# Time-dynamic evaluations under non-monotone information generated by marked point processes

**Marcus C. Christiansen[1]**

## Abstract

The information dynamics in finance and insurance applications is usually modelled by a filtration. This paper looks at situations where information restrictions apply so that the information dynamics may become non-monotone. A fundamental tool for calculating and managing risks in finance and insurance are martingale representations. We present a general theory that extends classical martingale representations to non-monotone information generated by marked point processes. The central idea is to focus only on those properties that martingales and compensators show on infinitesimally short intervals. While classical martingale representations describe innovations only, our representations have an additional symmetric counterpart that quantifies the effect of information loss. We exemplify the results with examples from life insurance and credit risk.

## 1 Introduction

The value at time $t \in [0, T]$ of a financial claim $\xi \in L^1(\Omega, \mathcal{A}, P)$ at time $T \in (0, \infty)$ is commonly calculated by

$$B(t)E_Q\left[\frac{\xi}{B(T)}\bigg|\mathcal{F}_t\right], \tag{1.1}$$

where $B$ is the value process of a risk-free asset, $(\mathcal{F}_t)_{t \geq 0}$ is a filtration that describes the available information at each time $t \geq 0$, and $Q$ is some equivalent measure.

✉ M.C. Christiansen
marcus.christiansen@uni-oldenburg.de

1    Institut für Mathematik, Carl von Ossietzky Universität Oldenburg, 26111 Oldenburg, Germany

For studying the time dynamics of the value process, we can exploit the fact that $t \mapsto E_Q[\xi/B(T)|\mathcal{F}_t]$ is always a martingale.

In this paper, we suppose that information restrictions apply and replace the filtration $(\mathcal{F}_t)_{t \geq 0}$ by a family of sub-sigma-algebras $(\mathcal{G}_t)_{t \geq 0}$ that may be non-monotone, i.e., we do not assume that $(\mathcal{G}_t)_{t \geq 0}$ is a filtration. We focus on modelling frameworks where $(\mathcal{G}_t)_{t \geq 0}$ is generated by a marked point process, because this allows us to calculate martingale representations explicitly. Our approach seems to work also in more general settings, but a general theory is left to future research.

Information restrictions can be motivated by legal restrictions, data privacy efforts, information summarisation or model simplifications. An example for a legal information restriction is the General Data Protection Regulation 2016/679 of the European Union, which includes in Article 17 a so-called 'right to erasure', causing possible information loss.

**Example 1.1** Consider *life insurance* contracts that are *evaluated* by using *big data*. Data from activity trackers, social media, etc., can improve individual forecasts of the mortality and morbidity of insured persons. By exercising the 'right to erasure' according to the General Data Protection Regulation of the European Union, the policyholder may ask the insurer to delete parts of the health-related data at discretion. Moreover, data providers might implement self-imposed information restrictions for data privacy reasons. For example, users of Google products can opt for an auto-delete of location history and activity data after a fixed time limit. As a result, the evaluation of an insurance liability $\xi$ according to (1.1) will be restricted to sub-sigma-algebras $(\mathcal{G}_t)_{t \geq 0}$ that are non-monotone in $t$ due to data deletions.

Examples of information summarisation can be found in Norberg [22], where summarised life insurance values (retrospective and prospective reserves) are defined that encompass non-monotone information. A popular model simplification is Markovian modelling even when the empirical data does not fully support the Markov assumption.

**Example 1.2** We consider a *credit rating* process. In the Jarrow–Lando–Turnbull model, the filtration $(\mathcal{F}_t)_{t \geq 0}$ is generated by a finite-state-space Markov chain $(R_t)_{t \geq 0}$ that represents credit ratings; cf. Jarrow et al. [15]. The Markov property makes it possible to equivalently replace $\mathcal{F}_t$ in (1.1) by the sub-sigma-algebra $\mathcal{G}_t := \sigma(R_t)$. The *Markov assumption* can be motivated by the theoretical idea that a credit rating should fully describe the current risk profile of a prospective debtor so that historical ratings can be ignored. However, empirical data does not always support the Markov property, so that $E_Q[\xi/B(T)|\mathcal{G}_t]$ may in fact differ from $E_Q[\xi/B(T)|\mathcal{F}_t]$; cf. Lando and Skodeberg [17]. The information dynamics of $\mathcal{G}_t = \sigma(R_t)$ is non-monotone in $t$.

Non-monotone information structures can also be found in Pardoux and Peng [24] and Tang and Wu [27], but in these papers, specific independence assumptions make it possible to go back to filtrations and work with classical martingale representations.

**From now on**, we skip the subscript $Q$ in (1.1) and all related expectations. Depending on the application, we interpret $P$ either as the real-world measure or as a risk-neutral measure.

When we replace the filtration $(\mathcal{F}_t)_{t \geq 0}$ in (1.1) by some non-monotone information $(\mathcal{G}_t)_{t \geq 0}$, all the powerful tools from martingale theory for studying the time dynamics of (1.1) are not available any more. In order to fill that gap, this paper derives general representations of the form

$$E[\xi|\mathcal{G}_t] - E[\xi|\mathcal{G}_0] = \sum_{I \in \mathcal{N}} \int_{(0,t] \times E_I} G_I(u-, u, e)\,(\mu_I - \nu_I)\big(\mathrm{d}(u, e)\big)$$
$$+ \sum_{I \in \mathcal{N}} \int_{(0,t] \times E_I} G_I(u, u, e)\,(\rho_I - \mu_I)\big(\mathrm{d}(u, e)\big), \qquad t \geq 0,$$
(1.2)

where $\xi$ is any integrable random variable, $(\mathcal{G}_t)_{t \geq 0}$ is a non-monotone family of sigma-algebras generated by an extended marked point process that involves information deletions, $(\mu_I)_{I \in \mathcal{N}}$ is a set of counting measures that uniquely corresponds to the extended marked point process, $(\nu_I)_{I \in \mathcal{N}}$ and $(\rho_I)_{I \in \mathcal{N}}$ are infinitesimal forward and backward compensators of $(\mu_I)_{I \in \mathcal{N}}$, and the integrands $G_I(u-, u, e)$ and $G_I(u, u, e)$ are adapted to the information at time $u-$ and time $u$, respectively. In case that $(\mathcal{G}_t)_{t \geq 0}$ is increasing, i.e., it is a filtration, the second line in (1.2) is zero and the first line conforms with classical martingale representations. The central idea in this paper is to focus on those properties only that martingales and compensators show on infinitesimally small intervals. We call this the 'infinitesimal approach'. In principle, the infinitesimal approach is not restricted to point process frameworks, but a fully general theory is beyond the scope of this paper. We further extend our representation results to processes of the form

$$t \mapsto E[X_t|\mathcal{G}_t], \qquad t \geq 0,$$
(1.3)

where $(X_t)_{t \geq 0}$ is a suitably integrable càdlàg process. In this case, an additional drift term appears on the right-hand side of (1.2).

Martingale representations have various applications in finance and insurance, and this is in particular true for marked point process frameworks:

– If a financial or insurance claim is hedgeable, then explicit hedges can be derived from martingale representations; see e.g. Norberg [23] and Last and Penrose [18].

– Martingale representations are a central tool for constructing and solving backward stochastic differential equations (BSDEs); see e.g. Cohen and Elliott [6], Bandini [1] and Confortola [8]. Many optimal control problems in finance and insurance correspond to a BSDE problem, see e.g. Cohen and Elliott [7] and Delong [11, Chap. 1].

– Martingale representations can serve as additive risk factor decompositions; see Schilling et al. [26]. An insurer needs to additively decompose the surplus from a policy or an insurance portfolio for regulatory reasons; see e.g. Møller and Steffensen [20, Sect. 6]. Additive risk factor decompositions are also used in finance; see e.g. Rosen and Saunders [25].

In all three applications, infinitesimal martingale representations according to (1.2) allow us to include information restrictions into the modelling. We study later a hedg-

ing application for the model in Example 1.2. We shall see that estimation and calculation of hedging strategies under inappropriate Markov assumptions may unintentionally replace classical martingales by infinitesimal forward martingales (the first line on the right-hand side of (1.2)), and then the implied hedging error is just the corresponding infinitesimal backward martingale part (the second line in (1.2)). The application of infinitesimal martingale representations in BSDE theory is exemplarily discussed for Example 1.1. We shall see that the integrands in (1.2) correspond to the so-called sum at risk, which is a central quantity in life insurance risk management. In Example 1.1, we also briefly discuss risk factor decompositions. Information deletions upon request for data privacy reasons can provoke arbitrage opportunities, and these can be split off as infinitesimal backward martingales, which is important for dealing with them.

The representation (1.2) implies that $t \mapsto E[\xi|\mathcal{G}_t]$ has a (unique) semimartingale modification. More generally, we show that $t \mapsto E[X_t|\mathcal{G}_t]$ has a (unique) semimartingale modification whenever $X$ is a semimartingale with integrable variation on compacts. The uniqueness and the semimartingale property are crucial in applications where the time dynamics need to be studied. For example, in life insurance, the differential $dE[X_t|\mathcal{G}_t]$ might describe the insurer's current surplus or loss at time $t$; cf. Norberg [21, 22].

The study of jump process martingales and their representations largely dates back to the 1970s; see e.g. Jacod [14], Boel et al. [2], Chou and Meyer [3], Davis [10] and Elliott [13]. Since then, extensions have been developed in different directions; see e.g. Last and Penrose [18] and Cohen [5]. All these papers stay within the framework of filtrations, i.e., the information dynamics is monotone. The infinitesimal approach we introduce here allows us to go beyond the framework of filtrations. An elegant way to derive the classical martingale representation is a bare-hands approach that starts with the Chou and Meyer construction of the martingale representation for a single jump process, followed by Elliott's extension to the case of ordered jumps. In this paper, we also use a bare-hands approach, but the classical stopping time concept is not applicable in our non-monotone information setting, so that we need to leave the common paths.

The paper is organised as follows. In Sect. 2, we explain the basic concepts of the infinitesimal approach but avoid technicalities. In Sect. 3, we add technical assumptions and narrow the modelling framework down to pure jump process drivers. Section 4 verifies that (1.2) is indeed a well-defined process. In Sect. 5, we identify infinitesimal compensators for a large class of jump processes. The central result (1.2) is proved in Sect. 6 and extended to processes of the form (1.3) in Sect. 7. In Sect. 8, we take a closer look at Examples 1.1 and 1.2.

## 2 The infinitesimal approach

The central idea of the infinitesimal approach is to focus only on those properties that martingales and compensators show on infinitesimally short intervals. **This section explains the basic ideas under the general assumption that all limits in this section actually exist**. Only from the next section on, we narrow the framework down

to pure jump process drivers, which is sufficient but not necessary to guarantee the existence of the limits. So in general, the infinitesimal approach is not restricted to jump process frameworks, but it is beyond the scope of this paper to find general conditions for the existence of the limits here.

Let $(\Omega, \mathcal{A}, P)$ be a complete probability space and let $\mathcal{Z} \subseteq \mathcal{A}$ be the family of its nullsets. Let $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ be a complete and right-continuous filtration on this probability space. We interpret $\mathcal{F}_t$ as the observable information on the time interval $[0, t]$. Suppose that certain pieces of information expire after a finite holding time. By subtracting from $\mathcal{F}_t$ all pieces of information that have expired until time $t$, we obtain the admissible information at time $t$. We assume that this admissible information is represented by a family $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ of complete sigma-algebras

$$\mathcal{G}_t \subseteq \mathcal{F}_t, \qquad t \geq 0,$$

which may be non-monotone in $t$.

A process $X$ is adapted to the filtration $\mathbb{F}$ if $X_t$ is $\mathcal{F}_t$-measurable for each $t \geq 0$. Likewise we say that a process $X$ is *adapted* to the possibly non-monotone information $\mathbb{G}$ if $X_t$ is $\mathcal{G}_t$-measurable for each $t \geq 0$. In addition to this classical concept, we also take an incremental perspective.

**Definition 2.1** We call a process $X$ *incrementally adapted to* $\mathbb{G}$ if the increment $X_t - X_s$ is $\sigma(\mathcal{G}_u, u \in (s, t])$-measurable for any interval $(s, t] \subseteq [0, \infty)$.

In finance and insurance applications, we think of $X$ as an aggregated cash flow where the aggregated payments $X_t - X_s$ on the interval $(s, t]$ should depend only on the admissible information on $(s, t]$. If $\mathbb{G}$ is a filtration, incremental adaptedness is equivalent to classical adaptedness, but the two concepts differ for non-monotone information.

An integrable process $X$ is a martingale with respect to $\mathbb{F}$ if it is $\mathbb{F}$-adapted and

$$E[X_t - X_s | \mathcal{F}_s] = 0$$

almost surely for each $0 \leq s \leq t$. Focusing on infinitesimally short intervals, in particular we have

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E[X_{t_{k+1}} - X_{t_k} | \mathcal{F}_{t_k}] = 0 \qquad (2.1)$$

a.s. for each $t \geq 0$, where $(\mathcal{T}_n^t)_{n \in \mathbb{N}}$ is any increasing sequence (i.e., $\mathcal{T}_n^t \subseteq \mathcal{T}_{n+1}^t$ for all $n$) of partitions $0 = t_0 < t_1 < \cdots < t_n = t$ of the interval $[0, t]$ such that the mesh size $|\mathcal{T}_n^t| := \max\{t_k - t_{k-1} : k = 1, 2, \ldots\}$ tends to 0 for $n \to \infty$. In the literature, we can find for (2.1) the intuitive notation $E[dX_t | \mathcal{F}_{t-}] = 0$.

**Definition 2.2** Let $X$ be incrementally adapted to $\mathbb{G}$. We say that $X$ is an *infinitesimal forward/backward martingale (IF/IB-martingale) with respect to* $\mathbb{G}$ if for each $t \geq 0$ and any increasing sequence $(\mathcal{T}_n^t)_{n \in \mathbb{N}}$ of partitions of $[0, t]$ with $\lim_{n \to \infty} |\mathcal{T}_n^t| = 0$,

we have

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E[X_{t_{k+1}} - X_{t_k} | \mathcal{G}_{t_k}] = 0$$

or

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E[X_{t_{k+1}} - X_{t_k} | \mathcal{G}_{t_{k+1}}] = 0,$$

respectively, assuming that the expectations and limits exist.

