Functional inequalities for marked point processes

Ian Flint[*]   Nicolas Privault[†]   Giovanni Luca Torrisi[‡]

Abstract

In recent years, a number of functional inequalities have been derived for Poisson random measures, with a wide range of applications. In this paper, we prove that such inequalities can be extended to the setting of marked temporal point processes, under mild assumptions on their Papangelou conditional intensity. First, we derive a Poincaré inequality. Second, we prove two transportation cost inequalities. The first one refers to functionals of marked point processes with a Papangelou conditional intensity and is new even in the setting of Poisson random measures. The second one refers to the law of marked temporal point processes with a Papangelou conditional intensity, and extends a related inequality which is known to hold on a general Poisson space. Finally, we provide a variational representation of the Laplace transform of functionals of marked point processes with a Papangelou conditional intensity. The proofs make use of an extension of the Clark-Ocone formula to marked temporal point processes. Our results are shown to apply to classes of renewal, nonlinear Hawkes and Cox point processes.

1 Introduction

Point processes with a Papangelou conditional intensity ([6], [9], [18], [20], [21]) constitute an important class of point process models, which generalizes the Poisson process. Roughly speaking, the intuitive meaning of this notion of conditional intensity, denoted by $\pi_x(\omega)$, is that, for a suitable state space S and reference measure σ on S, $\pi_x(\omega)\,\sigma(\mathrm{d}x)$, $x \in S$, is the conditional probability of having a particle in the infinitesimal region $\mathrm{d}x$ when the configuration ω is given outside $\mathrm{d}x$.

[*]Division of Mathematical Sciences, Nanyang Technological University, SPMS-MAS-04-02, 21 Nanyang Link Singapore 637371.

[†]Division of Mathematical Sciences, Nanyang Technological University, SPMS-MAS-05-43, 21 Nanyang Link Singapore 637371.

E-mail: nprivault@ntu.edu.sg

[‡]Istituto per le Applicazioni del Calcolo "Mauro Picone", CNR, Via dei Taurini 19, 00185 Roma, Italy.

E-mail: giovanniluca.torrisi@cnr.it

In this paper we provide several functional inequalities for marked temporal point processes having a Papangelou conditional intensity, with times in $\mathbb{R}_+$ and marks in a complete separable metric space E.

Our main achievements are (i) a Poincaré inequality for square-integrable functionals of marked point processes with a Papangelou conditional intensity (Theorem 3.1); (ii) transportation cost inequalities for the law of functionals of marked point processes with a Papangelou conditional intensity (Theorem 3.12) and for the law of the marked point process itself (Theorem 3.20); (iii) a variational representation of the Laplace transform of functionals, bounded from above, of marked point processes with a Papangelou conditional intensity (Theorem 3.26).

The Poincaré inequality and variational representations of the Laplace transform for functionals on the Poisson space have attracted a lot of interest (see [16], [22], [27] for the Poincaré inequality and [4], [29] for variational representations of the Laplace transform). Our results in this direction show that such functional relations hold in a non-Poissonian setting. We emphasize that the Poincaré inequality proved in this paper concerns one-dimensional marked point processes and it holds under different conditions than the Poincaré inequality for Gibbs point processes provided in [13] (we refer the reader to Remark 3.6 for a more detailed discussion). On the one hand, to the best of our knowledge, the transportation cost inequality for functionals of marked point processes with a Papangelou conditional intensity is new even in the Poisson setting. On the other hand, the transportation cost inequality for the law of the marked point process itself generalizes a related relation proved in [19].

A key ingredient in the proofs is a new Clark-Ocone formula for square-integrable functionals of marked point processes with a Papangelou conditional intensity (Theorem 3.28), which generalizes the corresponding formula in [8] in two directions. First, we allow for point processes with values on an unbounded time interval and marks in a complete separable metric space. Second, we prove the square integrability of the Clark-Ocone integrand, which is crucial in the proofs of the functional inequalities mentioned above since it enables the application of the isometry formula provided by Proposition 2.6.

Deviation bounds, which are a classical application of transportation cost inequalities, are presented in Remarks 3.14 and 3.21. The variational representation of the Laplace transform for functionals of marked point processes with a Papangelou conditional intensity can also be useful to derive large deviation principles for those functionals, along the lines of [4] and more generally relying on the theory developed in [7]. Further applications of our results to various classes of non-Poissonian point processes such as renewal, nonlinear Hawkes and Cox are presented in Corollaries 3.7, 3.9, 3.11, 3.17, 3.18, 3.19, 3.22, 3.23, 3.24 and in Remark 3.27.

The proof of the Poincaré inequality is based on the evaluation of the variance of the functional by a combination of the Clark-Ocone formula and the isometry formula for marked point processes with a stochastic intensity. The proof of the transportation cost inequality exploits its characterization via exponential moments proved in [10]. Such exponential moments are controlled by a stochastic convex inequality for functionals of marked point processes with a Papangelou conditional intensity (Proposition 4.2), which is based on the Clark-Ocone formula and generalizes the corresponding result in [12]. The proof of the variational representation of the Laplace transform uses a localization argument to deal with the unbounded case, which is out of the reach of the techniques in [29], and relies on the Clark-Ocone formula in order to take into account the non-Poissonian dynamics of the point process. The proof of the Clark-Ocone formula is based on the representation theorem for square-integrable martingales, on the use of an integration by parts formula for functionals of point processes with a Papangelou

conditional intensity, and on an isometry formula for point processes with stochastic intensity, which allows us to identify the integrand appearing in the representation theorem.

We note once again that, in contrast to the corresponding result in [8], the Clark-Ocone formula of Theorem 3.28 guarantees the square integrability of the integrand (3.36). This integrability property is crucial to the proofs of our main results due to a pervasive use of the isometry formula provided by Proposition 2.6.

The paper is organized as follows. In Section 2 we give some preliminaries on point processes including the notions of Papangelou conditional intensity, classical stochastic intensity and an important relation between them. In Section 3 we describe the main results of the paper and give their proofs in Section 4. We also include an appendix, where we prove a technical lemma and a technical proposition.

2 Preliminaries on point processes

Let E be a complete separable metric space and $\mathcal{E}$ the corresponding Borel σ-field. Let Ω be the set of all integer-valued measures ω on $(\mathbb{R}_+ \times E, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E})$, where $\mathcal{B}(\mathbb{R}_+)$ is the Borel σ-field on $\mathbb{R}_+$, such that $\omega(\{0\} \times E) = 0$, $\omega(\{t\} \times E) \leq 1$ for any $t \in \mathbb{R}_+$, and $\omega(K \times E) < \infty$ for any compact set $K \subset \mathbb{R}_+$ (in Remark 2.1, we shall explain in what sense our results apply under a more general definition of Ω). We define

$$N(\omega) := \omega, \quad \omega \in \Omega$$

and for a Borel set $B \in \mathcal{B}(\mathbb{R}_+)$ we shall consider the σ-field

$$\mathcal{F}_B := \sigma\{N(A \times D) : A \in \mathcal{B}(B), D \in \mathcal{E}\},$$

where $\mathcal{B}(B)$ denotes the restriction of $\mathcal{B}(\mathbb{R}_+)$ to B. For ease of notation we set $\mathcal{F}_t := \mathcal{F}_{[0,t]}$ and $\mathcal{F}_{t^-} := \mathcal{F}_{[0,t)} = \bigvee_{0 \leq s < t} \mathcal{F}_s$ for $t \in \mathbb{R}_+$.

We set $\mathcal{F}_\infty := \bigvee_{t \in \mathbb{R}_+} \mathcal{F}_t$, let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F}_\infty)$ and consider the canonical probability space $(\Omega, \mathcal{F}_\infty, \mathbb{P})$. The elements of $(\Omega, \mathcal{F}_\infty, \mathbb{P})$ are known in the literature as simple and locally finite marked point processes on $\mathbb{R}_+$ with marks in E. Throughout the paper we denote by $\mathbb{E}$ and $\mathbb{V}\mathrm{ar}$ the expectation and the variance operators with respect to $\mathbb{P}$, respectively.

By analogy with the un-marked setting, a mapping $X : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}$ which is measurable with respect to the σ-fields $(\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_\infty \otimes \mathcal{E}, \mathcal{B}(\mathbb{R}))$, where $\mathcal{B}(\mathbb{R})$ is the Borel σ-field on $\mathbb{R}$, is called a stochastic process. Let $\mathcal{G} := \{\mathcal{G}_t\}_{t \in \mathbb{R}_+}$ be a filtration such that $\mathcal{F}_\infty \supset \mathcal{G}_t \supseteq \mathcal{F}_t$, $t \in \mathbb{R}_+$. The $\mathcal{G}$-predictable σ-field on $\mathbb{R}_+ \times \Omega$, denoted by $\mathcal{P}(\mathcal{G})$, is the σ-field generated by the sets $(a, b] \times A$ with $A \in \mathcal{G}_a$, $a, b \in \mathbb{R}_+$. Throughout the paper, for any $\omega \in \Omega$ and $(t, x) \in \mathbb{R}_+ \times E$, we define $\omega - \varepsilon_{(t,x)} := \omega$ if $(t, x) \notin \mathrm{Supp}(\omega)$, where $\varepsilon_{(t,x)}$ is the Dirac measure at (t, x) and $\mathrm{Supp}(\omega)$ denotes the support of ω. A stochastic process $X : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}$ is said to be $\mathcal{G}$-predictable if it is measurable with respect to the σ-fields $(\mathcal{P}(\mathcal{G}) \otimes \mathcal{E}, \mathcal{B}(\mathbb{R}))$. We say that X is predictable if it is measurable with respect to the σ-fields $(\mathcal{P}(\mathcal{F}) \otimes \mathcal{E}, \mathcal{B}(\mathbb{R}))$, where $\mathcal{F} := \{\mathcal{F}_t\}_{t \in \mathbb{R}_+}$. For ease of notation, we set $X_{(t,x)}(\omega) := X(t, \omega, x)$ and for later purposes, we mention that if X is predictable, then for fixed $t \in \mathbb{R}_+$ and $x \in E$ the random variable $X_{(t,x)}$ is measurable with respect to $\mathcal{F}_{t^-}$ (and therefore with respect to $\mathcal{F}_t$), and so $X_{(t,x)}(\omega) = X_{(t,x)}(\omega - \varepsilon_{(t,x)})$. This claim follows by an obvious modification of the proof of Lemma A3.3.I p. 425 in [5], see also Proposition 3.3 in [15].

We shall consider two different notions of conditional intensity for marked point processes: the Papangelou conditional intensity and the classical stochastic intensity.

2.1 Marked point processes with a Papangelou conditional intensity

Let ν denote a locally finite measure on $(E, \mathcal{E})$. A non-negative stochastic process $\pi : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}_+$ is said to be a Papangelou conditional intensity of N with respect to $\mathrm{d}t\nu(\mathrm{d}x)$ if, for any non-negative stochastic process $X : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}_+$,

$$\mathbb{E}\left[\int_{\mathbb{R}_+ \times E} X_{(t,x)}(N - \varepsilon_{(t,x)}) \, N(\mathrm{d}t \times \mathrm{d}x)\right] = \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} X_{(t,x)}(N) \, \pi_{(t,x)}(N) \, \mathrm{d}t\nu(\mathrm{d}x)\right].$$
(2.1)

Intuitively, $\pi_{(t,x)}(\omega) \, \mathrm{d}t\nu(\mathrm{d}x)$ is the probability that the point process has a point in the infinitesimal region $\mathrm{d}t\mathrm{d}x$ given that it agrees with the configuration ω outside of $\mathrm{d}t\mathrm{d}x$. Point processes with a Papangelou conditional intensity are fully characterized in [20] (see Section 3) and [21] (see Theorem 2').

Remark 2.1. Some texts (e.g. [5, Definitions 6.4.I.]) use a more general definition of a marked point process which allows for N to have atoms $\varepsilon_{(t,x)}, \varepsilon_{(t,y)}$ for the same $t \in \mathbb{R}_+$ and $x \neq y$. We considered a more restrictive set Ω in Section 2 since all our main results hold for marked point processes with a Papangelou conditional intensity. In such case, there exists a version of the more general marked point process which takes its values in the set Ω, see the following Lemma 2.2.

Lemma 2.2. *Let Ω' be the set of all integer-valued measures ω on $(\mathbb{R}_+ \times E, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E})$, such that $\omega(\{0\} \times E) = 0$, $\omega(\{t\} \times \{x\}) \leq 1$ for any $(t,x) \in \mathbb{R}_+ \times E$, and $\omega(K \times E) < \infty$ for any compact set $K \subset \mathbb{R}_+$ (note that $\Omega \subset \Omega'$). Assume that the marked point process N is defined on Ω' instead of Ω and define the quantities from Section 2 in an analogous manner. In particular, if $\mathbb{P}$ is a probability measure on $(\Omega', \mathcal{F}_\infty)$, where $\mathcal{F}_\infty$ is appropriately defined and N has a Papangelou conditional intensity, then N takes its values in Ω $\mathbb{P}$-almost surely, i.e. $\mathbb{P}(\Omega) = 1$.*

We postpone the proof of this lemma to Section 5.1 in the appendix in order to improve the flow of the article.

Throughout this paper we shall often consider locally stable point processes, i.e. point processes N with a Papangelou conditional intensity π such that

$$\pi_{(s,x)}(\omega) \leq \beta(s,x), \quad \mathrm{d}s\nu(\mathrm{d}x)\mathbb{P}(\mathrm{d}\omega)\text{-almost everywhere}$$
(2.2)

for a function $\beta : \mathbb{R}_+ \times E \to \mathbb{R}$ which is integrable with respect to $\mathrm{d}s\mathrm{d}\nu$, on $[0,t] \times E$ for every $t \in \mathbb{R}_+$. The local stability is known to be satisfied by a wide range of point processes (see e.g. [18]).

In [17], the author provides a condition guaranteeing the existence of the predictable projection of a bounded stochastic process. In the following proposition, we specialize this result to our setting while relaxing the boundedness assumption.

Proposition 2.3. *Assume that N has a Papangelou conditional intensity π, i.e. (2.1) holds. Let $X : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}$ be a stochastic process which is such that either (i) $X \geq 0$ or (ii) for $\mathrm{d}t\nu(\mathrm{d}x)$-almost all $(t,x) \in \mathbb{R}_+ \times E$, $X_{(t,x)} \in L^1(\Omega, \mathcal{F}_\infty, \mathbb{P})$. Then there exists a predictable stochastic process $p(X) : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}$ called predictable projection, which is such that for $\mathrm{d}t\nu(\mathrm{d}x)$-almost all $(t,x) \in \mathbb{R}_+ \times E$, we have*

$$p(X)_{(t,x)} = \mathbb{E}\big[X_{(t,x)}\big|\mathcal{F}_{t^-}\big], \quad \mathbb{P}\text{-almost surely.}$$

Additionally, under (ii) the predictable projection $p(X)$ is such that, for $\mathrm{d}t\nu(\mathrm{d}x)$-almost all $(t,x) \in \mathbb{R}_+ \times E$,

$$p(X)_{(t,x)} < \infty, \quad \mathbb{P}\text{-almost surely.}$$

The proof is rather technical and therefore postponed to Section 5.2 in the appendix.

Throughout this paper we will use the discrete Malliavin derivative of $F : \Omega \to \mathbb{R}$, defined as

$$D_{(t,x)}F(\omega) := F^+_{(t,x)}(\omega) - F(\omega), \quad (t,x) \in \mathbb{R}_+ \times E, \ \omega \in \Omega,$$

where

$$F^+_{(t,x)}(\omega) := F(\omega + \varepsilon_{(t,x)}) = \begin{cases} F(\omega) & \text{if } (t,x) \in \text{Supp}(\omega), \\ F(\omega + \varepsilon_{(t,x)}) & \text{if } (t,x) \notin \text{Supp}(\omega). \end{cases}$$

Under suitable integrability conditions on X and π, we shall consider the stochastic integral

$$\Delta(X) := \int_{\mathbb{R}_+ \times E} X_{(t,x)} \left(N(\mathrm{d}t \times \mathrm{d}x) - \pi_{(t,x)} \, \mathrm{d}t\nu(\mathrm{d}x) \right), \quad \mathbb{P}\text{-almost surely.}$$

If X is predictable, then $X_{(t,x)} = X_{(t,x)}(N) = X_{(t,x)}(N - \varepsilon_{(t,x)})$, and so this integral can be rewritten as

$$\Delta(X) = \int_{\mathbb{R}_+ \times E} X_{(t,x)}(N - \varepsilon_{(t,x)}) \left(N(\mathrm{d}t \times \mathrm{d}x) - \pi_{(t,x)} \, \mathrm{d}t\nu(\mathrm{d}x) \right).$$

We conclude this paragraph with the following integration by parts formula, see Corollary 3.1 in [25].

Lemma 2.4. *Assume that N has a Papangelou conditional intensity π. Then, for any predictable stochastic process $X : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}$ and any random variable $G : \Omega \to \mathbb{R}$ such that*

$$\mathbb{E}\left[\int_{\mathbb{R}_+ \times E} |GX_{(t,x)}| \pi_{(t,x)} \, \mathrm{d}t\nu(\mathrm{d}x) \right] < \infty \ \text{ and } \ \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} |X_{(t,x)} D_{(t,x)} G| \pi_{(t,x)} \, \mathrm{d}t\nu(\mathrm{d}x) \right] < \infty,$$

(2.3)

we have

$$\mathbb{E}\left[G\Delta(X) \right] = \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} X_{(t,x)} \pi_{(t,x)} D_{(t,x)} G \, \mathrm{d}t\nu(\mathrm{d}x) \right].$$

2.2 Marked point processes with a classical stochastic intensity

A non-negative and $\mathcal{G}$-predictable stochastic process $\lambda : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}_+$ such that for all $t \in \mathbb{R}_+$ $\int_{[0,t] \times E} \lambda_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x) < \infty$ $\mathbb{P}$-almost surely is said to be a $\mathcal{G}$-stochastic intensity of N (see [1], [5] and [14]) if for any non-negative and $\mathcal{G}$-predictable stochastic process $X : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}_+$ we have

$$\mathbb{E}\left[\int_{\mathbb{R}_+ \times E} X_{(t,x)} N(\mathrm{d}t \times \mathrm{d}x) \right] = \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} X_{(t,x)} \lambda_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x) \right]. \tag{2.4}$$

Additionally, we call λ a (classical) stochastic intensity if $\mathcal{G} := \mathcal{F}$. Roughly speaking, the quantity $\lambda_{(t,x)}(\omega) \, \mathrm{d}t\nu(\mathrm{d}x)$ is the probability that the point process has a point in the infinitesimal region $\mathrm{d}t\mathrm{d}x$ given that it agrees with the configuration ω on $[0,t) \times E$.

Hereafter, we assume that N has a $\mathcal{G}$-stochastic intensity λ and, given two $\mathcal{G}$-predictable stochastic processes $X, Y : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}$, we write $X \sim Y$ if X and Y are equal $\lambda_{(t,x)}(\omega) \, \mathrm{d}t\mathbb{P}(\mathrm{d}\omega)\nu(\mathrm{d}x)$-almost everywhere on $\mathbb{R}_+ \times \Omega \times E$.

For $p \in [1, +\infty)$, we denote by $\mathcal{P}_p(\lambda)$ the family of equivalence classes (with respect to the equivalence relation $\sim$) formed by $\mathcal{G}$-predictable stochastic processes X such that

$$\|X\|^p_{\mathcal{P}_p(\lambda)} := \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} |X_{(t,x)}|^p \lambda_{(t,x)} \, \mathrm{d}t\nu(\mathrm{d}x) \right] < \infty$$

and, for ease of notation, we set $\mathcal{P}_{1,2}(\lambda) := \mathcal{P}_1(\lambda) \cap \mathcal{P}_2(\lambda)$. Note that $\| \cdot \|_{\mathcal{P}_p(\lambda)}$ is a norm on $\mathcal{P}_p(\lambda)$.

For any $X \in \mathcal{P}_1(\lambda)$ we define the stochastic integral

$$\delta(X) := \int_{\mathbb{R}_+ \times E} X_{(t,x)} \left(N(\mathrm{d}t \times \mathrm{d}x) - \lambda_{(t,x)} \, \mathrm{d}t\nu(\mathrm{d}x) \right), \quad \mathbb{P}\text{-almost surely,}$$

which is well defined $\mathbb{P}$-almost everywhere as the difference of two finite terms by (2.4). The next proposition provides a fundamental isometry formula for marked point processes with a $\mathcal{G}$-stochastic intensity.

Proposition 2.5 (Theorem 3 of [3] and p. 62 of [11])**.** *Assume that N has a $\mathcal{G}$-stochastic intensity λ. Then:*

(i) $\mathbb{E}\big[\delta(X)\big] = 0$, *for any* $X \in \mathcal{P}_1(\lambda)$;

(ii) $\mathbb{E}\big[\delta(X)\delta(Y)\big] = \mathbb{E}\big[\int_{\mathbb{R}_+ \times E} X_{(t,x)}Y_{(t,x)}\lambda_{(t,x)} \, \mathrm{d}t\nu(\mathrm{d}x)\big]$, *for any* $X, Y \in \mathcal{P}_{1,2}(\lambda)$;

(iii) *If* $\{\mathbb{1}_{[0,t]}(s)X_{(s,x)}\}_{(s,x)\in\mathbb{R}_+\times E}, \{\mathbb{1}_{[0,t]}(s)Y_{(s,x)}\}_{(s,x)\in\mathbb{R}_+\times E} \in \mathcal{P}_{1,2}(\lambda)$ *for any* $t \in \mathbb{R}_+$, *then the stochastic process*

$$\delta(\mathbb{1}_{[0,t]}X)\delta(\mathbb{1}_{[0,t]}Y) - \int_{[0,t]\times E} X_{(s,x)}Y_{(s,x)}\lambda_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x), \quad t \in \mathbb{R}_+$$

is a $\mathcal{G}$-martingale.

For later purposes, we extend the operator δ to $\mathcal{P}_1(\lambda) \cup \mathcal{P}_2(\lambda)$ and prove that the isometry formula of Proposition 2.5-*(ii)* holds on $\mathcal{P}_2(\lambda)$ for the extension of δ.

