under Assumptions A, D and E, we have that, for every $\omega \in F_2$, $\operatorname{card}(\mathbb{M}_n(\omega)) \leq 1$, $N_n(\omega) \prec \overline{N}(\omega)$, for all $n \in \mathbb{N}$, and*

$$\left\{ u > T_n(\omega) : \iint_{(T_n(\omega), u) \times \mathcal{M}} \int_{(0, \lambda_n(\omega, t, m)]} M(\omega, dt, dm, dz) = 0 \right\} \neq \varnothing \quad \text{for all } n \in \mathbb{N} \text{ s.t. } T_n(\omega) < \infty.$$

Hence, Algorithm 4.10 is well-defined under these assumptions.

Proof. We show the assertion by induction. Take any $\omega = (\omega_{\leq 0}, \omega_{>0}) \in F_2$. Let $n \in \mathbb{N}$ and, for all $i \in \mathbb{N}$ such that $i < n$ and $T_i(\omega) < \infty$, assume that

$$\left\{ u > T_i(\omega) : \iint_{(T_i(\omega), u) \times \mathcal{M}} \int_{(0, \lambda_i(\omega, t, m)]} M(\omega, dt, dm, dz) = 0 \right\} \neq \varnothing. \tag{4.10}$$

For all $i \in \mathbb{N}$ such that $i \leq n$, assume that $N_i(\omega) \prec \overline{N}$. If $T_n(\omega) = \infty$, then, by construction, this is also true for $n + 1$.

Now, assume that $T_n(\omega) < \infty$. We first show that (4.10) holds also for $i = n$. By adapting the proof of Proposition 4.15 and using Assumption D.(i), we get that $\lambda_n(\omega, t, m) \leq \overline{\lambda}(\omega, t, m)$, for all $t > T_n(\omega)$, $m \in \mathcal{M}$. Hence, for any $\varepsilon > 0$, we have that

$$\iint_{(T_n(\omega), T_n(\omega)+\varepsilon) \times \mathcal{M}} \int_{(0, \lambda_n(\omega, t, m)]} M(\omega, dt, dm, dz) \leq \iint_{(T_n(\omega), T_n(\omega)+\varepsilon) \times \mathcal{M}} \int_{(0, \overline{\lambda}(\omega, t, m)]} M(\omega, dt, dm, dz)$$

$$= \overline{N}(\omega, (T_n(\omega), T_n(\omega) + \varepsilon) \times \mathcal{M}) =: \overline{U}_n(\omega, \varepsilon) < \infty,$$

since $\overline{N}(\omega, \cdot \times \mathcal{M}) \in \mathcal{N}^{\#}_{\mathbb{R}}$. If $\overline{U}_n(\omega, \varepsilon) = 0$, then clearly (4.10) is satisfied for $i = n$. If not, $\overline{N}(\omega, \cdot \times \mathcal{M})$ has a finite number of points in $(T_n(\omega), T_n(\omega) + \varepsilon)$ and there exists $0 < \varepsilon' < \varepsilon$ such that $\overline{U}_n(\omega, \varepsilon') = 0$, in which case (4.10) is again satisfied for $i = n$.

Second, we show that $\mathrm{card}(\mathbb{M}_{n+1}) \leq 1$ and $N_{n+1}(\omega) \prec \overline{N}(\omega)$. If $T_{n+1}(\omega) = \infty$, then this is immediate. If not, as $\lambda_n(\omega, t, m) \leq \overline{\lambda}(\omega, t, m)$, for all $t > T_n(\omega)$, it is enough to notice that

$$
\begin{aligned}
\mathbb{M}_{n+1}(\omega) := & \{m \in \mathscr{M} \, : \, M(\omega, \{T_{n+1}(\omega)\} \times \{m\} \times \, (0, \lambda_n(\omega, T_{n+1}(\omega), m)]) > 0\} \\
\subset & \{m \in \mathscr{M} \, : \, M(\omega, \{T_{n+1}(\omega)\} \times \{m\} \times \, (0, \overline{\lambda}(\omega, T_{n+1}(\omega), m)]) > 0\} \\
= & \{m \in \mathscr{M} \, : \, \overline{N}(\omega, \{T_{n+1}(\omega)\} \times \{m\}) > 0\} \leq 1,
\end{aligned}
$$

since $\overline{N}(\omega) \in \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}$. As we already know that $N_n(\omega) \prec \overline{N}(\omega)$, looking at (4.3) and observing that $N_{n+1}(\omega)$ and $N_n(\omega)$ only differ by a mass at time $T_{n+1}(\omega)$, we further deduce that $N_{n+1}(\omega) \prec \overline{N}(\omega)$.

Regarding the basis of this induction, it is immediate that $N_0(\omega) \prec \overline{N}(\omega)$ since $N_0(\omega) = N_{\leq 0}(\omega_{\leq 0}) = \overline{N}^{\leq 0}(\omega)$. □

We are now in a position to finish the proof of Theorem 2.17 under Assumption A, D and E.

Proof of Theorem 2.17, Part 4. Let $N : \Omega \to \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\infty}$ be given by Algorithm 4.10 under Assumptions A, D and E, where F_1 is replaced by F_2. By Proposition 4.16, this mapping is well defined. Moreover, we notice that Propositions 4.12 and 4.13 still hold under the present assumptions. Hence, we can repeat the first step of Part 1 of the proof. By Proposition 4.16, we know that $N_n(\omega) \prec \overline{N}(\omega)$ for all $\omega \in F_2$, which implies by construction that $N(\omega) \prec \overline{N}(\omega)$ on $\mathbb{R}_{>0}$ for all $\omega \in \Omega$. Consequently, we have that $N(\omega) \in \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}$ and $T_\infty(\omega) = \infty$ for all $\omega \in \Omega$, which implies that N solves the Poisson-driven SDE. We conclude again by Corollary 4.8. □

4.3 Strong and weak uniqueness

Proof of Theorem 2.20. Let $\tilde{\Omega} \in \mathcal{F}$ be the almost sure event that both N and N' solve (2.11). Let $(T_n, M_n)_{n \in \mathbb{N}}$ and $(T'_n, M'_n)_{n \in \mathbb{N}}$ be the enumerations in $(0, \infty] \times \mathscr{M}$ to which N and N' are respectively equivalent. Now fix arbitrary $\omega \in \tilde{\Omega}$. We show by strong induction that $T_n(\omega) = T'_n(\omega)$ and $M_n(\omega) = M'_n(\omega)$ for all $n \in \mathbb{N}$.

