

J Theor Probab (2011) 24: 93–117  
DOI 10.1007/s10959-010-0292-9

---

# Initial Enlargement of Filtrations and Entropy of Poisson Compensators

Stefan Ankirchner · Jakub Zwieryz

Received: 26 January 2009 / Revised: 6 March 2010 / Published online: 30 June 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** Let  $\mu$  be a Poisson random measure, let  $\mathbb{F}$  be the smallest filtration satisfying the usual conditions and containing the one generated by  $\mu$ , and let  $\mathbb{G}$  be the initial enlargement of  $\mathbb{F}$  with the  $\sigma$ -field generated by a random variable  $G$ . In this paper, we first show that the *mutual information* between the enlarging random variable  $G$  and the  $\sigma$ -algebra generated by the Poisson random measure  $\mu$  is equal to the expected *relative entropy* of the  $\mathbb{G}$ -compensator relative to the  $\mathbb{F}$ -compensator of the random measure  $\mu$ . We then use this link to gain some insight into the changes of Doob–Meyer decompositions of stochastic processes when the filtration is enlarged from  $\mathbb{F}$  to  $\mathbb{G}$ . In particular, we show that if the mutual information between  $G$  and the  $\sigma$ -algebra generated by the Poisson random measure  $\mu$  is finite, then every square-integrable  $\mathbb{F}$ -martingale is a  $\mathbb{G}$ -semimartingale that belongs to the normed space  $\mathcal{S}^1$  relative to  $\mathbb{G}$ .

**Keywords** Initial enlargement of filtration · Poisson random measure · Entropy · Mutual information · Semimartingale · Embedding

**Mathematics Subject Classification (2000)** Primary 60G44 · 60H30 · Secondary 94A17

---

The second author acknowledges the support received from the European Science Foundation (ESF) within the activity *Advanced Mathematical Methods for Finance*.

S. Ankirchner  
Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin,  
Germany  
e-mail: [ankirchn@math.hu-berlin.de](mailto:ankirchn@math.hu-berlin.de)

J. Zwieryz (✉)  
Instytut Matematyczny, Uniwersytet Wrocławski, Pl. Grunwaldzki 2/4, 50-384 Wrocław, Poland  
e-mail: [jakub.zwieryz@math.uni.wroc.pl](mailto:jakub.zwieryz@math.uni.wroc.pl)

## 1 Introduction

The perception of a random process changes if events of its future development are anticipated. Such anticipations can be modeled in terms of *enlargements* of the filtration representing the perspective from which the random process is observed. Recall that an *enlargement* of a filtration  $\mathbb{F} = (\mathcal{F}_t)$  is a filtration  $\mathbb{G} = (\mathcal{G}_t)$  containing  $\mathbb{F}$ , i.e., a filtration satisfying  $\mathcal{G}_t \supset \mathcal{F}_t$  for all  $t \geq 0$ . An introduction into the theory of enlargements of filtrations is provided in the final chapter of [15]. Other standard references are [11, 12], and [14].

From the perspective of enlarged filtrations the course of a random process is distorted. For example, if one anticipates what value a Brownian motion  $B$  will take at time 1, then  $B$  will no longer appear as a martingale but rather as a stochastic process with the dynamics of a Brownian bridge. More generally, let  $M$  be an  $\mathbb{F}$ -martingale, and  $\mathbb{G}$  an enlargement of  $\mathbb{F}$ . Then the question arises whether  $M$  remains a semimartingale with respect to  $\mathbb{G}$ , and if yes, how the semimartingale decompositions relative to  $\mathbb{G}$  can be derived.

Suppose that  $M$  is a *continuous* martingale relative to a filtration  $\mathbb{F}$  and that one can compensate  $M$  in such a way that it remains a martingale with respect to an enlargement  $\mathbb{G}$ , by what we mean that one can subtract a  $\mathbb{G}$ -predictable finite-variation (FV) process  $A$  from  $M$  so that  $M - A$  is a martingale with respect to  $\mathbb{G}$ . Without making any assumption on the *kind* of enlargement, [2] provides sufficient conditions for the FV process  $A$  to possess a density with respect to the quadratic variation process  $[M, M]$ , defined as *information drift* in [8]. For initial enlargements of the Wiener filtration, it is shown in [8] that the information drift can be interpreted as a logarithmic stochastic derivative of conditional probabilities of the enlarging information with respect to the Wiener filtration. This has been generalized to arbitrary enlargements in [2]. Moreover, it is known that there exists a link between the square norm of the information drift and the so-called *mutual information* between the enlarging information and the smaller filtration. In [16] it is shown that given a Brownian motion with a Brownian filtration initially enlarged by a countable partition, the  $L^2$ -norm of the information drift of the Brownian motion is equal to the entropy of the partition. For arbitrary initial enlargements, the link has been analysed in [3], and for noninitial enlargements, in [2]. We recall that the mutual information coincides with the square norm of the information drift of *any* martingale, as long as it satisfies the predictable representation property (PRP).

So far only a few papers deal with how to compute semimartingale decomposition of *discontinuous* martingales with respect to enlarged filtrations. Suppose that  $M$  is a stochastic integral with respect to a compensated random measure and hence a purely discontinuous martingale relative to the filtration generated by the random measure. In [9] sufficient conditions are given for  $M$  to remain a semimartingale with respect to an *initial* enlargement  $\mathbb{G}$ , and the correcting  $\mathbb{G}$ -predictable drift  $A$ , making  $M - A$  into a  $\mathbb{G}$ -martingale, is characterized in terms of the density of the  $\mathbb{G}$ -compensator relative to the  $\mathbb{F}$ -compensator of the random measure. Reference [1] is concerned with *arbitrary* enlargements on the Poisson space and interprets the correction drift density as a logarithmic variational derivative of conditional probabilities of the enlarging information. In [6] and [7] initial enlargements of the Poisson filtration are interpreted

using a Bayesian approach, where the enlarging random variable acts as a random parameter in some statistical experiment.

In the present paper we will pursue the analysis of enlargements on the Poisson space and show how semimartingale norms with respect to enlarged filtrations are linked to the entropy of the new information. To get the foot in the door we mostly restrict the analysis to *initial enlargements*, say by the  $\sigma$ -algebra generated by a random variable  $G$ . The analysis of more general enlargements is planned for future research.