Proposition 2.6. *Assume that N has a $\mathcal{G}$-stochastic intensity λ and that $\mathbb{E}\big[N([0,t] \times E)\big] < \infty$, for all $t \in \mathbb{R}_+$. The operator δ with domain $\mathcal{P}_1(\lambda)$ can be uniquely extended to an operator with domain $\mathcal{P}_1(\lambda) \cup \mathcal{P}_2(\lambda)$, which we still denote by δ. Additionally, for any $X, Y \in \mathcal{P}_2(\lambda)$ we have*

$$\mathbb{E}\big[\delta(X)\delta(Y)\big] = \mathbb{E}\bigg[\int_{\mathbb{R}_+ \times E} X_{(t,x)}Y_{(t,x)}\lambda_{(t,x)} \, \mathrm{d}t\nu(\mathrm{d}x)\bigg].$$

The proof of this proposition follows a standard Cauchy sequence argument, along the lines of Section II.2 in [11]. For this reason, it is omitted.

2.3 A relation between the Papangelou conditional intensity and the classical stochastic intensity

In the next lemma we show that the Papangelou conditional intensity of a marked point process N determines its stochastic intensity.

Lemma 2.7. *If N has a Papangelou conditional intensity π and*

$$\int_{[0,t]\times E} p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x) < \infty, \quad \mathbb{P}\text{-almost surely, for all } t \in \mathbb{R}_+ \tag{2.5}$$

then N has a stochastic intensity $p(\pi)$.

Proof. Let X be a non-negative and predictable stochastic process. As recalled at the beginning of Section 2, we have $X_{(t,x)}(\omega) = X_{(t,x)}(\omega - \varepsilon_{(t,x)})$ and $X_{(t,x)}$ is $\mathcal{F}_{t^-}$-measurable, for any $t \in \mathbb{R}_+$ and $x \in E$. So by Fubini's theorem, standard properties of the conditional expectation and Proposition 2.3-*(i)*,

$$\mathbb{E}\left[\int_{\mathbb{R}_+ \times E} X_{(t,x)} \, N(\mathrm{d}t \times \mathrm{d}x)\right] = \int_{\mathbb{R}_+ \times E} \mathbb{E}\big[X_{(t,x)}\pi_{(t,x)}(N)\big] \, \mathrm{d}t\nu(\mathrm{d}x)$$

$$= \mathbb{E}\bigg[\int_{\mathbb{R}_+ \times E} X_{(t,x)}p(\pi)_{(t,x)} \, \mathrm{d}t\nu(\mathrm{d}x)\bigg]. \qquad \square$$

As already mentioned, the proofs of our main results are based on a Clark-Ocone formula for marked point processes, which, in turn, exploits the representation theorem of square integrable martingales (see e.g. [1]). For this reason, we shall consider the $\mathbb{P}$-completed filtration $\overline{\mathcal{F}} \equiv \{\overline{\mathcal{F}}_t\}_{t \in \mathbb{R}_+}$, defined by $\overline{\mathcal{F}}_t := \mathcal{F}_t \vee \mathcal{N} = \sigma(\{A \cup H : A \in \mathcal{F}_t, H \in \mathcal{N}\})$, where $\mathcal{N}$ is the family of $\mathbb{P}$-null events of $\mathcal{F}_\infty$.

The next lemma guarantees that the notion of stochastic intensity is equivalent to that of $\overline{\mathcal{F}}$-stochastic intensity. Its proof is standard, and therefore omitted.

Lemma 2.8. *Let $\lambda : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}_+$ be a predictable stochastic process. Then N has a stochastic intensity λ if and only if N has a $\overline{\mathcal{F}}$-stochastic intensity λ.*

3 Main results

In this section we state our main achievements, i.e. a Poincaré inequality for square-integrable functionals of marked point processes with a Papangelou conditional intensity, a transportation cost inequality for functionals of marked point processes with a Papangelou conditional intensity, a transportation cost inequality for the law of a marked point process with a Papangelou conditional intensity, and a variational representation formula for the Laplace transform of (bounded from above) functionals of marked point processes with a Papangelou conditional intensity. All these functional relations are obtained by applying a Clark-Ocone formula for square-integrable functionals of space-time point processes with a Papangelou conditional intensity, which generalizes in various directions the corresponding formula in [8].

Hereafter, we work under the convention $0/0 := 0$. Moreover, in order to be more precise, from now on we shall write $\mathcal{P}_p^{\mathcal{G}}(\lambda)$ in place of $\mathcal{P}_p(\lambda)$, $p \geq 1$, to stress that the stochastic processes therein are predictable with respect to some specific filtration $\mathcal{G}$.

3.1 Poincaré inequality

The following Poincaré inequality holds for functionals of marked point processes with a Papangelou conditional intensity.

Theorem 3.1. *Assume*

$$\mathbb{E}\big[N([0,t] \times E)^2\big] < \infty, \quad \text{for all } t \in \mathbb{R}_+, \tag{3.1}$$

and that N has a Papangelou conditional intensity π such that

$$\mathbb{E}\left[\left(\int_{[0,t] \times E} \pi_{(s,x)} \, ds\nu(dx)\right)^2\right] < \infty, \quad \text{for all } t \in \mathbb{R}_+, \tag{3.2}$$

and

$$\gamma := \left\|\int_{\mathbb{R}_+ \times E} \frac{\mathbb{V}\mathrm{ar}\big[\pi_{(t,x)} \mid \mathcal{F}_{t-}\big]}{\mathbb{E}\big[\pi_{(t,x)} \mid \mathcal{F}_{t-}\big]} \, dt\nu(dx)\right\|_{L^\infty(\Omega, \mathcal{F}_\infty, \mathbb{P})} < 1, \tag{3.3}$$

where $\mathbb{V}\mathrm{ar}\big[\pi_{(t,x)}\big|\mathcal{F}_{t-}\big] := \mathbb{E}\big[\pi_{(t,x)}^2\big|\mathcal{F}_{t-}\big] - \mathbb{E}\big[\pi_{(t,x)}\big|\mathcal{F}_{t-}\big]^2$. Then, for any $G \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$ we have

$$\mathbb{V}\mathrm{ar}(G) \leq \big(1 - \sqrt{\gamma}\big)^{-2} \, \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \pi_{(t,x)} \big|D_{(t,x)}G\big|^2 \, dt\nu(dx)\right]. \tag{3.4}$$

Remark 3.2. Note that if $\mathbb{E}\big[\pi_{(t,x)}\big|\mathcal{F}_{t-}\big] = 0$ $\mathbb{P}$-almost surely, then $\mathbb{P}(\pi_{(t,x)} = 0 \mid \mathcal{F}_{t-}) = 1$ $\mathbb{P}$-almost surely. Therefore $\mathbb{V}\mathrm{ar}\big[\pi_{(t,x)}\big|\mathcal{F}_{t-}\big] = 0$ $\mathbb{P}$-almost surely and the ratio $\mathbb{V}\mathrm{ar}[\pi_{(t,x)} \mid \mathcal{F}_{t-}]/\mathbb{E}[\pi_{(t,x)} \mid \mathcal{F}_{t-}]$ vanishes $\mathbb{P}$-almost surely by the convention $0/0 := 0$.

Remark 3.3. If N is a Poisson process with locally integrable intensity function $\sigma(t, x)$ then $\pi \equiv \sigma$ and all the assumptions of Theorem 3.1 hold. In particular, $\gamma = 0$ and we recover the well-known Poincaré inequality for Poisson functionals (see e.g. [15], [16] and [27]).

Remark 3.4. If N is locally stable in the sense of (2.2), then assumptions (3.1) and (3.2) hold. Indeed, it is clear that (2.2) implies (3.2). As far as the implication of (3.1), note that under assumption (2.2) we have

$$
\mathbb{E}\big[N([0,t] \times E)^2\big]
$$
$$
= \mathbb{E}\left[\iint_{([0,t] \times E)^2} \pi_{(u,x)}(N)\pi_{(s,y)}(N + \varepsilon_{(u,x)}) \, \mathrm{d}s\mathrm{d}u\nu(\mathrm{d}x)\nu(\mathrm{d}y)\right]
$$
$$
+ \mathbb{E}\left[\int_{[0,t] \times E} \pi_{(s,x)}(N) \, \mathrm{d}s\nu(\mathrm{d}x)\right]
$$
$$
\leq \int_{([0,t] \times E)^2} \beta(u,x)\beta(s,y) \, \mathrm{d}s\mathrm{d}u\nu(\mathrm{d}x)\nu(\mathrm{d}y) + \int_{[0,t] \times E} \beta(s,x) \, \mathrm{d}s\nu(\mathrm{d}x) < \infty, \qquad (3.5)
$$

where the first equality follows by the iterated Georgii-Nguyen-Zessin equation, see e.g. Proposition 15.5.II in [6]. We finally remark that one can check that (3.2) is verified if (3.1) holds and the marked point process is attractive (see e.g. equation (3.7) in [18] for a definition of this notion).

Remark 3.5. Let N be a point process on $\mathbb{R}_+ \times E$ with a Papangelou conditional intensity π such that (2.2) holds with dominating function β, for some measurable non-negative function $\alpha : \mathbb{R}_+ \times E \to \mathbb{R}$,

$$
p(\pi)_{(t,x)} \geq \alpha(t,x) \quad \mathrm{d}t\nu(\mathrm{d}x)\mathbb{P}(\mathrm{d}\omega)\text{-almost everywhere}
$$

and

$$
\int_{\mathbb{R}_+ \times E} (\beta(t,x) - \alpha(t,x)) \, \mathrm{d}t\nu(\mathrm{d}x) < 1. \qquad (3.6)
$$

Consequently, all assumptions of Theorem 3.1 are satisfied. In particular, we note that

$$
0 \leq \frac{\mathbb{V}\mathrm{ar}\big[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\big]}{\mathbb{E}\big[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\big]} \leq \beta(t,x) - \alpha(t,x), \quad \mathbb{P}\text{-almost surely}
$$

which, combined with (3.6), guarantees (3.3).

Remark 3.6. As mentioned in the introduction, a Poincaré inequality for Gibbs point processes was proved in [13] (see Corollaries 5.1 and 5.2 therein). Basically, Corollary 5.1 in [13] (of which Corollary 5.2 is a small improvement) states and proves the following. If N is a grand canonical Gibbs point process on $\mathbb{R}^d$ with activity parameter $z > 0$ and non-negative pair potential ϕ such that

$$
\delta := z \int_{\mathbb{R}^d} (1 - \mathrm{e}^{-\phi(x)}) \, \mathrm{d}x < 1,
$$

then, for any square-integrable functional F of the point process, we have

$$
\mathbb{V}ar(F) \leq (1 - \delta)^{-1}\mathbb{E}\left[\int_{\mathbb{R}^d} |D_x F|^2 \pi_x \, \mathrm{d}x\right],
$$

where D_x is the usual discrete Malliavin derivative,

$$
\pi_x(\omega) := z \exp(-\mathfrak{E}(x, \omega))
$$

is the Papangelou conditional intensity of N and $\mathfrak{E}$ denotes the relative energy. On the one hand, the Poincaré inequality provided by Theorem 3.1 can be applied e.g. to square-integrable functionals of renewal, non-linear Hawkes and Cox point processes, and these processes do not belong to the class of Gibbs point measures for which the inequality in [13] applies. On the other hand, we were not able to apply our Poincaré inequality to the Gibbs measures considered in [13].

The next three corollaries, whose proofs are given in Section 4, provide classes of non-Poissonian point processes which satisfy the Poincaré inequality (3.4).

Corollary 3.7 (Renewal point processes)**.** *Let N be a renewal point process on $[0, T]$, $T < \infty$, with points $T_0 := 0 < T_1 < T_2 < \cdots < T_{N([0,T])}$ and a spacing density f such that f is continuous on $[0, +\infty)$ and $f > 0$ on (a, C) for some $a \in [0, T]$ and $C \in (T, +\infty)$. Assume further that*

$$h(z) := \sup_{x \in [z, T]} \frac{f(x - z)}{f(x)} < \infty, \tag{3.7}$$

for any $z \in [0, T]$, as well as

$$\gamma := \int_0^T \sup_{z \in [0,t]} f(z) \left[\frac{\overline{F}(T - t)^2}{\overline{F}(T - t + z)} + h(z)^2 \big(\overline{F}(z) - \overline{F}(T - t + z) \big) - \frac{1}{\overline{F}(z)} \right] \mathrm{d}t < 1, \tag{3.8}$$

where $\overline{F}$ is the tail function of f. Then, for any square integrable functional G of N,

$$\mathbb{V}\mathrm{ar}(G) \leq \big(1 - \sqrt{\gamma} \big)^{-2} \mathbb{E} \left[\int_0^T \pi_t \, |D_t G|^2 \, \mathrm{d}t \right], \tag{3.9}$$

where the Papangelou conditional intensity π of N is given by

$$\pi_t = \begin{cases} f(T_i - t) f(t - T_{i-1}) \, \big/ \, f(T_i - T_{i-1}) & \text{if } T_{i-1} \leq t < T_i, \\ f(t - T_{N([0,T])}) \overline{F}(T - t) \, \big/ \, \overline{F}(T - T_{N([0,T])}) & \text{if } T_{N([0,T])} \leq t \leq T. \end{cases}$$

Next, we give some illustrating examples of renewal point processes which satisfy the assumptions of Corollary 3.7.

Example 3.8. First, assume that the spacing density function f is given by $f(x) := \lambda \exp(-\lambda x)$ for $\lambda > 0$, i.e. N is a Poisson process on $[0, T]$ with intensity $\lambda > 0$. Then all assumptions of Corollary 3.7 are satisfied. In particular $h(z) := \exp(\lambda z)$, $\pi_t \equiv \lambda$ and $\gamma = 0$.

Second, consider the Weibull spacing density function

$$f(x) := \beta x^{\beta - 1} \exp(-x^\beta), \quad \beta > 1, \ x \in \mathbb{R}_+.$$

The corresponding tail function is $\overline{F}(x) = \exp(-x^\beta)$, $x \in \mathbb{R}_+$, and since $x \mapsto x^\beta$ is Lipschitz continuous on $[0, T]$, for any $z \in [0, T]$ we have

$$h(z) = \sup_{x \in [z, T]} \left(1 - \frac{z}{x} \right)^{\beta - 1} \exp(x^\beta - (x - z)^\beta) \leq \exp(\beta T^{\beta - 1} z).$$

With the aim to check condition (3.8), we remark that by Lipschitz continuity, we have

$$\gamma \leq \int_0^T \sup_{z \in [0,t]} \beta z^{\beta - 1} \exp(-z^\beta) \Big[\exp(-2(T - t)^\beta + (T - t + z)^\beta)$$

$$+ \exp(2\beta T^{\beta - 1} z) \Big[\exp(-z^\beta) - \exp(-(T - t + z)^\beta) \Big] - \exp(z^\beta) \Big] \mathrm{d}t$$

$$\leq \int_0^T \sup_{z \in [0,t]} \beta z^{\beta - 1} \exp(-z^\beta) \Big[\exp(2\beta T^{\beta - 1} z - (T - t + z)^\beta)$$

$$+ \exp\left(2\beta T^{\beta-1}z\right)\left[\exp\left(-z^{\beta}\right) - \exp\left(-(T-t+z)^{\beta}\right)\right] - \exp\left(z^{\beta}\right)\Big] \mathrm{d}t$$

$$= \int_0^T \sup_{z\in[0,t]} \beta z^{\beta-1}\left[\exp\left(2z\left((\beta-1)T^{\beta-1} + T^{\beta-1} - z^{\beta-1}\right)\right) - 1\right]\mathrm{d}t$$

$$\leq \int_0^T \sup_{z\in[0,t]} \beta z^{\beta-1}\left[\exp\left(2z\left((\beta-1)T^{\beta-1} + (\beta-1)T^{\beta-2}(T-z)\right)\right) - 1\right]\mathrm{d}t$$

$$\leq \left[\sup_{u\in[0,T]} \exp\left(2(\beta-1)T^{\beta-2}u(2T-u)\right) - 1\right] \int_0^T \sup_{z\in[0,t]} \beta z^{\beta-1}\mathrm{d}t$$

$$= T^{\beta}\left[\exp\left(2(\beta-1)T^{\beta}\right) - 1\right] =: \gamma^*(\beta).$$

Hence for fixed $T > 0$ and β_0 satisfying the inequality $\gamma^*(\beta_0) < 1$ (in particular, note that it suffices to take β_0 close to one) the corresponding renewal point process satisfies the Poincaré inequality (3.9).

Next, consider the generalized Pareto spacing density function

$$f(x) := \lambda(1 + \xi\lambda x)^{-(1+1/\xi)}, \quad \lambda, \xi > 0, \ x \in \mathbb{R}_+,$$

whose tail function is $\overline{F}(x) = (1 + \xi\lambda x)^{-1/\xi}$, $x \in \mathbb{R}_+$. For any $z \in [0, T]$ we have

$$h(z) = \sup_{x\in[z,T]} \left(1 + \frac{\xi\lambda z}{1 + \xi\lambda(x-z)}\right)^{1+1/\xi} = (1 + \xi\lambda z)^{1+1/\xi}$$

and

$$\gamma = \int_0^T \sup_{z\in[0,t]} \lambda(1 + \xi\lambda z)^{-(1+1/\xi)}\left[\frac{(1 + \xi\lambda(T-t))^{-2/\xi}}{(1 + \xi\lambda(T-t+z))^{-1/\xi}}\right.$$

$$\left. + (1 + \xi\lambda z)^{2+2/\xi}\left((1 + \xi\lambda z)^{-1/\xi} - (1 + \xi\lambda(T-t+z))^{-1/\xi}\right) - \frac{1}{(1 + \xi\lambda z)^{-1/\xi}}\right]\mathrm{d}t$$

$$= \int_0^T \sup_{z\in[0,t]} \lambda(1 + \xi\lambda z)^{-1}\left[\left(\frac{1 + \frac{\xi\lambda(T-t)}{1+\xi\lambda z}}{1 + \xi\lambda(T-t)}\right)^{2/\xi}(1 + \xi\lambda z)^{1/\xi}(1 + \xi\lambda(T-t+z))^{-1/\xi}\right.$$

$$\left. + (1 + \xi\lambda z)^2\left(1 - (1 + \xi\lambda z)^{1/\xi}(1 + \xi\lambda(T-t+z))^{-1/\xi}\right) - 1\right]\mathrm{d}t$$

$$\leq \int_0^T \sup_{z\in[0,t]} \lambda(1 + \xi\lambda z)^{-1}\left[(1 + \xi\lambda z)^{1/\xi}(1 + \xi\lambda(T-t+z))^{-1/\xi}\right.$$

$$\left. + (1 + \xi\lambda z)^2\left(1 - (1 + \xi\lambda z)^{1/\xi}(1 + \xi\lambda(T-t+z))^{-1/\xi}\right) - 1\right]\mathrm{d}t$$

$$= \int_0^T \sup_{z\in[0,t]} \lambda(1 + \xi\lambda z)^{-1}\left(1 - (1 + \xi\lambda z)^{1/\xi}(1 + \xi\lambda(T-t+z))^{-1/\xi}\right)\left((1 + \xi\lambda z)^2 - 1\right)\mathrm{d}t$$

$$\leq \lambda \int_0^T \sup_{z\in[0,t]} \frac{\left((1 + \xi\lambda t)^2 - 1\right)}{1 + \xi\lambda z}\mathrm{d}t = \xi\lambda^2 T^2\left(1 + \frac{\xi\lambda T}{3}\right) =: \gamma^*(\xi).$$

Hence for fixed $T, \lambda > 0$ and ξ_0 satisfying the inequality $\gamma^*(\xi_0) < 1$ (in particular, note that it suffices to take ξ_0 close to zero) the corresponding renewal point process satisfies the Poincaré inequality (3.9).

A nonlinear Hawkes process on $[0, T]$ with parameters (h, ϕ) is a point process N on $[0, T]$ with stochastic intensity

$$\lambda_t := \phi\left(\int_{(0,t)} h(t-s)N(\mathrm{d}s)\right), \quad t \in [0, T],$$

where $\phi : \mathbb{R} \to \mathbb{R}_+$ and $h : [0, T] \to \mathbb{R}$ are two measurable functions, see e.g. [2], [3], [5] and [6].

Corollary 3.9 (Nonlinear Hawkes processes). *Assume that N is a nonlinear Hawkes process on $[0, T]$, $T < \infty$, with parameters (h, ϕ) such that h is non-negative and integrable on $[0, T]$, and ϕ is Lipschitz continuous and non-increasing with $\phi(0) > 0$. Additionally, assume*

$$\gamma := \phi(0) \int_0^T \left(\exp\left(2\|\phi\|_{\mathrm{Lip}} \int_0^\tau h(z) \, \mathrm{d}z \right) - 1 \right) \mathrm{d}\tau < 1,$$

where $\|\phi\|_{\mathrm{Lip}}$ denotes the Lipschitz constant of ϕ. Then, for any square-integrable functional G of N,

$$\mathbb{V}\mathrm{ar}(G) \leq \left(1 - \sqrt{\gamma} \right)^{-2} \mathbb{E}\left[\int_0^T \pi_t \, |D_t G|^2 \, \mathrm{d}t \right], \tag{3.10}$$

where the Papangelou conditional intensity π of N is given by

$$\pi_t := \lambda_t \mathrm{E}_t(N) \tag{3.11}$$

and

$$\begin{aligned}
\mathrm{E}_t(N) := &\exp\left(\int_t^T \left[\phi\left(\int_{(0,s)} h(s-u)N(\mathrm{d}u) \right) - \phi\left(h(s-t) + \int_{(0,s)} h(s-u)N(\mathrm{d}u) \right) \right] \mathrm{d}s \right) \\
&\times \exp\left(\int_t^T \left[\log \phi\left(h(s-t) + \int_{(0,s)} h(s-u)N(\mathrm{d}u) \right) \right.\right. \\
&\qquad\qquad\qquad\qquad \left.\left. - \log \phi\left(\int_{(0,s)} h(s-u)N(\mathrm{d}u) \right) \right] N(\mathrm{d}s) \right).
\end{aligned}$$

In particular, we have

$$\pi_t \leq \phi(0) \exp\left(\|\phi\|_{\mathrm{Lip}} \int_0^{T-t} h(s) \, \mathrm{d}s \right) \leq \phi(0) \exp\left(\|\phi\|_{\mathrm{Lip}} \|h\|_{L^1([0,T], \mathcal{B}([0,T]), \mathrm{d}t)} \right), \tag{3.12}$$

$\mathrm{d}t\mathbb{P}(\mathrm{d}\omega)$-*almost everywhere.*

Next, we give an example of nonlinear Hawkes processes which satisfies the assumptions of Corollary 3.9.