Let $n \in \mathbb{N}$ and assume that $T_i(\omega) = T'_i(\omega)$ and $M_i(\omega) = M'_i(\omega)$ for all $i = 1, \ldots, n-1$. By contradiction, assume that $T_n(\omega) \neq T'_n(\omega)$ and, moreover, without loss of generality, that $T_n(\omega) < T'_n(\omega)$. Then, this implies that

$$
N(\omega, (0, T_n(\omega)] \times \mathscr{M}) = \int_{(0, T_n(\omega)]} \int_{\mathscr{M}} \int_{(0, \lambda(\omega, t, m)]} M(\omega, dt, dm, dz) = n,
$$

$$
N'(\omega, (0, T_n(\omega)] \times \mathscr{M}) = \int_{(0, T_n(\omega)]} \int_{\mathscr{M}} \int_{(0, \lambda'(\omega, t, m)]} M(\omega, dt, dm, dz) = n - 1,
$$

where $\lambda(\omega, t, m) = \psi(m \,|\, \theta_t N(\omega)^{<0})$ and $\lambda'(\omega, t, m) = \psi(m \,|\, \theta_t N'(\omega)^{<0})$. But since $N(\omega)^{\leq 0} = N'(\omega)^{\leq 0}$ and also $T_i(\omega) = T'_i(\omega)$ and $M_i(\omega) = M'_i(\omega)$ for all $i = 1, \ldots, n-1$, we have that $\theta_t N(\omega)^{<0} = \theta_t N'(\omega)^{<0}$ for all $t \leq T_n(\omega)$ and, thus, $\lambda(\omega, t, m) = \lambda'(\omega, t, m)$ for all $t \leq T_n(\omega), m \in \mathscr{M}$. This implies that $n = n - 1$ which is a contradiction and, thus, necessarily, $T_n(\omega) = T'_n(\omega)$.

Similarly, if we assume that $M_n(\omega) \neq M'_n(\omega)$, then this implies that

$$
N(\omega, \{T_n(\omega)\} \times \{M_n(\omega)\}) = \int_{\{T_n(\omega)\}} \int_{\{M_n(\omega)\}} \int_{(0, \lambda(\omega, t, m)]} M(\omega, dt, dm, dz) = 1,
$$

$$
N'(\omega, \{T_n(\omega)\} \times \{M_n(\omega)\}) = \int_{\{T_n(\omega)\}} \int_{\{M_n(\omega)\}} \int_{(0, \lambda'(\omega, t, m)]} M(\omega, dt, dm, dz) = 0.
$$

But again, since $\lambda(\omega, t, m) = \lambda'(\omega, t, m)$ for all $t \leq T_n(\omega), m \in \mathscr{M}$, this leads to the contradiction $1 = 0$ and, thus, it follows that $M_n(\omega) = M_n'(\omega)$. The same reasoning allows us to prove the basis of the strong induction (i.e., to show that $T_1(\omega) = T_1'(\omega)$ and $M_1(\omega) = M_1'(\omega)$). $\qquad \square$

Proof of Theorem 2.21. Consider the canonical measurable space $(\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}, \mathcal{B}(\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}))$, where $\mathcal{B}(\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}) = \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g} \cap \mathcal{B}(\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#})$, and the canonical non-explosive marked point process N defined by $N(\omega) = \omega$ for all $\omega \in \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}$. Under both $\mathcal{P}^{N_1}$ and $\mathcal{P}^{N_2}$ (they only charge $\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}$), N satisfies the weak initial condition $N_{\leq 0}$ and admits an intensity given by (2.2). We will now apply Theorem 3.4 in Jacod (1975, p. 242) to show that $\mathcal{P}^{N_1} = \mathcal{P}^{N_2}$. By Lemma 3.1, we have that

$$\mathcal{F}_t^N = \mathcal{F}_0^N \vee \mathcal{F}_t^{\theta_0 N^{>0}} = \mathcal{F}_0^{N^{\leq 0}} \vee \mathcal{F}_t^{N^{>0}}, \quad t \in \mathbb{R}_{\geq 0},$$

and thus, Assumption A.1 of Jacod (1975) is satisfied (see Remark 4.17).

To apply Theorem 3.4 in Jacod (1975), it remains to verify that the restrictions of $\mathcal{P}^{N_1}$ and $\mathcal{P}^{N_2}$ coincide on $\mathcal{F}_0^N$. Note that $\mathcal{F}_0^N$ is generated by the π-system $\mathcal{C}$ of sets of the form

$$\{N \in \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g} : N(A_1 \times M_1) \geq n_1, \ldots, N(A_k \times M_k) \geq n_k\}, \quad n_1, \ldots, n_k \in \mathbb{N}, k \in \mathbb{N},$$

where $A_1, \ldots, A_k, \in \mathcal{B}(\mathbb{R}_{\leq 0})$ and $M_1, \ldots, M_k \in \mathcal{B}(\mathscr{M})$. For any such set $F \in \mathcal{C}$, setting $B_i := A_i \times M_i$ for $i = 1, \ldots, k$ and invoking the fact that both N_1 and N_2 satisfy the weak initial condition $N_{\leq 0}$, we deduce that