As mentioned above, in the *continuous* framework, such as provided by a Brownian motion, the entropy of an initially enlarging random variable equals the square norm of the information drift associated to any martingale verifying the PRP. On the Poisson space, all martingales can be represented as stochastic integrals relative to a compensated Poisson random measure but, in general, not as integrals with respect to a single reference martingale. Consequently, the entropy of an initially enlarging random variable can only match an entropy defined on the level of random measures. In the following we will show that this is indeed the case. To this end we assume that  $\mu$  is a Poisson random measure,  $\mathbb{F}$  is the smallest extension of the filtration generated by  $\mu$  with the usual conditions, and  $\mathbb{G} = (\mathcal{G}_t)$  is the initial enlargement  $\mathcal{G}_t = \bigcap_{s>t} \sigma(G) \vee \mathcal{F}_s$ . The mutual information between the enlarging random variable  $G$  and the  $\sigma$ -algebra generated by the Poisson random measure  $\mu$  is shown to be equal to the relative entropy of the  $\mathbb{G}$ -compensator relative to the  $\mathbb{F}$ -compensator. We thus manage to establish a link between two different ways of measuring the new information contained in the enlarged filtration. In many examples the mutual information between  $G$  and  $\mathcal{F}_\infty^\mu$ , the  $\sigma$ -algebra generated by the Poisson random measure  $\mu$ , can be easily calculated. On the other hand, the density of the  $\mathbb{G}$ -compensator relative to the  $\mathbb{F}$ -compensator is the fundamental ingredient to compute for any  $\mathbb{F}$ -martingale  $M$  its  $\mathbb{G}$ -predictable correcting density  $\alpha$  that turns  $M - \int_0^t \alpha_s ds$  into a  $\mathbb{G}$ -martingale. The link allows us to estimate the norm of the correction drift  $\mathbb{E} \int_0^\infty |\alpha_s| ds$  against the mutual information between  $G$  and  $\mathcal{F}_\infty^\mu$ , and the  $L^2$ -norm of the martingale  $M$ .

As a further consequence, we obtain that if the mutual information between  $\sigma(G)$  and  $\mathcal{F}_\infty^\mu$  is finite, then every  $\mathbb{F}$ -semimartingale is also a semimartingale with respect to  $\mathbb{G}$ . This conclusion can already be drawn from results of the seminal paper [9]. Indeed, if the mutual information between  $G$  and  $\mathcal{F}_\infty^\mu$  is finite, then a condition, referred to as *Jacod's condition* (see (5.2)) is satisfied, and consequently the semimartingale property is always preserved when passing from  $\mathbb{F}$  to  $\mathbb{G}$  (see the paragraph after Theorem 5.11 for details). However, assuming finiteness of the mutual information between  $G$  and  $\mathcal{F}_\infty^\mu$  allows us to go a step further and to prove that the set of square-integrable  $\mathbb{F}$ -martingales, denoted by  $\mathcal{H}^2(\mathbb{F})$ , can be continuously embedded into the space of  $\mathbb{G}$ -semimartingales with finite  $\mathcal{S}^1$  norm.<sup>1</sup>

The paper is organized as follows. In Sect. 2 we specify the basic setup and collect some definitions of notions used thereafter. In particular, we recall some properties of Poisson random measures and their compensators, which will be a central object we focus on. In Sect. 3 we investigate the relation of Radon–Nikodym densities of compensators and information drifts for enlarged filtrations. Section 3 deals mainly

<sup>1</sup>For a definition of the spaces  $\mathcal{H}^p$  and  $\mathcal{S}^p$ , see Definitions 5.3 and 5.6 or [15], p. 154.

with arbitrary enlargements, but starting from Sect. 4, we restrict the analysis to *initial* enlargements. A key theorem (Theorem 4.7) is shown, which links the expected relative entropy between compensators, with the mutual information between the enlarging  $\sigma$ -algebra and the original filtration. This allows us, in Sect. 5, to prove the inclusion  $\mathcal{H}^2(\mathbb{F}) \subset \mathcal{S}^1(\mathbb{G})$  if the mutual information is finite. We distinguish between finite and infinite time horizons and between finite and infinite jump measures  $\nu$ . We conclude with some examples, showing the usage of our results.

## 2 Preliminaries

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space, and  $\nu$  a  $\sigma$ -finite measure on a standard Borel space  $(E, \mathcal{E})$ . Denote by  $\lambda$  the Lebesgue measure, and let  $\mu$  be a homogeneous Poisson random measure on  $(\mathbb{R}_+ \times E, \lambda \otimes \nu)$ . By this we mean, in accordance with Definition 1.20 in Chap. II of [10], that  $\mu$  is an integer-valued random measure such that

- for all  $A \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$  with  $\lambda \otimes \nu(A) < \infty$ , the random variable  $\mu(\omega; A)$  is Poisson distributed with intensity  $\lambda \otimes \nu(A)$ ,
- for all  $A$  and  $B \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$  with  $A \cap B = \emptyset$ , the random variables  $\mu(\omega; A)$  and  $\mu(\omega; B)$  are independent.

Throughout by  $\mathbb{F}^0 = (\mathcal{F}_t^0)_{t \geq 0}$  we denote the filtration generated by the homogeneous Poisson random measure  $\mu$ , i.e.,

$$\mathcal{F}_t^0 = \sigma(\{\mu(\omega; [0, s] \times U) \leq c\} | c \in \mathbb{R}, s \leq t, U \in \mathcal{E}).$$

The smallest filtration satisfying the *usual conditions*<sup>2</sup> and containing  $\mathbb{F}^0$  will be denoted by  $\mathbb{F}$ .

In our study of the semimartingale spaces with respect to  $\mathbb{F}$  and their enlargements, the notion of a *compensator* of  $\mu$  will turn out to be of crucial importance. For the convenience of the reader, we next recall its definition.

**Lemma and Definition 2.1** *Let  $\mathbb{H}$  be a filtration containing  $\mathbb{F}$ . We denote by  $\mathcal{P}(\mathbb{H})$  the predictable  $\sigma$ -field on  $\Omega \times \mathbb{R}_+$  associated with  $\mathbb{H}$  and, for any  $t \geq 0$ , by  $\mathcal{P}_t(\mathbb{H})$  its restriction to  $\Omega \times [0, t]$ . Then, there exists a unique (up to a  $\mathbb{P}$ -null set) predictable random measure  $\pi^{\mathbb{H}}$ , called the compensator of  $\mu$  relative to  $\mathbb{H}$ , which satisfies the condition*

$$\mathbb{E} \left( \int_0^\infty \int_E \phi(s, z) \mu(dz, ds) \right) = \mathbb{E} \left( \int_0^\infty \int_E \phi(s, z) \pi^{\mathbb{H}}(dz, ds) \right) \tag{2.1}$$

for every nonnegative  $\mathcal{P}(\mathbb{H}) \otimes \mathcal{E}$ -measurable function  $\phi$ .

---

<sup>2</sup>Recall that a filtration  $\mathbb{F}$  is said to satisfy the usual conditions if

- (i)  $\mathcal{F}_0$  contains all  $\mathbb{P}$ -null sets of  $\mathcal{F}$ ,
- (ii)  $\mathbb{F}$  is right continuous, i.e.,  $\mathcal{F}_t = \bigcap_{u > t} \mathcal{F}_u$  for all  $0 \leq t < \infty$ .

It is straightforward to show that the compensator of  $\mu$  with respect to the filtration  $\mathbb{F}$  is the deterministic measure

$$\pi^{\mathbb{F}} = \lambda \otimes \nu.$$

In the following we will often use the abbreviation  $\mu^{\mathbb{H}} := \mu - \pi^{\mathbb{H}}$  for the  $\mathbb{H}$ -compensated measure. Moreover, we will use the following notation.