Example 3.10. Assume that N is a nonlinear Hawkes process with parameters $\phi(x) := \alpha \min(\max(K - x, 0), 1)$ and $h := \mathbb{1}_{[0,z]}$ for some $\alpha, K, z > 0$, K integer. This is a notable example of nonlinear Hawkes process since $N([0, t])$, $t \in [0, T]$, is the total number of customers who have entered, in the time interval $[0, t]$, the Erlang loss system (or M/D/K/0 queue) with arrival rate α, deterministic service time z and number of servers equal to K, see [2] for details. An easy computation shows that γ defined in Corollary 3.9 is equal to

$$\gamma(z) := \alpha \left[\frac{\mathrm{e}^{2\alpha \min(z,T)}}{2\alpha} - \frac{1}{2\alpha} - \min(z, T) + \left(\mathrm{e}^{2\alpha z} - 1 \right)(T - z)\mathbb{1}_{\{z \leq T\}} \right].$$

Therefore, for fixed $\alpha, T > 0$, and z_0 satisfying the inequality $\gamma(z_0) < 1$ (in particular, note that it suffices to take z_0 close to zero), the corresponding nonlinear Hawkes process satisfies the Poincaré inequality (3.10).

We conclude this subsection by providing a class of Cox processes which satisfy the Poincaré inequality.

Corollary 3.11 (Cox processes). *Assume $\nu(E) < \infty$ and let N be a Cox process on $[0,T] \times E$, $T < \infty$, with stochastic intensity $\{\lambda_{(t,x)}\}_{(t,x)\in[0,T]\times E}$ such that, for some non-negative functions $\alpha, \beta \in L^1([0,T] \times E, \mathcal{B}([0,T]) \otimes \mathcal{E}, \mathrm{d}t\nu(\mathrm{d}x))$,*

$$\alpha(t,x) \leq \lambda_{(t,x)} \leq \beta(t,x), \quad \mathrm{d}t\nu(\mathrm{d}x)\mathbb{P}(\mathrm{d}\omega)\text{-almost everywhere} \tag{3.13}$$

and

$$\gamma := \|\beta\|_{L^1([0,T]\times E, \mathcal{B}([0,T])\otimes\mathcal{E}, \mathrm{d}t\nu(\mathrm{d}x))} - \|\alpha\|_{L^1([0,T]\times E, \mathcal{B}([0,T])\otimes\mathcal{E}, \mathrm{d}t\nu(\mathrm{d}x))} < 1. \tag{3.14}$$

Then, for any square-integrable functional G of N we have

$$\mathbb{V}\mathrm{ar}(G) \leq \left(1 - \sqrt{\gamma}\right)^{-2} \mathbb{E}\left[\int_{[0,T]\times E} \pi_{(t,x)} \left|D_{(t,x)}G\right|^2 \mathrm{d}t\nu(\mathrm{d}x)\right],$$

where the Papangelou conditional intensity π of N is given, for fixed $\omega \in \Omega$, by

$$\pi_{(t,x)}(\omega) := \mathbb{E}[\mathcal{R}(\lambda,\omega)\lambda_{(t,x)}] \tag{3.15}$$

where

$$\mathcal{R}(\lambda,\omega) := \frac{\exp\left(-\int_{[0,T]\times E} \lambda_{(s,z)} \mathrm{d}s\nu(\mathrm{d}z)\right) \prod_{(s,z)\in\mathrm{Supp}(\omega)} \lambda_{(s,z)}}{\mathbb{E}\left[\exp\left(-\int_{[0,T]\times E} \lambda_{(s,z)} \mathrm{d}s\nu(\mathrm{d}z)\right) \prod_{(s,z)\in\mathrm{Supp}(\omega)} \lambda_{(s,z)}\right]}.$$

In particular, we have

$$\pi_{(t,x)}(\omega) \leq \beta(t,x)$$

$\mathrm{d}t\mathbb{P}(\mathrm{d}\omega)\nu(\mathrm{d}x)$-*almost everywhere.*

3.2 Transportation cost inequalities

Let χ be a Polish space equipped with its Borel σ-field $\mathcal{B}(\chi)$ and let d be a lower semi-continuous metric on χ (which does not necessarily generates the topology on χ). Letting σ_1 and σ_2 denote a couple of probability measures on $(\chi, \mathcal{B}(\chi))$, we define the transportation cost

$$\mathcal{T}_d(\sigma_1, \sigma_2) := \inf_\rho \int_{\chi\times\chi} d(x,y)\, \rho(\mathrm{d}x \times \mathrm{d}y) \in [0,\infty],$$

where the infimum is taken over all the probability measures ρ on $\chi \times \chi$ with first marginal σ_1 and second marginal σ_2. We denote by $M_1(\chi, d)$ the set of all probability measures σ on $(\chi, \mathcal{B}(\chi))$ such that $\int d(x_0, x)\, \sigma(\mathrm{d}x) < \infty$, for some $x_0 \in \chi$ and we remark that for $\sigma_1, \sigma_2 \in M_1(\chi, d)$, $\mathcal{T}_d(\sigma_1, \sigma_2) < \infty$. The relative entropy of σ_1 with respect to σ_2 is defined by

$$H(\sigma_1 \mid \sigma_2) := \int_\chi \log\left(\frac{\mathrm{d}\sigma_1}{\mathrm{d}\sigma_2}\right) \mathrm{d}\sigma_1$$

if σ_1 is absolutely continuous with respect to σ_2 and $H(\sigma_1 \mid \sigma_2) := +\infty$ otherwise (the reader is referred to [26] for more insight into the theory of optimal transportation).

In the following, we denote by

$$g^\odot(x) := \sup_{\theta \geq 0}(\theta x - g(\theta)), \quad x \in \mathbb{R}_+$$

the monotone conjugate of a measurable function $g : \mathbb{R}_+ \to [0,\infty]$.

3.2.1 A transportation cost inequality for functionals of N

In this subsection we take $\chi := \mathbb{R}$, suppose that there exists a norm on χ, say $\|\cdot\|_d$, such that $d(x, y) = \|x - y\|_d$, $x, y \in \mathbb{R}$, and denote by $\mathcal{L}(X)$ the law of a real-valued random variable X. The following theorem holds.

Theorem 3.12. *Suppose that N satisfies (2.2) for a measurable function β and let G be an integrable random variable such that*

$$\|D_{(t,x)}G(\omega)\|_d \le g_1(t, x) \quad \text{and} \quad \frac{\mathbb{E}\big[\|G - \mathbb{E}[G \mid \mathcal{F}_{t-}]\|_d |\pi_{(t,x)} - p(\pi)_{(t,x)}| \mid \mathcal{F}_{t-}\big](\omega)}{p(\pi)_{(t,x)}(\omega)} \le g_2(t, x)$$

for $\mathrm{d}t\mathrm{d}\mathbb{P}\mathrm{d}\nu$-almost all (t, ω, x) and some deterministic functions g_1, g_2 such that

$$\int_{\mathbb{R}_+ \times E} |h(t, x)|^2 \beta(t, x)\, \mathrm{d}t\nu(\mathrm{d}x) < \infty \tag{3.16}$$

where $h := g_1 + g_2$. Then

$$c(\mathcal{T}_d(\sigma, \mathcal{L}(G))) \le H(\sigma \mid \mathcal{L}(G)), \quad \text{for any } \sigma \in M_1(\mathbb{R}, d) \tag{3.17}$$

where

$$c(x) := \Lambda^{\odot}(x), \quad x \in \mathbb{R}_+ \tag{3.18}$$

and

$$\Lambda(\theta) := \int_{\mathbb{R}_+ \times E} \big(e^{\theta h(t,x)} - \theta h(t, x) - 1\big)\beta(t, x)\, \mathrm{d}t\nu(\mathrm{d}x), \quad \theta \in \mathbb{R}_+. \tag{3.19}$$

If additionally we assume that there exists $M > 0$ such that $h(t, x) \le M$ for $\mathrm{d}t\mathrm{d}\nu$-almost all (t, x), then

$$c(x) \ge \widetilde{c}(x), \quad x \in \mathbb{R}_+, \tag{3.20}$$

where

$$\widetilde{c}(x) := \frac{x + M^{-1}\int_{\mathbb{R}_+ \times E} h(t, z)^2 \beta(t, z)\mathrm{d}t\nu(\mathrm{d}z)}{M} \ln\left(1 + \frac{x}{M^{-1}\int_{\mathbb{R}_+ \times E} h(t, z)^2 \beta(t, z)\mathrm{d}t\nu(\mathrm{d}z)}\right) - \frac{x}{M}.$$

Remark 3.13. As noticed in Remark 3.2, if $\mathbb{E}\big[\pi_{(t,x)}\big|\mathcal{F}_{t-}\big] = 0$ $\mathbb{P}$-almost surely, then $\mathbb{P}(\pi_{(t,x)} = 0 \mid \mathcal{F}_{t-}) = 1$ $\mathbb{P}$-almost surely. Consequently, the ratio

$$\frac{\mathbb{E}\big[\|G - \mathbb{E}[G \mid \mathcal{F}_{t-}]\|_d |\pi_{(t,x)} - p(\pi)_{(t,x)}|\big|\mathcal{F}_{t-}\big]}{p(\pi)_{(t,x)}}$$

is always well-defined under the convention $0/0 := 0$.

Remark 3.14 (Deviation inequality). From the point of view of the applications, it is important to remark that a transportation cost inequality is often equivalent to a deviation bound. More precisely, in the context of Theorem 3.12, one has that the transportation cost inequality (3.17) is equivalent to the deviation bound

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n f(G_i) \ge \mathbb{E}\big[f(G)\big] + r\right) \le e^{-nc(r)}, \quad \text{for any } n \ge 1 \text{ and } r \in \mathbb{R}_+ \tag{3.21}$$

for any measurable function $f : \mathbb{R} \to \mathbb{R}$ which is Lipschitz continuous (with respect to d) with Lipschitz constant less than or equal to 1, i.e. such that

$$\sup_{x \ne y} \frac{|f(x) - f(y)|}{d(x, y)} \le 1,$$

where $\{G_n\}_{n \ge 1}$ is a sequence of independent random variables with the same law as G (see [10] and Theorem 1.1(c) in [19]).

To the best of our knowledge, the transportation cost inequality provided by Theorem 3.12 is new even in the Poisson case, which we state in a separate corollary.

Corollary 3.15 (Poisson processes)**.** *Suppose that N is a Poisson process on $\mathbb{R}_+ \times E$ with mean measure $\beta(t,x)\,\mathrm{d}t\nu(\mathrm{d}x)$ and let G be an integrable random variable such that*

$$\|D_{(t,x)}G(\omega)\|_d \le h(t,x)$$

for $\mathrm{d}t\mathrm{d}\mathbb{P}\mathrm{d}\nu$-almost all (t,ω,x) and some deterministic function h which satisfies (3.16)*. Then the transportation cost inequality* (3.17) *holds. If additionally we assume that there exists $M > 0$ such that $h(t,x) \le M$ for $\mathrm{d}t\mathrm{d}\nu$-almost all (t,x), then* (3.20) *holds, and provides a more explicit bound on the corresponding deviation inequality* (3.21)*.*

In the following proposition, we specialize Theorem 3.12 to first order integrals. The subsequent corollaries concern applications to renewal, nonlinear Hawkes and Cox point processes.

Proposition 3.16 (First order integrals)**.** *Suppose that N satisfies* (2.2) *for a measurable function β, $d(x,y) := |x - y|$, take*

$$G := \int_{\mathbb{R}_+ \times E} g(s,y)\,N(\mathrm{d}s \times \mathrm{d}y),$$

for some measurable and deterministic function g, and let g_1, g_2 be deterministic functions such that

$$|g(t,x)| \le g_1(t,x)$$

and

$$\sqrt{2}\left[\int_{[t,\infty)\times E} g(s,y)^2 \beta(s,y)\,\mathrm{d}s\nu(\mathrm{d}y) + \left(\int_{[t,\infty)\times E} |g(s,y)|\beta(s,y)\,\mathrm{d}s\nu(\mathrm{d}y)\right)^2\right]^{1/2}$$
$$\times \left\|\frac{\sqrt{\mathrm{Var}\big[\pi_{(t,x)} \mid \mathcal{F}_{t-}\big]}}{p(\pi)_{(t,x)}}\right\|_{L^\infty(\Omega,\mathcal{F}_\infty,\mathbb{P})} \le g_2(t,x), \quad (3.22)$$

for $(t,x) \in \mathbb{R}_+ \times E$. Assuming that $h := g_1 + g_2$ satisfies the corresponding assumption (3.16)*, the transportation cost inequality* (3.17) *holds. If additionally we assume that there exists $M > 0$ such that $h(t,x) \le M$ for $\mathrm{d}t\mathrm{d}\nu$-almost all (t,x), then* (3.20) *holds, and provides a more explicit bound on the corresponding deviation inequality* (3.21)*.*

In the next corollaries, we provide classes of point processes which satisfy the assumptions of Proposition 3.16. We point out that we do not aim to optimize the assumptions, favoring instead clarity and conciseness. The proofs are quite elementary, and provided in Section 4.4 for the readers' convenience.

Corollary 3.17 (Renewal point processes)**.** *Let N be a renewal point process on $[0,T]$, $T < \infty$, with a spacing density f such that f is continuous on $[0,+\infty)$ and $f > 0$ on (a,C) for some $a \in [0,T]$ and $C \in (T,+\infty]$, and let G be defined by*

$$G := \int_0^T g(s)\,N(\mathrm{d}s), \quad g \in L^2([0,T]). \tag{3.23}$$

Assume further that there exists $\overline{h} > 0$ such that

$$\sup_{z\in[0,T]} \sup_{x\in[z,T]} \frac{f(x-z)}{f(x)} \le \overline{h} \tag{3.24}$$

(note that this condition is always satisfied if $f > 0$ on $[0, T]$.) After straightforward adjustments due to the unmarked setting, the transportation cost inequality (3.17) holds with $g_1 \in L^2([0, T])$ such that $|g(t)| \leq g_1(t)$, $t \in [0, T]$, and $g_2 \in L^2([0, T])$ such that

$$\sqrt{2\left(\overline{h}^2 + \left(\int_T^C f(x)\,\mathrm{d}x\right)^{-2} - 1\right)}\left[\beta \int_t^T g(s)^2\,\mathrm{d}s + \beta^2\left(\int_t^T |g(s)|\,\mathrm{d}s\right)^2\right]^{1/2} \leq g_2(t), \quad (3.25)$$

for $t \in [0, T]$, where

$$\beta := \max\left(\frac{\sup_{x \in [0,T]} f(x)}{\int_T^C f(x)\,\mathrm{d}x}, \sup_{x \in [0,C)} f(x) + \frac{(\sup_{x \in [0,T]} f(x))^2}{\min_{x \in [a,T]} f(x)}\right). \quad (3.26)$$

For any $M > 0$ such that $\|g_1 + g_2\|_\infty \leq M$, we have that (3.20) holds, which yields a more explicit bound on the corresponding deviation inequality (3.21).

Corollary 3.18 (Nonlinear Hawkes processes)*. Assume that N is a nonlinear Hawkes process on $[0, T]$, $T < \infty$, with parameters (h, ϕ) such that h is non-negative and integrable on $[0, T]$, ϕ is Lipschitz continuous and non-increasing with $\phi(0) > 0$, and let G be defined by (3.23). After straightforward adjustments due to the unmarked setting, the transportation cost inequality (3.17) holds with $g_1 \in L^2([0, T])$ such that $|g(t)| \leq g_1(t)$, $t \in [0, T]$, and $g_2 \in L^2([0, T])$ such that*

$$g_2(t) \geq \sqrt{2}\left[\phi(0)\int_t^T g(s)^2 \exp\left(\|\phi\|_{\mathrm{Lip}}\int_0^{T-s} h(z)\,\mathrm{d}z\right)\mathrm{d}s\right.$$
$$\left. + \left(\phi(0)\int_t^T |g(s)|\exp\left(\|\phi\|_{\mathrm{Lip}}\int_0^{T-s} h(z)\,\mathrm{d}z\right)\mathrm{d}s\right)^2\right]^{1/2}\sqrt{\exp\left(2\|\phi\|_{\mathrm{Lip}}\int_0^{T-t} h(z)\mathrm{d}z\right) - 1},$$
$$(3.27)$$

$t \in [0, T]$. For any $M > 0$ such that $\|g_1 + g_2\|_\infty \leq M$, we have that (3.20) holds, which yields a more explicit bound on the corresponding deviation inequality (3.21).

Corollary 3.19 (Cox processes)*. Assume $\nu(E) < \infty$ and let N be a Cox process on $[0, T] \times E$, $T < \infty$, with stochastic intensity $\{\lambda_{(t,x)}\}_{(t,x) \in [0,T] \times E}$ satisfying (3.13), for some non-negative functions α, β with $\beta \in L^1([0, T] \times E, \mathrm{d}t\nu(\mathrm{d}x))$ and $\beta^3\alpha^{-2} \in L^1([0, T] \times E, \mathrm{d}t\nu(\mathrm{d}x))$. If G is defined by*

$$G := \int_{[0,T] \times E} g(s, y)\, N(\mathrm{d}s \times \mathrm{d}y), \quad g \in L^2([0, T] \times E, \beta(t, x)\mathrm{d}t\nu(\mathrm{d}x)),$$

then the transportation cost inequality (3.17) holds with $g_1 \in L^2([0, T] \times E, \beta(t, x)\mathrm{d}t\nu(\mathrm{d}x))$ such that $|g(t, x)| \leq g_1(t, x)$, $(t, x) \in [0, T] \times E$, and $g_2 \in L^2([0, T] \times E, \beta(t, x)\mathrm{d}t\nu(\mathrm{d}x))$ such that

$$\sqrt{2}\left[\int_{[t,T] \times E} g(s, y)^2 \beta(s, y)\,\mathrm{d}s\nu(\mathrm{d}y) + \left(\int_{[t,T] \times E} |g(s, y)|\beta(s, y)\,\mathrm{d}s\nu(\mathrm{d}y)\right)^2\right]^{1/2}$$
$$\times \sqrt{\frac{\beta(t, x)^2}{\alpha(t, x)^2} - 1} \leq g_2(t, x),$$

for $(t, x) \in [0, T] \times E$. For any $M > 0$ such that $\|g_1 + g_2\|_\infty \leq M$, we have that (3.20) holds, which yields a more explicit bound on the corresponding deviation inequality (3.21).

3.2.2 A transportation cost inequality for the law of N

In this subsection we take $\chi := \Omega$ and equip this set with the vague convergence topology, that is, the coarsest topology such that the map $\omega \mapsto \int_{\mathbb{R}_+ \times E} f(t, x)\, \omega(\mathrm{d}t \times \mathrm{d}x)$ is continuous, where $f : \mathbb{R}_+ \times E \to \mathbb{R}$ is a continuous function with compact support. It is well-known that this topology makes Ω a Polish space (see e.g. [5]). In this subsection, we let $\varphi : \mathbb{R}_+ \times E \to \mathbb{R}_+$ be a continuous function and define the following metric on Ω:

$$d_\varphi(\omega, \omega') := \int_{\mathbb{R}_+ \times E} \varphi(t, x)|\omega - \omega'|(\mathrm{d}t \times \mathrm{d}x), \quad \omega, \omega' \in \Omega$$

where, for $\kappa \in \Omega$, $|\kappa| := \kappa^+ + \kappa^-$, and κ^+ and κ^- denote respectively the positive and the negative parts of κ in the Hahn-Jordan decomposition. It is known that d_φ is a lower semi-continuous metric on Ω (see Lemma 2.2 in [19]). The following theorem holds.

Theorem 3.20. *Assume that N satisfies* (2.2) *for a measurable function β, that there exists a deterministic measurable function ψ which verifies* (3.22) *with φ in place of g and ψ in place of g_2 and that*

$$\int_{\mathbb{R}_+ \times E} |h_\varphi(t, x)|^2 \beta(t, x)\, \mathrm{d}t\nu(\mathrm{d}x) < \infty, \tag{3.28}$$

where $h_\varphi := \varphi + \psi$. Then

$$c(\mathcal{T}_{d_\varphi}(\mathbb{Q}, \mathbb{P})) \le H(\mathbb{Q} \mid \mathbb{P}), \quad \text{for any } \mathbb{Q} \in M_1(\Omega, d_\varphi) \tag{3.29}$$

where

$$c(x) := \Lambda_\varphi^\odot(x), \quad x \in \mathbb{R}_+$$

and

$$\Lambda_\varphi(\theta) := \int_{\mathbb{R}_+ \times E} \big(e^{\theta h_\varphi(t, x)} - \theta h_\varphi(t, x) - 1\big)\beta(t, x)\, \mathrm{d}t\nu(\mathrm{d}x), \quad \theta \in \mathbb{R}_+.$$

In particular, when N is a Poisson process with mean measure $\beta(t, x)\, \mathrm{d}t\nu(\mathrm{d}x)$ we recover the sharp transportation cost inequality in Remark 2.7 of [19].

Remark 3.21 (Deviation inequality)**.** Here again, it turns out (see [10] and Theorem 1.1(c) in [19]) that, in the context of Theorem 3.20, the transportation cost inequality (3.29) is equivalent to the deviation bound

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} F(N_i) \ge \mathbb{E}\big[F(N)\big] + r\right) \le e^{-nc(r)}, \quad \text{for any } n \ge 1 \text{ and } r \in \mathbb{R}_+ \tag{3.30}$$

for any measurable function $F : \Omega \to \mathbb{R}$ which is Lipschitz continuous (with respect to d_φ) with Lipschitz constant less than or equal to 1, i.e. such that

$$\sup_{\omega \ne \omega'} \frac{|f(\omega) - f(\omega')|}{d_\varphi(\omega, \omega')} \le 1,$$

where $\{N_n\}_{n \ge 1}$, $N_n : \Omega \to \Omega$, is a sequence of independent marked point processes with the same law as N. As in the previous Subsection 3.2.1, if additionally we assume that there exists $M > 0$ such that $h_\varphi(t, x) \le M$ for $\mathrm{d}t\mathrm{d}\nu$-almost all (t, x), then (3.20) with h_φ in place of h holds, and provides a more explicit bound on the deviation inequality (3.30).