Suppose now that $X$ is an $\mathbb{F}$-adapted and integrable counting process. The so-called compensator $C$ of $X$ is the unique $\mathbb{F}$-predictable finite-variation process starting from $C_0 = 0$ such that $X - C$ is an $\mathbb{F}$-martingale. In particular, $C$ satisfies the equation

$$C_t = \lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E[X_{t_{k+1}} - X_{t_k} | \mathcal{F}_{t_k}] \tag{2.2}$$

almost surely for each $t \geq 0$; see Karr [16, Theorem 2.17]. The intuitive notation for (2.2) is $E[dX_t | \mathcal{F}_{t-}] = dC_t$. Furthermore, one can show that the $\mathbb{F}$-predictability of $C$ implies that

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E[C_{t_{k+1}} - C_{t_k} | \mathcal{F}_{t_k}] = C_t - C_0 \tag{2.3}$$

almost surely for each $t \geq 0$, intuitively written as $E[dC_t | \mathcal{F}_{t-}] = dC_t$. The latter fact motivates the following definition.

**Definition 2.3** We call $X$ *infinitesimally forward/backward predictable (IF/IB-predictable)* with respect to $\mathbb{G}$ if for each $t \geq 0$ and any increasing sequence $(\mathcal{T}_n^t)_{n \in \mathbb{N}}$ of partitions of $[0, t]$ with $\lim_{n \to \infty} |\mathcal{T}_n^t| = 0$, we almost surely have

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E[X_{t_{k+1}} - X_{t_k} | \mathcal{G}_{t_k}] = X_t - X_0$$

or

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E[X_{t_{k+1}} - X_{t_k} | \mathcal{G}_{t_{k+1}}] = X_t - X_0,$$

respectively, assuming that the expectations and limits exist.

By combining (2.2) and (2.3), we obtain

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E[(X_{t_{k+1}} - C_{t_{k+1}}) - (X_{t_k} - C_{t_k}) | \mathcal{F}_{t_k}] = 0$$

almost surely for each $t \geq 0$, which means that the process $X - C$ is an IF-martingale with respect to $\mathbb{F}$ according to Definition 2.2.

**Definition 2.4** Let $X$ be incrementally adapted to $\mathbb{G}$. We say that a process $C$ is an *infinitesimal forward/backward compensator of $X$ (IF/IB-compensator) with respect to $\mathbb{G}$* if $C$ is incrementally adapted to $\mathbb{G}$ and IF/IB-predictable and $X - C$ is an IF/IB-martingale with respect to $\mathbb{G}$, respectively.

Let $\mathcal{G}_{[t_k, t_{k+1}]} := \sigma(\mathcal{G}_u, u \in [t_k, t_{k+1}])$ for any $t_{k+1} \geq t_k \geq 0$ and $\xi \in L^1(\Omega, \mathcal{A}, P)$. Then the construction

$$E[\xi|\mathcal{G}_t] - E[\xi|\mathcal{G}_0] = \lim_{n \to \infty} \sum_{\mathcal{T}_n^t} (E[\xi|\mathcal{G}_{t_{k+1}}] - E[\xi|\mathcal{G}_{t_k}])$$

$$= \lim_{n \to \infty} \sum_{\mathcal{T}_n^t} (E[\xi|\mathcal{G}_{[t_k, t_k+1]}] - E[\xi|\mathcal{G}_{t_k}])$$

$$- \lim_{n \to \infty} \sum_{\mathcal{T}_n^t} (E[\xi|\mathcal{G}_{[t_k, t_k+1]}] - E[\xi|\mathcal{G}_{t_{k+1}}])$$

may yield a decomposition of the process $t \mapsto E[\xi|\mathcal{G}_t]$ into the difference of an IF-martingale and an IB-martingale, since

$$E\big[E[\xi|\mathcal{G}_{[t_k, t_k+1]}] - E[\xi|\mathcal{G}_{t_k}]\big|\mathcal{G}_{t_k}\big] = 0,$$

$$E\big[E[\xi|\mathcal{G}_{[t_k, t_k+1]}] - E[\xi|\mathcal{G}_{t_{k+1}}]\big|\mathcal{G}_{t_{k+1}}\big] = 0.$$

**Definition 2.5** We say that $E[\xi|\mathcal{G}_t] - E[\xi|\mathcal{G}_0] = F_t - B_t$, $t \geq 0$, is an *infinitesimal martingale representation* if $F$ is an IF-martingale and $B$ is an IB-martingale with respect to $\mathbb{G}$.

Suppose now that $X$ describes a discounted claim process in a finance or insurance application. Then we are typically interested in the process $t \mapsto E[X_t|\mathcal{F}_t]$, which is not necessarily well defined. If $X$ is a càdlàg process whose suprema on compacts have finite expectations, then there exists a unique càdlàg process $X^{\mathbb{F}}$, the so-called optional projection of $X$ with respect to $\mathbb{F}$, such that

$$X_t^{\mathbb{F}} = E[X_t|\mathcal{F}_t]$$

almost surely for each $t \geq 0$. We say here that a process is unique if it is unique up to evanescence. We now expand the concept of optional projections to non-monotone information.

**Definition 2.6** Let $X$ be an integrable càdlàg process. If there exists a unique càdlàg process $X^{\mathbb{G}}$ such that

$$X_t^{\mathbb{G}} = E[X_t|\mathcal{G}_t]$$

almost surely for each $t \geq 0$, we call $X^{\mathbb{G}}$ the *optional projection of $X$ with respect to $\mathbb{G}$*.

The optional projection $X^{\mathbb{G}}$ can be decomposed to

$$E[X_t|\mathcal{G}_t] - E[X_0|\mathcal{G}_0] = \lim_{n\to\infty} \sum_{\mathcal{T}_n^t} (E[X_{t_{k+1}}|\mathcal{G}_{t_{k+1}}] - E[X_{t_k}|\mathcal{G}_{t_k}])$$

$$= \lim_{n\to\infty} \sum_{\mathcal{T}_n^t} (E[X_{t_k}|\mathcal{G}_{[t_k,t_k+1]}] - E[X_{t_k}|\mathcal{G}_{t_k}])$$

$$- \lim_{n\to\infty} \sum_{\mathcal{T}_n^t} (E[X_{t_k}|\mathcal{G}_{[t_k,t_k+1]}] - E[X_{t_k}|\mathcal{G}_{t_{k+1}}])$$

$$+ \lim_{n\to\infty} \sum_{\mathcal{T}_n^t} E[X_{t_{k+1}} - X_{t_k}|\mathcal{G}_{t_{k+1}}],$$

which may represent a sum of an IF-martingale, an IB-martingale and an IB-compensator with respect to $\mathbb{G}$. By switching the roles of $t_k$ and $t_{k+1}$, we can obtain a similar decomposition where the IB-compensator is replaced by an IF-compensator.

**Definition 2.7** We call $E[X_t|\mathcal{G}_t] - E[\xi|\mathcal{G}_0] = F_t - B_t + C_t$, $t \geq 0$, an *infinitesimal representation* if $F$ is an IF-martingale, $B$ is an IB-martingale and $C$ is either an IB-compensator or an IF-compensator with respect to $\mathbb{G}$.

As mentioned at the beginning of this section, we simply assumed so far that all the limits discussed here indeed exist. In the next section, we focus on a marked point process framework since this guarantees not only the existence of the limits, but also allows us to calculate the limits explicitly.

## 3 Jump process framework

In the literature, we can find different approaches for defining a jump process framework. One way is to start with a marked point process $(\tau_i, \zeta_i)_{i\in\mathbb{N}}$ on $(\Omega, \mathcal{A}, P)$ with some measurable mark space $(E, \mathcal{E})$, i.e.,

- the mappings $\tau_i : (\Omega, \mathcal{A}) \to ([0, \infty], \mathcal{B}([0, \infty]))$, $i \in \mathbb{N}$, are random times,
- the mappings $\zeta_i : (\Omega, \mathcal{A}) \to (E, \mathcal{E})$, $i \in \mathbb{N}$, are random marks.

Differently from the point process literature, we do not assume here that the random times $(\tau_i)_{i\in\mathbb{N}}$ are increasing or ordered in any specific way. This gives us useful modelling flexibility; see also the comments at the end of this section. Let $E$ be a Polish space and $\mathcal{E} := \mathcal{B}(E)$ its Borel sigma-algebra. For the sake of a simple notation, we moreover assume that $\Omega$ is a Polish space and $\mathcal{A}$ its Borel sigma-algebra. The latter assumption can actually be dropped by observing that all random activity in our model comes from a marked point process that can be embedded into a Polish space. We interpret each $\zeta_i$ as a piece of information that can be observed from time $\tau_i$ on. As motivated in the introduction, we additionally assume that the information pieces $\zeta_i$ are possibly deleted after a finite holding time. Therefore, we expand the marked point process $(\tau_i, \zeta_i)_{i\in\mathbb{N}}$ to $(\tau_i, \zeta_i, \sigma_i)_{i\in\mathbb{N}}$, where

– the mappings $\sigma_i : (\Omega, \mathcal{A}) \to ([0, \infty], \mathcal{B}([0, \infty]))$, $i \in \mathbb{N}$, are random times such that $\tau_i \leq \sigma_i$.

We interpret $\sigma_i$ as the deletion time of information piece $\zeta_i$. Note that the random times $(\sigma_i)_{i \in \mathbb{N}}$ are in general not ordered. For the sake of a more compact notation, we work in the following with the equivalent sequence $(T_i, Z_i)_{i \in \mathbb{N}}$ defined as

$$T_{2i-1} := \tau_i, \quad T_{2i} := \sigma_i, \quad Z_{2i-1} := \zeta_i, \quad Z_{2i} := \zeta_i, \qquad i \in \mathbb{N},$$

i.e., the random times $T_{2i-1}$ with odd indices refer to innovations and the consecutive random times $T_{2i}$ with even indices are the corresponding deletion times. We generally assume that

$$E\left[ \sum_{i=1}^{\infty} \mathbf{1}_{\{T_i \leq t\}} \right] < \infty, \qquad t \geq 0, \tag{3.1}$$

which will ensure the existence of (infinitesimal) compensators. Condition (3.1) implies that almost surely, there are at most finitely many random times on bounded intervals. Moreover, we assume that

$$T_{2i-1}(\omega) < T_{2i}(\omega) \qquad \text{for } \omega \in \{T_{2i} < \infty\}, i \in \mathbb{N},$$

i.e., a new piece of information is not instantaneously deleted but is available for at least a short amount of time. Based on the sequence $(T_i, Z_i)_{i \in \mathbb{N}}$, we generate random counting measures $\mu_I$ via

$$\mu_I([0, t] \times B) := \mathbf{1}_{\{t \geq T_i = T_j : i, j \in I\} \cap \{T_i \neq T_j : i \in I, j \notin I\}} \mathbf{1}_{\{Z_I \in B\}}$$

for $t \geq 0$, $B \in \mathcal{E}_I$ and finite subsets $I \subseteq \mathbb{N}$, where

$$\mathcal{E}_I := \mathcal{B}(E_I), \qquad E_I := E^{|I|}, \qquad Z_I := (Z_i)_{i \in I}.$$

If the different random times $T_i$ never coincide, then we just need to consider the counting measures $\mu_{\{i\}}$, $i \in \mathbb{N}$, which describe separate arrivals of the random times $T_i$ and their marks $Z_i$. But if random times can occur simultaneously, then we need the full scale of counting measures $\mu_I$, $I \subseteq \mathbb{N}$, $|I| < \infty$, which cover all kinds of separate and joint events. For each $I$, the measures $\{\mu_I(\cdot)(\omega) : \omega \in \Omega\}$ generated by their values on $[0, t] \times B$ form a random counting measure on $([0, \infty) \times E_I, \mathcal{B}([0, \infty) \times E_I))$, i.e.,

– for any fixed $A \in \mathcal{B}([0, \infty) \times E_I)$, the mapping $\omega \mapsto \mu_I(A)(\omega)$ is measurable from $(\Omega, \mathcal{A})$ to $(\overline{\mathbb{N}}_0, \mathcal{B}(\overline{\mathbb{N}}_0))$ with $\overline{\mathbb{N}}_0 := \mathbb{N}_0 \cup \{\infty\}$,
– for almost each $\omega \in \Omega$, the mapping $A \mapsto \mu_I(A)(\omega)$ is a locally finite measure on $([0, \infty) \times E_I, \mathcal{B}([0, \infty) \times E_I))$.

The observable information at time $t \geq 0$ is given by the complete filtration

$$\mathcal{F}_t := \sigma(\{T_{2i-1} \leq s < T_{2i}\} \cap \{Z_{2i} \in B\} : s \in [0, t], B \in \mathcal{E}, i \in \mathbb{N}) \vee \mathcal{Z},$$

which lets the random times $T_i$, $i \in \mathbb{N}$, be stopping times. Here the symbol $\vee$ denotes the sigma-algebra that is generated by the union of the involved sets. The admissible information at time $t \geq 0$ is given by the family of sub-sigma-algebras

$$\mathcal{G}_t = \sigma(\{T_{2i-1} \leq t < T_{2i}\} \cap \{Z_{2i} \in B\} : B \in \mathcal{E}, i \in \mathbb{N}) \vee \mathcal{Z}.$$

The admissible information immediately before time $t > 0$ is given by the family of sub-sigma-algebras

$$\mathcal{G}_t^- = \sigma(\{T_{2i-1} < t \leq T_{2i}\} \cap \{Z_{2i} \in B\} : B \in \mathcal{E}, i \in \mathbb{N}) \vee \mathcal{Z}.$$

Analogously to filtrations, we write $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ and $\mathbb{G}^- = (\mathcal{G}_t^-)_{t \geq 0}$.

**Remark 3.1** Recall that $T_{2i-1} \leq T_{2i}$, $i \in \mathbb{N}$, is the only kind of order that we assume to hold between the random times $T_i$, resulting from the natural assumption $\tau_i \leq \sigma_i$, $i \in \mathbb{N}$. This fact is relevant when an ordering unintentionally reveals additional information. For example, if we have a model where the innovation times $\tau_i$ are ordered, i.e., $T_1 < T_3 < T_5 < \cdots$, then $\mathcal{G}_t$ reveals among other things the exact number of deletions that have happened until $t$. This can be an unwanted feature if the number of past deletions is itself a non-admissible piece of information. In many situations, we can avoid such an implied information effect by ordering the pairs $(T_{2i-1}, T_{2i})$ in a non-informative way.