$$\mathcal{P}^{N_1}(F) = \mathcal{P}^{N_1}(N(B_1) \geq n_1, \ldots, N(B_k) \geq n_k) = \mathcal{P}^{N_1}(N^{\leq 0}(B_1) \geq n_1, \ldots, N^{\leq 0}(B_k) \geq n_k)$$
$$= \mathcal{P}^{N^{\leq 0}}(N^{\leq 0}(B_1) \geq n_1, \ldots, N^{\leq 0}(B_k) \geq n_k) = \mathcal{P}^{N_2}(N^{\leq 0}(B_1) \geq n_1, \ldots, N^{\leq 0}(B_k) \geq n_k)$$
$$= \mathcal{P}^{N_2}(N(B_1) \geq n_1, \ldots, N(B_k) \geq n_k) = \mathcal{P}^{N_2}(F).$$

Hence, $\mathcal{P}^{N_1}$ and $\mathcal{P}^{N_2}$ coincide on $\mathcal{C}$, a π-system that contains $\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}$. As a consequence, $\mathcal{P}^{N_1} = \mathcal{P}^{N_2}$ on $\mathcal{F}_0^N$, see for example Lemma 1.17 in Kallenberg (2002, p. 9), and we can apply Theorem 3.4 in Jacod (1975, p. 242) to deduce that $\mathcal{P}^{N_1} = \mathcal{P}^{N_2}$ on $(\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}, \mathcal{B}(\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}))$. $\qquad \square$

Remark 4.17. Let us clarify the relationship between our notations and those in Jacod (1975). Our canonical measurable space $(\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}, \mathcal{B}(\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}))$ plays the role of his measurable space $(\Omega, \mathcal{F}_\infty)$. Our marked point process $N^{>0}$ corresponds to his marked point process μ. Our probability measures $\mathcal{P}^{N_1}$ and $\mathcal{P}^{N_2}$ are the counterparts of P and P', respectively. Our filtrations $\mathbb{F}^N$ and $\mathbb{F}^{N^{>0}}$ correspond to his filtrations $(\mathcal{F}_t)_{t \geq 0}$ and $(\mathcal{G}_t)_{t \geq 0}$, respectively.

A   Appendix

A.1   The subspace $\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}$ is Borel

The following result is unlikely to be original, but we could not find it in Daley and Vere-Jones (2008).

Lemma A.1.1. *The set $\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g}$ is a Borel subset of $\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#}$, that is $\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#g} \in \mathcal{B}(\mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#})$.*

Proof. For all $n \in \mathbb{N}$, define the sets

$$F_n := \left\{ \xi \in \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#} : \xi([-n, n] \times \mathscr{M}) < \infty, \xi(\{t\} \times \mathscr{M}) \leq 1 \text{ for all } t \in [-n, n] \right\}$$

and notice that $\mathcal{N}^{\#g}_{\mathbb{R}\times\mathscr{M}} = \bigcap_{n\in\mathbb{N}} F_n$. Next, let $n \in \mathbb{N}$. By Proposition A2.1.IV in Daley and Vere-Jones (2003, p. 385), the interval $[-n, n]$ contains a dissecting system $((A_{ij})_{j\in\{1,\ldots,j_i\}})_{i\in\mathbb{N}}$ where $A_{ij} \in \mathcal{B}(\mathbb{R})$ for any $j \in \{1,\ldots,j_i\}$ and $i \in \mathbb{N}$, see Definition A1.6.1 in Daley and Vere-Jones (2003, p. 382). We show that

$$F_n = \left\{ \xi \in \mathcal{N}^{\#}_{\mathbb{R}\times\mathscr{M}} : \xi([-n,n]\times\mathscr{M}) < \infty, \limsup_{i\to\infty} \sup_{j\in\{1,\ldots,j_i\}} \xi(A_{ij}\times\mathscr{M}) \leq 1 \right\} =: G_n.$$

Let $\xi \in F_n$. Then $\xi(\cdot\times\mathscr{M})$ has finitely many atoms $t_1,\ldots,t_p$ in $[-n,n]$ for some $p \in \mathbb{N}$ number $p \in \mathbb{N}$ and their mass cannot exceed one. A key property of the dissecting system is that, for each pair of distinct atoms t_{q_1} and t_{q_2} with $q_1 \neq q_2$, there exists $n(q_1,q_2) \in \mathbb{N}$ such that, for all $i > n(q_1,q_2)$, $t_{q_1} \in A_{ij}$ implies $t_{q_2} \notin A_{ij}$. Thus, define

$$i^* := \max_{q_1,q_2\in\{1,\ldots,p\},\, q_1\neq q_2} n(q_1,q_2)$$

and then, $\xi(A_{ij}\times\mathscr{M}) \leq 1$ for all $j \in \{1,\ldots,j_i\}$ and $i > i^*$, which implies that

$$\limsup_{i\to\infty} \sup_{j\in\{1,\ldots,j_i\}} \xi(A_{ij}\times\mathscr{M}) \leq 1,$$

which in turn indicates that $\xi \in G_n$. Now, let $\xi \in G_n$ and $t \in [-n,n]$. Another salient property of the dissecting system is that there exists a sequence $(j_i)_{i\in\mathbb{N}}$ such that $\xi(\{t\}\times\mathscr{M}) = \lim_{i\to\infty} \xi(A_{ij_i}\times\mathscr{M})$. But since $\xi \in G_n$, we have that

$$\xi(\{t\}\times\mathscr{M}) = \lim_{i\to\infty} \xi(A_{ij_i}\times\mathscr{M}) \leq \limsup_{i\to\infty} \sup_{j\in\{1,\ldots,j_i\}} \xi(A_{ij}\times\mathscr{M}) \leq 1,$$