Arguing as in the proofs of Corollaries 3.17, 3.18 and 3.19, one can show that Theorem 3.20 and the deviation bound in Remark 3.21 apply to renewal, nonlinear Hawkes and Cox point processes. Similarly to the previous Subsection 3.2.1, we do not aim to optimize the assumptions, favoring instead clarity and conciseness.

Corollary 3.22 (Renewal point processes)**.** *Let N be a renewal point process on $[0,T]$, $T < \infty$, with a spacing density f such that f is continuous on $[0,+\infty)$ and $f > 0$ on (a,C) for some $a \in [0,T]$ and $C \in (T,+\infty]$. Assume further that there exists $\overline{h} > 0$ satisfying (3.24). Then, after straightforward adjustments due to the unmarked setting, the transportation cost inequality (3.29) holds with $\psi \in L^2([0,T])$ such that*

$$\sqrt{2\left(\overline{h}^2 + \left(\int_T^C f(x)\,\mathrm{d}x\right)^{-2} - 1\right)}\left[\beta \int_t^T \varphi(s)^2\,\mathrm{d}s + \beta^2\left(\int_t^T |\varphi(s)|\,\mathrm{d}s\right)^2\right]^{1/2} \leq \psi(t),$$

for $t \in [0,T]$, where β is defined by (3.26).

Corollary 3.23 (Nonlinear Hawkes processes)**.** *Assume that N is a nonlinear Hawkes process on $[0,T]$, $T < \infty$, with parameters (h,ϕ) such that h is non-negative and integrable on $[0,T]$, ϕ is Lipschitz continuous and non-increasing with $\phi(0) > 0$. Then, after straightforward adjustments due to the unmarked setting, the transportation cost inequality (3.29) holds with $\psi \in L^2([0,T])$ such that*

$$\psi(t) \geq \sqrt{2}\Bigg[\phi(0)\int_t^T \varphi(s)^2 \exp\left(\|\phi\|_{\mathrm{Lip}}\int_0^{T-s} h(z)\,\mathrm{d}z\right)\mathrm{d}s$$

$$+ \left(\phi(0)\int_t^T |\varphi(s)|\exp\left(\|\phi\|_{\mathrm{Lip}}\int_0^{T-s} h(z)\mathrm{d}z\right)\mathrm{d}s\right)^2\Bigg]^{1/2}\sqrt{\exp\left(2\|\phi\|_{\mathrm{Lip}}\int_0^{T-t} h(z)\mathrm{d}z\right) - 1},$$

for $t \in [0,T]$.

Corollary 3.24 (Cox point processes)**.** *Assume $\nu(E) < \infty$ and let N be a Cox process on $[0,T] \times E$, $T < \infty$, with stochastic intensity $\{\lambda_{(t,x)}\}_{(t,x) \in [0,T] \times E}$ satisfying (3.13), for some non-negative functions α, β with $\beta \in L^1([0,T] \times E, \mathrm{d}t\nu(\mathrm{d}x))$ and $\beta^3\alpha^{-2} \in L^1([0,T] \times E, \mathrm{d}t\nu(\mathrm{d}x))$. Assume further that $\varphi \in L^2([0,T] \times E, \beta(t,x)\mathrm{d}t\nu(\mathrm{d}x))$. Then, the transportation cost inequality (3.29) holds with $\psi \in L^2([0,T] \times E, \beta(t,x)\mathrm{d}t\nu(\mathrm{d}x))$ such that*

$$\sqrt{2}\Bigg[\int_{[t,T] \times E} \varphi(s,y)^2\beta(s,y)\,\mathrm{d}s\nu(\mathrm{d}y) + \left(\int_{[t,T] \times E} |\varphi(s,y)|\beta(s,y)\,\mathrm{d}s\nu(\mathrm{d}y)\right)^2\Bigg]^{1/2}$$

$$\times \sqrt{\frac{\beta(t,x)^2}{\alpha(t,x)^2} - 1} \leq \psi(t,x),$$

for $(t,x) \in [0,T] \times E$.

3.3 Variational representation of the Laplace transform

The following Theorem 3.26 generalizes to functionals of marked point processes with Papangelou conditional intensity the variational representation formula for the Laplace transform of Poisson functionals given by Theorem 4.4 of [29]. Note also that, in contrast to [29], here the marked point process is defined on the whole half-line.

Hereafter, we suppose that N satisfies (3.1) and (3.2), and denote by $\mathcal{H}$ the subset of $\mathcal{P}_2^{\mathcal{F}}(p(\pi))$ formed by the real-valued processes $\phi \in L^\infty(\mathbb{R}_+ \times \Omega \times E, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_\infty \otimes \mathcal{E}, \mathrm{d}t\mathrm{d}\mathbb{P}\mathrm{d}\nu)$ such that

$$\phi_{(t,x)} \geq c_\phi > -1, \quad \text{for some constant } c_\phi, \; \mathrm{d}t\mathrm{d}\mathbb{P}\nu(\mathrm{d}x)\text{-almost surely,} \tag{3.31}$$

$$\mathbb{E}\left[\left(\int_{\mathbb{R}_+ \times E} |\phi_{(s,x)}|^2\, p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\right)^2\right] < \infty \tag{3.32}$$

and

$$\text{for any } T > 0, \ (t,x) \mapsto \mathbb{1}_{[0,T]}(t)\mathcal{E}_t(\phi) \in \mathcal{P}_2^{\mathcal{F}}(p(\pi)), \tag{3.33}$$

where

$$\mathcal{E}_t(\phi) := \exp\bigg(\int_{[0,t]\times E} \log\big(1 + \phi_{(s,x)}\big)(N(\mathrm{d}s \times \mathrm{d}x) - p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x))$$

$$+ \int_{[0,t]\times E} \big(\log\big(1 + \phi_{(s,x)}\big) - \phi_{(s,x)}\big)p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\bigg), \quad t \in \mathbb{R}_+.$$

The following lemma ensures that $\{\mathcal{E}_t(\phi)\}_{t\in\mathbb{R}_+}$ is a square-integrable $\mathcal{F}$-martingale. Its proof is postponed to Section 4 (see Subsection 4.6).

Lemma 3.25. *Assume that N satisfies (3.1) and (3.2). Then, for any $\phi \in \mathcal{H}$, the stochastic process $\{\mathcal{E}_t(\phi)\}_{t\in\mathbb{R}_+}$ is a square-integrable $\mathcal{F}$-martingale.*

Under the assumptions of Lemma 3.25, for $\phi \in \mathcal{H}$, we define a new probability measure $\mathbb{P}_\phi$ on $(\Omega, \mathcal{F}_\infty)$ by

$$\frac{\mathrm{d}\mathbb{P}_\phi}{\mathrm{d}\mathbb{P}}\bigg|_{\mathcal{F}_t} = \mathcal{E}_t(\phi), \quad t \in \mathbb{R}_+. \tag{3.34}$$

The following variational representation of the Laplace transform holds.

Theorem 3.26. *Suppose that N satisfies (3.1), (3.2),*

$$K := \int_0^\infty \bigg\|\int_E p(\pi)_{(t,x)}\,\nu(\mathrm{d}x)\bigg\|_{L^\infty(\Omega,\mathcal{F}_\infty,\mathbb{P})}\,\mathrm{d}t < \infty, \tag{3.35}$$

and let G be a random variable on $(\Omega, \mathcal{F}_\infty, \mathbb{P})$ which is upper bounded. Then

$$-\log\big(\mathbb{E}\big[\mathrm{e}^{-G}\big]\big) = \inf_{\phi\in\mathcal{H}} \mathbb{E}_\phi\big[G + L(\phi)\big],$$

where $\mathbb{E}_\phi$ denotes the expectation under $\mathbb{P}_\phi$ and

$$L(\phi) := \int_{\mathbb{R}_+\times E} \big((1 + \phi_{(s,x)})\log\big(1 + \phi_{(s,x)}\big) - \phi_{(s,x)}\big)\,p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x).$$

Additionally, if $G \in L^\infty(\Omega, \mathcal{F}_\infty, \mathbb{P})$, then the infimum is uniquely attained at

$$\phi_{(t,x)}^{(F)} := \frac{\varphi_{(t,x)}^{(F)}}{p(F)_{(t,x)}}, \quad (t,x) \in \mathbb{R}_+ \times E$$

where $F := \mathrm{e}^{-G}$ and

$$\varphi_{(t,x)}^{(F)} := \frac{p(\pi F^+)_{(t,x)} - p(F)_{(t,x)}p(\pi)_{(t,x)}}{p(\pi)_{(t,x)}}, \quad (t,x) \in \mathbb{R}_+ \times E. \tag{3.36}$$

Here, we limit ourselves to note that $\varphi^{(F)}$ is well-defined and belongs to $\mathcal{P}_2^{\mathcal{F}}(p(\pi))$, and refer the reader to Theorem 3.28 and Remark 3.30 for details.

Remark 3.27. Assumptions (3.1), (3.2) and (3.35) are satisfied if the inequality (2.2) holds with $\beta \in L^1(\mathbb{R}_+ \times E, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}, \mathrm{d}t\nu(\mathrm{d}x))$, see also Remark 3.4. As such, Theorem 3.26 applies to (i) renewal point processes on $[0,T]$, $T < \infty$, with a spacing density f such that f is continuous on $[0, +\infty)$ and $f > 0$ on (a, C) for some $a \in [0,T]$ and $C \in (T, +\infty]$; (ii) nonlinear Hawkes processes on $[0,T]$, $T < \infty$, with parameters (h, ϕ) such that h is non-negative and integrable on $[0,T]$, ϕ is Lipschitz continuous and non-increasing with $\phi(0) > 0$; (iii) if $\nu(E) < \infty$, Cox processes on $[0,T] \times E$, $T < \infty$, with stochastic intensity $\{\lambda_{(t,x)}\}_{(t,x)\in[0,T]\times E}$ such that $\lambda_{(t,x)} \leq \beta(t,x)$, $(t,x) \in [0,T] \times E$, with $\beta \in L^1([0,T] \times E, \mathrm{d}t\nu(\mathrm{d}x))$.

$$= \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} (\varphi^{(G)}_{(t,x)})^2 p(\pi)_{(t,x)} \, dt\nu(dx)\right]$$

$$= \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \frac{\left(p(\pi G^+)_{(t,x)} - p(G)_{(t,x)} p(\pi)_{(t,x)}\right)^2}{p(\pi)_{(t,x)}} \, dt\nu(dx)\right]$$

$$= \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \frac{\left(\mathbb{E}\left[G^+_{(t,x)}\pi_{(t,x)} - G\,\mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right] \big|\mathcal{F}_{t^-}\right]\right)^2}{\mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]} \, dt\nu(dx)\right]. \qquad (4.1)$$

Note that

$$\mathbb{E}\left[G^+_{(t,x)}\pi_{(t,x)} - G\,\mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]\big|\mathcal{F}_{t^-}\right]$$
$$= \mathbb{E}\left[\pi_{(t,x)}\left(G^+_{(t,x)} - G\right)\big|\mathcal{F}_{t^-}\right] + \mathbb{E}\left[(G - \mathbb{E}[G])\left(\pi_{(t,x)} - \mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]\right)\big|\mathcal{F}_{t^-}\right].$$

Thus, for any $0 < q < 1 - \gamma$, by (4.1) and the convexity inequality

$$(a+b)^2 \leq a^2/q + b^2/(1-q), \quad a, b \in \mathbb{R},$$

we have

$$\mathbb{Var}(G) \leq \frac{1}{q}\,\mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \frac{\left(\mathbb{E}\left[\pi_{(t,x)}\left(G^+_{(t,x)} - G\right)\big|\mathcal{F}_{t^-}\right]\right)^2}{\mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]} \, dt\nu(dx)\right]$$
$$+ \frac{1}{1-q}\,\mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \frac{\left(\mathbb{E}\left[(G - \mathbb{E}[G])\left(\pi_{(t,x)} - \mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]\right)\big|\mathcal{F}_{t^-}\right]\right)^2}{\mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]} \, dt\nu(dx)\right]. \qquad (4.2)$$

By the Cauchy-Schwarz inequality

$$\left(\mathbb{E}\left[\pi_{(t,x)}\left(G^+_{(t,x)} - G\right)\big|\mathcal{F}_{t^-}\right]\right)^2 \leq \mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]\mathbb{E}\left[\pi_{(t,x)}\left(G^+_{(t,x)} - G\right)^2\big|\mathcal{F}_{t^-}\right]$$

and so

$$\mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \frac{\left(\mathbb{E}\left[\pi_{(t,x)}\left(G^+_{(t,x)} - G\right)\big|\mathcal{F}_{t^-}\right]\right)^2}{\mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]} \, dt\nu(dx)\right]$$
$$\leq \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \left|D_{(t,x)}G\right|^2 \pi_{(t,x)} \, dt\nu(dx)\right]. \qquad (4.3)$$

By the Cauchy-Schwarz inequality

$$\left(\mathbb{E}\left[(G - \mathbb{E}[G])\left(\pi_{(t,x)} - \mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]\right)\big|\mathcal{F}_{t^-}\right]\right)^2$$
$$\leq \mathbb{E}\left[(G - \mathbb{E}[G])^2\left(\mathbb{E}\left[\pi^2_{(t,x)}\big|\mathcal{F}_{t^-}\right] - \mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]^2\right)\big|\mathcal{F}_{t^-}\right].$$

So, by assumption (3.3) and the inequality

$$(\mathbb{E}\left[\pi^2_{(t,x)}\big|\mathcal{F}_{t^-}\right]/\mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]) - \mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right] \geq 0$$

which follows from Jensen's inequality, we have

$$\mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \frac{\left(\mathbb{E}\left[(G - \mathbb{E}[G])\left(\pi_{(t,x)} - \mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]\right)\big|\mathcal{F}_{t^-}\right]\right)^2}{\mathbb{E}\left[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\right]} \, dt\nu(dx)\right]$$

$$\leq \mathbb{E}\left[(G - \mathbb{E}[G])^2 \int_{\mathbb{R}_+ \times E} \left(\frac{\mathbb{E}\left[\pi^2_{(t,x)} \big| \mathcal{F}_{t-}\right]}{\mathbb{E}\left[\pi_{(t,x)} \big| \mathcal{F}_{t-}\right]} - \mathbb{E}\left[\pi_{(t,x)} \big| \mathcal{F}_{t-}\right]\right) \mathrm{d}t\nu(\mathrm{d}x)\right] \leq \gamma \,\mathbb{V}\mathrm{ar}(G).$$
$$(4.4)$$

Combining (4.2) with the bounds (4.3) and (4.4), we obtain

$$\mathbb{V}\mathrm{ar}(G) \leq q^{-1}\mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \left|D_{(t,x)}G\right|^2 \pi_{(t,x)} \,\mathrm{d}t\nu(\mathrm{d}x)\right] + \gamma(1-q)^{-1}\,\mathbb{V}\mathrm{ar}(G),$$

i.e.

$$\mathbb{V}\mathrm{ar}(G) \leq \frac{1/q}{1 - \gamma/(1-q)}\,\mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \left|D_{(t,x)}G\right|^2 \pi_{(t,x)} \,\mathrm{d}t\nu(\mathrm{d}x)\right], \quad \text{for any } 0 < q < 1 - \gamma.$$
$$(4.5)$$

Finally, we note that the choice $q = 1 - \sqrt{\gamma}$ minimizes the multiplicative constant appearing in (4.5), and the proof is complete.

4.2 Proofs of Corollaries 3.7, 3.9 and 3.11

Proof of Corollary 3.7. It is well-known that N has Papangelou conditional intensity $\{\pi_t\}_{t \in [0,T]}$ defined in the statement (see [24] and Corollary 2.9 in [8]). By Proposition 2.10 in [8] N is locally stable, and so by Remark 3.4 the corresponding assumptions (3.1) and (3.2) are satisfied. We recall that the points of N are denoted by $T_0 := 0 < T_1 < T_2 < \cdots < T_{N([0,T])}$. For any $t \in [0,T]$, we have

$$\mathbb{E}\left[\pi^2_t \mid \mathcal{F}_{t-}\right]$$
$$= \mathbb{E}\left[\frac{f(t - T_{N([0,t))})^2 f(T_{N([0,t))+1} - t)^2}{f(T_{N([0,t))+1} - T_{N([0,t))})^2}\mathbb{1}_{\{N((t,T])>0\}}\right.$$
$$\left. + \frac{f(t - T_{N([0,t))})^2 \overline{F}(T-t)^2}{\overline{F}(T - T_{N([0,t))})^2}\mathbb{1}_{\{N((t,T])=0\}} \,\Big|\, \mathcal{F}_{t-}\right]$$
$$= f(t - T_{N([0,t))})^2\mathbb{E}\left[\frac{f\left(T_{N([0,t))+1} - T_{N([0,t))} - (t - T_{N([0,t))})\right)^2}{f(T_{N([0,t))+1} - T_{N([0,t))})^2}\mathbb{1}_{\{N((t,T])>0\}} \,\Big|\, \mathcal{F}_{t-}\right]$$
$$+ \frac{f(t - T_{N([0,t))})^2 \overline{F}(T-t)^2}{\overline{F}(T - T_{N([0,t))})^2}\mathbb{P}\left(N((t,T]) = 0 \mid \mathcal{F}_{t-}\right)$$
$$\leq f(t - T_{N([0,t))})^2 h(t - T_{N([0,t))})^2\left(1 - \mathbb{P}\left(N((t,T]) = 0 \mid \mathcal{F}_{t-}\right)\right)$$
$$+ \frac{f(t - T_{N([0,t))})^2 \overline{F}(T-t)^2}{\overline{F}(T - T_{N([0,t))})^2}\mathbb{P}\left(N((t,T]) = 0 \mid \mathcal{F}_{t-}\right).$$
$$(4.6)$$

Hence by Lemma 2.7 and Lemma 1 in [2], we have

$$\int_0^T \frac{\mathbb{V}\mathrm{ar}[\pi_t \mid \mathcal{F}_{t-}]}{\mathbb{E}\left[\pi_t \mid \mathcal{F}_{t-}\right]}\,\mathrm{d}t$$
$$= \int_0^T \left(\frac{\mathbb{E}\left[\pi^2_t \mid \mathcal{F}_{t-}\right]}{\mathbb{E}\left[\pi_t \mid \mathcal{F}_{t-}\right]} - \mathbb{E}\left[\pi_t \mid \mathcal{F}_{t-}\right]\right)\mathrm{d}t$$
$$= \int_0^T \frac{f(t - T_{N([0,t))})}{\overline{F}(t - T_{N([0,t))})}\left[\frac{\overline{F}(t - T_{N([0,t))})^2\mathbb{E}\left[\pi^2_t \mid \mathcal{F}_{t-}\right]}{f(t - T_{N([0,t))})^2} - 1\right]\mathrm{d}t$$
$$\leq \int_0^T \frac{f(t - T_{N([0,t))})}{\overline{F}(t - T_{N([0,t))})}\left[\overline{F}(t - T_{N([0,t))})^2 h(t - T_{N([0,t))})^2\left(1 - \mathbb{P}\left(N((t,T]) = 0 \mid \mathcal{F}_{t-}\right)\right)\right.$$

$$+ \frac{\overline{F}(t - T_{N([0,t))})^2 \overline{F}(T - t)^2}{\overline{F}(T - t + (t - T_{N([0,t))}))^2} \mathbb{P}\big(N((t,T]) = 0 \mid \mathcal{F}_{t-}\big) - 1\bigg] \, \mathrm{d}t$$

$$= \int_0^T \frac{f(t - T_{N([0,t))})}{\overline{F}(t - T_{N([0,t))})} \bigg[\frac{\overline{F}(t - T_{N([0,t))})^2 \overline{F}(T - t)^2}{\overline{F}(T - t + (t - T_{N([0,t))}))^2} \exp\bigg(-\int_t^T \frac{f(s - t + (t - T_{N([0,t))}))}{\overline{F}(s - t + (t - T_{N([0,t))}))} \, \mathrm{d}s \bigg)$$

$$+ \overline{F}(t - T_{N([0,t))})^2 h(t - T_{N([0,t))})^2 \bigg(1 - \exp\bigg(-\int_t^T \frac{f(s - t + (t - T_{N([0,t))}))}{\overline{F}(s - t + (t - T_{N([0,t))}))} \, \mathrm{d}s \bigg) \bigg) - 1 \bigg] \, \mathrm{d}t$$

$$\leq \int_0^T \sup_{z \in [0,t]} \frac{f(z)}{\overline{F}(z)} \bigg[\frac{\overline{F}(z)^2 \overline{F}(T - t)^2}{\overline{F}(T - t + z)^2} \exp\bigg(-\int_t^T \frac{f(s - t + z)}{\overline{F}(s - t + z)} \, \mathrm{d}s \bigg)$$

$$+ \overline{F}(z)^2 h(z)^2 \bigg(1 - \exp\bigg(-\int_t^T \frac{f(s - t + z)}{\overline{F}(s - t + z)} \, \mathrm{d}s \bigg) \bigg) - 1 \bigg] \, \mathrm{d}t,$$

$\mathbb{P}$-almost surely. It remains to notice that since $f(u) = -\mathrm{d}\overline{F}(u) \, / \, \mathrm{d}u$, $u \geq 0$, we have

$$\exp\bigg(-\int_t^T \frac{f(s - t + z)}{\overline{F}(s - t + z)} \, \mathrm{d}s \bigg) = \frac{\overline{F}(T - t + z)}{\overline{F}(z)},$$

and so, by (3.8), the corresponding assumption (3.3) holds and the claim follows by Theorem 3.1. □

Proof of Corollary 3.9. By Lemma 2.7 in [8], N has a Papangelou conditional intensity π_t defined by (3.11). The assumptions on the parameters (h, ϕ) guarantee $\lambda_t \leq \phi(0)$ and

$$\mathrm{E}_t(N) \leq \mathrm{e}^{\|\phi\|_{\mathrm{Lip}} \int_0^{T-t} h(s) \, \mathrm{d}s} \leq \mathrm{e}^{\|\phi\|_{\mathrm{Lip}} \|h\|_{L^1([0,T], \mathcal{B}([0,T]), \mathrm{d}t)}}. \tag{4.7}$$

Consequently, the inequalities in (3.12) hold and by Remark 3.4 the corresponding conditions (3.1) and (3.2) are satisfied. A straightforward computation gives

$$\int_0^T \frac{\mathrm{Var}\big[\pi_t \big| \mathcal{F}_{t-}\big]}{\mathbb{E}\big[\pi_t \big| \mathcal{F}_{t-}\big]} \, \mathrm{d}t = \int_0^T \lambda_t \left(\mathbb{E}\big[(\mathrm{E}_t(N))^2 \big| \mathcal{F}_{t-} \big] - 1 \right) \mathrm{d}t$$

$$\leq \phi(0) \int_0^T \left(\mathrm{e}^{2\|\phi\|_{\mathrm{Lip}} \int_0^\tau h(z) \, \mathrm{d}z} - 1 \right) \mathrm{d}\tau < 1,$$

and therefore the corresponding assumption (3.3) holds. The claim follows by Theorem 3.1. □

Proof of Corollary 3.11. It is well-known (see e.g. [18]) that N has Papangelou conditional intensity $\pi_{(t,x)}$ defined by (3.15). The inequality $\pi_{(t,x)}(\omega) \leq \beta(t, x)$ easily follows by $\lambda_{(t,x)} \leq \beta(t, x)$ and so N is locally stable. Since $\lambda_{(t,x)} \geq \alpha(t, x)$, condition (3.6) is implied by (3.14). So, by Remark 3.5, Theorem 3.1 can be applied and the claim follows. □

4.3 Proof of Theorem 3.12

The proof of Theorem 3.12 is based on two preliminary propositions. The first one consists in a result from [10] (see Theorem 3 therein, as well as Theorem 1.1 in [19]). The second one, whose proof is given at the end of this subsection, provides a stochastic convex inequality for functionals of marked point processes and generalizes Theorem 4.1-(ii) in [12]. We recall that in general, χ denotes a Polish space equipped with its Borel σ-field $\mathcal{B}(\chi)$ and d denotes a lower semi-continuous metric on χ which does not necessarily generate the topology on χ.