**Remark 3.2** Without loss of generality, suppose here that $0 \notin E$. We define an infinite-dimensional process $(\Gamma_t)_{t \geq 0}$ by

$$\Gamma_t := (Z_{2i} \mathbf{1}_{\{T_{2i-1} \leq t < T_{2i}\}})_{i \in \mathbb{N}}.$$

Then, using the fact that the paths of $(\Gamma_t)_{t \geq 0}$ are componentwise càdlàg, the information $\mathcal{G}_t$ and $\mathcal{G}_t^-$ can be alternatively represented as

$$\mathcal{G}_t = \sigma(\Gamma_t) \vee \mathcal{Z}, \qquad t \geq 0,$$

$$\mathcal{G}_t^- = \sigma(\Gamma_{t-}) \vee \mathcal{Z}, \qquad t > 0,$$

where the left limit $\Gamma_{t-}$ is defined componentwise. However, $\mathcal{G}_t^-$ is usually different from the left set-limit $\mathcal{G}_{t-}$, and the latter set-limit might not even exist. For example, consider a model with only two jumps $T_1$, $T_2$ in finite time and a trivial mark $Z_2 = \mathrm{const}$. It is not difficult to choose $T_1, T_2$ in such a way that the events $\{T_1 \leq s < T_2\}$ and $\{T_1 \leq u < T_2\}$ are different for each $s < u \leq t$. In this case, $\liminf_{s \uparrow t} \mathcal{G}_s = \bigvee_{s<t} \bigcap_{s<u<t} \mathcal{G}_u$ equals the completed trivial sigma-algebra, whereas $\limsup_{s \uparrow t} \mathcal{G}_s = \bigcap_{s<t} \bigvee_{s<u<t} \mathcal{G}_u$ equals $\mathcal{G}_t^-$.

## 4 Optional projections

In this section, we study existence and path properties of optional projections. Note that **this and all following sections generally assume that we are in the marked point process framework of Sect. 3**. Recall also our specific definition of $\mathcal{G}_t^-$.

**Theorem 4.1** *Suppose that $X = (X_t)_{t \geq 0}$ is a càdlàg process that satisfies*

$$E\left[\sup_{0 \leq s \leq t} |X_s|\right] < \infty, \qquad t \geq 0. \tag{4.1}$$

*Then the optional projection $X^{\mathbb{G}}$ according to Definition 2.6 exists, and we have $X_{t-}^{\mathbb{G}} = E[X_{t-}|\mathcal{G}_t^-]$ almost surely for each $t > 0$. If $X$ has integrable variation on compacts, then $X^{\mathbb{G}}$ has paths of finite variation on compacts.*

It might be surprising here that $X^{\mathbb{G}}$ is always a càdlàg process, but note that condition (3.1) rules out clusters of jump times in our marked point process framework. Before we turn to the proof of Theorem 4.1, we develop several auxiliary results. Let

$$\mathcal{N} := \{M \subseteq \mathbb{N} : |M| < \infty\},$$
$$\mathcal{M} := \{M \subseteq \{1, 3, 5, \ldots\} : |M| < \infty\}$$

be all finite subsets of the natural or the odd natural numbers, and define

$$R_I := \big(Q_I, (Z_i)_{i \in I}\big), \qquad I \in \mathcal{N},$$

where $Q_I := \sup\{t \geq 0 : \mu_I([0, t] \times E_I) = 0\}$.

Since $\Omega$ is a Polish space and $\mathcal{A}$ its Borel sigma-algebra, there exist regular conditional probabilities $P[\cdot|Z_M]$ and $P[\cdot|Z_M, R_I]$ on $(\Omega, \mathcal{A})$ for each $M \in \mathcal{M}$ and $I \in \mathcal{N}$. As the sets $\mathcal{M}$ and $\mathcal{N}$ are countable, all these conditional probabilities are simultaneously unique up to a joint exception nullset. In this paper, the notation

$$P_{M,R_I}[\cdot] = P[\cdot|Z_M, R_I]$$

refers to an arbitrary but fixed regular version of the conditional probability on the right-hand side, and for any integrable random variable $Z$, we set

$$E_{M,R_I}[Z] := \int Z \, dP_{M,R_I},$$

i.e., $E_{M,R_I}[Z]$ is the specific version of the conditional expectation $E[Z|Z_M, R_I]$ that we obtain by integrating $Z$ with respect to the specific regular versions that we picked for $P[\cdot|Z_M, R_I]$. In case of $I = \emptyset$, we also use the short forms $P_M = P_{M,R_\emptyset}$ and $E_M = E_{M,R_\emptyset}$ since $P_{M,R_\emptyset}$ is a version of $P[\cdot|Z_M]$.

Moreover, with defining $I - 1 := \{i - 1 : i \in I\}$, the mappings

$$P_{M,R_I=r}[\cdot] := P[\cdot|Z_{M_I} = z, R_I = r]\big|_{z=Z_{M_I}},$$
$$M_I := M \setminus \big(I \cup (I - 1)\big), \tag{4.2}$$

refer to arbitrary but fixed regular versions of the factorised conditional expectations on the right-hand side, and for any integrable random variable $Z$, we define

$$E_{M,R_I=r}[Z] := \int Z \, dP_{M,R_I=r}.$$

By reducing $M$ down to $M_I$, we leave out exactly those random variables in $Z_M$ that are already covered by $R_I$. Note that the mapping $P_{M,R_I=r}[\,\cdot\,]|_{r=R_I}$ equals $P_{M,R_I}[\,\cdot\,]$. For $M \in \mathcal{M}$ and $t \geq 0$, we define the $\mathcal{G}_t$-measurable sets

$$A_t^M := \bigcap_{i \in M}\{T_i \leq t < T_{i+1}\} \cap \bigcap_{i \notin M}(\Omega \setminus \{T_i \leq t < T_{i+1}\})$$

and corresponding $\mathbb{G}$-adapted stochastic processes $\mathbb{I}^M = (\mathbb{I}_t^M)_{t \geq 0}$ via

$$\mathbb{I}_t^M := \mathbf{1}_{A_t^M}, \qquad t \geq 0.$$

Because of the assumption (3.1), the paths of $\mathbb{I}^M$ have finitely many jumps on compacts only, so that they have left and right limits. Moreover, they are right-continuous by construction, so that the processes $\mathbb{I}^M$ are càdlàg. The left limits can be represented as $\mathbb{I}_{t-}^M = \mathbf{1}_{A_{t-}^M}$, where

$$A_{t-}^M := \bigcap_{i \in M}\{T_i < t \leq T_{i+1}\} \cap \bigcap_{i \notin M}(\Omega \setminus \{T_i < t \leq T_{i+1}\}).$$

**Proposition 4.2** *For any integrable random variable $\xi$ and any sets $M \in \mathcal{M}$ and $I \in \mathcal{N}$, we almost surely have*

$$\mathbb{I}_t^M E[\xi | \mathcal{G}_t \vee \sigma(R_I)] = \mathbb{I}_t^M \frac{E_{M,R_I}[\xi \mathbb{I}_t^M]}{E_{M,R_I}[\mathbb{I}_t^M]},$$

$$\mathbb{I}_{t-}^M E[\xi | \mathcal{G}_t^- \vee \sigma(R_I)] = \mathbb{I}_{t-}^M \frac{E_{M,R_I}[\xi \mathbb{I}_{t-}^M]}{E_{M,R_I}[\mathbb{I}_{t-}^M]}, \qquad (4.3)$$

*with the convention that $0/0 := 0$.*

Note here that $\sigma(R_I)$ equals the trivial sigma-algebra if $I = \emptyset$. Whenever we have $E_{M,R_I}[\mathbb{I}_t^M] = 0$ and $E_{M,R_I}[\mathbb{I}_{t-}^M] = 0$, we necessarily have $E_{M,R_I}[\xi \mathbb{I}_t^M] = 0$ and $E_{M,R_I}[\xi \mathbb{I}_{t-}^M] = 0$, respectively, so that the right-hand sides of (4.3) are well defined.

***Proof of Proposition 4.2*** The left-hand sides of (4.3) almost surely equal the conditional expectations that one obtains when the families $\mathbb{G}$ and $\mathbb{G}^-$ of sigma-algebras are replaced by their non-completed versions. Therefore, in the remaining proof, we ignore the extension by $\mathcal{Z}$ in the definitions of $\mathbb{G}$ and $\mathbb{G}^-$.

For each $H \in \sigma(Z_M)$, there exists a $G \in \mathcal{G}_t$ such that $H \cap A_t^M = G \cap A_t^M$, and for each $G \in \mathcal{G}_t$, there exists an $H \in \sigma(Z_M)$ such that $G \cap A_t^M = H \cap A_t^M$. Thus

$$\big(\sigma(Z_M) \vee \sigma(R_I)\big) \cap A_t^M = \big(\mathcal{G}_t \vee \sigma(R_I)\big) \cap A_t^M \subseteq \mathcal{G}_t \vee \sigma(R_I), \qquad t \geq 0. \quad (4.4)$$

This implies that the random variable $\mathbb{I}_t^M \frac{E_{M,R_I}[\xi \mathbb{I}_t^M]}{E_{M,R_I}[\mathbb{I}_t^M]}$ is $(\mathcal{G}_t \vee \sigma(R_I))$-measurable, and for each $G \in \mathcal{G}_t \vee \sigma(R_I)$, we obtain

$$E\left[\mathbf{1}_G \mathbb{I}_t^M \frac{E_{M,R_I}[\xi\mathbb{I}_t^M]}{E_{M,R_I}[\mathbb{I}_t^M]}\right] = E\left[E_{M,R_I}\left[\mathbf{1}_H\mathbb{I}_t^M \frac{E_{M,R_I}[\xi\mathbb{I}_t^M]}{E_{M,R_I}[\mathbb{I}_t^M]}\right]\right]$$

$$= E\left[\mathbf{1}_H E_{M,R_I}[\xi\mathbb{I}_t^M]\right]$$

$$= E[\mathbf{1}_G \mathbb{I}_t^M \xi]$$

$$= E\left[\mathbf{1}_G \mathbb{I}_t^M E[\xi|\mathcal{G}_t \vee \sigma(R_I)]\right],$$

i.e., the first equation in (4.3) holds. By replacing (4.4) by

$$\left(\sigma(Z_M) \vee \sigma(R_I)\right) \cap A_{t-}^M = \left(\mathcal{G}_t^- \vee \sigma(R_I)\right) \cap A_{t-}^M \subseteq \mathcal{G}_t^- \vee \sigma(R_I), \quad t \geq 0, \quad (4.5)$$

we can analogously show that the second equation in (4.3) holds. □

**Lemma 4.3** *For each $M \in \mathcal{M}, I \in \mathcal{N}, r \in [0,\infty) \times E_I$ and each càdlàg process $X$ that satisfies condition* (4.1), *the stochastic processes*

$$t \mapsto E_{M,R_I}[X_t\mathbb{I}_t^M],$$

$$t \mapsto E_{M,R_I=r}[X_t\mathbb{I}_t^M]$$

*have càdlàg paths. Moreover, their left limits can be obtained by replacing $X_t\mathbb{I}_t^M$ by $X_{t-}\mathbb{I}_{t-}^M$.*

**Proof** Apply the dominated convergence theorem. □

**Proposition 4.4** *With the conventions $0/0 := 0$ and $1/0 := \infty$, we have for each $M \in \mathcal{M}$ almost surely that*

$$\sup_{t\in[0,\infty)} \frac{\mathbb{I}_t^M}{E_M[\mathbb{I}_t^M]} < \infty.$$

**Proof** Let $\tau$ and $\sigma$ be any two nonnegative random times such that $\tau \leq \sigma$. At first we are going to show that

$$Z := \sup_{t\in[0,\infty)} \frac{\mathbf{1}_{\{\tau \leq t < \sigma\}}}{E[\mathbf{1}_{\{\tau \leq t < \sigma\}}]} < \infty \qquad (4.6)$$

almost surely. For each $(t, s) \in [0, \infty)^2$, we define the unbounded rectangles

$$A_{(t,s)} := \{(t', s') : t' \leq t, s < s'\}$$

and the countably generated set

$$B := \bigcup_{(t,s)\in\beta} A_{(t,s)}, \qquad \beta := \{(t,s) \in \mathbb{Q}_+^2 : t < s, P[(\tau,\sigma) \in A_{(t,s)}] = 0\}.$$

Let $\partial B$ and $B^\circ$ be the boundary and the interior of $B$. Any line of the form

$$L_x := \{(x, x) + \lambda(1, -1) : \lambda \in \mathbb{R}\}$$

intersects $\partial B$ at most at one point, since for any two points $y, y' \in L_x$ with $y \neq y'$, we either have $y \in A_{y'}^\circ$ or $y' \in A_y^\circ$. Therefore the set

$$\gamma := \left( \bigcup_{x \in \mathbb{Q}_+} L_x \right) \cap \partial B \cap \{(t, s) \in [0, \infty)^2 : P[(\tau, \sigma) \in A_{(t,s)}] = 0\}$$

is countable, and

$$C := \bigcup_{(t,s) \in \gamma} A_{(t,s)}$$

is countably generated. The sets $N_B = \{(\tau, \sigma) \in B\}$ and $N_C = \{(\tau, \sigma) \in C\}$ are both nullsets since they equal countable unions of nullsets.

Suppose now that $Z(\omega) = \infty$ for an arbitrary but fixed $\omega \in \Omega$. We necessarily have $\tau(\omega) < \sigma(\omega)$. Since $t \mapsto E[\mathbf{1}_{\{\tau \leq t < \sigma\}}]$ is a càdlàg function, at least one of the following statements is true:

(1) $E[\mathbf{1}_{\{\tau \leq u < \sigma\}}] = 0$ for some $u \in (\tau(\omega), \sigma(\omega))$.
(2) $E[\mathbf{1}_{\{\tau < u \leq \sigma\}}] = 0$ for some $u \in (\tau(\omega), \sigma(\omega))$.
(3) $E[\mathbf{1}_{\{\tau \leq u < \sigma\}}] = 0$ for $u = \tau(\omega)$.
(4) $E[\mathbf{1}_{\{\tau < u \leq \sigma\}}] = 0$ for $u = \sigma(\omega)$.

In case (1), we have $P[(\tau, \sigma) \in A_{(u,u)}] = E[\mathbf{1}_{\{\tau \leq u < \sigma\}}] = 0$ and $(\tau(\omega), \sigma(\omega))$ is in $A_{(u,u)}^\circ$, so that we can conclude that $\omega \in N_B$.

In case (2), we can argue analogously to case (1), but need to replace the definition of $A_{(t,s)}$ by $\{(t', s') : t' < t, s \leq s'\}$ and define a corresponding nullset $N_B'$. We obtain that $\omega \in N_B'$.

In case (3), we have $P[(\tau, \sigma) \in A_{(\tau(\omega), \tau(\omega))}] = E[\mathbf{1}_{\{\tau \leq \tau(\omega) < \sigma\}}] = 0$ as well as $(\tau(\omega), \sigma(\omega)) \in A_{(\tau(\omega), \tau(\omega))} \subseteq B \cup \partial B$. If $(\tau(\omega), \sigma(\omega)) \in B$, then $\omega \in N_B$. If $(\tau(\omega), \sigma(\omega)) \in \partial B$, then the whole line segment

$$\{\theta(\tau(\omega), \tau(\omega)) + (1 - \theta)(\tau(\omega), \sigma(\omega)) : \theta \in (0, 1)\}$$

is in $\partial B$ because of $(\tau(\omega), \tau(\omega)) \in \partial B$ and the rectangular shape of the sets $A_{(t,s)}$. On this line, there is at least one intersection with $C$, so that we can conclude that $\omega \in N_C$.