**Definition 2.2** Let  $\eta$  be a random measure. Then we denote by  $\eta \times \mathbb{P}$  the measure on  $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$  defined by

$$(\eta \times \mathbb{P})(A) = \mathbb{E} \iint 1_A(\omega, s, z) \eta(\omega, dz, ds), \quad A \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}.$$

Moreover, given a measurable process  $\psi$ , we will sometimes write  $(\psi * \eta)_t = \int_0^t \int_E \psi(s, z) \eta(dz, ds)$  if the integral is defined either in the Lebesgue or stochastic sense.

In the case  $\nu(E) = \infty$ , it will be sometimes necessary to approximate  $E$  with measurable subsets of finite measure. Since  $\nu$  is  $\sigma$ -finite, there exist  $E_n \in \mathcal{E}$ ,  $n = 1, 2, \dots$ , such that  $E_n \uparrow E$  and  $\nu(E_n) < \infty$ . We define  $\mathbb{F}^n = (\mathcal{F}_t^n)_{t \geq 0}$ ,  $n = 1, 2, \dots$ , as the smallest filtration satisfying the usual conditions and containing the filtration  $\mathbb{F}^{0,n} = (\mathcal{F}_t^{0,n})_{t \geq 0}$ , where

$$\mathcal{F}_t^{0,n} = \sigma(\{\mu(\omega; [0, s] \times A) \leq c \mid c \in \mathbb{R}, s \leq t, A \in \mathcal{E} \cap E_n\}). \tag{2.2}$$

It is straightforward to see that  $\bigvee_{n \geq 1} \mathcal{F}_t^{0,n} = \mathcal{F}_t^0$  and that the  $\mathbb{F}$ -compensators of the Poisson random measures

$$\mu_n(\cdot) := \mu(\cdot \cap \mathbb{R}_+ \times E_n)$$

have the following form:

$$\pi^{\mathbb{F},n}(\cdot) = \pi^{\mathbb{F}}(\cdot \cap \mathbb{R}_+ \times E_n). \tag{2.3}$$

We will adopt the analogous notation of approximating filtrations and compensators for enlarged filtrations  $\mathbb{G}$ .

Finally, we recall the well-known fact that the compensated Poisson random measure  $\mu^{\mathbb{F}}$  satisfies the *predictable representation property (PRP)* with respect to the filtration  $\mathbb{F}$ , i.e., for any locally square-integrable  $(\mathbb{F}, \mathbb{P})$ -martingale  $M = (M_t)_{t \geq 0}$ , there exists a  $\mathcal{P}(\mathbb{F}) \otimes \mathcal{E}$ -measurable process  $\psi \in L^2(\pi \times \mathbb{P})$ , such that

$$M_t = M_0 + \int_0^t \int_E \psi(s, z) \mu^{\mathbb{F}}(dz, ds) = M_0 + (\psi * \mu^{\mathbb{F}})_t.$$

### 3 Information Drifts for Pure Jump Martingales

In this section we address the question under which conditions some given  $(\mathbb{F}, \mathbb{P})$ -local martingale remains a semimartingale with respect to an arbitrary enlarged filtration, without making any assumption on the *kind* of the enlargement. Throughout let  $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$  be an enlargement of the filtration  $\mathbb{F}$  that satisfies the usual conditions.

During the whole section, we assume that

$$\pi^{\mathbb{G}} \times \mathbb{P} \ll \pi^{\mathbb{F}} \times \mathbb{P} \quad \text{on } \mathcal{P}(\mathbb{G}) \otimes \mathcal{E}. \tag{3.1}$$

In accordance with the literature (see [8]), we will use the following notion of *information drift*:

**Definition 3.1** Let  $X = (X_t)_{t \geq 0}$  be an  $(\mathbb{F}, \mathbb{P})$ -local martingale, and let  $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$  be a filtration containing  $\mathbb{F}$ . An *information drift* of  $X$  with respect to  $\mathbb{G}$  is a  $\mathbb{G}$ -predictable process  $\alpha$  such that the Lebesgue integral  $\int_0^t \alpha_s ds$  is a.s. defined for all  $t \geq 0$  and

$$X_t - \int_0^t \alpha_s ds$$

is a  $(\mathbb{G}, \mathbb{P})$ -local martingale.

We will now define the process which will turn out to be the crucial object for determining the information drifts with respect to enlarged filtrations.

**Definition 3.2** Let  $\pi^{\mathbb{G}} \times \mathbb{P}$  be absolutely continuous with respect to  $\pi^{\mathbb{F}} \times \mathbb{P}$  on  $\mathcal{P}(\mathbb{G}) \otimes \mathcal{E}$ , and let  $\frac{d(\pi^{\mathbb{G}} \times \mathbb{P})}{d(\pi^{\mathbb{F}} \times \mathbb{P})} \Big|_{\mathcal{P}(\mathbb{G}) \otimes \mathcal{E}}$  be the Radon–Nikodym density. Throughout by  $\delta$  we denote the process given by

$$\delta(\omega, s, z) = \frac{d(\pi^{\mathbb{G}} \times \mathbb{P})}{d(\pi^{\mathbb{F}} \times \mathbb{P})} \Big|_{\mathcal{P}(\mathbb{G}) \otimes \mathcal{E}} - 1. \tag{3.2}$$

The following lemma gives a first result concerning the form of the information drift with respect to the filtration  $\mathbb{G}$ .

**Lemma 3.3** Let  $\psi \in L^2(\pi^{\mathbb{F}} \times \mathbb{P}) \cap L^1(\pi^{\mathbb{F}} \times \mathbb{P})$ , and  $X$  be the square-integrable martingale defined by  $X_t = \int_0^t \int_E \psi(s, z) \mu^{\mathbb{F}}(dz, ds)$ . Then  $\psi \delta \in L^1(\pi^{\mathbb{F}} \times \mathbb{P})$ , and the process defined by

$$M_t = X_t - \int_0^t \int_E \psi(s, z) \delta(s, z) \pi^{\mathbb{F}}(dz, ds)$$

is a  $(\mathbb{G}, \mathbb{P})$ -martingale.