Proposition 4.1. *Let $c : [0, \infty) \to [0, \infty]$ be a non-decreasing, left-continuous and convex function with $c(0) = 0$, and let $\mu \in M_1(\chi, d)$. Then*

$$c(\mathcal{T}_d(\sigma, \mu)) \leq H(\sigma \mid \mu), \quad \text{for any } \sigma \in M_1(\chi, d)$$

if and only if, for any function $f : \chi \to \mathbb{R}$ which is measurable, bounded, Lipschitz continuous (with respect to the metric d) with Lipschitz constant less than or equal to 1, i.e. $\sup_{x \neq y} \frac{|f(x)-f(y)|}{d(x,y)} \leq 1$, we have

$$\int_\chi e^{\theta\left(f(x) - \int_\chi f(z)\mu(\mathrm{d}z)\right)}\mu(\mathrm{d}x) \leq e^{c^\odot(\theta)}, \quad \text{for any } \theta \in \mathbb{R}_+.$$

Proposition 4.2. *Let the notation of Theorem 3.28 prevail. Assume that N satisfies (2.2) with a dominating function β, $G \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$ and $\left|\varphi_{(t,x)}^{(G)}\right| \leq h(t,x)$, $\mathrm{d}t\mathbb{P}(\mathrm{d}\omega)\nu(\mathrm{d}x)$-almost everywhere, for some deterministic function h such that*

$$\int_{\mathbb{R}_+ \times E} |h(t,x)|^2 \beta(t,x)\,\mathrm{d}t\nu(\mathrm{d}x) < \infty.$$

Then, letting $\mathbb{E}'$ denote the expectation corresponding to a probability measure $\mathbb{P}'$ on $(\Omega, \mathcal{F}_\infty)$ under which N is a Poisson process on $\mathbb{R}_+ \times E$ with intensity function β,

$$\mathbb{E}\big[\phi(G - \mathbb{E}[G])\big] \leq \mathbb{E}'\left[\phi\left(\int_{\mathbb{R}_+ \times E} h(t,x)\big(N(\mathrm{d}t \times \mathrm{d}x) - \beta(t,x)\,\mathrm{d}t\nu(\mathrm{d}x)\big)\right)\right], \qquad (4.8)$$

for all twice continuously differentiable convex functions $\phi : \mathbb{R} \to \mathbb{R}$ such that ϕ' is convex.

Proof of Theorem 3.12. We take $\chi = \mathbb{R}$ and let $f : \mathbb{R} \to \mathbb{R}$ be bounded and Lipschitz continuous with Lipschitz constant less than or equal to 1 (i.e $\sup_{x \neq y} |f(x) - f(y)|/d(x,y) \leq 1$). We note that for $\mathrm{d}t\nu(\mathrm{d}x)\mathbb{P}(\mathrm{d}\omega)$-almost every (t, x, ω),

$$p(\pi f \circ G^+)_{(t,x)} - p(f \circ G)_{(t,x)} p(\pi)_{(t,x)}$$
$$= \mathbb{E}[\pi_{(t,x)} D_{(t,x)}(f \circ G) \mid \mathcal{F}_{t^-}] + \mathbb{E}\big[\big(f \circ G - f(\mathbb{E}[G \mid \mathcal{F}_{t^-}])\big)\big(\pi_{(t,x)} - \mathbb{E}\big[\pi_{(t,x)}\big|\mathcal{F}_{t^-}\big]\big) \mid \mathcal{F}_{t^-}\big],$$

and so

$$\left|\varphi_{(t,x)}^{(f \circ G)}(\omega)\right|$$
$$\leq \frac{\mathbb{E}\left[\pi_{(t,x)}\left|f \circ G_{(t,x)}^+ - f \circ G\right| + \left|f \circ G - f(\mathbb{E}[G \mid \mathcal{F}_{t^-}])\right|\left|\pi_{(t,x)} - \mathbb{E}\left[\pi_{(t,x)} \mid \mathcal{F}_{t^-}\right]\right| \mid \mathcal{F}_{t^-}\right](\omega)}{\mathbb{E}\left[\pi_{(t,x)} \mid \mathcal{F}_{t^-}\right](\omega)}$$
$$\leq \frac{\mathbb{E}\left[\pi_{(t,x)}\|D_{(t,x)}G\|_d + \|G - \mathbb{E}[G \mid \mathcal{F}_{t^-}]\|_d\left|\pi_{(t,x)} - \mathbb{E}\left[\pi_{(t,x)} \mid \mathcal{F}_{t^-}\right]\right| \mid \mathcal{F}_{t^-}\right](\omega)}{\mathbb{E}\left[\pi_{(t,x)} \mid \mathcal{F}_{t^-}\right](\omega)}$$
$$\leq h(t,x).$$

So, for any $\theta \geq 0$, by Proposition 4.2 with $\phi(x) := e^{\theta x}$ and $f(G)$ in place of G, we have

$$\mathbb{E}\left[e^{\theta(f(G) - \mathbb{E}[f(G)])}\right] \leq \mathbb{E}'\left[\exp\left(\theta \int_{\mathbb{R}_+ \times E} h(t,x)\big(N(\mathrm{d}t \times \mathrm{d}x) - \beta(t,x)\,\mathrm{d}t\nu(\mathrm{d}x)\big)\right)\right]$$
$$= e^{\Lambda(\theta)},$$

where Λ is defined by (3.19). A straightforward computation shows that Λ is non-negative, non-decreasing, left-continuous and convex, with $\Lambda(0) = 0$. Thus, by Proposition 1 in [10] $\Lambda^{\odot\odot} = \Lambda$. The transportation cost inequality (3.17) then follows by Proposition 4.1. Now, assuming that h is bounded by $M > 0$, we have

$$\Lambda(\theta) \leq \frac{e^{\theta M} - \theta M - 1}{M^2} \int_{\mathbb{R}_+ \times E} h(t,z)^2 \beta(t,z)\,\mathrm{d}t\nu(\mathrm{d}z),$$

and as in the proof of Theorem 2.6 in [19], we get the inequality (3.20). $\qquad\square$

Proof of Proposition 4.2. This proof is inspired by that of Lemma 3.2 of [28]. Throughout this proof all the random quantities are defined on the product probability space $(\Omega^2, \mathcal{F}_\infty \otimes \mathcal{F}_\infty, \mathbb{P} \otimes \mathbb{P}')$, and we let $N'(\omega, \omega') = \omega'$ and $N(\omega, \omega') = \omega$. With an abuse of notation, we set $p(\pi)_{(t,x)}(\omega, \omega') := p(\pi)_{(t,x)}(\omega)$, $\varphi^{(G)}_{(t,x)}(\omega, \omega') := \varphi^{(G)}_{(t,x)}(\omega)$, and we denote by $\widetilde{\mathbb{E}}$ the expectation with respect to $\mathbb{P} \otimes \mathbb{P}'$.

Let $\{M_t\}_{t \in \mathbb{R}_+}$ and $\{M_t^*\}_{t \in \mathbb{R}_+}$ be the stochastic processes defined, respectively, by

$$M_t(\omega, \omega') := \int_{[0,t] \times E} \varphi^{(G)}_{(s,x)}(\omega)(\omega(\mathrm{d}s \times \mathrm{d}x) - p(\pi)_{(s,x)}(\omega)\,\mathrm{d}s\nu(\mathrm{d}x))$$

and

$$M_t^*(\omega, \omega') := \int_{[t,\infty) \times E} h(s,x)(\omega'(\mathrm{d}s \times \mathrm{d}x) - \beta(s,x)\,\mathrm{d}s\nu(\mathrm{d}x)).$$

Let $\{\mathcal{H}_t\}_{t \geq 0}$ and $\{\mathcal{H}_t^*\}_{t \geq 0}$ be, respectively, the forward and backward filtrations defined by

$$\mathcal{H}_t := \mathcal{F}_t \otimes \mathcal{F}_\infty \quad \text{and} \quad \mathcal{H}_t^* := \mathcal{F}_\infty \otimes \mathcal{F}_{[t,\infty)}, \quad t \in \mathbb{R}_+.$$

By Corollary C4 p. 235 in [1] and standard properties of the conditional expectation, we have that $\{M_t\}_{t \geq 0}$ is an $\mathcal{H}^*$-adapted $\mathcal{H}$-martingale and $\{M_t^*\}_{t \geq 0}$ is an $\mathcal{H}$-adapted $\mathcal{H}^*$-backward martingale. Letting $\varepsilon_{(t,u)}$ and ε_t denote, respectively, the Dirac measure at $(t,u) \in \mathbb{R}_+ \times \mathbb{R}$ and at $t \in \mathbb{R}_+$, we define the jump measures of $\{M_t\}_{t \in \mathbb{R}_+}$ and $\{M_t^*\}_{t \in \mathbb{R}_+}$ respectively by

$$\mu(\mathrm{d}s \times \mathrm{d}\tau) := \sum_{t > 0 \,:\, \Delta M_t \neq 0} \varepsilon_{(t, \Delta M_t)}(\mathrm{d}s \times \mathrm{d}\tau) = \sum_{(t,x) \in \mathrm{Supp}(N)} \varepsilon_{\left(t, \varphi^{(G)}_{(t,x)}\right)}(\mathrm{d}s \times \mathrm{d}\tau)$$

and

$$\mu^*(\mathrm{d}s \times \mathrm{d}\tau) := \sum_{(t,x) \in \mathrm{Supp}(N')} \varepsilon_{(t, h(t,x))}(\mathrm{d}s \times \mathrm{d}\tau),$$

where $\Delta M_t := M_t - M_{t^-}$. For any fixed $t \in \mathbb{R}_+$, denote by $\nu_t(\mathrm{d}\tau)$ the (random) image measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ of $p(\pi)_{(t,x)}\nu(\mathrm{d}x)$ by the mapping $E \ni x \mapsto \varphi^{(G)}_{(t,x)} \in \mathbb{R}$, i.e. for any bounded and measurable $f : \mathbb{R} \to \mathbb{R}$,

$$\int_{\mathbb{R}} f(\tau)\,\nu_t(\mathrm{d}\tau) := \int_E f\left(\varphi^{(G)}_{(t,x)}(\omega)\right) p(\pi)_{(t,x)}(\omega)\,\nu(\mathrm{d}x),$$

and similarly let ν_t^* be the measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ defined by

$$\int_{\mathbb{R}} f(\tau)\,\nu_t^*(\mathrm{d}\tau) := \int_E f\left(h(t,x)\right)\beta(t,x)\,\nu(\mathrm{d}x).$$

It turns out that $\nu_t(\mathrm{d}\tau)\,\mathrm{d}t$ is the dual $\mathcal{H}$-predictable projection of μ and $\nu_t^*(\mathrm{d}\tau)\,\mathrm{d}t$ is the dual $\mathcal{H}^*$-predictable projection of μ^*. Indeed focusing e.g. on μ, again by Corollary C4 p. 235 in [1] and standard properties of the conditional expectation, for any $t, \Delta t \geq 0$ and $A \in \mathcal{B}(\mathbb{R})$,

$$\widetilde{\mathbb{E}}\left[\mu([t, t+\Delta t] \times A) - \int_t^{t+\Delta t} \nu_s(A)\,\mathrm{d}s \Big| \mathcal{H}_t\right]$$
$$= \widetilde{\mathbb{E}}\left[\int_{[t,t+\Delta t] \times E} \mathbb{1}_{\left\{\varphi^{(G)}_{(s,x)} \in A\right\}}(N(\mathrm{d}s \times \mathrm{d}x) - p(\pi)_{(s,x)}\mathrm{d}s\nu(\mathrm{d}x)) \Big| \mathcal{H}_t\right] = 0.$$

Consequently, conditions (3.1), (3.2) and (3.3) of [12] are verified. We also note that condition (3.4) of [12] is trivially satisfied with $H \equiv H^* \equiv 0$. For any $t \in \mathbb{R}_+$, we define

the following (random) measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$:

$$\widetilde{\nu}_t(\mathrm{d}\tau) := |\tau|^2 \nu_t(\mathrm{d}\tau) \quad \text{and} \quad \widetilde{\nu}_t^*(\mathrm{d}\tau) := |\tau|^2 \nu_t^*(\mathrm{d}\tau).$$

For any $u \in \mathbb{R}$, we have

$$
\begin{aligned}
\widetilde{\nu}_t([u,\infty)) &= \int_E \mathbb{1}_{[u,\infty)}\big(\varphi_{(t,x)}^{(G)}\big)\big|\varphi_{(t,x)}^{(G)}\big|^2 p(\pi)_{(t,x)}\,\nu(\mathrm{d}x) \\
&\leq \int_E \mathbb{1}_{[u,\infty)}\big(h(t,x)\big)\big|h(t,x)\big|^2 \beta(t,x)\,\nu(\mathrm{d}x) \\
&= \widetilde{\nu}_t^*([u,\infty)).
\end{aligned}
$$

Furthermore, for any $u \in \mathbb{R}$,

$$\int_0^\infty \widetilde{\nu}_t^*([u,\infty))\,\mathrm{d}t \leq \int_{\mathbb{R}_+ \times E} \big|h(t,x)\big|^2 \beta(t,x)\,\mathrm{d}t\nu(\mathrm{d}x) < \infty.$$

Therefore, for any $u \in \mathbb{R}$ and any $(\omega, \omega') \in \Omega^2$, $\widetilde{\nu}_t^*([u,\infty)) < \infty$ $\mathrm{d}t$-almost everywhere, and so by Theorem 3.3 in [12] we have

$$\widetilde{\mathbb{E}}[\phi(M_t + M_t^*)] \leq \widetilde{\mathbb{E}}[\phi(M_0^*)],$$

for all $t \in \mathbb{R}_+$ and any function ϕ as in the statement. By Theorem 3.28 and Proposition 2.5 we have

$$
\begin{aligned}
\widetilde{\mathbb{E}}[(M_t + M_t^* + \mathbb{E}[G] - G)^2] &\leq 2\bigg(\widetilde{\mathbb{E}}\bigg[\bigg(\int_{[t,\infty)\times E} h(s,x)\big(N'(\mathrm{d}s \times \mathrm{d}x) - \beta(s,x)\,\mathrm{d}s\nu(\mathrm{d}x)\big)\bigg)^2\bigg] \\
&\quad + \widetilde{\mathbb{E}}\bigg[\bigg(\int_{[t,\infty)\times E} \varphi_{(s,x)}^{(G)}\big(N(\mathrm{d}s \times \mathrm{d}x) - p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\big)\bigg)^2\bigg]\bigg) \\
&= 2\int_{[t,\infty)\times E} |h(s,x)|^2 \beta(s,x)\,\mathrm{d}s\nu(\mathrm{d}x) \\
&\quad + 2\mathbb{E}\bigg[\int_{[t,\infty)\times E} \big|\varphi_{(s,x)}^{(G)}\big|^2 p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\bigg] \xrightarrow[t\to\infty]{} 0.
\end{aligned}
$$

Thus, there exists a sequence $\{t_n\}_{n\geq 1}$ such that $M_{t_n} + M_{t_n}^* \to G - \mathbb{E}G$, $\widetilde{\mathbb{P}}$-almost surely, and therefore by Fatou's lemma

$$\widetilde{\mathbb{E}}[\phi(G - \mathbb{E}[G])] = \widetilde{\mathbb{E}}\Big[\liminf_{n\to\infty} \phi(M_{t_n} + M_{t_n}^*)\Big] \leq \liminf_{n\to\infty} \widetilde{\mathbb{E}}[\phi(M_{t_n} + M_{t_n}^*)] \leq \widetilde{\mathbb{E}}[\phi(M_0^*)],$$

which is exactly (4.8). $\qquad\square$

4.4 Proofs of Proposition 3.16 and Corollaries 3.17, 3.18, 3.19

Proof of Proposition 3.16. We shall apply Theorem 3.12. For any $(t,x) \in \mathbb{R}_+ \times E$, we have $\big|D_{(t,x)}G\big| = |g(t,x)| \leq g_1(t,x)$, and by the Cauchy-Schwarz inequality

$$
\begin{aligned}
\frac{\mathbb{E}\big[\big|G - \mathbb{E}[G \mid \mathcal{F}_{t^-}]\big|\big|\pi_{(t,x)} - p(\pi)_{(t,x)}\big| \mid \mathcal{F}_{t^-}\big](\omega)}{p(\pi)_{(t,x)}(\omega)} & \\
&\hspace{-6cm}\leq \mathbb{E}\big[(G - \mathbb{E}[G \mid \mathcal{F}_{t^-}])^2 \mid \mathcal{F}_{t^-}\big]^{1/2}\left\|\frac{\sqrt{\mathbb{V}\mathrm{ar}\big[\pi_{(t,x)} \mid \mathcal{F}_{t^-}\big]}}{p(\pi)_{(t,x)}}\right\|_{L^\infty(\Omega,\mathcal{F}_\infty,\mathbb{P})}.
\end{aligned}
$$

Additionally, by (2.4)

$$\mathbb{E}[G \mid \mathcal{F}_{t^-}] = \int_{(0,t)\times E} g(s,y)\,N(\mathrm{d}s \times \mathrm{d}y) + \int_{[t,\infty)\times E} g(s,y)\mathbb{E}[\lambda_{(s,y)} \mid \mathcal{F}_{t^-}]\,\mathrm{d}s\nu(\mathrm{d}y),$$

and so by Proposition 2.5*(ii)*

$$\mathbb{E}\big[(G - \mathbb{E}[G \mid \mathcal{F}_{t-}])^2 \mid \mathcal{F}_{t-}\big]$$

$$= \mathbb{E}\Bigg[\bigg(\int_{[t,\infty)\times E} g(s,y)\,N(\mathrm{d}s \times \mathrm{d}y) - \int_{[t,\infty)\times E} g(s,y)\mathbb{E}[\lambda_{(t,y)} \mid \mathcal{F}_{t-}]\,\mathrm{d}s\nu(\mathrm{d}y)\bigg)^2 \,\bigg|\, \mathcal{F}_{t-}\Bigg]$$

$$\leq 2\mathbb{E}\Bigg[\bigg(\int_{[t,\infty)\times E} g(s,y)\,(N(\mathrm{d}s \times \mathrm{d}y) - \lambda_{(s,y)}\mathrm{d}s\nu(\mathrm{d}y)\bigg)^2$$

$$+ \bigg(\int_{[t,\infty)\times E} g(s,y)(\lambda_{(s,y)} - \mathbb{E}[\lambda_{(s,y)} \mid \mathcal{F}_{t-}])\,\mathrm{d}s\nu(\mathrm{d}y)\bigg)^2 \,\bigg|\, \mathcal{F}_{t-}\Bigg]$$

$$= 2\Bigg[\iint_{[t,\infty)\times E} g(s,y)^2\mathbb{E}[\lambda_{(s,y)} \mid \mathcal{F}_{t-}]\,\mathrm{d}s\nu(\mathrm{d}y)$$

$$+ \mathbb{E}\Bigg[\bigg(\int_{[t,\infty)\times E} g(s,y)(\lambda_{(s,y)} - \mathbb{E}[\lambda_{(s,y)} \mid \mathcal{F}_{t-}])\,\mathrm{d}s\nu(\mathrm{d}y)\bigg)^2 \,\bigg|\, \mathcal{F}_{t-}\Bigg]\Bigg]$$

$$\leq 2\Bigg[\int_{[t,\infty)\times E} g(s,y)^2\beta(s,y)\,\mathrm{d}s\nu(\mathrm{d}y) + \bigg(\int_{[t,\infty)\times E} |g(s,y)|\beta(s,y)\,\mathrm{d}s\nu(\mathrm{d}y)\bigg)^2\Bigg]. \qquad (4.9)$$

The claim follows by Theorem 3.12. $\qquad\square$

Proof of Corollary 3.17. By the proof of Proposition 2.10 in [8] we have $\pi_t \leq \beta$, where β is given by (3.26). Additionally, by (3.24), (4.6) and the form of the stochastic intensity for renewal processes,

$$\frac{\sqrt{\mathbb{V}\mathrm{ar}[\pi_t \mid \mathcal{F}_{t-}]}}{\lambda_t} = \sqrt{\mathbb{E}\left[\frac{\pi_t^2}{\lambda_t^2} \,\Big|\, \mathcal{F}_{t-}\right] - 1}$$

$$\leq \sqrt{\overline{h}^2\overline{F}(t - T_{N([0,t))})^2 + \frac{\overline{F}(T-t)^2\overline{F}(t - T_{N([0,t))})^2}{\overline{F}(T - T_{N([0,t))})^2} - 1}$$

$$\leq \sqrt{\overline{h}^2 + \left(\int_C^T f(x)\,\mathrm{d}x\right)^{-2} - 1}.$$