In case (4), we can argue similarly to case (3), but need to replace the definition of $A_{(t,s)}$ by $\{(t', s') : t' < t, s \leq s'\}$ and define corresponding nullsets $N_B'$ and $N_C'$.

All in all, we have $P[Y = \infty] \leq P[N_B \cup N_C \cup N_B' \cup N_C'] = 0$, i.e., (4.6) holds.

Now let $M \in \mathcal{M}$ be arbitrary but fixed and choose $\tau$ and $\sigma$ as the random times where $\mathbb{I}_t^M$ jumps from zero to one and jumps back to zero, respectively. Suppose that $P_{Z_M = z}$ is a regular version of $P[\cdot | Z_M = z]$ with corresponding expectation $E_{Z_M = z}[\cdot]$. Then from (4.6), we can conclude that

$$P_{Z_M = z}\left[ \sup_{t \in [0, \infty)} \frac{\mathbb{I}_t^M}{E_{Z_M = z}[\mathbb{I}_t^M]} = \infty \right] = 0$$

for each choice of $z$. Replacing both $z$ by $Z_M$, where we use the insertion rule for conditional expectations for the inner $z$, and taking the unconditional expectation on both sides of the equation, we end up with

$$P\left[\sup_{t\in[0,\infty)} \frac{\mathbb{I}_t^M}{E_{Z_M}[\mathbb{I}_t^M]} = \infty\right] = 0.$$

$\square$

***Proof of Theorem 4.1*** Motivated by Proposition 4.2, we set

$$Y_t := \sum_{M\in\mathcal{M}} \mathbb{I}_t^M \frac{E_M[X_t\mathbb{I}_t^M]}{E_M[\mathbb{I}_t^M]}, \qquad t \geq 0,$$

since this process almost surely equals $X_t^{\mathbb{G}}$ for each $t \geq 0$. Note that there are at most a countable number of conditional expectations involved; so the corresponding regular versions are simultaneously unique up to evanescence. For each compact interval $[0, t]$ and almost each $\omega \in \Omega$, the set

$$\mathcal{M}_t(\omega) := \{M \in \mathcal{M} : \mathbb{I}_u^M(\omega) = 1 \text{ for at least one } u \in [0, t]\} \tag{4.7}$$

is finite due to the assumption (3.1). If $E_M[\mathbb{I}_t^M](\omega) \neq 0$, Lemma 4.3 yields that

$$\lim_{\varepsilon\downarrow0} Y_{t+\varepsilon}(\omega) = \sum_{M\in\mathcal{M}_{t+1}(\omega)} \lim_{\varepsilon\downarrow0}\mathbb{I}_{t+\varepsilon}^M(\omega) \frac{E_M[X_{t+\varepsilon}\mathbb{I}_{t+\varepsilon}^M](\omega)}{E_M[\mathbb{I}_{t+\varepsilon}^M](\omega)}$$

$$= \sum_{M\in\mathcal{M}_{t+1}(\omega)} \mathbb{I}_t^M(\omega) \frac{E_M[X_t\mathbb{I}_t^M](\omega)}{E_M[\mathbb{I}_t^M](\omega)}$$

$$= Y_t(\omega). \tag{4.8}$$

If $E_M[\mathbb{I}_t^M](\omega) = 0$, Proposition 4.4 implies that $\mathbb{I}_t^M = 0$ for almost all $\omega \in \Omega$, where the exception nullset does not depend on the choice of $t$. So (4.8) is almost surely true on $[0, \infty)$, since $\mathbb{I}_t^M(\omega) = 0$ implies that there is a whole interval $[t, t + \epsilon_\omega)$ where the right-continuous jump path $s \mapsto \mathbb{I}_s^M(\omega)$ is constantly zero. Similarly, we can show that the process $Y$ almost surely has left limits, which are of the form

$$Y_{t-} = \sum_{M\in\mathcal{M}} \mathbb{I}_{t-}^M \frac{E_M[X_{t-}\mathbb{I}_{t-}^M]}{E_M[\mathbb{I}_{t-}^M]}, \qquad t > 0.$$

According to Proposition 4.2, $Y_{t-}$ almost surely equals $E[X_{t-}|\mathcal{G}_t^-]$. As càdlàg processes are uniquely defined by their values on countable dense subsets of the time line, our choice for $X^{\mathbb{G}}$ is almost surely the only possible modification of $(E[X_t|\mathcal{G}_t])_{t\geq0}$.

The variation of $Y$ on $[0, t]$ is bounded by

$$
\sum_{M \in \mathcal{M}_t} \sup_{\mathcal{T}^t} \sum_{\mathcal{T}^t} \left| \mathbb{I}_{t_{k+1}}^M \frac{E_M[X_{t_{k+1}} \mathbb{I}_{t_{k+1}}^M]}{E_M[\mathbb{I}_{t_{k+1}}^M]} - \mathbb{I}_{t_k}^M \frac{E_M[X_{t_k} \mathbb{I}_{t_k}^M]}{E_M[\mathbb{I}_{t_k}^M]} \right|
$$

$$
\leq \sum_{M \in \mathcal{M}_t} \sup_{\mathcal{T}^t} \sum_{\mathcal{T}^t} \left( \left| \frac{\mathbb{I}_{t_{k+1}}^M}{E_M[\mathbb{I}_{t_{k+1}}^M]} - \frac{\mathbb{I}_{t_k}^M}{E_M[\mathbb{I}_{t_k}^M]} \right| E_M[|X_{t_{k+1}}| \mathbb{I}_{t_{k+1}}^M] \right.
$$

$$
\left. + \frac{\mathbb{I}_{t_k}^M}{E_M[\mathbb{I}_{t_k}^M]} E_M[|X_{t_{k+1}} \mathbb{I}_{t_{k+1}}^M - X_{t_k} \mathbb{I}_{t_k}^M|] \right),
$$

where $\mathcal{T}^t$ is any partition of $[0, t]$. As $C_M(\omega) := \sup_t \mathbb{I}_t^M(\omega)/E_M[\mathbb{I}_t^M](\omega)$ is finite for almost each $\omega \in \Omega$ (see Proposition 4.4) and the variation of $L_M(s) := E_M[\mathbb{I}_s^M]$ is bounded by 2, the latter bound is dominated by

$$
\sum_{M \in \mathcal{M}_t} \left( \left( 2 C_M + \int_{[0,t]} \mathbb{I}_s^M \frac{1}{L_M(s) L_M(s-)} d|L_M|(s) \right) E_M \left[ \sup_{0 \leq s \leq t} |X_s| \right] \right.
$$

$$
\left. + C_M E_M \left[ 2 \sup_{0 \leq s \leq t} |X_s| + \int_{[0,t]} \mathbb{I}_s^M d|X|_s \right] \right)
$$

$$
\leq \sum_{M \in \mathcal{M}_t} \left( (2 C_M + 2 t\, C_M^2) E_M \left[ \int_{[0,t]} d|X|_s \right] + 3\, C_M E_M \left[ \int_{[0,t]} d|X|_s \right] \right),
$$

which is finite for almost each $\omega \in \Omega$ since $X$ has integrable variation on compacts and $\mathcal{M}_t(\omega)$ is finite. $\qquad \square$

## 5 Infinitesimal compensators

In this section, we derive infinitesimal compensators for a large class of incrementally adapted jump processes, in particular for the counting processes $t \mapsto \mu_I([0, t] \times B)$ for any $I \in \mathcal{N}$ and $B \in \mathcal{E}_I$. Under the conventions $0/0 := 0$ and (4.2), let

$$
\nu_I([0, t] \times B) := \sum_{M \in \mathcal{M}} \int_{(0,t] \times B} \mathbb{I}_{u-}^M \frac{P_{M, R_I = (u,e)}[A_{u-}^M]}{P_M[A_{u-}^M]} P_M^{R_I} (d(u, e)),
$$

$$
\rho_I([0, t] \times B) := \sum_{M \in \mathcal{M}} \int_{(0,t] \times B} \mathbb{I}_u^M \frac{P_{M, R_I = (u,e)}[A_u^M]}{P_M[A_u^M]} P_M^{R_I} (d(u, e))
$$

for $t \geq 0$, $B \in \mathcal{E}_I$, $I \in \mathcal{N}$.

**Proposition 5.1** *For each $I \in \mathbb{N}$, the mappings $\nu_I$ and $\rho_I$ can be uniquely extended to random measures on $([0, \infty) \times E_I, \mathcal{B}([0, \infty) \times E_I))$.*

The proof of the proposition is given below. In the following, we use the notation

$$F \bullet \kappa \big((0,t] \times B\big) := \int_{(0,t] \times B} F(u,e)\,\kappa\big(\mathrm{d}(u,e)\big)$$

for random measures $\kappa$ and integrable random functions $F$.

**Theorem 5.2** *Suppose that the mappings $(t,e,\omega) \mapsto F_I(t,e)(\omega)$, $I \in \mathcal{N}$, are jointly measurable and satisfy*

$$E\left[\int_{(0,t] \times E_I} |F_I(u,e)|\,\mu_I\big(\mathrm{d}(u,e)\big)\right] < \infty. \tag{5.1}$$

*If $F_I(t,e)$ is $\mathcal{G}_t^-$-measurable for each $(t,e)$, then for each $B \in \mathcal{E}_I$, the jump process*

$$t \mapsto F_I \bullet \mu_I\big((0,t] \times B\big)$$

*has the IF-compensator*

$$t \mapsto F_I \bullet \nu_I\big((0,t] \times B\big).$$

*If $F_I(t,e)$ is $\mathcal{G}_t$-measurable for each $(t,e)$, then for each $B \in \mathcal{E}_I$, the jump process*

$$t \mapsto F_I \bullet \mu_I\big((0,t] \times B\big)$$

*has the IB-compensator*

$$t \mapsto F_I \bullet \rho_I\big((0,t] \times B\big).$$

By choosing $F_I \equiv 1$, Theorem 5.2 yields in particular that $\nu_I$ is the IF-compensator and $\rho_I$ is the IB-compensator of the counting process $\mu_I$. In intuitive notation, we write this fact as

$$E[\mu_I(\mathrm{d}t \times B)|\mathcal{G}_t^-] = \nu_I(\mathrm{d}t \times B), \qquad B \in \mathcal{E}_I,$$
$$E[\mu_I(\mathrm{d}t \times B)|\mathcal{G}_t] = \rho_I(\mathrm{d}t \times B), \qquad B \in \mathcal{E}_I.$$

The proofs of Proposition 5.1 and Theorem 5.2 follow now in several steps.

**Lemma 5.3** *For each $M \in \mathcal{M}$ and $t \geq 0$, we almost surely have*

$$\sum_{I \in \mathcal{N}} \int_{[0,t] \times E_I} P_M^{R_I}\big(\mathrm{d}(u,e)\big) < \infty. \tag{5.2}$$

***Proof*** For each $t \geq 0$ and $M \in \mathcal{M}$, (3.1) implies that

$$E_M\left[\sum_{j=1}^{\infty} \mathbf{1}_{\{T_j \leq t\}}\right] < \infty$$

almost surely. Therefore, applying the monotone convergence theorem yields

$$\infty > E_M\left[\sum_{i=1}^{\infty} \mathbf{1}_{\{T_i \le t\}}\right] \ge E_M\left[\sum_{I \in \mathcal{N}} \mu_I([0,t] \times E_I)\right]$$

$$= \sum_{I \in \mathcal{N}} E_M\left[\mu_I([0,t] \times E_I)\right]$$

$$= \sum_{I \in \mathcal{N}} \int_{[0,t] \times E_I} P_M^{R_I}\big(\mathrm{d}(u,e)\big)$$

almost surely for each $M \in \mathcal{M}$ and $t \ge 0$. □

**Proof of Proposition 5.1** The processes $(\mathbb{I}_{u-}^M)$ and $(P_M[A_{u-}^M])$ are jointly measurable with respect to $(u, \omega)$ since they are left-continuous in $u$; see Lemma 4.3. The mapping $(u, e, \omega) \mapsto P_{M, R_I=(u,e)}[A_{u-}^M]$ is jointly measurable with respect to $(u, e, \omega)$ since $P_{M, R_I=(u,e)}[A_{s-}^M]$ is left-continuous in $s$ and jointly measurable with respect to $(u, e, \omega) \in [0, \infty)^{|I|} \times E_I \times \Omega$; see Lemma 4.3. Thus for any fixed $A \in \mathcal{B}([0, \infty) \times E_I)$, the mapping $\omega \mapsto \nu_I(A)(\omega)$ is measurable. Moreover, for almost each $\omega \in \Omega$, the mapping $A \mapsto \nu_I(A)(\omega)$ is a locally finite measure on $([0, \infty) \times E_I, \mathcal{B}([0, \infty) \times E_I))$. This can be seen by combining Proposition 4.4 and (5.2) and using the fact that $P_{M, R_I=(u,e)}[A_{u-}^M]$ is bounded by 1. Hence $\nu_I$ has a unique extension to a random measure on $([0, \infty) \times E_I, \mathcal{B}([0, \infty) \times E_I))$. Similar conclusions hold for the mappings $\rho_I$. □

**Proposition 5.4** *Suppose that the mappings* $(t, e, \omega) \mapsto F_I(t, e)(\omega)$, $I \in \mathcal{N}$, *are jointly measurable and satisfy* (5.1). *For each* $t > 0$ *and* $B \in \mathcal{E}_I$, *we almost surely have*

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E\big[F_I \bullet \mu_I\big((t_k, t_{k+1}] \times B\big)\big|\mathcal{G}_{t_k}\big] = G_I \bullet \nu_I\big((0,t] \times B\big),$$

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E\big[F_I \bullet \mu_I\big((t_k, t_{k+1}] \times B\big)\big|\mathcal{G}_{t_{k+1}}\big] = H_I \bullet \rho_I\big((0,t] \times B\big)$$

*for any increasing sequence* $(\mathcal{T}_n^t)_{n \in \mathbb{N}}$ *of partitions of* $[0, t]$ *with* $\lim_{n \to \infty} |\mathcal{T}_n^t| = 0$ *and for* $G_I$ *and* $H_I$ *defined by*

$$G_I(u, e) := \sum_{M \in \mathcal{M}} \mathbb{I}_{u-}^M \frac{E_{M, R_I=(u,e)}[\mathbb{I}_{u-}^M F_I(u, e)]}{E_{M, R_I=(u,e)}[\mathbb{I}_{u-}^M]},$$

$$H_I(u, e) := \sum_{M \in \mathcal{M}} \mathbb{I}_u^M \frac{E_{M, R_I=(u,e)}[\mathbb{I}_u^M F_I(u, e)]}{E_{M, R_I=(u,e)}[\mathbb{I}_u^M]}.$$

**Proof** By decomposing $F$ into a positive part $F^+$ and a negative part $F^-$, it suffices to prove the first equation for the nonnegative mappings $F^+$ and $F^-$ only. Therefore, without loss of generality, we suppose from now on that $F$ is nonnegative.