*Proof* Let  $0 \leq s \leq t$  and  $B \in \mathcal{G}_s$ . Notice that the  $\mathbb{G}$ -predictable process  $1_B 1_{[s,t]} \psi$  belongs to  $L^1(\mu \times \mathbb{P})$  and hence also to  $L^1(\pi^{\mathbb{G}} \times \mathbb{P})$  and that

$$\mathbb{E} \left[ 1_B \int_{s+}^t \int_E \psi(s, z) \mu(dz, ds) \right] = \mathbb{E} \left[ 1_B \int_{s+}^t \int_E \psi(s, z) \pi^{\mathbb{G}}(dz, ds) \right]$$

(see (2.1)). As a consequence, we may write

$$\begin{aligned} \mathbb{E}[1_B(X_t - X_s)] &= \mathbb{E} \left[ 1_B \int_s^t \int_E \psi(s, z) (\pi^{\mathbb{G}} - \pi^{\mathbb{F}})(dz, ds) \right] \\ &= \mathbb{E} \left[ 1_B \int_s^t \int_E \psi(s, z) \delta(s, z) \pi^{\mathbb{F}}(dz, ds) \right], \end{aligned}$$

which implies the result. □

If we drop the assumption of Lemma 3.3 that  $\psi \in L^1(\pi^{\mathbb{F}} \times \mathbb{P})$ , then  $\psi \delta$  is not necessarily integrable with respect to  $\pi^{\mathbb{F}} \times \mathbb{P}$ , and hence the Lebesgue integral  $\int_0^t \int_E \psi(s, z) \delta(s, z) \pi^{\mathbb{F}}(dz, ds)$  may not be defined. However, one can subtract from any square-integrable  $\mathbb{F}$ -martingale  $M$ , a continuous square-integrable and  $\mathbb{G}$ -predictable process so that  $M$  remains a martingale with respect to the enlarged filtration  $\mathbb{G}$ .

**Lemma 3.4** *Let  $\psi \in L^2(\pi^{\mathbb{F}} \times \mathbb{P})$  be  $\mathbb{F}$ -predictable and put  $M_t = (\psi * \mu^{\mathbb{F}})_t$ . Then there exists a continuous square-integrable and  $\mathbb{G}$ -predictable process  $A$  such that  $M_t - A_t$  is a  $\mathbb{G}$ -martingale.*

*Proof* Let  $\gamma(\omega, s, z) = \frac{d(\pi^{\mathbb{G}} \times \mathbb{P})}{d(\pi^{\mathbb{F}} \times \mathbb{P})} |_{\mathcal{P}(\mathbb{G}) \otimes \mathcal{E}}(\omega, s, z)$ , and let  $E_n$  be an increasing sequence of measurable sets in  $E$  such that  $\bigcup_n E_n = E$  and  $\nu(E_n) < \infty$  for all  $n \geq 1$ .

We define the integrands  $\psi_n(\omega, s, z) = 1_{E_n}(z) \psi(\omega, s, z)$  and set  $M_t^n = (\psi^n * \mu^{\mathbb{F}})_t$ . Note that  $\psi^n(s, z) 1_{[0,t]}(s) \in L^1(\pi^{\mathbb{F}} \times \mathbb{P})$ , and since it is  $\mathbb{F}$ -predictable, it is also in  $L^1(\pi^{\mathbb{G}} \times \mathbb{P})$ . Therefore, we obtain from the equality  $\psi^n \delta = \psi^n \gamma - \psi^n$  that  $\psi^n(s, z) \delta(s, z) 1_{[0,t]}(s) \in L^1(\pi^{\mathbb{F}} \times \mathbb{P})$ . Thus, we may define the Lebesgue integrals

$$A_t^n = \int_0^t \int_E \psi^n(s, z) \delta(s, z) \pi^{\mathbb{F}}(ds, dz).$$

By Lemma 3.3 we know that  $\tilde{M}_t^n = M_t^n - A_t^n$  is a martingale with respect to the filtration  $\mathbb{G}$ .

Notice that for any pair  $n, m \in \mathbb{N}$ ,

$$\mathbb{E} \sup_{t \in [0, T]} (A_t^n - A_t^m)^2 \leq 2 \mathbb{E} \left[ \sup_{t \in [0, T]} (M_t^n - M_t^m)^2 + \sup_{t \in [0, T]} (\tilde{M}_t^n - \tilde{M}_t^m)^2 \right],$$

and hence the Burkholder–Davis–Gundy inequality implies that for some constant  $C \in \mathbb{R}_+$ ,

$$\mathbb{E} \sup_{t \in [0, T]} (A_t^n - A_t^m)^2 \leq C \mathbb{E} \int_{0+}^T \int_E (\psi^n - \psi^m)^2 \pi^{\mathbb{F}}(ds, dz). \tag{3.3}$$

The right hand side of (3.3) converges to 0 as  $n, m \rightarrow \infty$ . Consequently, the sequence  $A^n$  is a Cauchy sequence converging to a continuous  $\mathbb{G}$ -predictable process  $A$  satisfying  $\mathbb{E} \sup_{t \in [0, T]} A_t^2 < \infty$ . Moreover, for  $0 \leq s \leq t$  and  $B \in \mathcal{G}_s$ , we have

$$\begin{aligned} \mathbb{E}[1_B(M_t - M_s)] &= \lim_n \mathbb{E}[1_B(M_t^n - M_s^n)] = \lim_n \mathbb{E}[1_B(A_t^n - A_s^n)] \\ &= \mathbb{E}[1_B(A_t - A_s)], \end{aligned}$$

which proves that  $M_t - A_t$  is a  $\mathbb{G}$ -martingale. □

If we assume that  $\psi\delta \in L^1(\pi^{\mathbb{F}} \times \mathbb{P})$ , then Lemma 3.4 immediately implies the following result that provides our first general criterion for square integrable martingales to possess information drifts with respect to enlarged filtrations.

**Theorem 3.5** *Let  $\psi \in L^2(\pi^{\mathbb{F}} \times \mathbb{P})$  be  $\mathbb{F}$ -predictable and put  $M_t = (\psi * \mu^{\mathbb{F}})_t$ . If  $\psi\delta \in L^1(\pi^{\mathbb{F}} \times \mathbb{P})$ , then  $M$  is a  $(\mathbb{G}, \mathbb{P})$ -semimartingale, and the process  $\int_E \psi(t, z) \times \delta(t, z)v(dz)$  is its information drift with respect to  $\mathbb{G}$ .*

In general it may be very difficult to verify directly whether  $\psi\delta \in L^1(\pi^{\mathbb{F}} \times \mathbb{P})$ . In case of initial enlargements we will show that there is a very simple criterion that allows one to deduce whether  $\psi\delta \in L^1(\pi^{\mathbb{F}} \times \mathbb{P})$  or not. More precisely, if  $\mathbb{G}$  is an initial enlargement of  $\mathbb{F}$  by a random variable  $G$ , and the mutual information (see Definition 4.2) between  $\sigma(G)$  and  $\mathcal{F}_T$  is finite, then for all  $\psi \in L^2(\pi^{\mathbb{F}} \times \mathbb{P})$ , we have that  $\psi\delta \in L^1(\pi^{\mathbb{F}} \times \mathbb{P})$ . Before we can show this, we need some further auxiliary results which are gathered in the remaining of this section.

Notice that  $\delta$  is not necessarily integrable with respect to  $\pi^{\mathbb{F}} \times \mathbb{P}$ . However, we have the following property of the process  $\delta$ , stated as a lemma since it will be needed later.