which means that $\xi \in F_n$. Now that we have shown that $F_n = G_n$, we invoke Theorem A2.6.III in Daley and Vere-Jones (2003, p. 404) to deduce that $\xi \mapsto \xi([-n,n]\times\mathscr{M})$ and $\xi \mapsto \xi(A_{ij}\times\mathscr{M})$, for any $j \in \{1,\ldots,j_i\}$ and $i \in \mathbb{N}$, are measurable and use Lemma 1.9 in Kallenberg (2002, p. 6) to conclude that $\xi \mapsto \limsup_{i\to\infty} \sup_{j\in\{1,\ldots,j_i\}} \xi(A_{ij}\times\mathscr{M})$ is measurable. It then follows that $F_n \in \mathcal{B}(\mathcal{N}^{\#}_{\mathbb{R}\times\mathscr{M}})$, whence $\mathcal{N}^{\#g}_{\mathbb{R}\times\mathscr{M}} = \bigcap_{n\in\mathbb{N}} F_n \in \mathcal{B}(\mathcal{N}^{\#}_{\mathbb{R}\times\mathscr{M}})$. □

A.2 Measurability and continuity properties of shifts and resitrictions

From Daley and Vere-Jones (2008, p. 178, Lemma 12.1.I), we know the shift operators are continuous under the $w^{\#}$-topology. We are able to go further and show that $\theta_t\xi$ is actually jointly continuous in ξ and t. We also prove that taking the restriction to the positive or negative real line of a boundedly finite measure is a measurable operation. Moreover, we show that $\theta_t\xi^{<0}$ is left-continuous as a function of $t \in \mathbb{R}$ for any $\xi \in \mathcal{N}^{\#}_{\mathcal{U}}$, which is crucial in our proof that an intensity functional applied to the history of a point process generates a predictable process (Lemma 3.4). Before giving the formal statements, we define some notations. Recall that $d^{\#}$ is the $w^{\#}$-distance (Daley and Vere-Jones, 2003, p. 403) on the space $\mathcal{N}^{\#}_{\mathcal{U}}$. The open ball with centre $u \in \mathcal{U}$ and radius r is denoted by $B_r(u)$. For any subset $A \subset \mathcal{U}$ and $\varepsilon > 0$, the ε-neighbourhood of A is defined by $A^\varepsilon := \bigcup_{a\in A} B_\varepsilon(a)$ and the boundary of A is denoted by ∂A.

Lemma A.2.1. *When $\mathcal{N}^{\#}_{\mathbb{R}\times\mathcal{U}}$ is equipped with the $w^{\#}$-distance $d^{\#}$ and $\mathcal{N}^{\#}_{\mathbb{R}\times\mathcal{U}}\times\mathbb{R}$ is equipped with the product metric, the mapping*

$$\mathcal{N}^{\#}_{\mathbb{R}\times\mathcal{U}} \times \mathbb{R} \to \mathcal{N}^{\#}_{\mathbb{R}\times\mathcal{U}}$$
$$(\xi, t) \mapsto \theta_t\xi$$

is continuous.

Proof. Let $\xi \in \mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#}$, $t \in \mathbb{R}$ and let $(\xi_n, t_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#} \times \mathbb{R}$ such that $d^{\#}(\xi_n, \xi) \to 0$ and $t_n \to t$ as $n \to \infty$. By Proposition A2.6.II in Daley and Vere-Jones (2003, p. 403), it is enough to show that $\xi_n(A + t_n) \to \xi(A + t)$ as $n \to \infty$ for any bounded $A \in \mathcal{B}(\mathbb{R} \times \mathcal{U})$ such that $\xi(\partial(A + t)) = 0$. For such a set A, which we can assume without loss of generality to be non-empty, there exists $\delta > 0$ such that $B_{2\delta}(a_n) \subset (A + t)$ and $\xi(\partial B_\delta(a_n)) = 0$, $n = 1, \ldots, N$, where $a_1, \ldots, a_N$ are the atoms of ξ in $A + t$, and such that $\xi((A + t)^\delta) = \xi(A + t)$ with $\xi(\partial((A + t)^\delta)) = 0$. Introduce the two bounded sets

$$S_1 := (A + t)^\delta \setminus (A + t) \quad \text{and} \quad S_2 := (A + t) \setminus \left(\bigcup_{n=1}^{N} B_\delta(a_n) \right)$$

and notice that $\xi(S_1) = \xi(S_2) = \xi(\partial S_1) = \xi(\partial S_2) = 0$. Since $d^{\#}(\xi_n, \xi) \to 0$ as $n \to \infty$, we have that $\xi_n(S_1) = \xi_n(S_2) = \xi_n(\partial S_1) = \xi_n(\partial S_2) = 0$ for n large enough. This implies that, for n large enough, all the atoms of ξ_n in $(A + t)^\delta$ actually lie in $(A + t)$ and their distance to the boundary of $(A + t)$ is bigger than δ (all the atoms are in the balls $B_\delta(a_n)$). This means that, for n large enough, $\xi_n((A + t) \setminus (A + s)) = \xi_n((A + s) \setminus (A + t)) = 0$ for all $s \in \mathbb{R}$ such that $|t - s| < \delta$, implying that $\xi_n(A + t) = \xi_n(A + s)$ for all such n and s. But for n large enough, we also have that $|t_n - t| < \delta$ and $\xi_n(A + t) = \xi(A + t)$, which finally gives that, for such large enough n,

$$\xi_n(A + t_n) = \xi_n(A + t + (t_n - t)) = \xi_n(A + t) = \xi(A + t). \qquad \square$$

Lemma A.2.2. *The restrictions $\xi^{<0}$, $\xi^{\le 0}$, $\xi^{>0}$ and $\xi^{\ge 0}$ are measurable mappings from $\mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#}$ into itself.*