Let $\mathcal{M}_t = \mathcal{M}_t(\omega)$ be defined as in (4.7). In the following, we use the notation $J_k := (t_k, t_{k+1}]$. Since $\sum_{M \in \mathcal{M}_t} \mathbb{I}_{t_k}^M = 1$ for any $t_k$, applying (4.3), the monotone convergence theorem and the law of total probability gives

$$E[F_I \bullet \mu_I (J_k \times B) | \mathcal{G}_{t_k}]$$

$$= \sum_{M \in \mathcal{M}_t} \mathbb{I}_{t_k}^M \frac{E_M[\mathbb{I}_{t_k}^M F_I \bullet \mu_I (J_k \times B)]}{E_M[\mathbb{I}_{t_k}^M]}$$

$$= \sum_{M \in \mathcal{M}_t} \int_{J_k \times E_I} \mathbb{I}_{t_k}^M \frac{E_{M, R_I = (u,e)}[\mathbb{I}_{t_k}^M F_I \bullet \mu_I (J_k \times B)]}{P_M[A_{t_k}^M]} P_M^{R_I}(\mathrm{d}(u, e))$$

for almost each $\omega \in \Omega$. For $u \in (0, t]$, let $J^u$ be the unique interval $(t_k, t_{k+1}]$ from $\mathcal{T}_n^t$ such that $t_k < u \le t_{k+1}$, and let $t(u)$ be the left end point of $J^u$. Then we can write

$$\sum_{\mathcal{T}_n^t} E[F_I \bullet \mu_I (J_k \times B) | \mathcal{G}_{t_k}]$$

$$= \sum_{M \in \mathcal{M}_t} \int_{(0,t] \times E_I} \mathbb{I}_{t(u)}^M \frac{E_{M, R_I = (u,e)}[\mathbb{I}_{t(u)}^M F \bullet \mu_I (J^u \times B)]}{P_M[A_{t(u)}^M]} P_M^{R_I}(\mathrm{d}(u, e)).$$

Taking the limit for $n \to \infty$, we obtain for almost each $\omega \in \Omega$ that

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E[F_I \bullet \mu_I (J_k \times B) | \mathcal{G}_{t_k}]$$

$$= \sum_{M \in \mathcal{M}_t} \int_{(0,t] \times E_I} \lim_{n \to \infty} \mathbb{I}_{t(u)}^M \frac{E_{M, R_I = (u,e)}[\mathbb{I}_{t(u)}^M F_I \bullet \mu_I (J^u \times B)]}{P_M[A_{t(u)}^M]} P_M^{R_I}(\mathrm{d}(u, e)),$$

$$(5.3)$$

using that $\mathcal{M}_t$ is finite for almost each $\omega$ and applying the monotone convergence and the dominated convergence theorem. Note that Proposition 4.4, the assumption (5.1) and $0 \le \mathbb{I}_{t(u)}^M F_I \bullet \mu_I (J^u \times B) \le F_I \bullet \mu_I ((0, t] \times B)$ ensure the existence of an integrable majorant. For $n \to \infty$, we have $t(u) \uparrow u$ and $J^u \downarrow \{u\}$; so the dominated convergence theorem implies that

$$\lim_{n \to \infty} E_{M, R_I = (u,e)}[\mathbb{I}_{t(u)}^M F_I \bullet \mu_I (J^u \times B)]$$

$$= E_{M, R_I = (u,e)}[\mathbb{I}_{u-}^M \mathbf{1}_B(e) F_I(u, e) \mu_I(\{u\} \times \{e\})]$$

$$= \mathbf{1}_B(e) E_{M, R_I = (u,e)}[\mathbb{I}_{u-}^M F_I(u, e)].$$

In summary, the right-hand side of (5.3) equals the integral $G_I \bullet \nu_I ((0, t] \times B)$, and we can conclude that the first equation in Proposition 5.4 holds. The proof of the second is similar. $\qquad\square$

**Proposition 5.5** *Under the assumptions of Proposition* 5.4, *for each $t \geq 0$ and $B \in \mathcal{E}_I$, we almost surely have*

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E\big[G_I \bullet \nu_I\big((t_k, t_{k+1}] \times B\big)\big|\mathcal{G}_{t_k}\big] = G_I \bullet \nu_I\big((0, t] \times B\big),$$

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E\big[H_I \bullet \rho_I\big((t_k, t_{k+1}] \times B\big)\big|\mathcal{G}_{t_{k+1}}\big] = H_I \bullet \rho_I\big((0, t] \times B\big)$$

*for any increasing sequence $(\mathcal{T}_n^t)_{n \in \mathbb{N}}$ of partitions of $[0, t]$ with $\lim_{n \to \infty} |\mathcal{T}_n^t| = 0$.*

**Proof** By decomposing $G$ into a positive part $G^+$ and a negative part $G^-$, it suffices to prove the first equation for the nonnegative mappings $G^+$ and $G^-$ only. Therefore, without loss of generality, we suppose from now on that $G$ is nonnegative.

From the definition of $\nu_I$ and the monotone convergence theorem, we get

$$E\big[G_I \bullet \nu_I\big((0, t] \times E_I\big)\big]$$

$$= \sum_{M \in \mathcal{M}} E\bigg[E_M\bigg[\int_{(0,t] \times E_I} G_I(u, e) \mathbb{I}_{u-}^M \frac{P_{M, R_I=(u,e)}[A_{u-}^M]}{P_M[A_{u-}^M]} P_M^{R_I}\big(\mathrm{d}(u, e)\big)\bigg]\bigg].$$

From Proposition 4.2, we know that $G_I(u, e)$ is $\mathcal{G}_u^-$-measurable for each $(u, e)$. This fact and (4.5) imply that

$$G_I(u, e) \mathbb{I}_{u-}^M P_{M, R_I=(u,e)}[A_{u-}^M] = \mathbb{I}_{u-}^M E_{M, R_I=(u,e)}[\mathbb{I}_{u-}^M G_I(u, e)].$$

Applying the Fubini–Tonelli theorem and the monotone convergence theorem gives

$$E[G_I \bullet \nu_I((0, t] \times E_I)]$$

$$= \sum_{M \in \mathcal{M}} E\bigg[\int_{(0,t] \times E_I} E_M[\mathbb{I}_{u-}^M] \frac{E_{M, R_I=(u,e)}[\mathbb{I}_{Q_I-}^M G_I(Q_I, Z_I)]}{P_M[A_{u-}^M]} P_M^{R_I}\big(\mathrm{d}(u, e)\big)\bigg]$$

$$= \sum_{M \in \mathcal{M}} E\bigg[E_M\big[\mathbb{I}_{Q_I-}^M G_I(Q_I, Z_I) \mu_I\big((0, t] \times E_I\big)\big]\bigg]$$

$$= E\bigg[G_I \bullet \mu_I\big((0, t] \times E_I\big)\bigg].$$

The latter expectation is finite according to (5.1). Hence for each $M \in \mathcal{M}$, we almost surely have

$$E_M\big[G_I \bullet \nu_I\big((0, t] \times E_I\big)\big] < \infty,$$

$$G_I \bullet \nu_I\big((0, t] \times E_I\big) < \infty. \tag{5.4}$$

Let $J_k := (t_k, t_{k+1}]$. From the dominated convergence theorem, we obtain

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} \mathbb{I}_{t_k}^M G_I \bullet \nu_I(J_k \times B) = (\mathbb{I}_{\cdot-}^M G_I) \bullet \nu_I\big((0, t] \times B\big)$$

since $\mathbb{I}^M$ is bounded by 1 and because of the second line in (5.4). By using the first line in (5.4), the dominated convergence theorem moreover yields

$$\lim_{n\to\infty} \sum_{\mathcal{T}_n^t} E_M[\mathbb{I}_{t_k}^M G_I \bullet \nu_I(J_k \times B)] = E_M\big[(\mathbb{I}_{.-}^M G_I) \bullet \nu\big((0,t] \times B\big)\big].$$

By applying the Fubini–Tonelli theorem, we can show that the last term equals

$$\int_{(0,t]\times B} E_M[\mathbb{I}_{u-}^M G_I(u,e)] \frac{P_{M,R_I=(u,e)}[A_{u-}^M]}{P_M[A_{u-}^M]} P_M^{R_I}\big(\mathrm{d}(u,e)\big).$$

Using Proposition 4.4 and the dominated convergence theorem, we therefore obtain

$$\lim_{n\to\infty} \sum_{\mathcal{T}_n^t} \mathbb{I}_{t_k}^M \frac{E_M[\mathbb{I}_{t_k}^M G_I \bullet \nu_I(J_k \times B)]}{E_M[\mathbb{I}_{t_k}^M]}$$

$$= \int_{(0,t]\times B} \frac{\mathbb{I}_{u-}^M}{E[\mathbb{I}_{u-}^M]} E_M[\mathbb{I}_{u-}^M G_I(u,e)] \frac{P_{M,R_I=(u,e)}[A_{u-}^M]}{P_M[A_{u-}^M]} P_M^{R_I}\big(\mathrm{d}(u,e)\big)$$

$$= \int_{(0,t]\times B} \mathbb{I}_{u-}^M G_I(u,e) \frac{P_{M,R_I=(u,e)}[A_{u-}^M]}{P_M[A_{u-}^M]} P_M^{R_I}\big(\mathrm{d}(u,e)\big),$$

where the second equality uses that (4.5) and the $\mathcal{G}_u^-$-measurability of $G_I(u,e)$ allows us to pull $G_I(u,e)$ out of the conditional expectation $E_M[\mathbb{I}_{u-}^M G_I(u,e)]$. Summing the latter equation over $M \in \mathcal{M}_t$ for $\mathcal{M}_t$ as in (4.7) and applying Proposition 4.2, we obtain

$$\lim_{n\to\infty} \sum_{\mathcal{T}_n^t} E\big[G_I \bullet \nu_I\big((t_k, t_{k+1}] \times B\big)\big|\mathcal{G}_{t_k}\big] = G_I \bullet \nu_I\big((0,t] \times B\big)$$

almost surely. Thus we can conclude that the first equation in Proposition 5.5 holds. The proof of the second is similar. □

***Proof of Theorem 5.2*** Let $G_I$ and $H_I$ be defined as in Proposition 5.4. If $F_I(t,e)$ is $\mathcal{G}_t^-$-measurable for each $(t,e)$, then Proposition 4.2 implies that $G_I(t,e) = F_I(t,e)$ almost surely. Similarly, if $F_I(t,e)$ is $\mathcal{G}_t$-measurable for each $(t,e)$, we have almost surely that $H_I(t,e) = F_I(t,e)$. With this fact and by subtracting the limit equations in Propositions 5.4 and 5.5, we obtain that

$$G_I \bullet \mu_I([0,t] \times B) - G_I \bullet \nu_I([0,t] \times B),$$

$$H_I \bullet \mu_I([0,t] \times B) - H_I \bullet \rho_I([0,t] \times B)$$

satisfy the defining limit equations for IF/IB-martingales. IF/IB-predictability of the compensators follows from Proposition 5.5. Note that all involved processes are incrementally adapted to $\mathbb{G}$ because of (4.4) and (4.5). □

## 6 Infinitesimal martingale representations

Suppose that $\lambda_I$ is the compensator of $\mu_I$ with respect to $\mathbb{F}$. For each integrable random variable $\xi$, the classical martingale representation theorem yields that the martingale $X_t = E[\xi|\mathcal{F}_t]$, $t \geq 0$, can be represented as

$$X_t = X_0 + \sum_{I \in \mathcal{N}} \int_{(0,t] \times E_I} F_I(u,e)\Big(\mu_I\big(\mathrm{d}(u,e)\big) - \lambda_I\big(\mathrm{d}(u,e)\big)\Big), \qquad (6.1)$$

where the mapping $(u,e,\omega) \mapsto F(u,e)(\omega)$ is jointly measurable and the mapping $\omega \mapsto F(u,e)(\omega)$ is $\mathcal{F}_{u-}$-measurable for each $(u,e)$; see e.g. Karr [16, Theorem 2.34]. We now extend this result to the non-monotone information $\mathbb{G}$.

**Theorem 6.1** *Let $\xi$ be an integrable random variable. Then for each $t \geq 0$, (1.2) holds almost surely for*

$$G_I(s,u,e) := \sum_{M \in \mathcal{M}} \mathbb{I}_s^M \left( \frac{E_{M,R_I=(u,e)}[\mathbb{I}_s^M \xi]}{E_{M,R_I=(u,e)}[\mathbb{I}_s^M]} - \frac{E_M[\mathbb{I}_{u-}^M \mathbb{I}_u^M \xi]}{E_M[\mathbb{I}_{u-}^M \mathbb{I}_u^M]} \right). \qquad (6.2)$$

*For each $I \in \mathcal{N}$ and $e \in E_I$, the process $u \mapsto G_I(u-,u,e)$ is $\mathbb{G}^-$-adapted and the process $u \mapsto G_I(u,u,e)$ is $\mathbb{G}$-adapted.*

If the mappings $F_I(u,e) = G_I(u-,u,E)$ and $F_I(u,e) = G_I(u,u,e)$ both satisfy the integrability condition in Theorem 5.2, then the representation (1.2) is a sum of IF-martingales and IB-martingales with respect to $\mathbb{G}$. In the case of $\mathbb{F} = \mathbb{G}$, we have $\nu_I = \lambda_I$, $\rho_I = \mu_I$ and (1.2) equals (6.1); so (1.2) is a generalisation of (6.1).

The proof of Theorem 6.1 is given below. Recall that our notation uses the convention (4.2).