**Lemma 3.6** *Assume that the jump measure  $\nu$  satisfies  $\nu(E) < \infty$ . Then, for all  $t \geq 0$ , the process  $(\delta * \pi^{\mathbb{F}})_t$  is integrable, and*

$$\mathbb{E} \left( \int_0^t \int_E \delta_s(z, \omega, \omega) \pi^{\mathbb{F}}(dz, ds) \right) = 0.$$

*Proof* Let  $M_t := \mu^{\mathbb{F}}(\omega; [0, t] \times E)$ . Since  $\nu(E) < \infty$ , we have that  $M$  is square integrable. We also know that

$$\tilde{M}_t = M_t - \int_0^t \int_E \delta_s(z, \omega, \omega) \pi^{\mathbb{F}}(dz, ds)$$

is a  $(\mathbb{G}, \mathbb{P})$ -martingale. Thus,

$$\int_0^t \int_E \delta_s(z, \omega, \omega) \pi^{\mathbb{F}}(dz, ds) \in L^1(\mathbb{P}),$$

so that, finally,

$$0 = \mathbb{E}(M_t) = \mathbb{E} \left( \int_0^t \int_E \delta_s(z, \omega, \omega) \pi^{\mathbb{F}}(dz, ds) \right). \quad \square$$



We close this section by recalling that the process  $\delta$ , defined in (3.2), can be interpreted as a logarithmic variational derivative of conditional probabilities of  $\mathbb{P}$ . This alternative representation has been employed in [1] as a mean to calculate explicitly information drifts for examples of enlargements.

If we write  $P_t$  for the regular conditional probability relative to  $\mathcal{F}_t^0$ , then for any  $A \in \mathcal{F}_T$ , the process  $(t, \omega) \rightarrow P_t(\omega, A)$  is an  $\mathbb{F}^0$ -martingale, and by the representation property quoted above, there exists a  $\mathcal{P}(\mathbb{F}) \otimes \mathcal{E}$ -measurable process  $k_s(\omega, z, A)$  such that for all  $t \geq 0$ , we have  $P$ -a.s.

$$P_t(\cdot, A) = P(A) + \int_0^t \int_E k_s(\cdot, z, A) \mu^{\mathbb{F}}(dz, ds). \tag{3.4}$$

We first consider the case where  $\mathbb{G}$  is an initial enlargement of  $\mathbb{F}$  by a finite partition. More precisely, let  $\mathcal{P} = \{A_1, \dots, A_n\}$  be a finite partition of  $\Omega$  into measurable sets, and let

$$\mathcal{G}_t = \bigcap_{s>t} \mathcal{F}_s \vee \sigma(\mathcal{P}). \tag{3.5}$$

Then one has the following simple representation of  $\delta$ .

**Lemma 3.7** *The process  $\delta$  defined in (3.2) satisfies,  $\pi^{\mathbb{F} \times \mathbb{P}}$ -a.s.,*

$$\delta(\omega, s, z) = \sum_{i=1}^n 1_{A_i} \frac{k(s, z, A_i)}{\mathbb{P}_{s-}(A_i)}.$$

*Proof* Let  $\psi$  be  $\mathcal{P}(\mathbb{F}) \otimes \mathcal{E}$ -measurable such that  $\mathbb{E}[ (|\psi| + |\psi|^2) * \mu ] < \infty$ . Then for any  $i \in \{1, \dots, n\}$ , we have

$$\begin{aligned} & \mathbb{E}[(1_{A_i} \psi * \mu^{\mathbb{F}})] \\ &= \mathbb{E} \left[ \left( \mathbb{P}(A_i) + \int_0^\infty \int_E k(s, z, A_i) \mu^{\mathbb{F}}(ds, dz) \right) \int_0^\infty \int_E \psi \mu^{\mathbb{F}}(ds, dz) \right] \\ &= \mathbb{E} \left[ \int_0^\infty \int_E k(s, z, A_i) \psi(s, z) \pi^{\mathbb{F}}(ds, dz) \right] \\ &= \mathbb{E} \left[ \int_0^\infty \int_E \int 1_{A_i}(\omega') \frac{k(s, z, A_i)}{\mathbb{P}_s(A_i)} \mathbb{P}_s(\omega, d\omega') \psi(s, z) \pi^{\mathbb{F}}(ds, dz) \right]. \end{aligned}$$

Fubini’s theorem further yields

$$\mathbb{E}[(1_{A_i} \psi * \mu^{\mathbb{F}})] = \mathbb{E} \left[ 1_{A_i} \int_0^\infty \int_E \frac{k(s, z, A_i)}{\mathbb{P}_s(A_i)} \psi(s, z) \pi^{\mathbb{F}}(ds, dz) \right],$$

which implies

$$\mathbb{E}[(1_{A_i} \psi * \mu)] = \mathbb{E} \left[ 1_{A_i} \psi(s, z) \left( 1 + \sum_{i=1}^n 1_{A_i} \frac{k(s, z, A_i)}{\mathbb{P}_{s-}(A_i)} \right) * \pi^{\mathbb{F}}(ds, dz) \right].$$

Finally a monotone class argument shows the result. □

Now suppose again that  $\mathbb{G}$  is an *arbitrary* enlargement. Moreover assume that there exists a countably generated filtration  $\mathbb{G}^0 = (\mathcal{G}_t^0)_{t \geq 0}$  such that  $\mathbb{G}$  is the smallest filtration containing  $\mathbb{G}^0$  and satisfying the usual conditions. The next lemma, a generalization of Lemma 3.7, summarizes some results shown in [1].

**Lemma 3.8** *If  $k_t(\omega, z, \cdot)$  is a signed measure on  $\mathcal{G}_{t-}^0$ , such that  $k_t(\omega, z, \cdot)|_{\mathcal{G}_{t-}^0} \ll P_t(\omega, \cdot)|_{\mathcal{G}_{t-}^0}$ , ( $\pi^{\mathbb{F}} \times \mathbb{P}$ )-a.s., then there exists an  $\mathbb{F} \otimes \mathbb{G}$ -predictable process  $\beta$  such that*

$$\beta(\omega, \omega', t, z) = \frac{k_t(\omega, z, d\omega')}{P_t(\omega, d\omega')} \Big|_{\mathcal{G}_{t-}^0}$$

for  $\pi^{\mathbb{F}} \times \mathbb{P}$ -a.a.  $(\omega, t, z)$ . Moreover, the process  $\delta$  defined in (3.2) is equal to the diagonal of  $\beta$ , i.e., we have,  $\pi^{\mathbb{F}} \times \mathbb{P}$ -a.s.,

$$\beta(\omega, \omega, t, z) = \delta(\omega, t, z).$$

*Proof* See Lemma 2.5 and Theorem 2.6 in [1]. □

#### 4 Entropy of Compensators Relative to Initially Enlarged Filtrations

In this section we will have a closer look at *initial* enlargements. Throughout let  $G$  be a random variable and assume that  $\mathbb{G} = (\mathcal{G}_t)$  is obtained by enlarging  $\mathbb{F}$  initially with  $\sigma(G)$ , i.e., that  $\mathcal{G}_t = \bigcap_{s > t} (\mathcal{F}_s \vee \sigma(G))$ . We will see that in this case the relative entropy of the compensator  $\pi^{\mathbb{G}}$  with respect to the compensator  $\pi^{\mathbb{F}}$  coincides with the *mutual information* between the enlarging random variable  $G$  and the  $\sigma$ -field generated by the Poisson random measure  $\mu$ . Thus, two different ways of quantifying the new relative information contained in the enlarging random variable  $G$  are shown to provide the same result.