The claim follows by Proposition 3.16. $\qquad\square$

Proof of Corollary 3.18. By the first inequality in (3.12) we can take as dominating function of π_t

$$\beta(t) = \phi(0)\exp\left(\|\phi\|_{\mathrm{Lip}}\int_0^{T-t} h(z)\,\mathrm{d}z\right).$$

Additionally, by the first inequality in (4.7),

$$\frac{\sqrt{\mathbb{V}\mathrm{ar}[\pi_t \mid \mathcal{F}_{t-}]}}{\lambda_t} = \sqrt{\mathbb{E}\left[\frac{\pi_t^2}{\lambda_t^2} \,\Big|\, \mathcal{F}_{t-}\right] - 1} \leq \sqrt{\exp\left(2\|\phi\|_{\mathrm{Lip}}\int_0^{T-t} h(z)\,\mathrm{d}z\right) - 1}.$$

The claim follows by Proposition 3.16. $\qquad\square$

Proof of Corollary 3.19. We already noticed in the proof of Corollary 3.11 that β is a dominating function of the Papangelou conditional intensity. The claim easily follows by Proposition 3.16 noticing that

$$\frac{\sqrt{\mathbb{V}\mathrm{ar}[\pi_{(t,x)} \mid \mathcal{F}_{t-}]}}{\lambda_{(t,x)}} = \sqrt{\mathbb{E}\left[\frac{\pi_{(t,x)}^2}{\lambda_{(t,x)}^2} \,\Big|\, \mathcal{F}_{t-}\right] - 1} \leq \sqrt{\frac{\beta(t,x)^2}{\alpha(t,x)^2} - 1}. \qquad\square$$

4.5 Proof of Theorem 3.20

Since the proof is conceptually similar to that of Theorem 3.12, we only emphasize the main differences. We take $\chi = \Omega$, $d := d_\varphi$ and let $F : (\Omega, d_\varphi) \to \mathbb{R}$ be bounded and Lipschitz continuous with Lipschitz constant less than or equal to one. We have for $\mathrm{d}t\nu(\mathrm{d}x)$-almost every (t, x),

$$
\left| \varphi_{(t,x)}^{(F)} \right|
$$

$$
\leq \frac{\mathbb{E}\left[\pi_{(t,x)} \left| F(N + \varepsilon_{(t,x)}) - F(N) \right| + \left| F(N) - F(N|_{[0,t)\times E}) \right| \left| \pi_{(t,x)} - \mathbb{E}\left[\pi_{(t,x)} \mid \mathcal{F}_{t^-} \right] \right| \mid \mathcal{F}_{t^-} \right]}{\mathbb{E}\left[\pi_{(t,x)} \mid \mathcal{F}_{t^-} \right]}
$$

$$
\leq \frac{\mathbb{E}\left[\pi_{(t,x)} d_\varphi(N + \varepsilon_{(t,x)}, N) + d_\varphi(N, N|_{[0,t)\times E}) \left| \pi_{(t,x)} - \mathbb{E}\left[\pi_{(t,x)} \mid \mathcal{F}_{t^-} \right] \right| \mid \mathcal{F}_{t^-} \right]}{\mathbb{E}\left[\pi_{(t,x)} \mid \mathcal{F}_{t^-} \right]}
$$

$$
= \varphi(t,x) + \frac{\mathbb{E}\left[\left(\int_{[t,\infty)\times E} \varphi(s,y) N(\mathrm{d}s \times \mathrm{d}y) \right) \left| \pi_{(t,x)} - \mathbb{E}\left[\pi_{(t,x)} \mid \mathcal{F}_{t^-} \right] \right| \mid \mathcal{F}_{t^-} \right]}{\mathbb{E}\left[\pi_{(t,x)} \mid \mathcal{F}_{t^-} \right]}
$$

$$
\leq \varphi(t,x) + \left\| \frac{\sqrt{\mathbb{V}\mathrm{ar}\left[\pi_{(t,x)} \mid \mathcal{F}_{t^-} \right]}}{p(\pi)_{(t,x)}} \right\|_{L^\infty(\Omega, \mathcal{F}_\infty, \mathbb{P})} \sqrt{\mathbb{E}\left[\left(\int_{[t,\infty)\times E} \varphi(s,y) N(\mathrm{d}s \times \mathrm{d}y) \right)^2 \mid \mathcal{F}_{t^-} \right]},
$$

$\mathbb{P}(\mathrm{d}\omega)$-almost everywhere. Proceeding similarly to the series of inequalities (4.9), we obtain

$$
\left| \varphi_{(t,x)}^{(F)} \right| \leq \varphi(t,x) + \sqrt{2} \Bigg[\int_{[t,\infty)\times E} \varphi(s,y)^2 \beta(s,y)\, \mathrm{d}s\nu(\mathrm{d}y)
$$

$$
+ \left(\int_{[t,\infty)\times E} \varphi(s,y)\beta(s,y)\, \mathrm{d}s\nu(\mathrm{d}y) \right)^2 \Bigg]^{1/2} \times \left\| \frac{\sqrt{\mathbb{V}\mathrm{ar}\left[\pi_{(t,x)} \mid \mathcal{F}_{t^-} \right]}}{p(\pi)_{(t,x)}} \right\|_{L^\infty(\Omega, \mathcal{F}_\infty, \mathbb{P})}.
$$

So, for any $\theta \geq 0$, by Proposition 4.2 with $\phi(x) := \mathrm{e}^{\theta x}$, we have

$$
\mathbb{E}\left[\mathrm{e}^{\theta(F - \mathbb{E}[F])} \right] \leq \mathbb{E}'\left[\exp\left(\theta \int_{\mathbb{R}_+ \times E} h_\varphi(t,x)\big(N(\mathrm{d}t \times \mathrm{d}x) - \beta(t,x)\, \mathrm{d}t\nu(\mathrm{d}x) \big) \right) \right] = \mathrm{e}^{\Lambda_\varphi(\theta)}.
$$

Note also that

$$
\int_\Omega d_\varphi(\omega, \mathbf{0})\, \mathbb{P}(\mathrm{d}\omega) = \mathbb{E}\left[d_\varphi(N, \mathbf{0}) \right] = \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \varphi(s,y)\, N(\mathrm{d}s \times \mathrm{d}y) \right]
$$

$$
\leq \int_{\mathbb{R}_+ \times E} \varphi(s,y)\beta(s,y)\, \mathrm{d}s\nu(\mathrm{d}y) < \infty.
$$

The remainder of the proof is similar to that of the final part of Theorem 3.12.

4.6 Proof of Theorem 3.26

The proof of Theorem 3.26 is based on two preliminary propositions, which extend to our setting Propositions 4.2 and 4.3 in [29], respectively.

Proposition 4.3. *Let the assumptions and notation of Theorem 3.28 prevail and let G be a random variable on $(\Omega, \mathcal{F}_\infty, \mathbb{P})$ which satisfies $0 < c_0 \leq G \leq c_1$ almost surely, for some positive constants $c_0, c_1 > 0$. Setting*

$$
\phi_{(t,x)}^{(G)} := \frac{\varphi_{(t,x)}^{(G)}}{p(G)_{(t,x)}}, \quad t \in \mathbb{R}_+, x \in E, \tag{4.10}
$$

we have $\phi^{(G)} \in \mathcal{H}$. Additionally, for any $t \in \mathbb{R}_+$ we have

$$\mathbb{E}\big[G\big|\mathcal{F}_t\big] = \mathbb{E}[G] \cdot \mathcal{E}_t(\phi^{(G)}). \tag{4.11}$$

Proposition 4.4. *Let the assumptions and notation of Theorem 3.28 prevail and, for $\phi \in \mathcal{H}$, let $\mathbb{P}_\phi$ be defined by (3.34). For any predictable stochastic process ψ such that*

$$\mathbb{E}\Bigg[\bigg(\int_{[0,T]\times E} |\psi_{(s,x)}|^2\, p(\pi)_{(s,x)}\, \mathrm{d}s\nu(\mathrm{d}x)\bigg)^2\Bigg] < \infty, \quad \text{for any } T > 0 \tag{4.12}$$

we have that, under $\mathbb{P}_\phi$,

$$\bigg\{\int_{[0,t]\times E} \psi_{(s,x)}(N(\mathrm{d}s\times\mathrm{d}x) - p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)) - \int_{[0,t]\times E} \psi_{(s,x)}\phi_{(s,x)}p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\bigg\}_{t\in\mathbb{R}_+}$$

is a square-integrable $\mathcal{F}$-martingale with null mean.

These propositions are proved at the end of this subsection, and we start proving Lemma 3.25.

Proof of Lemma 3.25. By Itô's formula (see e.g. Theorem 5.1 p. 66 of [11]) we have

$$\mathcal{E}_t(\phi) = 1 + \int_{[0,t]\times E} \mathcal{E}_{s^-}(\phi)\phi_{(s,x)}(N(\mathrm{d}s\times\mathrm{d}x) - p(\pi)_{(s,x)}\mathrm{d}s\nu(\mathrm{d}x)), \quad t \in \mathbb{R}_+. \tag{4.13}$$

Moreover, for any $T > 0$ we have

$$\{\mathbb{1}_{[0,T]}(t)\mathcal{E}_t(\phi)\phi_{(t,x)}\}_{(t,x)\in\mathbb{R}_+\times E} \in \mathcal{P}_{1,2}(p(\pi)). \tag{4.14}$$

Indeed, letting $M > 0$ denote a constant such that $|\phi_{(t,x)}| \leq M$ $\mathrm{d}t\nu(\mathrm{d}x)\mathrm{d}\mathbb{P}$-almost everywhere, by the Cauchy-Schwarz inequality, (3.33) and the square integrability of $N([0,T]\times E)$, it follows

$$\mathbb{E}\bigg[\int_{\mathbb{R}_+\times E} |\mathbb{1}_{[0,T]}(t)\mathcal{E}_t(\phi)\phi_{(t,x)}|p(\pi)_{(t,x)}\,\mathrm{d}t\nu(\mathrm{d}x)\bigg]$$
$$\leq M\sqrt{\mathbb{E}\bigg[\int_{[0,T]\times E} \mathcal{E}_t(\phi)^2 p(\pi)_{(t,x)}\,\mathrm{d}t\nu(\mathrm{d}x)\bigg]}\sqrt{\mathbb{E}\big[N([0,T]\times E)\big]}$$
$$< \infty \tag{4.15}$$

and

$$\mathbb{E}\bigg[\int_{\mathbb{R}_+\times E} |\mathbb{1}_{[0,T]}(t)\mathcal{E}_t(\phi)\phi_{(t,x)}|^2 p(\pi)_{(t,x)}\,\mathrm{d}t\nu(\mathrm{d}x)\bigg]$$
$$\leq M^2\mathbb{E}\bigg[\int_{[0,T]\times E} \mathcal{E}_t(\phi)^2 p(\pi)_{(t,x)}\,\mathrm{d}t\nu(\mathrm{d}x)\bigg]$$
$$< \infty. \tag{4.16}$$

By (4.13) and Proposition 2.5-(ii) we then have

$$\mathbb{E}\big[\mathcal{E}_t(\phi)^2\big] = 1 + \mathbb{E}\bigg[\int_{[0,t]\times E} \phi_{(s,x)}^2 \mathcal{E}_s(\phi)^2 p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\bigg], \tag{4.17}$$

which gives the square integrability of $\{\mathcal{E}_t(\phi)\}_{t\in\mathbb{R}_+}$. The martingale property follows by (4.13) and Corollary C4 p. 235 of [1]. □

Proof of Theorem 3.26. We divide the proof in three steps. In the first step we identify $d\mathbb{P}_\phi/d\mathbb{P}$, in the second step we prove the claim when $G \in L^\infty(\Omega, \mathcal{F}_\infty, \mathbb{P})$, in the third step we prove the variational representation in the more general case of functionals G which are bounded from above.

Step1.

Letting $M > 0$ denote a constant such that $|\phi_{(s,x)}| \leq M \, ds\nu(dx)d\mathbb{P}$-almost everywhere, by (4.17) we have

$$\mathbb{E}[\mathcal{E}_t(\phi)^2] \leq 1 + M^2 \int_0^t \mathbb{E}[\mathcal{E}_s(\phi)^2]\left\|\int_E p(\pi)_{(s,x)}\,\nu(dx)\right\|_{L^\infty(\Omega,\mathcal{F}_\infty,\mathbb{P})} ds, \quad t \in \mathbb{R}_+.$$

Let $T > 0$ be arbitrarily fixed. Again by (4.17) we have that $s \mapsto \mathbb{E}[\mathcal{E}_s(\phi)^2]$ is non-decreasing and continuous on $[0, T]$ and by (3.35)

$$s \mapsto \left\|\int_E p(\pi)_{(s,x)}\,\nu(dx)\right\|_{L^\infty(\Omega,\mathcal{F}_\infty,\mathbb{P})}$$

is integrable on $[0, T]$. Therefore, by Grönwall's lemma

$$\mathbb{E}[\mathcal{E}_t(\phi)^2] \leq \exp\left(M^2 \int_0^T \left\|\int_E p(\pi)_{(s,x)}\,\nu(dx)\right\|_{L^\infty(\Omega,\mathcal{F}_\infty,\mathbb{P})} ds\right) \leq \mathrm{e}^{M^2 K}, \quad \text{for any } t \in \mathbb{R}_+$$

and so $\sup_{t\in\mathbb{R}_+}\mathbb{E}[\mathcal{E}_t(\phi)^2] \leq \mathrm{e}^{M^2 K}$. By this latter relation, (4.13), Proposition 2.5-(ii) and (3.35), we have

$$\mathbb{E}[(\mathcal{E}_{t+h}(\phi) - \mathcal{E}_t(\phi))^2] \leq M^2 \int_t^{t+h} \mathbb{E}[\mathcal{E}_s(\phi)^2]\left\|\int_E p(\pi)_{(s,x)}\,\nu(dx)\right\|_{L^\infty(\Omega,\mathcal{F}_\infty,\mathbb{P})} ds$$

$$\leq M^2\mathrm{e}^{M^2 K} \int_t^{t+h} \left\|\int_E p(\pi)_{(s,x)}\,\nu(dx)\right\|_{L^\infty(\Omega,\mathcal{F}_\infty,\mathbb{P})} ds \xrightarrow[t,h\to\infty]{} 0,$$

and thus $\mathcal{E}_t(\phi)$ converges in $L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$ to a random variable X. Letting $\mathcal{E}_\infty(\phi)$ denote the $\mathbb{P}$-almost sure limit of $\mathcal{E}_t(\phi)$ as $t \to \infty$, we necessarily have $X = \mathcal{E}_\infty(\phi)$ almost surely, and so

$$\mathcal{E}_\infty(\phi) \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P}). \tag{4.18}$$

By the martingale property it follows that $\mathcal{E}_t(\phi) = \mathbb{E}[\mathcal{E}_n(\phi)|\mathcal{F}_t]$ for any $t \in \mathbb{R}_+$ and any integer $n > t$. By the L^2-convergence of $\mathcal{E}_t(\phi)$ to $\mathcal{E}_\infty(\phi)$ we easily have that $\mathbb{E}[\mathcal{E}_n(\phi)|\mathcal{F}_t]$ converges to $\mathbb{E}[\mathcal{E}_\infty(\phi)|\mathcal{F}_t]$ in L^1 as $n \to \infty$. This convergence holds almost surely for a suitable subsequence $\{n'\}$ and passing to the limit as $n' \to \infty$ in the equality $\mathcal{E}_t(\phi) = \mathbb{E}[\mathcal{E}_{n'}(\phi)|\mathcal{F}_t]$ we get $\mathcal{E}_t(\phi) = \mathbb{E}[\mathcal{E}_\infty(\phi)|\mathcal{F}_t]$. By this relation we finally deduce $d\mathbb{P}_\phi/d\mathbb{P} = \mathcal{E}_\infty(\phi)$.

Step2.

Using the elementary inequality $|(1 + x)\log(1 + x) - x| \leq x^2/2$, $x > -1$, for any $\phi \in \mathcal{H}$, we have

$$\mathbb{E}_\phi[|L(\phi)|] \leq \frac{1}{2}\,\mathbb{E}\left[\mathcal{E}_\infty(\phi)\int_{\mathbb{R}_+\times E} |\phi_{(s,x)}|^2\,p(\pi)_{(s,x)}\,ds\nu(dx)\right]$$

$$\leq \frac{1}{2}\,\|\mathcal{E}_\infty(\phi)\|_{L^2(\Omega,\mathcal{F}_\infty,\mathbb{P})}\left\|\int_{\mathbb{R}_+\times E} |\phi_{(s,x)}|^2\,p(\pi)_{(s,x)}\,ds\nu(dx)\right\|_{L^2(\Omega,\mathcal{F}_\infty,\mathbb{P})} < \infty,$$

$$\tag{4.19}$$

where the finiteness of the L^2-norms follows by (3.32) and (4.18). Let $\phi \in \mathcal{H}$ and $T > 0$ be arbitrarily fixed, and set $\psi_{(s,x)} := \log(1 + \phi_{(s,x)})$, $(s,x) \in [0,T] \times E$. Clearly ψ is predictable. Additionally, we note that there exists $C > 0$ such that $\left| \log(1 + \phi_{(s,x)}) \right| \leq C |\phi_{(s,x)}|$, $(s,x) \in [0,T] \times E$ (since there exists $c_\phi > -1$ such that $\phi_{(s,x)} \geq c_\phi$). Therefore, by (3.32) we have (4.12). Note that

$$
\begin{aligned}
\mathbb{E}_\phi\big[\log\big(\mathcal{E}_\infty(\phi)\big)\big] &= \mathbb{E}_\phi\bigg[\iint_{\mathbb{R}_+ \times E} \log\big(1 + \phi_{(s,x)}\big)(N(\mathrm{d}s \times \mathrm{d}x) - p(\pi)_{(s,x)}\mathrm{d}s\nu(\mathrm{d}x)) \\
&\qquad + \int_{\mathbb{R}_+ \times E} \big(\log\big(1 + \phi_{(s,x)}\big) - \phi_{(s,x)}\big) \, p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x) \bigg] \\
&= \mathbb{E}_\phi[L(\phi)] + \mathbb{E}_\phi\bigg[\iint_{\mathbb{R}_+ \times E} \log\big(1 + \phi_{(s,x)}\big)(N(\mathrm{d}s \times \mathrm{d}x) - p(\pi)_{(s,x)}\mathrm{d}s\nu(\mathrm{d}x)) \\
&\qquad - \int_{\mathbb{R}_+ \times E} \phi_{(s,x)} \log\big(1 + \phi_{(s,x)}\big) \, p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x) \bigg] \\
&= \mathbb{E}_\phi[L(\phi)], \tag{4.20}
\end{aligned}
$$

where we have used that the martingale provided by Proposition 4.4 has null mean. By Jensen's inequality and this relation it follows

$$
\begin{aligned}
-\log\big(\mathbb{E}\big[\mathrm{e}^{-G}\big]\big) &= -\log\Big(\mathbb{E}_\phi\big[\exp\big(-G - \log\big(\mathcal{E}_\infty(\phi)\big)\big)\big]\Big) \\
&\leq \mathbb{E}_\phi\big[G + \log\big(\mathcal{E}_\infty(\phi)\big)\big] = \mathbb{E}_\phi[G + L(\phi)] < \infty, \tag{4.21}
\end{aligned}
$$

and so

$$
-\log\big(\mathbb{E}\big[\mathrm{e}^{-G}\big]\big) \leq \inf_{\phi \in \mathcal{H}} \mathbb{E}_\phi\big[G + L(\phi)\big]. \tag{4.22}
$$

Setting $F := \mathrm{e}^{-G}$, we have $0 < \mathrm{e}^{-\|G\|_{L^\infty(\Omega, \mathcal{F}_\infty, \mathbb{P})}} \leq F \leq \mathrm{e}^{\|G\|_{L^\infty(\Omega, \mathcal{F}_\infty, \mathbb{P})}}$, $\mathbb{P}$-almost surely. Therefore, by Proposition 4.3, $\mathcal{E}_t(\phi^{(F)}) = \mathbb{E}\big[F \big| \mathcal{F}_t\big]/\mathbb{E}[F]$, for any $t \in \mathbb{R}_+$, and letting t go to infinity we deduce

$$
\mathcal{E}_\infty(\phi^{(F)}) = \frac{\mathbb{E}\big[F \big| \mathcal{F}_\infty\big]}{\mathbb{E}[F]} = \frac{F}{\mathbb{E}[F]},
$$

where $\phi^{(F)}_{(t,x)} := \varphi^{(F)}_{(t,x)}/p(F)_{(t,x)} \in \mathcal{H}$, $(t,x) \in \mathbb{R}_+ \times E$. Therefore

$$
\begin{aligned}
-\log\big(\mathbb{E}\big[\mathrm{e}^{-G}\big]\big) &= -\mathbb{E}_{\phi^{(F)}}\big[\log\big(F \mathcal{E}_\infty(\phi^{(F)})^{-1}\big)\big] = \mathbb{E}_{\phi^{(F)}}\big[G + \log\big(\mathcal{E}_\infty(\phi^{(F)})\big)\big] \\
&= \mathbb{E}_{\phi^{(F)}}\big[G + L(\phi^{(F)})\big], \tag{4.23}
\end{aligned}
$$

where the latter equality follows by (4.20). Combining (4.23) and (4.22) we deduce

$$
-\log\big(\mathbb{E}\big[\mathrm{e}^{-G}\big]\big) = \inf_{\phi \in \mathcal{H}} \mathbb{E}_\phi\big[G + L(\phi)\big]
$$

and the infimum is attained at $\phi^{(F)}$. It remains to show that the infimum is uniquely attained at $\phi^{(F)}$, i.e. if $\phi \in \mathcal{H}$ is a stochastic process at which the infimum is attained then necessarily $\phi_{(t,x)}(\omega) = \phi^{(F)}_{(t,x)}(\omega)$ for $p(\pi)_{(t,x)}(\omega)\mathrm{d}t\nu(\mathrm{d}x)\mathrm{d}\mathbb{P}(\omega)$-almost all (t,x,ω). So let $\phi \in \mathcal{H}$ be a process at which the infimum is attained. Then Jensen's inequality (4.21) holds as an equality and therefore we have

$$
\exp\big(-G - \log\big(\mathcal{E}_\infty(\phi)\big)\big) = \mathbb{E}[F], \quad \mathbb{P}_\phi\text{-almost surely.}
$$

Similarly

$$\exp\big(-G - \log\big(\mathcal{E}_\infty(\phi^{(F)})\big)\big) = \mathbb{E}[F], \quad \mathbb{P}_\phi\text{-almost surely.}$$

Therefore $\mathcal{E}_\infty(\phi) = \mathcal{E}_\infty(\phi^{(F)})$ $\mathbb{P}_\phi$-almost surely. Since the probability measures $\mathbb{P}_\phi$ and $\mathbb{P}$ are equivalent it follows that $\mathcal{E}_\infty(\phi) = \mathcal{E}_\infty(\phi^{(F)})$ $\mathbb{P}$-almost surely. Consequently, for any $t \in \mathbb{R}_+$ we have $\mathcal{E}_t(\phi) = \mathcal{E}_t(\phi^{(F)})$ $\mathbb{P}$-almost everywhere, and so by (4.13) we find

$$\int_{[0,t]\times E} \mathcal{E}_{s-}(\phi)\big(\phi_{(s,x)} - \phi_{(s,x)}^{(F)}\big)(N(\mathrm{d}s \times \mathrm{d}x) - p(\pi)_{(s,x)}\mathrm{d}s\nu(\mathrm{d}x)) = 0.$$

Taking the expectation of the square of this quantity, by Proposition 2.6 we have

$$\mathbb{E}\bigg[\int_{[0,t]\times E} \mathcal{E}_{s-}(\phi)^2\big(\phi_{(s,x)} - \phi_{(s,x)}^{(F)}\big)^2 p(\pi)_{(s,x)}\mathrm{d}s\nu(\mathrm{d}x)\bigg] = 0.$$

Since $\mathcal{E}_{s-}(\phi) > 0$ $\mathrm{d}s\mathrm{d}\mathbb{P}$-almost everywhere, we get $\phi = \phi^{(F)}$ in $\mathcal{P}_2(p(\pi))$, and the proof is complete.