**Lemma 6.2** *Let $\xi$ be an integrable random variable. Then for each $t \geq 0$, we have*

$$E_M[\mathbb{I}_t^M \xi] - E_M[\mathbb{I}_0^M \xi]$$
$$= \sum_{I \in \mathcal{N}} \int_{(0,t] \times E_I} E_{M,R_I=(u,e)}[(\mathbb{I}_u^M - \mathbb{I}_{u-}^M)\xi] P_M^{R_I}\big(\mathrm{d}(u,e)\big). \qquad (6.3)$$

**Proof** As (6.3) is additive in $\xi$, it suffices to show the equation for nonnegative and bounded random variables $\xi$ only. The general case then follows from monotone convergence applied to both parts of the sequence $\xi_n := (\xi_n \wedge n)^+ - (-\xi_n \wedge n)^+$, $n \in \mathbb{N}$. Therefore, in the remaining proof, we suppose that $0 \leq \xi \leq C$ for a finite real number $C$.

Let $U_{t_k}(\omega) := \sup\{s \in (t_k, \infty) : T_j(\omega) \notin (t_k, s), j \in \mathbb{N}\}$, i.e., $U_{t_k}$ is the time of the first occurrence of a random time strictly after $t_k$. Since $1 = \sum_{I \in \mathcal{N}} \mathbf{1}_{\{U_{t_k}=Q_I\}}$, we can conclude that

$$E_M[\mathbb{I}_t^M \xi] - E_M[\mathbb{I}_0^M \xi]$$

$$= \lim_{n \to \infty} \sum_{\mathcal{T}_n^t} \sum_{I \in \mathcal{N}} E_M \big[ \mathbf{1}_{\{U_{t_k} = Q_I\}} (\mathbb{I}_{t_{k+1}}^M - \mathbb{I}_{t_k}^M) \xi \big]$$

$$= \lim_{n \to \infty} \sum_{I \in \mathcal{N}} \sum_{\mathcal{T}_n^t} \int_{B_{I,k}} E_{M, R_I = (u,e)} \big[ \mathbf{1}_{\{U_{t_k} = Q_I\}} (\mathbb{I}_{t_{k+1}}^M - \mathbb{I}_{t_k}^M) \xi \big] P_M^{R_I} (\mathrm{d}(u, e)), \quad (6.4)$$

for $B_{I,k} = (t_k, t_{k+1}] \times E_I$, where we use the fact that

$$\mathbf{1}_{\{U_{t_k} = Q_I\}} (\mathbb{I}_{t_{k+1}}^M - \mathbb{I}_{t_k}^M) = 0$$

unless $t_k < Q_I \le t_{k+1}$. Because of (5.2) and

$$\big| E_{M, R_I = (u,e)} \big[ \mathbf{1}_{\{U_{t_k} = Q_I\}} (\mathbb{I}_{t_{k+1}}^M - \mathbb{I}_{t_k}^M) \xi \big] \big| \le 2C,$$

we can apply the dominated convergence theorem on the last line in (6.4), which leads to (6.3). Note here that

$$\mathbf{1}_{\{U_{t_k} = Q_I = u\}} (\mathbb{I}_{t_{k+1}}^M - \mathbb{I}_{t_k}^M) \longrightarrow \mathbf{1}_{\{Q_I = u\}} (\mathbb{I}_u^M - \mathbb{I}_{u-}^M)$$

for $t_{k+1} \downarrow u$ and $t_k \uparrow u$ implies that

$$E_{M, R_I = (u,e)} \big[ \mathbf{1}_{\{U_{t_k} = Q_I\}} (\mathbb{I}_{t_{k+1}}^M - \mathbb{I}_{t_k}^M) \xi \big] \longrightarrow E_{M, R_I = (u,e)} [(\mathbb{I}_u^M - \mathbb{I}_{u-}^M) \xi].$$

$$\square$$

**Proof of Theorem 6.1** Fix $M \in \mathcal{M}$ and define $M + 1 := \{i + 1 : i \in M\}$. If $\mathbb{I}_{u-}^M = 1$, then only random times from the index set

$$M' := (\{1, 3, \ldots\} \setminus M) \cup (M + 1)$$

can occur at time $u$. If $\mathbb{I}_u^M = 1$, then only random times from the index set

$$M'' := M \cup \big( \{2, 4, \ldots\} \setminus (M + 1) \big)$$

can occur at time $u$. Therefore (6.3) can be represented as

$$E_M[\mathbb{I}_t^M \xi] = K_t + L_t, \qquad t \ge 0,$$

where

$$K_t := \sum_{I \subseteq M'} \int_{(0,t] \times E_I} E_{M, R_I = (u,e)} [(\mathbb{I}_u^M - \mathbb{I}_{u-}^M) \xi] P_M^{R_I} (\mathrm{d}(u, e)) + E_M[\mathbb{I}_0^M \xi]$$

$$= - \sum_{I \subseteq M'} \int_{(0,t] \times E_I} E_{M, R_I = (u,e)} [\mathbb{I}_{u-}^M \xi] P_M^{R_I} (\mathrm{d}(u, e)) + E_M[\mathbb{I}_0^M \xi],$$

$$L_t := \sum_{I \subseteq M''} \int_{(0,t] \times E_I} E_{M, R_I = (u,e)} [(\mathbb{I}_u^M - \mathbb{I}_{u-}^M) \xi] P_M^{R_I} (\mathrm{d}(u, e))$$

$$= \sum_{I \subseteq M''} \int_{(0,t] \times E_I} E_{M, R_I = (u,e)} [\mathbb{I}_u^M \xi] P_M^{R_I} (\mathrm{d}(u, e)).$$

Furthermore, using $M' \cap M'' = \emptyset$, we can show that

$$E_M[\mathbb{I}_t^M \xi] - E_M[\mathbb{I}_{t-}^M \mathbb{I}_t^M \xi] = \sum_{I \subseteq M''} E_M[\mathbb{I}_t^M \mathbf{1}_{\{Q_I = t\}} \xi]$$

$$= \sum_{I \subseteq M''} \int_{\{t\} \times E_I} E_{M, R_I = (u,e)} [\mathbb{I}_u^M \xi] P_M^{R_I} \big( \mathrm{d}(u, e) \big)$$

$$= \Delta L_t$$

for $t > 0$, which implies that $E_M[\mathbb{I}_{t-}^M \mathbb{I}_t^M \xi] = K_t + L_{t-}$. Analogously we obtain

$$E_M[\mathbb{I}_{t-}^M \xi] - E_M[\mathbb{I}_{t-}^M \mathbb{I}_t^M \xi] = -\Delta K_t.$$

In the specific case $\xi = 1$, we write $k_t$ and $\ell_t$ instead of $K_t$ and $L_t$. By applying integration by parts pathwise for each $\omega \in \Omega$, we get that

$$(K_t + L_{t-}) \, \mathrm{d}\mathbb{I}_t^M + \mathbb{I}_{t-}^M \mathrm{d}K_t + \mathbb{I}_t^M \mathrm{d}L_t$$

$$= \mathrm{d}(\mathbb{I}_t^M E_M[\mathbb{I}_t^M \xi])$$

$$= \mathrm{d}\left( \frac{\mathbb{I}_t^M E_M[\mathbb{I}_t^M \xi]}{E_M[\mathbb{I}_t^M]} E_M[\mathbb{I}_t^M] \right)$$

$$= (k_t + \ell_{t-}) \, \mathrm{d}\left( \frac{\mathbb{I}_t^M E_M[\mathbb{I}_t^M \xi]}{E_M[\mathbb{I}_t^M]} \right) + \frac{\mathbb{I}_{t-}^M E_M[\mathbb{I}_{t-}^M \xi]}{E_M[\mathbb{I}_{t-}^M]} \mathrm{d}k_t + \frac{\mathbb{I}_t^M E_M[\mathbb{I}_t^M \xi]}{E_M[\mathbb{I}_t^M]} \mathrm{d}\ell_t$$

$$= (k_t + \ell_{t-}) \, \mathrm{d}\left( \frac{\mathbb{I}_t^M E_M[\mathbb{I}_t^M \xi]}{E_M[\mathbb{I}_t^M]} \right) + \mathbb{I}_{t-}^M \frac{K_{t-} + L_{t-}}{k_{t-} + \ell_{t-}} \mathrm{d}k_t + \mathbb{I}_t^M \frac{K_t + L_t}{k_t + \ell_t} \mathrm{d}\ell_t.$$

The equation formed by the first and last line can be rewritten as

$$(K_t + L_{t-}) \, \mathrm{d}\mathbb{I}_t^M + \frac{k_t + \ell_{t-}}{k_{t-} + \ell_{t-}} \mathbb{I}_{t-}^M \mathrm{d}K_t + \frac{k_t + \ell_{t-}}{k_t + \ell_t} \mathbb{I}_t^M \mathrm{d}L_t$$

$$= (k_t + \ell_{t-}) \, \mathrm{d}\left( \frac{\mathbb{I}_t^M E_M[\mathbb{I}_t^M \xi]}{E_M[\mathbb{I}_t^M]} \right) + \mathbb{I}_{t-}^M \frac{K_t + L_{t-}}{k_{t-} + \ell_{t-}} \mathrm{d}k_t + \mathbb{I}_t^M \frac{K_t + L_{t-}}{k_t + \ell_t} \mathrm{d}\ell_t,$$

because of

$$\left( \frac{k_t + \ell_{t-}}{k_{t-} + \ell_{t-}} - 1 \right) \mathbb{I}_{t-}^M \Delta K_t + \left( \frac{k_t + \ell_{t-}}{k_t + \ell_t} - 1 \right) \mathbb{I}_t^M \Delta L_t$$

$$= \mathbb{I}_{t-}^M \left( \frac{K_t + L_{t-}}{k_{t-} + \ell_{t-}} - \frac{K_{t-} + L_{t-}}{k_{t-} + \ell_{t-}} \right) \Delta k_t + \mathbb{I}_t^M \left( \frac{K_t + L_{t-}}{k_t + \ell_t} - \frac{K_t + L_{t-}}{k_t + \ell_t} \right) \Delta k_t.$$

With the convention $0/0 := 0$ and by using the Radon–Nikodým theorem, we may multiply by $(k_t + \ell_{t-})^{-1}$ on both sides, which leads to

$$\frac{E_M[\mathbb{I}_{t-}^M \mathbb{I}_t^M \xi]}{E_M[\mathbb{I}_{t-}^M \mathbb{I}_t^M]} \, d\mathbb{I}_t^M + \frac{\mathbb{I}_{t-}^M}{E_M[\mathbb{I}_{t-}^M]} dK_t + \frac{\mathbb{I}_t^M}{E_M[\mathbb{I}_t^M]} dL_t$$

$$= d\left(\frac{\mathbb{I}_t^M E_M[\mathbb{I}_t^M \xi]}{E_M[\mathbb{I}_t^M]}\right) + \frac{E_M[\mathbb{I}_{t-}^M \mathbb{I}_t^M \xi]}{E_M[\mathbb{I}_{t-}^M \mathbb{I}_t^M]} \left(\frac{\mathbb{I}_{t-}^M}{E_M[\mathbb{I}_{t-}^M]} dk_t + \frac{\mathbb{I}_t^M}{E_M[\mathbb{I}_t^M]} d\ell_t\right).$$

Because of (6.3) and $d\mathbb{I}_t^M = \sum_{I \in \mathcal{N}} (\mathbb{I}_t^M - \mathbb{I}_{t-}^M)\mu_I(dt \times E_I)$, the last equation can be rewritten to

$$\sum_{I \in \mathcal{N}} \frac{E_M[\mathbb{I}_{t-}^M \mathbb{I}_t^M \xi]}{E_M[\mathbb{I}_{t-}^M \mathbb{I}_t^M]} (\mathbb{I}_t^M - \mathbb{I}_{t-}^M)\mu_I(dt \times E_I)$$

$$- \sum_{I \in \mathcal{N}} \int_{E_I} \mathbb{I}_{t-}^M \frac{E_{M,R_I=(t,e)}[\mathbb{I}_{t-}^M \xi]}{E_{M,R_I=(t,e)}[\mathbb{I}_{t-}^M]} \nu_I(dt \times de)$$

$$+ \sum_{I \in \mathcal{N}} \int_{E_I} \mathbb{I}_t^M \frac{E_{M,R_I=(t,e)}[\mathbb{I}_t^M \xi]}{E_{M,R_I=(t,e)}[\mathbb{I}_t^M]} \rho_I(dt \times de)$$

$$= d\left(\frac{\mathbb{I}_t^M E_M[\mathbb{I}_t^M \xi]}{E_M[\mathbb{I}_t^M]}\right)$$

$$+ \sum_{I \in \mathcal{N}} \frac{E_M[\mathbb{I}_{t-}^M \mathbb{I}_t^M \xi]}{E_M[\mathbb{I}_{t-}^M \mathbb{I}_t^M]} \left(-\mathbb{I}_{t-}^M \nu_I(dt \times E_I) + \mathbb{I}_t^M \rho_I(dt \times E_I)\right). \tag{6.5}$$

Let $I \in \mathcal{N}$ be arbitrary but fixed. Then for each $M \in \mathcal{M}$, there exists an $\tilde{M} \in \mathcal{M}$, and for each $\tilde{M} \in \mathcal{M}$, there exists an $M \in \mathcal{M}$ such that

$$\mathbb{I}_{t-}^M \mu_I(dt \times de) = \mathbb{I}_t^{\tilde{M}} \mu_I(dt \times de).$$

As a consequence, for almost each $\omega \in \Omega$, we have

$$0 = \sum_{M \in \mathcal{M}} \sum_{I \in \mathcal{N}} \int_{E_I} \left(\mathbb{I}_t^M \frac{E_{M,R_I=(t,e)}[\mathbb{I}_t^M \xi]}{E_{M,R_I=(t,e)}[\mathbb{I}_t^M]} - \mathbb{I}_{t-}^M \frac{E_{M,R_I=(t,e)}[\mathbb{I}_{t-}^M \xi]}{E_{M,R_J=(t,e)}[\mathbb{I}_{t-}^M]}\right) \mu_I(dt \times de).$$
$$\tag{6.6}$$

Because of

$$dE[\xi|\mathcal{G}_t] = \sum_{M \in \mathcal{M}} d\left(\frac{\mathbb{I}_t^M E_M[\mathbb{I}_t^M \xi]}{E_M[\mathbb{I}_t^M]}\right),$$

summing (6.5) over $M \in \mathcal{M}$ and adding (6.6) yields for almost each $\omega \in \Omega$ that we have (1.2) and (6.2) after rearranging the addends. By applying Proposition 4.2, we can see that $G_I(u-, u, e)$ is $\mathcal{G}_{u-}$-measurable and $G_I(u, u, e)$ is $\mathcal{G}_u$-measurable for each $(I, e)$. $\qquad\square$

## 7 Infinitesimal representations for optional projections

Suppose that $X$ is a càdlàg process that satisfies (4.1) and such that $X_t - X_0$ is $\mathcal{F}_t$-measurable for each $t \geq 0$. Then the optional projection of $X$ with respect to $\mathbb{F}$ can be represented as

$$\mathrm{d}E[X_t | \mathcal{F}_t] = \mathrm{d}X_t + \sum_{I \in \mathcal{N}} \int_{E_I} F_I(t, e) \big( \mu_I(\mathrm{d}t \times \mathrm{d}e) - \lambda_I(\mathrm{d}t \times \mathrm{d}e) \big) \qquad (7.1)$$

for random mappings $F_I(t, e)$ that are $\mathcal{F}_{t-}$-measurable for each $(t, I, e)$. In order to see this, apply the classical martingale representation theorem on the $\mathbb{F}$-martingale

$$E[X_0 | \mathcal{F}_t] - E[X_0 | \mathcal{F}_0] = E[X_t | \mathcal{F}_t] - E[X_0 | \mathcal{F}_0] - (X_t - X_0), \qquad t \geq 0,$$

and rearrange the addends. The following theorem extends (7.1) to non-monotone information settings.