We start by recalling some definitions and basic results that will be used later. Apart from that, we will use the same notation as in the previous sections.

**Definition 4.1** Let  $\mathbb{Q}^1$  and  $\mathbb{Q}^2$  be finite measures on a measure space  $(A, \mathcal{A})$ , and let  $\mathcal{B}$  be a sub- $\sigma$ -field of  $\mathcal{A}$ . The relative entropy of  $\mathbb{Q}^1$  with respect to  $\mathbb{Q}^2$  on the  $\sigma$ -field  $\mathcal{B}$  is defined by

$$\mathcal{H}_{\mathcal{B}}(\mathbb{Q}^1 \parallel \mathbb{Q}^2) = \begin{cases} \int \log \left( \frac{d\mathbb{Q}^1}{d\mathbb{Q}^2} \Big|_{\mathcal{B}} \right) d\mathbb{Q}^1 & \text{if } \mathbb{Q}^1 \ll \mathbb{Q}^2 \text{ on } \mathcal{B}, \\ \infty & \text{else.} \end{cases}$$

**Definition 4.2** Let  $\mathcal{A}, \mathcal{B}$  be two sub- $\sigma$ -algebras of  $\mathcal{F}$ . Suppose that there exists a regular conditional probability of  $\mathbb{P}$  with respect to  $\mathcal{B}$ , denoted by  $\mathbb{P}(\cdot | \mathcal{B})$ . Then, the *mutual information* between  $\mathcal{A}$  and  $\mathcal{B}$  is defined as

$$I(\mathcal{A} \parallel \mathcal{B}) = \mathbb{E} \mathcal{H}_{\mathcal{A}}(\mathbb{P}(\cdot | \mathcal{B}) \parallel \mathbb{P}).$$

Consequently, if  $\mathbb{P}(\cdot|\mathcal{B}) \ll \mathbb{P}$  on  $\mathcal{A}$  a.s., then

$$I(\mathcal{A}||\mathcal{B}) = \int \int \log \left( \frac{\mathbb{P}(d\omega'|\mathcal{B})(\omega)}{\mathbb{P}(d\omega')} \Big|_{\mathcal{A}} \right) \mathbb{P}(d\omega'|\mathcal{B}) \mathbb{P}(d\omega).$$

**Lemma 4.3** *Let  $G, H$  be two random variables, and denote by  $\mathbb{P}_G$  and  $\mathbb{P}_H$  their distributions. Moreover let  $\mathbb{P}_{G,H}$  be the joint distribution of  $G$  and  $H$ . Setting  $\mathcal{A} = \sigma(G)$  and  $\mathcal{B} = \sigma(H)$ , we have*

$$I(\mathcal{A}||\mathcal{B}) = \mathcal{H}(\mathbb{P}_{G,H} || \mathbb{P}_G \otimes \mathbb{P}_H).$$

*Proof* The statement follows from the fact that

$$\frac{\mathbb{P}(H \in dx|G = y)}{\mathbb{P}(H \in dx)} = \frac{\mathbb{P}(H \in dx, G \in dy)}{\mathbb{P}(H \in dx) \otimes \mathbb{P}(G \in dy)}. \quad \square$$

It is well known that the relative entropy satisfies the following monotonicity property.

**Lemma 4.4** (See, e.g., Theorem 1.30 in [13]) *Let  $\mathbb{Q}^1$  and  $\mathbb{Q}^2$  be two probability measures defined on a measure space  $(A, \mathcal{A})$ , and let  $(\mathcal{A}_n)$  be a sequence of increasing sub- $\sigma$ -fields such that  $\mathcal{A} = \bigvee_n \mathcal{A}_n$ . Then  $(\mathcal{H}_{\mathcal{A}_n}(\mathbb{Q}^1 || \mathbb{Q}^2))$  is an increasing sequence, and*

$$\lim_n \mathcal{H}_{\mathcal{A}_n}(\mathbb{Q}^1 || \mathbb{Q}^2) = \mathcal{H}_{\mathcal{A}}(\mathbb{Q}^1 || \mathbb{Q}^2).$$

Lemma 4.4 immediately yields that the mutual information between two  $\sigma$ -algebras can be approximated by mutual information between finite sub- $\sigma$ -algebras.

**Lemma 4.5** *Let  $\Delta = \{A_1, \dots, A_n\}$  be a partition of  $\Omega$  into measurable sets. Then, for all  $t \geq 0$ ,*

$$I(\sigma(\Delta) || \mathcal{F}_t) = \sum_{i=1}^n \mathbb{E}[1_{A_i} (\log(\mathbb{P}_t(\cdot, A_i)) - \log(\mathbb{P}(A_i)))].$$

Moreover, if  $\tau_m = t \wedge \inf\{s \geq 0 : \mathbb{P}_s(\cdot, A_i) \leq \frac{1}{m} \text{ for one } i\}$ , then

$$I(\sigma(\Delta) || \mathcal{F}_t) = \sup_m I(\sigma(\Delta) || \mathcal{F}_{\tau_m}).$$

*Proof* We may assume that  $\mathbb{P}(A_i) > 0$  for all  $i \in \{1, \dots, n\}$ . Observe that

$$\begin{aligned} I(\sigma(\Delta) || \mathcal{F}_{\tau_m}) &= \int_{\Omega} \log \left( \frac{\mathbb{P}(d\omega'|\mathcal{F}_{\tau_m})}{\mathbb{P}(d\omega')} \Big|_{\sigma(\Delta)} \right) \mathbb{P}(d\omega'|\mathcal{F}_{\tau_m}) d\mathbb{P}(\omega) \\ &= \int_{\Omega} \sum_{i=1}^n 1_{A_i}(\omega') \log \left( \frac{\mathbb{P}(A_i|\mathcal{F}_{\tau_m})(\omega)}{\mathbb{P}(A_i)} \right) \mathbb{P}(d\omega'|\mathcal{F}_{\tau_m}) d\mathbb{P}(\omega) \\ &= \sum_{i=1}^n \mathbb{E}[\mathbb{P}(A_i|\mathcal{F}_{\tau_m}) \log \mathbb{P}(A_i|\mathcal{F}_{\tau_m}) - \mathbb{P}(A_i) \log \mathbb{P}(A_i)]. \end{aligned}$$

Since  $\{\mathbb{P}(A_i|\mathcal{F}_{\tau_m}) = 0\} \subset \{\mathbb{P}(A_i|\mathcal{F}_t) = 0\}$  a.s., we have  $\lim_m \mathbb{P}(A_i|\mathcal{F}_{\tau_m}) = \mathbb{P}(A_i|\mathcal{F}_t)$  a.s. Consequently, we obtain  $\lim_m I(\sigma(\Delta)\|\mathcal{F}_{\tau_m}) = I(\sigma(\Delta)\|\mathcal{F}_t)$ , and hence the result.  $\square$

The next lemma guarantees that we may also approximate the density of  $\pi^{\mathbb{G}} \times \mathbb{P}$  relative to  $\pi^{\mathbb{F}} \times \mathbb{P}$  by using finite partitions.