Proof. We prove the assertion for $\xi^{<0}$, the other three restrictions can be treated similarly. Consider the function $f : \mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#} \ni \xi \mapsto \xi^{<0} \in \mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#}$. Remember that, by Theorem A2.6.III in Daley and Vere-Jones (2003, p. 404), the Borel σ-algebra $\mathcal{B}(\mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#})$ is generated by the sets

$$F_{A,n} := \{ \xi \in \mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#} : \xi(A) \in [n, \infty] \}, \quad A \in \mathcal{B}(\mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#}), \, n \in \mathbb{R}.$$

Since

$$f^{-1}(F_{A,n}) = \{ \xi \in \mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#} : \xi(A \cap \mathbb{R}_{<0} \times \mathcal{U}) \in [n, \infty] \} \in \mathcal{B}(\mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#}),$$

we conclude that f is measurable by Lemma 1.4 in Kallenberg (2002, p. 4). $\qquad \square$

Lemma A.2.3. *Let $\xi \in \mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#}$. Then the mapping*

$$\mathbb{R} \to \mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#}$$
$$t \mapsto (\theta_t \xi)^{<0}$$

is left continuous when $\mathcal{N}_{\mathbb{R} \times \mathcal{U}}^{\#}$ is equipped with the $w^{\#}$-distance $d^{\#}$.

Proof. Fix $t \in \mathbb{R}$ and take any non-decreasing sequence $(t_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$ such that $t_n \uparrow t$ as $n \to \infty$. By Proposition A2.6.II in Daley and Vere-Jones (2003, p. 403), it is enough to show that $(\theta_{t_n} \xi)^{<0}(A) \to (\theta_t \xi)^{<0}(A)$ as $n \to \infty$ for all bounded $A \in \mathcal{B}(\mathbb{R} \times \mathcal{U})$ such that $(\theta_t \xi)^{<0}(\partial A) = 0$. Clearly, it suffices to consider bounded Borel sets A such that $A \subset \mathbb{R}_{<0} \times \mathcal{U}$. First, consider the case where $\partial A \subset \mathbb{R}_{<0} \times \mathcal{U}$. This implies that

$$(\theta_t \xi)(\partial A) = (\theta_t \xi)^{<0}(\partial A) = 0.$$

By Lemma A.2.1, and using again the characterisation of Proposition A2.6.II in Daley and Vere-Jones (2003, p. 403), this implies that

$$(\theta_{t_n}\xi)^{<0}(A) = (\theta_{t_n}\xi)(A) \to (\theta_t\xi)(A) = (\theta_t\xi)^{<0}(A), \quad n \to \infty.$$

Second, consider the remaining case where $\partial A \cap \{0\} \times \mathcal{U} \neq \varnothing$. Then, $(\theta_t\xi)^{<0}(\partial A) = 0$ does not imply anymore that $(\theta_t\xi)(\partial A) = 0$. However, let ξ_- be the measure ξ that omits all atoms with time coordinate t. Then, for this measure ξ_-, we have again that

$$(\theta_t\xi_-)(\partial A) = (\theta_t\xi)^{<0}(\partial A) = 0.$$

Since $t_n \leq t$, and adapting the preceding argument for ξ_-, we finally find that

$$(\theta_{t_n}\xi)^{<0}(A) = (\theta_{t_n}\xi)(A) = (\theta_{t_n}\xi_-)(A) \to (\theta_t\xi_-)(A) = (\theta_t\xi)^{<0}(A), \quad n \to \infty. \qquad \square$$

A.3 Enumeration representation of marked point processes

The following result confirms that a non-explosive enumeration in $\mathbb{R}_{>0} \times \mathcal{M}$ corresponds indeed to a non-explosive marked point process.

Lemma A.3.1. *Let $(T_n, M_n)_{n\in\mathbb{N}}$ be an enumeration in $\mathbb{R}_{>0} \times \mathcal{M}$ such that $\lim_{n\to\infty} T_n = \infty$ a.s. Let $F \in \mathcal{F}$ be the almost sure event that $\lim_{n\to\infty} T_n = \infty$ and define*

$$N(\omega) := \begin{cases} \sum_n \delta_{(T_n(\omega), M_n(\omega))} \mathbb{1}_{\{T_n(\omega) < \infty\}}, & \text{if } \omega \in F, \\ 0, & \text{if } \omega \notin F. \end{cases}$$

Then, N defines a non-explosive marked point process on $\mathbb{R}_{\geq 0} \times \mathcal{M}$.

Proof. By Proposition 9.1.X in Daley and Vere-Jones (2008, p. 13), N defines a non-explosive point process on $\mathbb{R}_{\geq 0} \times \mathcal{M}$. Moreover, the monotonicity of the sequence $(T_n)_{n\in\mathbb{N}}$ implies that $N(\{t\} \times \mathcal{M}) = 0$ or 1 for all $\omega \in \Omega$. Also, using that $\lim_{n\to\infty} T_n = \infty$ on F, notice that $N(\omega, A \times \mathcal{M}) < \infty$, for every bounded set $A \in \mathcal{B}(\mathbb{R}_{\geq 0})$, for all $\omega \in \Omega$. This means that $N \in \mathcal{N}^{\#g}_{\mathbb{R}_{\geq 0} \times \mathcal{M}}$ and, thus, N defines a non-explosive marked point process. $\qquad \square$

Conversely, every non-explosive marked point process generates an enumeration.