**Theorem 7.1** *Let $X$ be a càdlàg process that satisfies* (4.1) *and has an IB-compensator with respect to $\mathbb{G}$, denoted as $X^{IB}$. Then*

$$E[X_t | \mathcal{G}_t] - E[X_0 | \mathcal{G}_0] = X_t^{IB} + \sum_{I \in \mathcal{N}} \int_{(0,t] \times E_I} G_I(u-, u, e) (\mu_I - \nu_I) \big( \mathrm{d}(u, e) \big)$$

$$+ \sum_{I \in \mathcal{N}} \int_{(0,t] \times E_I} G_I(u, u, e) (\rho_I - \mu_I) \big( \mathrm{d}(u, e) \big) \qquad (7.2)$$

*almost surely with*

$$G_I(s, u, e) := \sum_{M \in \mathcal{M}} \mathbb{I}_s^M \left( \frac{E_{M, R_I = (u,e)}[\mathbb{I}_s^M X_{u-}]}{E_{M, R_I = (u,e)}[\mathbb{I}_s^M]} - \frac{E_M[\mathbb{I}_{u-}^M \mathbb{I}_u^M X_{u-}]}{E_M[\mathbb{I}_{u-}^M \mathbb{I}_u^M]} \right). \qquad (7.3)$$

*If $X$ has an IF-compensator with respect to $\mathbb{G}$, denoted as $X^{IF}$, then* (7.2) *still holds but with $X_t^{IB}$ replaced by $X_t^{IF}$ and $X_{u-}$ replaced by $X_u$ in* (7.3).

By applying Proposition 4.2, we can see that $G_I(u-, u, e)$ is $\mathcal{G}_u^-$-measurable and $G_I(u, u, e)$ is $\mathcal{G}_u$-measurable. Hence the integrals in the first and second line of (7.2) describe IF-martingales and IB-martingales with respect to $\mathbb{G}$ if the mappings $F_I(u, e) = G_I(u-, u, e)$ and $F_I = G_I(u, u, e)$ both satisfy the integrability condition (5.1); see the comments below Theorem 5.2.

In the special case $\mathbb{G} = \mathbb{F}$, we have $\nu_I = \lambda_I$, $\rho_I = \mu_I$, $X = X^{IB}$ and the representations (7.2) and (7.1) are equivalent, i.e., (7.2) is a generalisation of (7.1).

Even if $\mathbb{G} \neq \mathbb{F}$, we can still have $X = X^{IB}$ or $X = X^{IB}$. The following example presents non-trivial processes $X$ that equal their IB-compensators or their IF-compensators.

**Example 7.2** Let $h(M, t)(\omega) : \mathcal{M} \times [0, \infty) \times \Omega \to \mathbb{R}$ be measurable and suppose that $|h(M, t)| \leq Z$ for an integrable majorant $Z$. Let $\gamma$ be the sum of the Lebesgue

measure and a countable number of Dirac measures,

$$\gamma(B) = \lambda(B) + \sum_{i=1}^{\infty} \delta_{t_i}(B), \qquad B \in \mathcal{B}\big([0, \infty)\big),$$

for deterministic time points $0 \le t_1 < t_2 < \cdots$ that increase to infinity. Then the càdlàg process $X$ defined by

$$X_t := \sum_{M \in \mathcal{M}} \int_{[0,t]} \mathbb{I}_s^M h(M, s) \, \gamma(\mathrm{d}s)$$

has the IB-compensator

$$X_t^{IB} = \int_{(0,t]} \sum_{M \in \mathcal{M}} \mathbb{I}_s^M E[h(M, s)|\mathcal{G}_s] \, \gamma(\mathrm{d}s).$$

In order to see this, apply Proposition 4.2, the dominated convergence theorem, Proposition 4.4 and Lemma 4.3 in order to obtain that

$$\lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E[X_{t_{k+1}} - X_{t_k}|\mathcal{G}_{t_{k+1}}]$$

$$= \lim_{n \to \infty} \sum_{\mathcal{T}_n^t} \int_{(t_k, t_{k+1}]} \sum_{M \in \mathcal{M}_t} \mathbb{I}_{t_{k+1}}^M E\left[ \sum_{\tilde{M} \in \mathcal{M}} h(\tilde{M}, s) \mathbb{I}_s^{\tilde{M}} \bigg| \mathcal{G}_{t_{k+1}} \right] \gamma(\mathrm{d}s)$$

$$= \sum_{M \in \mathcal{M}_t} \lim_{n \to \infty} \sum_{\mathcal{T}_n^t} \int_{(t_k, t_{k+1}]} \mathbb{I}_{t_{k+1}}^M \frac{E_M[\sum_{\tilde{M} \in \mathcal{M}} \mathbb{I}_{t_{k+1}}^M h(\tilde{M}, s) \mathbb{I}_s^{\tilde{M}}]}{E_M[\mathbb{I}_{t_{k+1}}^M]} \gamma(\mathrm{d}s)$$

$$= \sum_{M \in \mathcal{M}_t} \int_{(0,t]} \mathbb{I}_s^M \frac{E_M[h(M, s) \mathbb{I}_s^M]}{E_M[\mathbb{I}_s^M]} \gamma(\mathrm{d}s)$$

$$= \sum_{M \in \mathcal{M}} \int_{(0,t]} \mathbb{I}_s^M E[h(M, s)|\mathcal{G}_s] \, \gamma(\mathrm{d}s)$$

almost surely, where $\mathcal{M}_t$ is defined as in (4.7). If $s \mapsto h(M, s)$ is $\mathbb{G}$-adapted for each $M$, we have $X = X^{IB}$. Likewise we can show that the càdlàg process

$$Y_t := \sum_{M \in \mathcal{M}} \int_{[0,t]} \mathbb{I}_{s-}^M h(M, s) \, \gamma(\mathrm{d}s)$$

has the IF-compensator

$$Y_t^{IF} = \int_{(0,t]} \sum_{M \in \mathcal{M}} \mathbb{I}_{s-}^M E[h(M, s)|\mathcal{G}_{s-}] \, \gamma(\mathrm{d}s).$$

If $s \mapsto h(M, s)$ is $\mathbb{G}^-$-adapted for each $M$, we have $Y = Y^{IF}$.

***Proof of Theorem 7.1*** The theorem follows from the additive decomposition

$$E[X_t|\mathcal{G}_t] - E[X_0|\mathcal{G}_0] = \lim_{n \to \infty} \sum_{\mathcal{T}_n^t} (E[X_{t_{k+1}}|\mathcal{G}_{t_{k+1}}] - E[X_{t_k}|\mathcal{G}_{t_k}])$$

$$= \lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E[X_{t_{k+1}} - X_{t_k}|\mathcal{G}_{t_{k+1}}]$$

$$+ \lim_{n \to \infty} \sum_{\mathcal{T}_n^t} (E[X_{t_k}|\mathcal{G}_{t_{k+1}}] - E[X_{t_k}|\mathcal{G}_{t_k}])$$

and from applying Theorem 6.1 for each summand $E[X_{t_k}|\mathcal{G}_{t_{k+1}}] - E[X_{t_k}|\mathcal{G}_{t_k}]$. The sum $\sum_{\mathcal{T}_n^t}(E[X_{t_k}|\mathcal{G}_{t_{k+1}}] - E[X_{t_k}|\mathcal{G}_{t_k}])$ has a representation of the form (1.2) in case of $t_k < s \le t_{k+1}$ for $G_I(s, u, e)$ defined by (7.3). Because of the càdlàg property of $X$, by applying the dominated convergence theorem pathwise for almost each $\omega \in \Omega$, we end up with (7.2) and (7.3). The alternative decomposition

$$E[X_t|\mathcal{G}_t] - E[X_0|\mathcal{G}_0] = \lim_{n \to \infty} \sum_{\mathcal{T}_n^t} E[X_{t_{k+1}} - X_{t_k}|\mathcal{G}_{t_k}]$$

$$+ \lim_{n \to \infty} \sum_{\mathcal{T}_n^t} (E[X_{t_{k+1}}|\mathcal{G}_{t_{k+1}}] - E[X_{t_{k+1}}|\mathcal{G}_{t_k}])$$

leads to the second variant with $X^{IB}$ replaced by $X^{IF}$ and $X_{u-}$ by $X_u$ in (7.3). □

***Remark 7.3*** Without loss of generality, suppose here that $0 \notin E$. Motivated by Remark 3.2, for any $t > 0$ and any integrable random variable $\xi$, define for $e \in E_I$,

$$E[\xi|\mathcal{G}_t, R_I = (t, e)] := E[\xi|(\Gamma_t, R_I) = \cdot](\Gamma_t, (t, e)),$$

$$E[\xi|\mathcal{G}_{t-}, R_I = (t, e)] := E[\xi|(\Gamma_{t-}, R_I) = \cdot](\Gamma_{t-}, (t, e)),$$

$$E[\xi|\mathcal{G}_t, J_t = 0] := E[\xi|(\Gamma_t, J_t) = \cdot](\Gamma_t, 0),$$

$$E[\xi|\mathcal{G}_{t-}, J_t = 0] := E[\xi|(\Gamma_{t-}, J_t) = \cdot](\Gamma_{t-}, 0),$$

where $J_t := \sum_{I \in \mathcal{N}} \mu_I(\{t\} \times E_I)$ indicates whether there is a stopping event at time $t$. One can then show that the integrands in (7.2) are almost surely equal to

$$G_I(t-, t, e) = E[X_{t-}|\mathcal{G}_{t-}, R_I = (t, e)] - E[X_{t-}|\mathcal{G}_{t-}, J_t = 0],$$

$$G_I(t, t, e) = E[X_{t-}|\mathcal{G}_t, R_I = (t, e)] - E[X_{t-}|\mathcal{G}_t, J_t = 0]$$

for each $t > 0$, $I \in \mathcal{N}$ and $e \in E_I$. The differences on the right-hand side have intuitive interpretations. The first line describes the difference in expectation between a change scenario and a remain scenario if we are currently at time $t-$ and are looking forward in time. Similarly, the second line describes the difference in expectation between a change scenario and a remain scenario if we are currently at time $t$ and are looking backward in time. In (7.2), these differences in expectation are integrated with respect to the compensated forward and backward scenario dynamics.

## 8 Examples

Here we come back to Examples 1.1 and 1.2 and show how our infinitesimal martingale representations can be applied in life insurance and credit risk modelling.

***Example 8.1*** Consider a life insurance contract where the insurer collects health-related information about the insured with the aim to improve forecasts of the individual future insurance liabilities. For example, this can involve data from activity trackers or social media. Here, the marked point process includes the time of death $\tau_1$, which is recorded as $\zeta_1 := \tau_1$, and further health-related information $(\tau_i, \zeta_i)_{i \geq 2}$. Upon request of the policyholder with reference to the 'right to erasure' according to the General Data Protection Regulation of the European Union, or as a self-imposed data privacy effort of the data provider, the insurer deletes parts of the health-related data at certain time points, i.e., we expand $(\tau_i, \zeta_i)_{i \geq 2}$ by deletion times $(\sigma_i)_{i \geq 2}$. For completeness, we define $\sigma_1 := \infty$.

In the classical insurance modelling without data deletion, the time dynamics of the expected future insurance payments is commonly described by Thiele's equation; see e.g. Møller [19] and Djehiche and Löfdahl [12]. Suppose that $A_t$ gives the aggregated benefit cash flow of the life insurance contract on $[0, t]$, including survival benefits with rate $a(t)$ and a death benefit of $\alpha(t)$ upon death at time $t$, i.e.,

$$A_t = \int_0^t \mathbf{1}_{\{\tau_1 > s\}} a(s) \mathrm{d}s + \int_{[0,t] \times E} \alpha(s) \, \mu_{\{1\}}\big(\mathrm{d}(s, e)\big), \qquad t \geq 0.$$

We assume here that $a : [0, \infty) \to \mathbb{R}$ and $\alpha : [0, \infty) \to \mathbb{R}$ are bounded. For a given interest intensity $\phi : [0, \infty) \to [0, \infty)$ and a finite contract horizon $T$, the process

$$X_t := \int_{(t,T]} e^{-\int_t^s \phi(u)\mathrm{d}u} \mathrm{d}A_s$$

describes the discounted future liabilities of the insurer seen from time $t$. As the càdlàg process $X = (X_t)_{t \geq 0}$ is neither adapted to $\mathbb{F}$ nor to $\mathbb{G}$, an insurer has to work with the optional projection instead (the so-called *prospective reserve*), i.e., the insurer aims to calculate

$$X_t^{\mathbb{F}} = E[X_t | \mathcal{F}_t], \qquad t \geq 0,$$

in case that there is no data deletion and

$$X_t^{\mathbb{G}} = E[X_t | \mathcal{G}_t], \qquad t \geq 0,$$

in case that information deletions may occur. The process $X^{\mathbb{G}}$ is a well-defined càdlàg process according to Theorem 4.1.

By applying (7.1) and Itô's lemma, we can derive the so-called *stochastic Thiele equation*

$$\mathrm{d}X_t^{\mathbb{F}} = -\mathrm{d}A_t + \phi(t) \, X_{t-}^{\mathbb{F}} \mathrm{d}t + \sum_I \int_{E_I} F_I(t, e) \, (\mu_I - \lambda_I)(\mathrm{d}t \times \mathrm{d}e) \qquad (8.1)$$

with terminal condition $X_T^{\mathbb{F}} = 0$; cf. Møller [19, Eq. (2.17)]. The integrand $F_I(t, e)$, which almost surely equals

$$F_I(t, e) = E[X_{t-}|\mathcal{F}_{t-}, R_I = (t, e)] - E[X_{t-}|\mathcal{F}_{t-}, J_t = 0]$$

according to Remark 7.3, is a key quantity in life insurance risk management and is known as *sum at risk*. Equation (8.1) can be interpreted as a backward stochastic differential equation (BSDE) with solution $(X^{\mathbb{F}}, (F_I)_I)$; see Djehiche and Löfdahl [12] for Markovian and Christiansen and Djehiche [4] for non-Markovian models. The BSDE (8.1) is in particular relevant if the life insurance payments $a$ and $\alpha$ depend on the current policy value so that the insurance cash flow $A$ is only implicitly defined.