Step3.
For any integer $n \geq 1$, define $G^{(n)} := \max\{G, -n\}$. Since G is upper bounded, we have $G^{(n)} \in L^\infty(\Omega, \mathcal{F}_\infty, \mathbb{P})$. Therefore, by Step 2 we have

$$-\log\big(\mathbb{E}\big[\mathrm{e}^{-G^{(n)}}\big]\big) = \inf_{\phi\in\mathcal{H}} \mathbb{E}_\phi\big[G^{(n)} + L(\phi)\big] \geq \inf_{\phi\in\mathcal{H}} \mathbb{E}_\phi\big[G + L(\phi)\big].$$

In addition, since the sequence $\{G^{(n)}\}_{n\geq 1}$ is non-increasing, taking the limit as $n \to \infty$, by the monotone convergence theorem we obtain $-\log\big(\mathbb{E}\big[\mathrm{e}^{-G}\big]\big) \geq \inf_{\phi\in\mathcal{H}} \mathbb{E}_\phi\big[G + L(\phi)\big]$. For the reversed inequality we note that, for any $\psi \in \mathcal{H}$,

$$-\log\big(\mathbb{E}\big[\mathrm{e}^{-G}\big]\big) = \lim_{n\to\infty} \inf_{\phi\in\mathcal{H}} \mathbb{E}_\phi\big[G^{(n)} + L(\phi)\big] \leq \lim_{n\to\infty} \mathbb{E}_\psi\big[G^{(n)} + L(\psi)\big] = \mathbb{E}_\psi\big[G + L(\psi)\big], \tag{4.24}$$

where the latter equality follows again by the monotone convergence theorem (note that each $G^{(n)}$ is (for n large enough) bounded above by a same constant). Taking the infimum on $\mathcal{H}$ in (4.24) yields the reversed inequality and the proof is complete. $\qquad\square$

Proof of Proposition 4.3. Since N has stochastic intensity $p(\pi)$, the stochastic process

$$\left\{\int_{[0,t]\times E} \varphi_{(s,x)}^{(G)}\,(N(\mathrm{d}s \times \mathrm{d}x) - p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x))\right\}_{t\in\mathbb{R}_+}$$

is an $\mathcal{F}$-martingale, and by Theorem 3.28 we have

$$\mathbb{E}\big[G\big|\mathcal{F}_t\big] = \mathbb{E}[G] + \int_{[0,t]\times E} \varphi_{(s,x)}^{(G)}\,(N(\mathrm{d}s\times\mathrm{d}x) - p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)), \quad \mathbb{P}\text{-almost surely.} \tag{4.25}$$

Letting ϕ be defined by (4.10) and suppressing the dependence of ϕ on G for ease of notation we have, since $0 < c_0 \leq G \leq c_1$,

$$\phi_{(t,x)} = \frac{\varphi_{(t,x)}^{(G)}}{p(G)_{(t,x)}} = \frac{p(\pi G^+)_{(t,x)} - p(G)_{(t,x)}p(\pi)_{(t,x)}}{p(G)_{(t,x)}p(\pi)_{(t,x)}}$$

$$\geq \frac{c_0}{c_1} - 1 > -1, \quad \mathbb{P}\text{-almost surely,}$$

and ϕ is predictable and bounded with $\sup_{(t,x)\in\mathbb{R}_+\times E} |\phi_{(t,x)}| \leq (c_1/c_0) + 1$, $\mathbb{P}$-almost surely. Set

$$X_t := \int_{[0,t]\times E} \phi_{(s,x)}(N(\mathrm{d}s \times \mathrm{d}x) - p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)), \quad t \in \mathbb{R}_+.$$

By Proposition 2.5-(ii) we have

$$\sup_{t \in \mathbb{R}_+} \mathbb{E}\big[X_t^2\big] = \mathbb{E}\bigg[\int_{\mathbb{R}_+ \times E} |\phi_{(s,x)}|^2 p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x)\bigg] < \infty, \tag{4.26}$$

where the finiteness of the latter quantity follows noticing that $|\phi_{(t,x)}| \leq |\varphi_{(t,x)}^{(G)}|/c_0$, $t \in \mathbb{R}_+$, $x \in E$, and $\varphi^{(G)} \in \mathcal{P}_2(p(\pi))$. So $\{X_t\}_{t \in \mathbb{R}_+}$ is a square-integrable $\mathcal{F}$-martingale. By the definition of ϕ, the relation $p(G)_t = \mathbb{E}\big[G\big|\mathcal{F}_{t-}\big]$ $\mathbb{P}$-almost surely and (4.25), we have

$$\mathbb{E}\big[G\big|\mathcal{F}_t\big] = \mathbb{E}[G] + \int_0^t \mathbb{E}\big[G\big|\mathcal{F}_{s-}\big] \, \mathrm{d}X_s, \quad \mathbb{P}\text{-almost surely.} \tag{4.27}$$

Note that

$$(c_1/c_0) + 1 \geq \Delta X_s := X_s - X_{s-} \geq (c_0/c_1) - 1 > -1,$$

and so X is of finite variation (and càdlàg). Therefore, its quadratic variation process is given by

$$[X, X]_t = \sum_{0 < s \leq t} (\Delta X_s)^2 = \int_{[0,t] \times E} \phi_{(s,x)}^2 N(\mathrm{d}s \times \mathrm{d}x) \tag{4.28}$$

and thus the path-by-path continuous part of $[X, X]$ is equal to zero. By (4.27) and Theorem 37 p. 84 in [23], we have

$$\frac{\mathbb{E}\big[G\big|\mathcal{F}_t\big]}{\mathbb{E}[G]} = \exp(X_t) \prod_{0 < s \leq t} (1 + \Delta X_s) \exp(-\Delta X_s)$$

$$= \exp\bigg(X_t + \sum_{0 < s \leq t} (\log(1 + \Delta X_s) - \Delta X_s)\bigg). \tag{4.29}$$

We note that

$$\sum_{0 < s \leq t} (\log(1 + \Delta X_s) - \Delta X_s)$$

$$= \int_{[0,t] \times E} (\log(1 + \phi_{(s,x)}) - \phi_{(s,x)}) N(\mathrm{d}s \times \mathrm{d}x)$$

$$= \int_{[0,t] \times E} (\log(1 + \phi_{(s,x)}) - \phi_{(s,x)})(N(\mathrm{d}s \times \mathrm{d}x) - p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x))$$

$$+ \int_{[0,t] \times E} (\log(1 + \phi_{(s,x)}) - \phi_{(s,x)}) p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x)$$

$$= -X_t + \log \mathcal{E}_t(\phi).$$

Substituting this expression into (4.29) we deduce $\mathbb{E}\big[G\big|\mathcal{F}_t\big] = \mathbb{E}[G]\mathcal{E}_t(\phi)$, $t \in \mathbb{R}_+$. In particular, for any $T > 0$, we have $(t, x) \mapsto \mathbb{1}_{[0,T]}(t)\mathcal{E}_t(\phi) \in \mathcal{P}_2(p(\pi))$ since

$$\mathbb{E}\bigg[\int_{[0,T] \times E} \mathcal{E}_s(\phi)^2 p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x)\bigg] \leq \bigg(\frac{c_1}{c_0}\bigg)^2 \mathbb{E}\big[N([0,T] \times E)\big] < \infty.$$

Finally, we prove that (3.32) holds (so that $\phi \in \mathcal{H}$). First, set

$$Y_t := \int_{[0,t] \times E} \varphi_{(s,x)}^{(G)}(N(\mathrm{d}s \times \mathrm{d}x) - p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x)), \quad t \in \mathbb{R}_+,$$

and note that (by the same computation as in (4.28)) its quadratic variation process is

$$[Y, Y]_t = \int_{[0,t] \times E} \big|\varphi_{(s,x)}^{(G)}\big|^2 N(\mathrm{d}s \times \mathrm{d}x), \quad t \in \mathbb{R}_+.$$

By Burkholder-Davis-Gundy's and Doob's inequalities, there exists a positive constant $C > 0$ such that, for all $t \geq 0$, we have

$$\mathbb{E}\bigg[\bigg(\int_{[0,t]\times E}|\varphi^{(G)}_{(s,x)}|^2\,N(\mathrm{d}s\times\mathrm{d}x)\bigg)^2\bigg] = \mathbb{E}[[Y,Y]_t^2] \leq C\mathbb{E}\bigg[\Big(\sup_{s\in[0,t]}|Y_s|\Big)^4\bigg] \leq C\Big(\frac{4}{3}\Big)^4\mathbb{E}[|Y_t|^4].$$

By (4.25) the right-most term of this relation is equal to $C(4/3)^4\mathbb{E}\big[|\mathbb{E}[G|\mathcal{F}_t] - \mathbb{E}[G]|^4\big]$, which is in turn less than or equal to a positive constant, say $C' > 0$, which is independent of t. Hence, for any $t \in \mathbb{R}_+$, we have

$$\mathbb{E}\bigg[\bigg(\int_{[0,t]\times E}|\phi_{(s,x)}|^2\,p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\bigg)^2\bigg]$$

$$\leq 2\,\mathbb{E}\bigg[\bigg(\int_{[0,t]\times E}|\phi_{(s,x)}|^2\,N(\mathrm{d}s\times\mathrm{d}x)\bigg)^2\bigg]$$

$$+ 2\,\mathbb{E}\bigg[\bigg(\int_{[0,t]\times E}|\phi_{(s,x)}|^2(N(\mathrm{d}s\times\mathrm{d}x) - p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x))\bigg)^2\bigg]$$

$$\leq \frac{2}{c_0^4}\,\mathbb{E}\bigg[\bigg(\int_{[0,t]\times E}|\varphi^{(G)}_{(s,x)}|^2\,N(\mathrm{d}s\times\mathrm{d}x)\bigg)^2\bigg]$$

$$+ 2\,\mathbb{E}\bigg[\bigg(\int_{[0,t]\times E}|\phi_{(s,x)}|^2(N(\mathrm{d}s\times\mathrm{d}x) - p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x))\bigg)^2\bigg]$$

$$\leq C'' + 2\,\mathbb{E}\bigg[\bigg(\int_{[0,t]\times E}|\phi_{(s,x)}|^2(N(\mathrm{d}s\times\mathrm{d}x) - p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x))\bigg)^2\bigg]$$

$$= C'' + 2\,\mathbb{E}\bigg[\int_{[0,t]\times E}|\phi_{(s,x)}|^4 p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\bigg]$$

$$\leq C'' + 2\Big(\frac{c_1}{c_0}+1\Big)^2\mathbb{E}\bigg[\int_{\mathbb{R}_+\times E}|\phi_{(s,x)}|^2 p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\bigg] < \infty,$$

where $C'' := (2C')/c_0^4 > 0$ and we have applied Proposition 2.5 and (4.26). Taking the limit as t goes to infinity in the above relations finally yields (3.32). $\qquad\square$

Proof of Proposition 4.4. For $t \in \mathbb{R}_+$ we put

$$\delta(\mathbb{1}_{[0,t]}\mathcal{E}(\phi)\phi) := \int_{[0,t]\times E}\mathcal{E}_{s^-}(\phi)\phi_{(s,x)}(N(\mathrm{d}s\times\mathrm{d}x) - p(\pi)_{(s,x)}\mathrm{d}s\nu(\mathrm{d}x)),$$

$$\delta(\mathbb{1}_{[0,t]}\psi) := \int_{[0,t]\times E}\psi_{(s,x)}(N(\mathrm{d}s\times\mathrm{d}x) - p(\pi)_{(s,x)}\mathrm{d}s\nu(\mathrm{d}x)),$$

and note that by Proposition 2.5-(iii) we have that

$$\delta(\mathbb{1}_{[0,t]}\psi)\delta(\mathbb{1}_{[0,t]}\mathcal{E}(\phi)\phi) - \int_{[0,t]\times E}\psi_{(s,x)}\mathcal{E}_s(\phi)\phi_{(s,x)}p(\pi)_{(s,x)}\mathrm{d}s\nu(\mathrm{d}x), \quad t\in\mathbb{R}_+$$

is an $\mathcal{F}$-martingale since (4.12) implies $(t,x)\mapsto\mathbb{1}_{[0,T]}(t)\psi_{(t,x)}\in\mathcal{P}_2(p(\pi))$ for any $T > 0$, along a similar computation as in (4.15) we have $(t,x)\mapsto\mathbb{1}_{[0,T]}(t)\psi_{(t,x)}\in\mathcal{P}_1(p(\pi))$ for any $T > 0$, and (4.15) and (4.16) guarantee (4.14). Using the "angle bracket" notation (see p. 53 in [11] and pp. 122-123 in [23]), we have

$$\langle Z,\mathcal{E}(\phi)\rangle_t = \langle Z,\mathcal{E}(\phi)-1\rangle_t = \int_{[0,t]\times E}\mathcal{E}_s(\phi)\phi_{(s,x)}\psi_{(s,x)}p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x), \quad t\in\mathbb{R}_+,$$

where we put $Z_t := \delta(\mathbb{1}_{[0,t]}\psi)$ and used (4.13), i.e. $\mathcal{E}_t(\phi) - 1 = \delta(\mathbb{1}_{[0,t]}\mathcal{E}(\phi)\phi)$. Since $\{Z_t\}_{t\in\mathbb{R}_+}$ is an $\mathcal{F}$-martingale under $\mathbb{P}$, with $Z_0 = 0$, and $\mathcal{E}_t(\phi)$ is the Radon-Nikodym derivative of $\mathbb{P}_\phi$ with respect to $\mathbb{P}$ on $\mathcal{F}_t$, by the Meyer-Girsanov theorem, see Theorem 36 p. 133 in [23], we have that

$$Z_t - \int_{[0,t]\times E} \mathcal{E}_{s-}(\phi)^{-1} \, \mathrm{d}\langle Z, \mathcal{E}(\phi)\rangle_s = Z_t - \int_{[0,t]\times E} \phi_{(s,x)}\psi_{(s,x)}p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x), \quad t \in \mathbb{R}_+$$

(4.30)

is a local $\mathcal{F}$-martingale under $\mathbb{P}_\phi$. In the following, $\mathbb{E}_\phi$ denotes the expectation with respect to $\mathbb{P}_\phi$. In order to prove that the process defined in (4.30) is an $\mathcal{F}$-martingale under $\mathbb{P}_\phi$, it suffices to prove that the stochastic process $\{(1+\phi_{(t,x)})p(\pi)_{(t,x)}\}_{(t,x)\in\mathbb{R}_+\times E}$ is a stochastic intensity of N under $\mathbb{P}_\phi$ i.e., for any non-negative and predictable stochastic process $X : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}_+$ we have

$$\mathbb{E}_\phi\left[\int_{\mathbb{R}_+\times E} X_{(t,x)} \, N(\mathrm{d}t \times \mathrm{d}x)\right] = \mathbb{E}_\phi\left[\int_{\mathbb{R}_+\times E} X_{(t,x)}(1 + \phi_{(s,x)})p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x)\right], \quad (4.31)$$

and the integrability condition

$$\mathbb{E}_\phi\left[\int_{[0,T]\times E} |\psi_{(t,x)}|(1 + \phi_{(s,x)})p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x)\right] < \infty, \quad T > 0. \qquad (4.32)$$

Indeed, the martingale property follows by Corollary C4 p. 235 in [1]. In order to prove that the process defined in (4.30) is square integrable under $\mathbb{P}_\phi$, it suffices to prove

$$\mathbb{E}_\phi\left[\int_{[0,T]\times E} |\psi_{(t,x)}|^2(1 + \phi_{(s,x)})p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x)\right] < \infty, \quad T > 0. \qquad (4.33)$$

The square integrability then follows by Proposition 2.5-(ii). We start by proving (4.31). Setting $X_{(t,x)} := \mathbb{1}_{(a,b]}(t)\mathbb{1}_A(\omega)\mathbb{1}_L(x)$, $a, b \in \mathbb{R}_+$, $A \in \mathcal{F}_a$, $L \in \mathcal{E}$, and reasoning exactly as in the first part of the proof we have that the process defined by (4.30) with X in place of ψ is a local $\mathcal{F}$-martingale under $\mathbb{P}_\phi$, as $X \in \mathcal{P}_2(p(\pi))$ since $\mathbb{E}\big[N((a,b]\times L)\big] < \infty$. Therefore there exists a sequence of $\mathcal{F}$-stopping times $\{T_n\}_{n\geq 0}$ increasing to infinity such that

$$\mathbb{E}_\phi\left[\int_{[0,\min\{T,T_n\}]\times E} X_{(t,x)} \, N(\mathrm{d}t \times \mathrm{d}x)\right]$$
$$= \mathbb{E}_\phi\left[\int_{[0,\min\{T,T_n\}]\times E} X_{(t,x)}(1 + \phi_{(s,x)})p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x)\right],$$

for any $T \in \mathbb{R}_+$ and $n \geq 0$. Letting n and T go to infinity in the above equation, by the monotone convergence theorem we obtain (4.31) for simple predictable stochastic processes. The result follows for a general non-negative predictable stochastic process by a standard application of the monotone class theorem, see e.g. [1], Theorem T1 p. 260. Finally we prove (4.32) and (4.33). Letting $M > 0$ denote a constant such that $|\phi_{(t,x)}| \leq M \, \mathrm{d}t\nu(\mathrm{d}x)\mathrm{d}\mathbb{P}$-almost everywhere, for any $T > 0$, since the probability measures $\mathbb{P}$ and $\mathbb{P}_\phi$ are equivalent we also have $|\phi_{(t,x)}| \leq M \, \mathrm{d}t\nu(\mathrm{d}x)\mathrm{d}\mathbb{P}_\phi$-almost everywhere, for any $T > 0$. Therefore,

$$\mathbb{E}_\phi\left[\int_{[0,T]\times E} |\psi_{(t,x)}|^2(1 + \phi_{(s,x)})p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x)\right]$$
$$\leq (1 + M) \, \|\mathcal{E}_T(\phi)\|_{L^2(\Omega,\mathcal{F}_\infty,\mathbb{P})} \sqrt{\mathbb{E}\left[\left(\int_{[0,T]\times E} |\psi_{(t,x)}|^2 p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x)\right)^2\right]} < \infty,$$

and

$$\mathbb{E}_\phi\Big[\int_{[0,T]\times E}|\psi_{(t,x)}|(1+\phi_{(s,x)})p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\Big]$$

$$\leq \sqrt{(1+M)\mathbb{E}_\phi\Big[\int_{[0,T]\times E}|\psi_{(t,x)}|^2(1+\phi_{(s,x)})p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\Big]}\sqrt{\mathbb{E}_\phi[N([0,T]\times E)]}$$

$$\leq \sqrt{(1+M)\mathbb{E}_\phi\Big[\int_{[0,T]\times E}|\psi_{(t,x)}|^2(1+\phi_{(s,x)})p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\Big]}$$

$$\times \mathbb{E}\big[\mathcal{E}_T(\phi)^2\big]^{1/4}\mathbb{E}\big[N([0,T]\times E)^2\big]^{1/4}$$

$$< \infty,$$

which concludes the proof. $\qquad\square$

4.7 Proof of Theorem 3.28

We divide the proof into two steps. In the first step we derive a predictable representation of G, in the second step we identify the integrand of the predictable representation.
Step 1.
By (3.38) we have that for all $t\in\mathbb{R}_+$, $\int_{[0,t]\times E}p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)<\infty$ $\mathbb{P}$-almost surely, so by Lemma 2.7, $p(\pi)$ is a stochastic intensity of N. Hence by Lemma 2.8, $p(\pi)$ is an $\overline{\mathcal{F}}$-stochastic intensity of N. Therefore, by the representation theorem for square integrable functionals (see e.g. [1])

$$G=\mathbb{E}[G]+\int_{\mathbb{R}_+\times E}u^{(G)}_{(s,x)}\left(N(\mathrm{d}s\times\mathrm{d}x)-p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\right),\quad \mathbb{P}\text{-almost surely,}$$

for some $u^{(G)}\in\mathcal{P}_2^{\overline{\mathcal{F}}}(p(\pi))$. By Proposition 2.3 there exists an $\mathcal{F}$-predictable stochastic process $p(u^{(G)})$ such that

$$p(u^{(G)})_{(t,x)}=\mathbb{E}\big[u^{(G)}_{(t,x)}|\overline{\mathcal{F}}_{t^-}\big]=u^{(G)}_{(t,x)},\quad \mathbb{P}\text{-almost surely,} \tag{4.34}$$

where the latter equality follows by the $\overline{\mathcal{F}}$-predictability of $u^{(G)}$ which guarantees that $u^{(G)}_{(t,x)}$ is $\overline{\mathcal{F}}_{t^-}$-measurable (as already noticed in Section 2, this follows by an obvious modification of the proof of Lemma A3.3.I p. 425 in [5]). So finally, we deduce the predictable representation

$$G=\mathbb{E}[G]+\int_{\mathbb{R}_+\times E}p(u^{(G)})_{(s,x)}\left(N(\mathrm{d}s\times\mathrm{d}x)-p(\pi)_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\right),\quad \mathbb{P}\text{-almost surely.} \tag{4.35}$$