**Lemma 4.6** *Let  $v(E) < \infty$  and  $T \in \mathbb{R}_+$ . Let  $\mathcal{P}_n$  be a sequence of finite partitions of  $\Omega$  such that  $\mathcal{P}_n \subset \mathcal{P}_{n+1}$  and  $\mathcal{A} = \bigvee_{n \geq 1} \sigma(\mathcal{P}_n)$ . Let  $\mathcal{G}_t^n = \bigcap_{s > t} (\mathcal{F}_t \vee \sigma(\mathcal{P}_n))$  and set  $\mathcal{G}_t = \bigcap_{s > t} (\mathcal{F}_t \vee \mathcal{A})$ . If  $\mathcal{H}_{\mathcal{P}_T(\mathbb{G}) \otimes \mathcal{E}}(\pi^{\mathbb{G}} \times \mathbb{P} \|\pi^{\mathbb{F}} \times \mathbb{P}) < \infty$ , then the densities of  $\pi^{\mathbb{G}_n} \times \mathbb{P}$  relative to  $\pi^{\mathbb{F}} \times \mathbb{P}$  converge along a subsequence weakly in  $L^1(\pi^{\mathbb{F}} \times \mathbb{P})$  to the density of  $\pi^{\mathbb{G}} \times \mathbb{P}$  relative to  $\pi^{\mathbb{F}} \times \mathbb{P}$ .*

*Proof* Let  $\gamma_n(s, z)$  be a  $\mathcal{P}(\mathbb{G}) \otimes \mathcal{E}$ -measurable process such that  $\gamma_n(s, z) = \frac{\pi^{\mathbb{G}_n}(ds, dz)}{\pi^{\mathbb{F}}(ds, dz)}$  a.s. Then  $\sup_n E \int_0^T \int_E \gamma_n \log(\delta_n) \pi^{\mathbb{F}}(ds, dz) \leq \mathcal{H}_{\mathcal{P}_T(\mathbb{G}) \otimes \mathcal{E}}(\pi^{\mathbb{G}} \|\pi^{\mathbb{F}}) < \infty$ , and hence the sequence  $(\gamma_n)$  is uniformly integrable in  $L^1(\pi^{\mathbb{F}} \times \mathbb{P})$  (see, e.g., Theorem 4.5.9 in [4]). Consequently, there exists a subsequence, also denoted by  $(\gamma_n)$ , that converges weakly in  $L^1(\pi^{\mathbb{F}} \times \mathbb{P})$  to a function  $\gamma \in L^1(\pi^{\mathbb{F}} \times \mathbb{P})$  (see, e.g., Theorem 4.7.19 in [4]). Now let  $N \in \mathbb{N}$ , and let  $f$  be a bounded and nonnegative  $\mathcal{P}(\mathbb{G}^N) \otimes \mathcal{E}$ -measurable process. Then, due to the weak convergence,

$$\begin{aligned} \mathbb{E}(f * \mu)_T &= \lim_n \mathbb{E}(f * \pi^{\mathbb{G}_n})_T = \mathbb{E} \int_0^T \int_E f \gamma_n \pi^{\mathbb{F}}(ds, dz) \\ &= \mathbb{E} \int_0^T \int_E f \gamma \pi^{\mathbb{F}}(ds, dz). \end{aligned}$$

Moreover, by the definition of the  $\mathbb{G}$ -compensator  $\pi^{\mathbb{G}}$ , we have  $\mathbb{E}(f * \mu)_T = \mathbb{E}(f * \pi^{\mathbb{G}})_T$ , and hence

$$\mathbb{E}(f * \pi^{\mathbb{G}})_T = \mathbb{E} \int_0^T \int_E f \gamma \pi^{\mathbb{F}}(ds, dz). \tag{4.1}$$

With a monotone limit theorem one can show that (4.1) holds for any nonnegative  $\mathcal{P}(\mathbb{G}) \otimes \mathcal{E}$ -measurable process  $f$ , from which we obtain that  $\gamma$  is the density of  $\pi^{\mathbb{G}} \times \mathbb{P}$  relative to  $\pi^{\mathbb{F}} \times \mathbb{P}$ .  $\square$

We are now in a position to state the main result of this section, establishing the link between the relative entropy of  $\pi^{\mathbb{G}}$  and  $\pi^{\mathbb{F}}$ , and the mutual information between the enlarging information  $G$  and the old information represented by  $\mathbb{F}$ .

**Theorem 4.7**

(i) *If  $T \in \mathbb{R}_+$  and  $v(E) < \infty$ , then*

$$\mathcal{H}_{\mathcal{P}_T(\mathbb{G}) \times \mathcal{E}}(\pi^{\mathbb{G}} \times \mathbb{P} \|\pi^{\mathbb{F}} \times \mathbb{P}) = I(\sigma(G)\|\mathcal{F}_T).$$

(ii) Let  $T_n \in \mathbb{R}_+$  such that  $T_n \uparrow \infty$ , and  $E_1, E_2, \dots$  an increasing sequence of sets in  $\mathcal{E}$  with  $E_n \uparrow E$  and  $\nu(E_n) < \infty$ . Then we have

$$\sup_n \mathcal{H}_{\mathcal{P}_{T_n}(\mathbb{G}) \times (\mathcal{E} \cap E_n)}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}) = I(\sigma(G) \| \mathcal{F}_\infty).$$

We first show the theorem for filtrations that are initially enlarged by finitely many measurable sets. The proof of the general result will then be based on the fact that any initial enlargement can be approximated by initial enlargements with finite partitions.