By applying Theorem 7.1 and Itô's lemma and using the fact that the process $A$ equals its own IB-compensator (since $\sigma_1 = \infty$), we are able to derive an analogous equation for $X^{\mathbb{G}}$, namely

$$dX_t^{\mathbb{G}} = -dA_t + \phi(t) X_{t-}^{\mathbb{G}} dt + \sum_I \int_{E_I} G_I(t-, t, e)(\mu_I - \nu_I)(dt \times de)$$

$$+ \sum_I \int_{E_I} G_I(t, t, e)(\rho_I - \mu_I)(dt \times de) \tag{8.2}$$

with terminal condition $X_T^{\mathbb{G}} = 0$. This equation can be interpreted as a new type of BSDE with solution $(X^{\mathbb{G}}, (G_I)_I)$, featuring an IF-martingale and an IB-martingale instead of a classical martingale. The IF-martingale in the first line describes the impact of new information on the optional projection $X^{\mathbb{G}}$. The IB-martingale in the second line quantifies the effect on $X^{\mathbb{G}}$ of information deletions. The integrands $G_I(t-, t, e)$ and $G_I(t, t, e)$, which are almost surely equal to

$$G_I(t-, t, e) = E[X_{t-}|\mathcal{G}_{t-}, R_I = (t, e)] - E[X_{t-}|\mathcal{G}_{t-}, J_t = 0],$$

$$G_I(t, t, e) = E[X_{t-}|\mathcal{G}_t, R_I = (t, e)] - E[X_{t-}|\mathcal{G}_t, J_t = 0]$$

according to Remark 7.3, generalise the classical definition of the sum at risk. They are needed in life insurance risk management for sensitivity analyses, safe-side calculations, contract modifications and surplus decompositions.

If the policyholder may decide about data deletions at discretion, then the resulting value changes of the insurance contract can be systematically exploited by the policyholder, leading to a kind of data privacy arbitrage. Since it is the IB-martingale in (8.2) that measures the value changes due to data deletions at times $(\sigma_i)_{i \geq 2}$, it represents the potential data privacy arbitrage. A simple solution for avoiding data privacy arbitrage could be to charge the IB-martingale as a fee upon a data deletion request. The fee can also be negative and represents then a bonus payment. However, more complex risk sharing schemes will be needed in insurance practice that moreover distinguish between different causes for data deletions. By following the concept of Schilling et al. [26] to interpret martingale representations as risk factor decompositions, we may interpret the infinitesimal martingale parts in (8.2) as an additive surplus decomposition that can distinguish between numerous kinds of jump events $\mu_I$, $I \subseteq \mathbb{N}$, $|I| < \infty$. Such an additive decomposition of the insurer's surplus is an important step for aligning insurance risk management to the digital age.

**Example 8.2** A popular approximation concept in credit rating modelling is to pretend that the credit rating process is Markovian even if the empirical data does not fully support this assumption. Suppose that credit ratings are updated at integer times only. By setting $\tau_i := i - 1$ and $\sigma_i := i$ for $i \in \mathbb{N}$ and defining $\zeta_i$ as the credit rating at time $\tau_i$, the rating process $R = (R_t)_{t \geq 0}$ has the representation

$$R_t = \sum_{i=1}^{\infty} \zeta_i \mathbf{1}_{\{\tau_i \leq t < \sigma_i\}}, \qquad t \geq 0,$$

and satisfies

$$\mathcal{G}_t = \sigma(R_t) \vee \mathcal{Z}, \qquad t \geq 0.$$

The jumps of the process $R$ correspond to the random counting measures $\mu_I$. In the Jarrow–Lando–Turnbull model, the rating space $E$ is finite, $(R_i)_{i \in \mathbb{N}_0}$ is assumed to be a Markov chain, and

$$Q[R_{i+1} = r_{i+1} | R_i = r_i] = \pi(i, r_i) \, P[R_{i+1} = r_{i+1} | R_i = r_i]$$

for $r_i, r_{i+1} \in E$ and $i \in \mathbb{N}_0$, where $Q$ is the risk-neutral measure and $\pi$ is a deterministic function on $\mathbb{N}_0 \times E$. The latter formula allows us to estimate $Q$ from market data by a two step method. First, the transition probabilities $P[R_{i+1} = r_{i+1} | R_i = r_i]$ are estimated from observed credit rating time series. Then the function $\pi$ is calibrated such that the risk-neutral values of credit rating derivatives conform with observed market prices. Once we have $Q$, we can use the (classical) martingale representation (6.1) in order to explicitly construct hedges for financial claims $\xi$; see e.g. Last and Penrose [18]. For example, by arguing analogously to (8.1), the claim $\xi = h(R_T)$ has the martingale representation

$$h(R_T) - B(0) E_Q\left[\frac{h(R_T)}{B(T)} \bigg| \mathcal{F}_0\right] = \int_{(0,T]} B(t-) E_Q\left[\frac{h(R_T)}{B(T)} \bigg| \mathcal{F}_{t-}\right] \frac{B(\mathrm{d}t)}{B(t-)}$$
$$+ \sum_{I \in \mathcal{N}} \int_{(0,T] \times E_I} F_I(t, e)(\mu_I - \lambda_I)(\mathrm{d}t \times \mathrm{d}e).$$
$$(8.3)$$

The integral in the first line describes the investments in the risk-free asset $B$. The second line corresponds to risky investments. It can be rewritten in terms of the tradable assets in a complete financial market; cf. Last and Penrose [18, Sect. 5], which yields a trading strategy that can be used to replicate the claim $\xi$.

A standard estimator for the state occupation probabilities of the Markov chain $R$ with respect to $P$ is the Aalen–Johansen estimator, which directly corresponds to the Nelson–Aalen estimator for the compensators $\lambda_I$ of the random counting measures $\mu_I$. Under the assumption that $R$ is Markovian, the Nelson–Aalen estimator consistently estimates $\lambda_I = \nu_I$. If $R$ is not Markovian, then the Nelson–Aalen estimator still consistently estimates $\nu_I$, see Datta and Satten [9], but now $\nu_I \neq \lambda_I$. In other words, if we ignore the information beyond $\mathbb{G}$ in the estimation of $\lambda_I$ due to

an incorrect Markov assumption, then we actually estimate the infinitesimal forward compensator $\nu_I$ instead of the classical compensator $\lambda_I$. Similarly, ignoring the information beyond $\mathbb{G}$ upon estimating $F_I$ and (1.1) from market data means that we unintentionally end up with the integrands

$$G_I(t-,t,e) = \frac{B(t)}{B(T)}\big(E_Q[h(R_T)|\mathcal{G}_{t-}, R_I = (t,e)] - E_Q[h(R_T)|\mathcal{G}_{t-}, J_t = 0]\big)$$

and $B(t-)E_Q[h(R_T)/B(T)|\mathcal{G}_t]$ rather than the integrands

$$F_I(t,e) = \frac{B(t)}{B(T)}\big(E_Q[h(R_T)|\mathcal{F}_{t-}, R_I = (t,e)] - E_Q[h(R_T)|\mathcal{F}_{t-}, J_t = 0]\big)$$

and $B(t-)E_Q[h(R_T)/B(T)|\mathcal{F}_t]$. For the correct interpretation of the latter conditional expectations with mixed conditions, see Remark 7.3. All in all, by ignoring the information beyond $\mathbb{G}$ in the estimation and calculation of (8.3) due to an incorrect Markov assumption, we unintentionally end up with

$$\int_{(0,T]} B(t-)E_Q\left[\frac{h(R_T)}{B(T)}\bigg|\mathcal{G}_{t-}\right]\frac{B(\mathrm{d}t)}{B(t-)}$$
$$+ \sum_{I\in\mathcal{N}}\int_{(0,T]\times E_I} G_I(t-,t,e)(\mu_I - \nu_I)(\mathrm{d}t \times \mathrm{d}e) \qquad (8.4)$$

instead of the right-hand side of (8.3). This unintentional modification distorts the replicating trading strategy for the claim $h(R_T)$ which was included in (8.3). Do we still correctly replicate $h(R_T)$? By applying Theorem 6.1 instead of (6.1) and using that $\mathcal{F}_0 = \mathcal{G}_0$ and $\mathcal{G}_T = \sigma(R_T) \vee \mathcal{Z}$, we get analogously to (8.3) that

$$h(R_T) - B(0)E_Q\left[\frac{h(R_T)}{B(T)}\bigg|\mathcal{F}_0\right] = \int_{(0,T]} B(t-)E_Q\left[\frac{h(R_T)}{B(T)}\bigg|\mathcal{G}_{t-}\right]\frac{B(\mathrm{d}t)}{B(t-)}$$
$$+ \sum_{I\in\mathcal{N}}\int_{(0,T]\times E_I} G_I(t-,t,e)(\mu_I - \nu_I)(\mathrm{d}t \times \mathrm{d}e)$$
$$+ \sum_{I\in\mathcal{N}}\int_{(0,T]\times E_I} G_I(t,t,e)(\rho_I - \mu_I)(\mathrm{d}t \times \mathrm{d}e).$$
$$(8.5)$$

Equation (8.5) implies that the distorted trading strategy we might derive from (8.4) is actually not a hedge for $\xi = h(R_T)$. The hedging error is given by the third line in (8.5). To sum up, by estimating and calculating (8.3) under an incorrect Markov assumption for $R$, we unintentionally replace the (classical) $\mathbb{F}$-martingale in (8.3) by the $\mathbb{G}$-IF-martingale in (8.4) (the risk-free investment is also affected), and the corresponding $\mathbb{G}$-IB-martingale is just the hedging error.

Schilling et al. [26] interpret martingale representations as additive risk factor decompositions. Likewise we can read the (infinitesimal) martingale parts in (8.3) and (8.5) as linear risk factor decompositions. The relevance of such decompositions in credit risk modelling is explained in Rosen and Saunders [25].

# References

1. Bandini, E.: Existence and uniqueness for backward stochastic differential equations driven by a random measure, possibly non quasi-left continuous. Electron. Commun. Probab. **20**, 1–13 (2015)
2. Boel, R., Varaiya, P., Wong, E.: Martingales on jump processes I: representation results. SIAM J. Control **13**, 999–1021 (1975)
3. Chou, C.-S., Meyer, P.A.: Sur la répresentation des martingales comme intégrales stochastiques dans les processus ponctuels. In: Meyer, P.A. (ed.) Séminaire de Probabilités IX. Lecture Notes in Mathematics, vol. 465, pp. 226–236. Springer, Berlin (1975)
4. Christiansen, M.C., Djehiche, B.: Nonlinear reserving and multiple contract modifications in life insurance. Insur. Math. Econ. **93**, 187–195 (2020)
5. Cohen, S.N.: A martingale representation theorem for a class of jump processes. Preprint (2013). arXiv:1310.6286
6. Cohen, S.N., Elliott, R.J.: Solutions of backward stochastic differential equations on Markov chains. Commun. Stoch. Anal. **2**, 251–262 (2008)
7. Cohen, S.N., Elliott, R.J.: Comparisons for backward stochastic differential equations on Markov chains and related no-arbitrage conditions. Ann. Appl. Probab. **20**, 267–311 (2010)
8. Confortola, F.: $L^p$ solution of backward stochastic differential equations driven by a marked point process. Math. Control Signals Syst. **31**, 1–32 (2019)
9. Datta, S., Satten, G.A.: Validity of the Aalen–Johansen estimators of stage occupation probabilities and Nelson–Aalen estimators of integrated transition hazards for non-Markov models. Stat. Probab. Lett. **55**, 403–411 (2001)
10. Davis, M.H.A.: The representation of martingales of jump processes. SIAM J. Control Optim. **14**, 623–638 (1976)
11. Delong, Ł.: Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications. Springer, London (2013)
12. Djehiche, B., Löfdahl, B.: Nonlinear reserving in life insurance: aggregation and mean-field approximation. Insur. Math. Econ. **69**, 1–13 (2016)
13. Elliott, R.J.: Stochastic integrals for martingales of a jump process with partially accessible jump times. Z. Wahrscheinlichkeitstheor. Verw. Geb. **36**, 213–266 (1976)
14. Jacod, J.: Multivariate point processes: predictable projections, Radon–Nikodým derivatives, representation of martingales. Z. Wahrscheinlichkeitstheor. Verw. Geb. **31**, 235–253 (1975)
15. Jarrow, R.A., Lando, D., Turnbull, S.M.: A Markov model for the term structure of credit risk spreads. Rev. Financ. Stud. **10**, 481–523 (1997)
16. Karr, A.: Point Processes and Their Statistical Inference. Dekker, New York (1986)
17. Lando, D., Skodeberg, T.: Analyzing rating transitions and rating drift with continuous observations. J. Bank. Finance **26**, 423–444 (2002)
18. Last, G., Penrose, M.D.: Martingale representation for Poisson processes with applications to minimal variance hedging. Stoch. Process. Appl. **121**, 1588–1606 (2011)
19. Møller, C.M.: A stochastic version of Thiele's differential equation. Scand. Actuar. J. **1**, 1–16 (1993)
20. Møller, T., Steffensen, M.: Market-Valuation Methods in Life and Pension Insurance. Cambridge University Press, Cambridge (2007)
21. Norberg, R.: Hattendorff's theorem and Thiele's differential equation generalized. Scand. Actuar. J. **1992**, 2–14 (1992)

22. Norberg, R.: A theory of bonus in life insurance. Finance Stoch. **3**, 373–390 (1999)
23. Norberg, R.: The Markov chain market. ASTIN Bull. **33**, 265–287 (2003)
24. Pardoux, É., Peng, S.: Backward doubly stochastic differential equations and systems of quasilinear SPDEs. Probab. Theory Relat. Fields **98**, 209–227 (1994)
25. Rosen, D., Saunders, D.: Risk factor contributions in portfolio credit risk models. J. Bank. Finance **34**, 336–349 (2010)
26. Schilling, K., Bauer, D., Christiansen, M.C., Kling, A.: Decomposing dynamic risks into risk components. Manag. Sci. **66**, 5485–6064 (2020)
27. Tang, H., Wu, Z.: Backward stochastic differential equations with Markov chains and related asymptotic properties. Adv. Differ. Equ. **285**, 1–17 (2013)