Lemma A.3.2. *Let N be a non-explosive marked point process on $\mathbb{R}_{\geq 0} \times \mathcal{M}$ such that $N(\{0\} \times \mathcal{M}) = 0$ a.s. Define the sequence $(T_n)_{n\in\mathbb{N}}$ by $T_n := \sup\{t > 0 : N((0, t) \times \mathcal{M}) \leq n\}$. Then $(T_n)_{n\in\mathbb{N}}$ is a non-decreasing sequence of random variables in $(0, \infty]$. Moreover, for each $n \in \mathbb{N}$, on $\{T_n < \infty\}$, one can define M_n as the unique element in $\mathcal{M}$ such that $N(\{T_n\} \times \{M_n\}) > 0$. On $\{T_n = \infty\}$, simply set $M_n = m_\infty$ for some fixed $m_\infty \in \mathcal{M}$. Then, $(T_n, M_n)_{n\in\mathbb{N}}$ is an enumeration in $\mathbb{R}_{\geq 0} \times \mathcal{M}$ such that*

$$N = \sum_{n\in\mathbb{N}} \delta_{(T_n, M_n)} \mathbb{1}_{\{T_n < \infty\}} \tag{A.1}$$

and $\lim_{n\to\infty} T_n(\omega) = \infty$ for all $\omega \in \Omega$.

Proof. We proceed in several steps.

(i) For each $n \in \mathbb{N}$, the mapping $\omega \mapsto T_n(\omega)$ is measurable. Indeed, notice that $\{T_n < t\} = \{N((0,t) \times \mathcal{M}) > n\}$. Then recall that, by Theorem A2.6.III in Daley and Vere-Jones (2003, p. 404), $\mathcal{N}^{\#}_{\mathbb{R}_{\geq 0} \times \mathcal{M}} \ni \xi \mapsto \xi((0,t) \times \mathcal{M})$ is measurable and, thus, as a composition (Kallenberg, 2002, Lemma 1.7, p. 5), the mapping $\omega \mapsto N(\omega, (0,t) \times \mathcal{M})$ is measurable. Consequently, $\{T_n < t\} \in \mathcal{F}$. We conclude using Lemma 1.4 in Kallenberg (2002, p. 4) that the mapping $\omega \mapsto T_n(\omega)$ is measurable.

(ii) Using the fact that the ground measure is simple, $T_n < \infty$ implies that $T_n < T_{n+1}$. Also, it is easy to check that when $T_n = \infty$, then $T_{n+1} = \infty$. Hence, $(T_n)_{n \in \mathbb{N}}$ is sequence of random variables in $(0, \infty]$ satisfying the monotonicity of an enumeration.

(iii) For each $n \in \mathbb{N}$, the mapping $\omega \mapsto M_n(\omega)$ is well defined (using again the fact that the ground measure is simple). Also, this mapping is measurable. Indeed, let $A \in \mathcal{B}(\mathcal{M})$ and consider the most delicate case where $m_\infty \in A$. Notice that

$$\{M_n \in A\} = (\{T_n < \infty\} \cap \{N(\{T_n\} \times A) > 0\}) \cup \{T_n = \infty\}.$$

Based on what we have seen so far, we know that $\{T_n = \infty\} \in \mathcal{F}$. Therefore, it suffices to show that the set $\{T_n < \infty\} \cap \{N(\{T_n\} \times A) > 0\}$ is measurable. To this end, notice that

$$\{T_n < \infty\} \cap \{N(\{T_n\} \times A) > 0\} = \{T_n < \infty\} \cap \{\theta_{T_n} N(\{0\} \times A) > 0\},$$

where θ_{T_n} is the shift operator defined in Subsection 2.1.6. Then, by Lemma A.2.1, we know that the mapping $\mathcal{N}^{\#}_{\mathbb{R}_{\geq 0} \times \mathcal{M}} \times \mathbb{R}_{\geq 0} \ni (\xi, t) \mapsto \theta_t \xi \in \mathcal{N}^{\#}_{\mathbb{R}_{\geq 0} \times \mathcal{M}}$ is continuous and thus, by Lemma 1.5 in Kallenberg (2002, p. 4), measurable. Also, by Lemma 1.8 in Kallenberg (2002, p. 5), the mapping $\omega \mapsto (N(\omega), T_n(\omega))$ is measurable, and thus, as a composition (Kallenberg, 2002, Lemma 1.7, p. 5), the mapping $\omega \mapsto \theta_{T_n(\omega)} N(\omega)$ is measurable. Using again Theorem A2.6.III in Daley and Vere-Jones (2003, p. 404), we conclude that $\{T_n < \infty\} \cap \{\theta_{T_n} N(\{0\} \times A) > 0\} \in \mathcal{F}$ and, thus, the mapping $\omega \mapsto M_n(\omega)$ is measurable. So far, these first three steps establish that $(T_n, M_n)_{n \in \mathbb{N}}$ is an enumeration. Moreover, (A.1) holds by construction.

(iv) Since $N(\cdot \times \mathcal{M}) \in \mathcal{N}^{\#}_{\mathbb{R}_{\geq 0}}$, we have that $\lim_{n \to \infty} T_n(\omega) = \infty$ for all $\omega \in \Omega$. $\qquad \square$

Remark A.3.3. On the one hand, Lemma A.3.1 gives us a mapping that generates a non-explosive marked point process out of a non-explosive enumeration. One can see that if two non-explosive enumerations are not almost surely equal, then the corresponding non-explosive marked point processes cannot be almost surely equal either. In other words, the mapping of Lemma A.3.1 is injective. On the other hand, Lemma A.3.2 tells us that this mapping is surjective. As a consequence, the above two lemmas tell us that non-explosive enumerations and non-explosive marked point processes are two equivalent ways of looking at the same object.