Step 2.
Let $u:\mathbb{R}_+\times\Omega\times E\to\mathbb{R}$ be a predictable stochastic process and, for any $n\geq 0$, define the (predictable) stochastic process $u^{(n)}$ by $u^{(n)}_{(s,x)}:=u_{(s,x)}\mathbb{1}_{\{u_{(s,x)}\in[-n,n]\}}$. For any $n\geq 0$ and $t\in\mathbb{R}_+$, by the square integrability of G and $N([0,t]\times E)$ and (3.2), we have

$$\mathbb{E}\Big[\int_{\mathbb{R}_+\times E}|u^{(n)}_{(s,x)}\mathbb{1}_{[0,t]}(s)G|\pi_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\Big]$$

$$\leq n\mathbb{E}\Big[|G|\int_{[0,t]\times E}\pi_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\Big]$$

$$\leq n\|G\|_{L^2(\Omega,\mathcal{F}_\infty,\mathbb{P})}\Big\|\int_{[0,t]\times E}\pi_{(s,x)}\,\mathrm{d}s\nu(\mathrm{d}x)\Big\|_{L^2(\Omega,\mathcal{F}_\infty,\mathbb{P})}$$

$$<\infty$$

and

$$\mathbb{E}\left[\int_{\mathbb{R}_+ \times E} |u^{(n)}_{(s,x)} \mathbb{1}_{[0,t]}(s) G^+_{(s,x)}| \pi_{(s,x)} \, \mathrm{d}s \nu(\mathrm{d}x)\right] \leq n\mathbb{E}\left[\int_{[0,t] \times E} |G^+_{(s,x)}| \pi_{(s,x)} \, \mathrm{d}s \nu(\mathrm{d}x)\right]$$

$$= n\mathbb{E}\left[\int_{[0,t] \times E} |G(N)| N(\mathrm{d}s \times \mathrm{d}x)\right]$$

$$\leq n\|G\|_{L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})} \|N([0,t] \times E)\|_{L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})}$$

$$< \infty,$$

thus both conditions in (2.3) are verified for the predictable $X^{(n)}_{(s,x)} := u^{(n)}_{(s,x)} \mathbb{1}_{[0,t]}(s)$ and the random variable $F := G$. The stochastic process $X^{(n)}$ defined above is clearly in $\mathcal{P}^{\mathcal{F}}_{1,2}(p(\pi)) \subset \mathcal{P}^{\overline{\mathcal{F}}}_{1,2}(p(\pi))$, since it is bounded and $\mathbb{E}[N([0,t] \times E)] < \infty$. Hence by Lemma 2.4, Proposition 2.6 and (4.34), we have

$$\mathbb{E}\left[\int_{\mathbb{R}_+ \times E} X^{(n)}_{(s,x)} \pi_{(s,x)} D_{(s,x)} G \, \mathrm{d}s \nu(\mathrm{d}x)\right]$$

$$= \mathbb{E}\left[G\Delta(X^{(n)})\right]$$

$$= \mathbb{E}\left[G \int_{[0,t] \times E} u^{(n)}_{(s,x)} (N(\mathrm{d}s \times \mathrm{d}x) - p(\pi)_{(s,x)} \, \mathrm{d}s \nu(\mathrm{d}x))\right]$$

$$+ \mathbb{E}\left[G \int_{[0,t] \times E} u^{(n)}_{(s,x)} (p(\pi)_{(s,x)} - \pi_{(s,x)}) \, \mathrm{d}s \nu(\mathrm{d}x)\right]$$

$$= \mathbb{E}\left[\int_{[0,t] \times E} u^{(n)}_{(s,x)} u^{(G)}_{(s,x)} p(\pi)_{(s,x)} \, \mathrm{d}s \nu(\mathrm{d}x)\right]$$

$$+ \mathbb{E}\left[G \int_{[0,t] \times E} u^{(n)}_{(s,x)} (p(\pi)_{(s,x)} - \pi_{(s,x)}) \, \mathrm{d}s \nu(\mathrm{d}x)\right]$$

$$= \mathbb{E}\left[\int_{[0,t] \times E} u^{(n)}_{(s,x)} p(u^{(G)})_{(s,x)} p(\pi)_{(s,x)} \, \mathrm{d}s \nu(\mathrm{d}x)\right]$$

$$+ \mathbb{E}\left[G \int_{[0,t] \times E} u^{(n)}_{(s,x)} (p(\pi)_{(s,x)} - \pi_{(s,x)}) \, \mathrm{d}s \nu(\mathrm{d}x)\right].$$

Therefore we have

$$0 = \mathbb{E}\left[\int_{[0,t] \times E} u^{(n)}_{(s,x)} (\pi_{(s,x)} D_{(s,x)} G - p(u^{(G)})_{(s,x)} p(\pi)_{(s,x)} - Gp(\pi)_{(s,x)} + G\pi_{(s,x)}) \, \mathrm{d}s \nu(\mathrm{d}x)\right]$$

$$= \mathbb{E}\left[\int_{[0,t] \times E} u^{(n)}_{(s,x)} \left(\mathbb{E}\left[\pi_{(s,x)} G^+_{(s,x)} \mid \mathcal{F}_{s^-}\right] - p(u^{(G)})_{(s,x)} p(\pi)_{(s,x)}\right.\right.$$

$$\left.\left. - \mathbb{E}\left[G \mid \mathcal{F}_{s^-}\right] p(\pi)_{(s,x)}\right) \, \mathrm{d}s \nu(\mathrm{d}x)\right].$$

As discussed in Remark 3.30 the predictable projections $p(\pi G^+)$ and $p(G)$ exist and so

$$\mathbb{E}\left[\int_{[0,t] \times E} u^{(n)}_{(s,x)} \left(p(\pi G^+)_{(s,x)} - p(u^{(G)})_{(s,x)} p(\pi)_{(s,x)} - p(G)_{(s,x)} p(\pi)_{(s,x)}\right) \, \mathrm{d}s \nu(\mathrm{d}x)\right] = 0,$$

for all $n \geq 0$ and $t \in \mathbb{R}_+$. As shown in Remark 3.30 the stochastic process $\varphi^{(G)}$ in (3.36) is well-defined and predictable. So the above relation can be rewritten as

$$\mathbb{E}\left[\int_0^t \int_E u^{(n)}_{(s,x)} \left(\varphi^{(G)}_{(s,x)} - p(u^{(G)})_{(s,x)}\right) p(\pi)_{(s,x)} \, \mathrm{d}s \nu(\mathrm{d}x)\right] = 0, \quad \forall n \geq 0, \ \forall t \in \mathbb{R}_+. \quad (4.36)$$

Since u is an arbitrary predictable stochastic process, by choosing $u = \varphi^{(G)} - p(u^{(G)})$, equation (4.36) reads as

$$\mathbb{E}\left[\int_0^t \int_E \left(\varphi^{(G)}_{(s,x)} - p(u^{(G)})_{(s,x)}\right)^2 \mathbb{1}_{\{\varphi^{(G)}_{(s,x)} - p(u^{(G)})_{(s,x)} \in [-n,n]\}} p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x)\mathrm{d}s\right] = 0,$$

for all $n \geq 0$ and $t \in \mathbb{R}_+$. By the monotone convergence theorem, letting n tend to infinity, we get

$$\mathbb{E}\left[\int_0^t \int_E \left(\varphi^{(G)}_{(s,x)} - p(u^{(G)})_{(s,x)}\right)^2 p(\pi)_{(s,x)} \, \mathrm{d}s\nu(\mathrm{d}x)\right] = 0, \quad \text{for all } t \in \mathbb{R}_+$$

which implies that $p(u^{(G)})_{(s,x)}(\omega) = \varphi^{(G)}_{(s,x)}(\omega) \; \mathbb{P}(\mathrm{d}\omega)p(\pi)_{(s,x)}(\omega) \, \mathrm{d}s\nu(\mathrm{d}x)$-almost everywhere on $(\mathbb{R}_+ \times \Omega \times E, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_\infty \otimes \mathcal{E})$. Combining this with (4.35) we finally deduce (3.37). Since $u^{(G)}$ is square integrable with respect to $\mathbb{P}(\mathrm{d}\omega)p(\pi)_{(s,x)}(\omega) \, \mathrm{d}s\nu(\mathrm{d}x)$ by Jensen's inequality we easily have that $p(u^{(G)})$ is square integrable with respect to $\mathbb{P}(\mathrm{d}\omega)p(\pi)_{(s,x)}(\omega) \, \mathrm{d}s\nu(\mathrm{d}x)$ and so $\varphi^{(G)} \in \mathcal{P}_2^{\mathcal{F}}(p(\pi))$.

5 Appendix

5.1 Proof of Lemma 2.2

We have

$$\begin{aligned}
1 - \mathbb{P}(\Omega) &= \mathbb{P}(\{\omega \in \Omega' : \exists(t,x) \in \mathrm{Supp}(\omega), \; \omega(\{t\} \times E) \geq 2\}) \\
&= \mathbb{E}\left[\mathbb{1}_{\{\exists(t,x) \in \mathrm{Supp}(N), \; N(\{t\} \times E) \geq 2\}}\right] \\
&\leq \mathbb{E}\left[\int \mathbb{1}_{\{N(\{t\} \times E) \geq 2\}} N(\mathrm{d}t \times \mathrm{d}x)\right] \\
&= \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \mathbb{1}_{\{(N+\varepsilon_{(t,x)})(\{t\} \times E) \geq 2\}} \pi_{(t,x)} \, \mathrm{d}t\nu(\mathrm{d}x)\right] \\
&= \mathbb{E}\left[\int_{\mathbb{R}_+ \times E} \mathbb{1}_{\{N(\{t\} \times E) \geq 1\}} \pi_{(t,x)} \, \mathrm{d}t\nu(\mathrm{d}x)\right] \\
&= 0,
\end{aligned}$$

where the last equality follows since for any $\omega \in \Omega'$, we have that $\mathbb{1}_{\{\omega(\{t\} \times E) \geq 1\}} = 0$ $\mathrm{d}t$-almost surely. Indeed, for any $\omega \in \Omega'$ and any $T > 0$, we have $\omega([0,T] \times E) < \infty$, therefore the cardinal of $\{t \in [0,T] : \omega(\{t\} \times E) \geq 1\}$ is finite, and thus

$$\int_0^T \mathbb{1}_{\{t \in [0,T] \, : \, \omega(\{t\} \times E) \geq 1\}} \, \mathrm{d}t = 0.$$

5.2 Proof of Proposition 2.3

The proof uses the following lemma which guarantees the existence of a predictable version of a bounded stochastic process X.

Lemma 5.1. *Assume that N has a Papangelou conditional intensity π, i.e. (2.1) holds, and let $X \in L^\infty(\mathbb{R}_+ \times \Omega \times E, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}_\infty \times \mathcal{E}, \mathrm{d}t\nu(\mathrm{d}x)\mathbb{P}(\mathrm{d}\omega))$ be a real-valued stochastic process. Then there exists a predictable stochastic process $p(X) : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R}$ such that, for all $(t,x) \in \mathbb{R}_+ \times E$, we have*

$$p(X)_{(t,x)} = \mathbb{E}\left[X_{(t,x)} \big| \mathcal{F}_{t-}\right], \quad \mathbb{P}\text{-almost surely.}$$

Proof of Proposition 2.3. If X is assumed to be non-negative, we set $X^{(n)} := \min(X, n)$, $n \geq 0$. By Lemma 5.1, for any n and (t,x),

$$p(X^{(n)}_{(t,x)}) = \mathbb{E}\left[X^{(n)}_{(t,x)} \big| \mathcal{F}_{t-}\right], \quad \mathbb{P}\text{-almost surely.}$$

By the monotone convergence theorem

$$\lim_{n \to \infty} p(X^{(n)}_{(t,x)}) = \mathbb{E}\big[X_{(t,x)}\big|\mathcal{F}_{t-}\big], \quad \mathbb{P}\text{-almost surely.}$$

Since $p(X^{(n)})$ is predictable for each n its limit exists and is predictable, completing the proof under the assumption that X is non-negative.

If one assumes that, for $\mathrm{d}t\nu(\mathrm{d}x)$-almost all $(t,x) \in \mathbb{R}_+ \times E$, $X_{(t,x)} \in L^1(\Omega, \mathcal{F}_\infty, \mathbb{P})$, then we write $X = X^+ - X^-$, where $X^+ := \max(X,0)$ and $X^- := -\min(X,0)$. Applying the first part of the proposition to X^+ and X^- we have that there exist two predictable stochastic processes $p(X^+)$ and $p(X^-)$ such that

$$p(X^+)_{(t,x)} = \mathbb{E}\big[X^+_{(t,x)}\big|\mathcal{F}_{t-}\big] \quad \text{and} \quad p(X^-)_{(t,x)} = \mathbb{E}\big[X^-_{(t,x)}\big|\mathcal{F}_{t-}\big], \quad \mathbb{P}\text{-almost surely.}$$

By taking the expectation of these two equalities, one has, for $\mathrm{d}t\nu(\mathrm{d}x)$-almost all (t,x),

$$p(X^+)_{(t,x)} < \infty \quad \text{and} \quad p(X^-)_{(t,x)} < \infty,$$

and therefore

$$\mathbb{E}\big[X_{(t,x)}\big|\mathcal{F}_{t-}\big] = \mathbb{E}\big[X^+_{(t,x)}\big|\mathcal{F}_{t-}\big] - \mathbb{E}\big[X^-_{(t,x)}\big|\mathcal{F}_{t-}\big] = p(X^+)_{(t,x)} - p(X^-)_{(t,x)}, \qquad (5.1)$$

$\mathbb{P}$-almost surely. The claim follows by noticing that the right-hand side of (5.1) is a predictable stochastic process. $\qquad \square$

Proof of Lemma 5.1. The idea is to apply the existence part of Theorem 3.3 in [17] and the last displayed formula on p. 368 again in [17]. Following the notation of [17], we consider the locally compact second countable Hausdorff space $\mathbf{X} := \mathbb{R}_+ \times E$ and the DC-semiring $\mathcal{S} := \{(a,b] \times B : a,b \in \mathbb{R}_+,\ a < b,\ B \in \mathcal{E},\ B \text{ relatively compact}\}$. Moreover, we consider the system $\Gamma := \{\Gamma_{(t,x)}, \Gamma_S :\ (t,x) \in \mathbb{R}_+ \times E,\ S \in \mathcal{S}\}$ defined by $\Gamma_{(t,x)} := [0,t) \times E$ and $\Gamma_{(a,b] \times B} := [0,a] \times E$. It is readily checked that it satisfies conditions (2.3) and (2.4) in [17]. In this setting, for $(t,x) \in \mathbb{R}_+ \times E$, $a,b \in \mathbb{R}_+$: $a < b$, $B \in \mathcal{E}$ a relatively compact set, the σ-fields $\mathcal{F}((t,x))$ and $\mathcal{F}((a,b] \times B)$ defined on p. 364 of [17] are given by

$$\mathcal{F}((t,x)) := \sigma(N|_{[0,t) \times E}) = \mathcal{F}_{t-} \quad \text{and} \quad \mathcal{F}((a,b] \times B) := \sigma(N|_{[0,a] \times E}) = \mathcal{F}_a,$$

where $N|_A$ denotes the restriction of the random measure N to $A \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$. One may easily check that the predictable σ-algebra $\mathcal{P}$ in [17] coincides with the σ-field $\mathcal{P}(\mathcal{F}) \otimes \mathcal{E}$, and that the point process N satisfies condition (2.5) on p. 364 of [17]. We note that if condition $\Sigma(\Lambda)$ on p. 367 of [17] holds, then the claim of the lemma follows by the existence part of Theorem 3.3 in [17] and the last displayed formula on p. 368 again in [17]. In our setting, condition $\Sigma(\Lambda)$ on p. 367 of [17] reads

$$\Sigma(\Lambda): \ \mathbb{P}(N((a,b] \times E) = 0 \mid \mathcal{F}_a) > 0 \quad \mathbb{P}\text{-a.s.,}\ a,b \in \mathbb{R}_+: a < b.$$

Since N has a Papangelou conditional intensity it satisfies condition (Σ) from Remark 2.5(c) in [9] which, in our setting, reads

$$(\Sigma): \ \mathbb{P}(N((a,b] \times E) = 0 \mid \mathcal{F}_a \vee \mathcal{F}_{(b,\infty)}) > 0 \quad \mathbb{P}\text{-a.s.,}\ a,b \in \mathbb{R}_+: a < b,$$

where $\mathcal{F}_{(b,\infty)} := \sigma(N|_{(b,\infty) \times E})$. Condition $\Sigma(\Lambda)$ easily follows by (Σ) and the properties of the conditional expectation, indeed

$$\mathbb{P}(N((a,b] \times E) = 0 \mid \mathcal{F}_a) = \mathbb{E}\big[\mathbb{P}(N((a,b] \times E) = 0\big|\mathcal{F}_a \vee \mathcal{F}_{(b,\infty)})\big] > 0$$

$\mathbb{P}$-a.s., for any $a,b \in \mathbb{R}_+$ such that $a < b$. $\qquad \square$

References

[1] P. Brémaud. *Point Processes and Queues.* Springer-Verlag, New York, 1981. MR-0636252

[2] P. Brémaud and L. Massoulié. Stability of nonlinear Hawkes processes. *The Annals of Probability*, 24: 1563–1588, 1996. MR-1411506

[3] P. Brémaud and L. Massoulié. Power spectra of general shot noises and Hawkes point processes with a random excitation. *Advances in Applied Probability*, 34: 205–222, 2002. MR-1895338

[4] A. Budhiraja, P. Dupuis and V. Maroulas. Variational representations for continuous time processes. *Annales de l'Institut Henri Poincaré – Probabilités et Statistiques*, 47: 725–747, 2011. MR-2841073

[5] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Vol. I.* Springer-Verlag, New York, 2003. MR-1950431

[6] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Vol. II.* Springer-Verlag, New York, 2008. MR-2371524

[7] P. Dupuis and R.S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations.* Wiley, New York, 1997. MR-1431744

[8] I. Flint and G.L. Torrisi. A Clark-Ocone formula for temporal point processes, and applications. *The Annals of Probability*, 45: 3266–3292, 2017. MR-3706743

[9] H. Georgii and H.J. Yoo. Conditional Intensity and Gibbsianness of Determinantal Point Processes. *J. Statist. Phys.*, 118: 55–84, 2005. MR-2122549

[10] N. Gozlan and C. Léonard. A large deviation approach to some transportation cost inequalities. *Probability Theory and Related Fields*, 139: 235–283, 2007. MR-2322697

[11] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam, 1989. MR-1011252

[12] T. Klein, Y. Ma and N. Privault. Convex concentration inequalities and forward-backward stochastic calculus. *Electronic Journal of Probability*, 11: 486–512, 2006. MR-2242653

[13] Y. Kondratiev and E. Lytvynov. Glauber dynamics of continuous particle systems. *Annales de l'I.H.P. Probabilités et statistiques*, 41(4): 685–702, 2005. MR-2144229

[14] G. Last and A. Brandt. *Marked point processes on the real line.* Springer-Verlag, New York, 1995. MR-1353912

[15] G. Last and M.D. Penrose. Martingale representation for Poisson processes with applications to minimal variance hedging. *Stochastic Processes and their Applications*, 121: 1588–1606, 2011. MR-2802467

[16] G. Last and M.D. Penrose. *Lectures on the Poisson process.* Cambridge University Press, Cambridge, 2017. MR-3791470

[17] G. Last. Predictable projections for point process filtrations. *Probability Theory and Related Fields*, 99: 361–388, 1994. MR-1283117

[18] M.N.M. van Lieshout. *Markov Point Processes and their Applications.* Imperial College Press, London, 2000. MR-1789230

[19] Y. Ma, S. Shen, X. Wang and L. Wu. Transportation inequalities: from Poisson to Gibbs measures. *Bernoulli*, 17: 155–169, 2011. MR-2797986

[20] K. Matthes, W. Warmuth and J. Mecke. Bemerkungen zu einer arbeit von Nguyen Xuan Xanh une Hans Zessin. *Math. Nachr.*, 88: 117–127, 1979. MR-0543397

[21] X.X. Nguyen and H. Zessin. Integral and differential characterizations of the Gibbs process. *Math. Nachr.*, 88: 105–115, 1979. MR-0543396

[22] N. Privault. *Stochastic Analysis in Discrete and Continuous Settings: with Normal Martingales.* Springer-Verlag, Berlin, 2009. MR-2531026

[23] P.E. Protter. *Stochastic integration and differential equations.* Springer, New York, 2004. MR-2020294

[24] R. Siegmund-Schultze, E. Warmuth and W. Warmuth. A Gibbsean description for a class of simple stationary recurrent point processes. *Colloquia Mathematica Societatis János Bolyai*, 27: 967-987, 1982. MR-0712721

[25] G.L. Torrisi. Point processes with Papangelou conditional intensity: from the Skorohod integral to the Dirichlet form. *Markov Processes and Related Fields*, 19: 195–248, 2013. MR-3113944

[26] C. Villani. *Optimal Transport. Old and New*. Springer-Verlag, Berlin, 2009. MR-2459454

[27] L. Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probability Theory and Related Fields*, 118: 427-438, 2000. MR-1800540

[28] L. Wu. Transportation inequalities for stochastic differential equations of pure jumps. *Annales de l'Institut Henri-Poincaré – Probabilités et Statistiques*, 46: 465-479, 2010. MR-2667706

[29] X. Zhang. Clark-Ocone formula and variational representation for Poisson functionals. *The Annals of Probability*, 37(2): 506–529, 2009. MR-2510015