A.4 Hawkes functionals

One can generalise the multivariate linear Hawkes processes presented in the introduction by defining Hawkes functionals as intensity functionals $\psi : \mathcal{M} \times \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M}} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ of the form

$$\psi(m \,|\, \xi) = \nu(m) + \iint_{(-\infty, 0) \times \mathcal{M}} k(-t', m', m) \xi(dt', dm'), \quad m \in \mathcal{M}, \xi \in \mathcal{N}^{\#}_{\mathbb{R} \times \mathcal{M}}, \tag{A.2}$$

36

where $\nu : \mathscr{M} \to \mathbb{R}_{\geq 0}$ and $k : \mathbb{R} \times \mathscr{M} \times \mathscr{M} \to \mathbb{R}_{\geq 0}$ are non-negative measurable functions. We show that such event functionals are measurable, so that they are admissible in our framework.

Proposition A.4.1. *Hawkes functionals of the form* (A.2) *are jointly measurable in* $m \in \mathscr{M}$ *and* $\xi \in \mathcal{N}_{\mathcal{U}}^{\#}$.

Proof. It will be enough to show that the integral term in (A.2), now denoted by $I(m, \xi)$, is measurable as a function of $m \in \mathscr{E}$ and $\xi \in \mathcal{N}_{\mathbb{R} \times \mathscr{M}}^{\#}$. First, consider the functions k of the form $k(t', m', m) = \mathbb{1}_S(t', m', m)$ where $S \in \mathcal{B}(\mathbb{R} \times \mathscr{M} \times \mathscr{M})$ and let $\mathcal{C}$ be the class of sets $S \in \mathcal{B}(\mathbb{R} \times \mathscr{M} \times \mathscr{M})$ such that $(m, \xi) \mapsto I(m, \xi)$ is measurable. By monotone convergence, the class $\mathcal{C}$ is a monotone class (i.e., it is closed under monotonically increasing sequences). Denote by $\mathcal{R}$ the class of sets of the form $\bigcup_{i=1}^n A_i \times M_i' \times M_i$ where $A_i \in \mathcal{B}(\mathbb{R})$, $M_i' \in \mathcal{B}(\mathscr{M})$, $M_i \in \mathcal{B}(\mathscr{M}), n \in \mathbb{N}$. This class forms a ring (i.e., it is closed under finite intersections and symmetric differences). Indeed, the difference of unions of Cartesian products is a union of Cartesian product. Moreover, since any union of Cartesian products can be decomposed as a union of disjoint Cartesian products, we have that $\mathcal{R} \subset \mathcal{C}$. Indeed, by Theorem A2.6.III in Daley and Vere-Jones (2003, p. 404), for any $A \in \mathcal{B}(\mathbb{R})$, $M' \in \mathcal{B}(\mathscr{M})$, $M \in \mathcal{B}(\mathscr{M})$, the function

$$(m, \xi) \mapsto \iint_{(-\infty,0) \times \mathscr{M}} \mathbb{1}_A(-t')\mathbb{1}_{M'}(m')\mathbb{1}_M(m)\xi(dt', dm') = \mathbb{1}_M(m)\xi((-A) \cap (-\infty, 0) \times M')$$

is measurable. Then, by the monotone class theorem (Daley and Vere-Jones, 2003, p. 369), we have that $\mathcal{B}(\mathbb{R} \times \mathscr{M} \times \mathscr{M}) = \sigma(\mathcal{R}) \subset \mathcal{C}$. The linearity of the integral implies that $(m, \xi) \mapsto I(m, \xi)$ is measurable for all simple functions k and, by monotone convergence, for all non-negative measurable functions k (Kallenberg, 2002, Lemma 1.11, p. 7). $\square$

Acknowledgements

References

Bacry, E., Jaisson, T., and Muzy, J.-F. (2016). Estimation of slowly decreasing Hawkes kernels: application to high-frequency order book dynamics. *Quantitative Finance*, 16(8):1179–1201.

Bacry, E., Mastromatteo, I., and Muzy, J.-F. (2015). Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(1):1550005, 59 pages.

Blundell, C., Beck, J., and Heller, K. A. (2012). Modelling reciprocating relationships with Hawkes processes. In Pereira, F., Burges, C. J. C., Bottou, L., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 25*, pages 2600–2608.

Bowsher, C. G. (2007). Modelling security market events in continuous time: intensity based, multivariate point process models. *Journal of Econometrics*, 141(2):876–912.

Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics*. Springer, New York.

Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *Annals of Probability*, 24(3):1563–1588.

Cartea, A., Donnelly, R., and Jaimungal, S. (2018). Enhancing trading strategies with order book signals. *Applied Mathematical Finance*, 25(1):1–35.

Chevallier, J. (2017). Mean-field limit of generalized Hawkes processes. *Stochastic Processes and their Applications*, 127(12):3870–3912.

Chevallier, J., Caceres, M. J., Doumic, M., and Reynaud-Bouret, P. (2015). Microscopic approach of a time elapsed neural model. *Mathematical Models and Methods in Applied Sciences*, 25(14):2669–2719.

Chornoboy, E. S., Schramm, L. P., and Karr, A. F. (1988). Maximum likelihood identification of neural point process systems. *Biological Cybernetics*, 59(4):265–275.

Çınlar, E. (2011). *Probability and Stochastics*. Springer, New York.

Cohen, S. N. and Elliott, R. J. (2013). Filters and smoothers for self-exciting Markov modulated counting processes. Preprint, available at: http://arxiv.org/abs/1311.6257.

Cont, R. and de Larrard, A. (2013). Price dynamics in a Markovian limit order market. *SIAM Journal on Financial Mathematics*, 4(1):1–25.

Cont, R., Stoikov, S., and Talreja, R. (2010). A stochastic model for order book dynamics. *Operations Research*, 58(3):549–563.

Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes. Vol. I*. Springer, New York, second edition.

Daley, D. J. and Vere-Jones, D. (2008). *An Introduction to the Theory of Point Processes. Vol. II*. Springer, New York, second edition.

Delattre, S., Fournier, N., and Hoffmann, M. (2016). Hawkes processes on large networks. *Annals of Applied Probability*, 26(1):216–261.

Farajtabar, M., Yang, J., Ye, X., Xu, H., Trivedi, R., Khalil, E., Li, S., Song, L., and Zha, H. (2017). Fake news mitigation via point process based intervention. In Precup, D. and Teh, Y. W., editors, *Proceedings of Machine Learning Research 70*, pages 1097–1106.

Fox, E. W., Schoenberg, F. P., and Gordon, J. S. (2016). Spatially inhomogeneous background rate estimators and uncertainty quantification for nonparametric Hawkes point process models of earthquake occurrences. *Annals of Applied Statistics*, 10(3):1725–1756.

Gerhard, F., Deger, M., and Truccolo, W. (2017). On the stability and dynamics of stochastic spiking neuron models: Nonlinear Hawkes process and point process GLMs. *PLoS Computational Biology*, 13(2):e1005390, 31 pages.

Gonzalez, F. and Schervish, M. (2017). Instantaneous order impact and high-frequency strategy optimization in limit order books. *Market Microstructure and Liquidity*, 3(2):1850001, ?? pages.

Grigelionis, B. (1971). On the representation of integer valued measures by means of stochastic integrals with respect to Poisson measures. *Litovskij Matematicheskij Sbornik*, 11:93–108.

Hawkes, A. G. (1971). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 33(3):438–443.

Hawkes, A. G. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability*, 11(3):493–503.

Huang, W., Lehalle, C.-A., and Rosenbaum, M. (2015). Simulating and analyzing order book data: The queue-reactive model. *Journal of the American Statistical Association*, 110(509):107–122.

Huang, W. and Rosenbaum, M. (2017). Ergodicity and diffusivity of Markovian order book models: a general framework. *SIAM Journal on Financial Mathematics*, 8(1):874–900.

Jacobsen, M. (2006). *Point Process theory and Applications: Marked Point and Piecewise Deterministic Processes*. Birkhäuser, Boston.

Jacod, J. (1975). Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 31(3):235–253.

Jacod, J. (1979). *Calcul stochastique et problèmes de martingales*. Springer, Berlin.

Jaisson, T. and Rosenbaum, M. (2016). Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *Annals of Applied Probability*, 26(5):2860–2882.

Kallenberg, O. (2002). *Foundations of Modern Probability*. Springer, New York, second edition.

Kallenberg, O. (2017). *Random Measures, Theory and Applications*. Springer, Cham.

Kerstan, J. (1964). Teilprozesse Poissonscher Prozese. In *Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions and Random Processes (Liblice, 1962)*, pages 377–403.

Large, J. (2007). Measuring the resiliency of an electronic limit order book. *Journal of Financial Markets*, 10(1):1–25.

Last, G. and Brandt, A. (1995). *Marked Point Processes on the Real Line: The Dynamical Approach*. Springer, New York.

Laub, P. J., Taimre, T., and Pollett, P. K. (2015). Hawkes processes. Preprint, available at: http://arxiv.org/abs/1507.02822.

Lewis, E., Mohler, G., Brantingham, P. J., and Bertozzi, A. L. (2012). Self-exciting point process models of civilian deaths in Iraq. *Security Journal*, 25(3):244–264.

Lewis, P. A. W. and Shedler, G. S. (1976). Simulation of nonhomogeneous Poisson processes with log linear rate function. *Biometrika*, 63(3):501–505.

Lindvall, T. (1988). Ergodicity and inequalities in a class of point processes. *Stochastic Processes and their Applications*, 30(1):121–131.

Loeffler, C. and Flaxman, S. (2018). Is gun violence contagious? A spatiotemporal test. *Journal of Quantitative Criminology*, 34(4):999–1017.

Massoulié, L. (1998). Stability results for a general class of interacting point processes dynamics, and applications. *Stochastic Processes and their Applications*, 75(1):1–30.

Mohler, G. (2013). Modeling and estimation of multi-source clustering in crime and security data. *Annals of Applied Statistics*, 7(3):1525–1539.

Morariu-Patrichi, M. (2018). On the weak-hash metric for boundedly finite integer-valued measures. *Bulletin of the Australian Mathematical Society*, 98(2):265–276.

Morariu-Patrichi, M. and Pakkanen, M. S. (2018). State-dependent Hawkes processes and their application to limit order book modelling. Preprint, available at: http://arxiv.org/abs/1809.08060.

Ogata, Y. (1981). On Lewis' simulation method for point processes. *IEEE Transactions on Information Theory*, 27(1):23–31.

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, 50(2):379–402.

Rambaldi, M., Bacry, E., and Lillo, F. (2017). The role of volume in order book dynamics: a multivariate Hawkes process analysis. *Quantitative Finance*, 17(7):999–1020.

Sigman, K. (1995). *Stationary Marked Point Processes: An Intuitive Approach*. Chapman & Hall, New York.

Sokol, A. and Hansen, N. R. (2015). Exponential martingales and changes of measure for counting processes. *Stochastic Analysis and Applications*, 33(5):823–843.

Swishchuk, A. (2017). General compound Hawkes processes in limit order books. Preprint, available at: http://arxiv.org/abs/1706.07459.

Torrisi, G. L. (2016). Gaussian approximation of nonlinear Hawkes processes. *Annals of Applied Probability*, 26(4):2106–2140.

Türkyilmaz, K., van Lieshout, M. N. M., and Stein, A. (2013). Comparing the Hawkes and trigger process models for aftershock sequences following the 2005 Kashmir earthquake. *Mathematical Geosciences*, 45(2):149–164.

Vinkovskaya, E. (2014). *A Point Process Model for the Dynamics of Limit Order Books*. PhD thesis, Columbia University.

Zhou, K., Zha, H., and Song, L. (2013). Learning social infectivity in sparse low-rank networks using multi-dimensional Hawkes processes. *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, pages 641–649.