**Lemma 4.8** Assume that  $T \in \mathbb{R}_+$  and  $\nu(E) < \infty$ . Let  $\Delta = \{A_1, \dots, A_n\}$  be a partition of  $\Omega$  into measurable sets, and consider the enlargement of the form

$$\mathcal{G}_t^0 = \mathcal{F}_t^0 \vee \sigma(A_1, \dots, A_n).$$

Again, let  $\mathbb{G}$  be the smallest filtration containing  $\mathbb{G}^0$  and satisfying the usual conditions. Then, the relative entropy of  $\pi^{\mathbb{G}} \times \mathbb{P}$  with respect to  $\pi^{\mathbb{F}} \times \mathbb{P}$  on  $\mathcal{P}_T(\mathbb{G}) \otimes \mathcal{E}$  is given by

$$\mathcal{H}_{\mathcal{P}_T(\mathbb{G}) \times \mathcal{E}}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}) = I(\sigma(\Delta) \| \mathcal{F}_T).$$

*Proof* Let  $\epsilon > 0$ . Then for any  $i = 1, \dots, n$ , from Itô’s formula we obtain

$$\begin{aligned} & \log(\mathbb{P}_t(\cdot, A_i) + \epsilon) \\ &= \log(\mathbb{P}(A_i) + \epsilon) + \int_{0+}^t \frac{d\mathbb{P}_s(\cdot, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon} \\ & \quad + \sum_{0 < s \leq t} \left\{ \log(\mathbb{P}_s(\cdot, A_i) + \epsilon) - \log(\mathbb{P}_{s-}(\cdot, A_i) + \epsilon) - \frac{\Delta\mathbb{P}_s(\cdot, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon} \right\}. \end{aligned}$$

With (3.4) we further get

$$\begin{aligned} & \log(\mathbb{P}_t(\cdot, A_i) + \epsilon) \\ &= \log(\mathbb{P}(A_i) + \epsilon) + \int_{0+}^t \int_E \frac{k_s(\cdot, z, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon} \mu^{\mathbb{F}}(ds, dz) \\ & \quad + \sum_{0 < s \leq t} \left\{ \log\left(1 + \frac{\Delta\mathbb{P}_s(\cdot, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon}\right) - \frac{\Delta\mathbb{P}_s(\cdot, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon} \right\} \\ &= \log(\mathbb{P}(A_i) + \epsilon) + \int_{0+}^t \int_E \frac{k_s(\cdot, z, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon} \mu^{\mathbb{F}}(ds, dz) \\ & \quad + \int_{0+}^t \int_E \left\{ \log\left(1 + \frac{k_s(\cdot, z, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon}\right) - \frac{k_s(\cdot, z, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon} \right\} \mu(ds, dz). \end{aligned}$$

Using Itô’s isometry and the fact that  $\mathbb{P}_t(\cdot, A_i)$  is an  $\mathbb{L}^2$ -martingale, one can deduce that for all  $i \in \{1, \dots, n\}$ , we have  $k_s(\omega, z, A_i) \in \mathbb{L}^2(\pi^{\mathbb{F}} \times \mathbb{P})$ , and therefore also  $\frac{k_s(\omega, z, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon} \in \mathbb{L}^2(\pi^{\mathbb{F}} \times \mathbb{P})$ .

Consider, for all  $m \geq 1$ , the stopping times

$$\tau_m = T \wedge \inf \left\{ t \geq 0 : \mathbb{P}_t(\cdot, A_i) \leq \frac{1}{m} \text{ for one } i \right\}.$$

By Lemma 3.7,

$$\delta_{S \wedge \tau_m} = \sum_{i=1}^n 1_{A_i} \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i)} 1_{[0, \tau_m]}(s),$$

which yields that  $1_{[0, \tau_m]}(s) \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i) + \epsilon} \delta_S(z, \cdot)$  belongs to  $\mathbb{L}^1(\pi^{\mathbb{F}} \times \mathbb{P})$ . In particular, the integral  $\int_0^{\tau_m} \int_E \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i) + \epsilon} (\pi^{\mathbb{G}} - \pi^{\mathbb{F}})(ds, dz)$  is defined. Denoting by  $M_t^{\mathbb{G}, i}$  the  $(\mathbb{G}, \mathbb{P})$ -local martingale  $\int_0^t \int_E \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i) + \epsilon} \mu^{\mathbb{G}}(ds, dz)$ , we get

$$\begin{aligned} & \log(\mathbb{P}_{\tau_m}(\cdot, A_i) + \epsilon) - \log(\mathbb{P}(A_i) + \epsilon) \\ &= M_{\tau_m}^{\mathbb{G}, i} + \int_{0+}^{\tau_m} \int_E \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i) + \epsilon} (\pi^{\mathbb{G}} - \pi^{\mathbb{F}})(ds, dz) \\ & \quad + \int_{0+}^{\tau_m} \int_E \left\{ \log \left( 1 + \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i) + \epsilon} \right) - \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i) + \epsilon} \right\} \mu(ds, dz). \end{aligned}$$

Taking the expectations, we obtain the equality

$$\begin{aligned} & \mathbb{E} \left( \sum_{i=1}^n 1_{A_i} [\log(\mathbb{P}_{\tau_m}(\cdot, A_i) + \epsilon) - \log(\mathbb{P}(A_i) + \epsilon)] \right) \\ &= \mathbb{E} \left( \sum_{i=1}^n \int_{0+}^{\tau_m} \int_E 1_{A_i} \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i) + \epsilon} \delta_S(z, \omega, \omega) \pi^{\mathbb{F}}(dz, ds) \right) \\ & \quad + \mathbb{E} \left( \sum_{i=1}^n \int_{0+}^{\tau_m} \int_E 1_{A_i} \left\{ \log \left( 1 + \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i) + \epsilon} \right) \right. \right. \\ & \quad \left. \left. - \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i) + \epsilon} \right\} \pi^{\mathbb{G}}(ds, dz) \right), \end{aligned}$$

where  $\delta$  is defined as in (3.2). Observe that the expectation can be further simplified to

$$\begin{aligned} & \mathbb{E} \left[ \sum_{i=1}^n \int_{0+}^{\tau_m} \int_E 1_{A_i} \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i) + \epsilon} \delta_S(z, \omega, \omega) \pi^{\mathbb{F}}(dz, ds) \right. \\ & \quad + \int_{0+}^{\tau_m} \int_E 1_{A_i} \left\{ \log \left( 1 + \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i) + \epsilon} \right) - \frac{k_S(\cdot, z, A_i)}{\mathbb{P}_{S-}(\cdot, A_i) + \epsilon} \right\} \\ & \quad \left. \times (1 + \delta_S(z, \omega, \omega)) \pi^{\mathbb{F}}(dz, ds) \right] \end{aligned}$$

$$\begin{aligned}
 &= +\mathbb{E}\left(\sum_{i=1}^n \int_{0+}^{\tau_m} \int_E 1_{A_i} \log\left(1 + \frac{k_s(\cdot, z, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon}\right) (1 + \delta_s(z, \omega, \omega)) \pi^{\mathbb{F}}(dz, ds)\right) \\
 &\quad - \mathbb{E}\left(\sum_{i=1}^n \int_{0+}^{\tau_m} \int_E 1_{A_i} \frac{k_s(\cdot, z, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon} \pi^{\mathbb{F}}(dz, ds)\right).
 \end{aligned}$$

We now let  $\epsilon$  converge to 0 on both sides of the previous equation. First note that, as  $\epsilon \downarrow 0$ , the left-hand side converges to

$$\mathbb{E}\left(\sum_{i=1}^n 1_{A_i} [\log(\mathbb{P}_{\tau_m}(\cdot, A_i)) - \log(\mathbb{P}(A_i))]\right) = I(\sigma(\Delta) \|\mathcal{F}_{\tau_m}).$$

Moreover, it follows from dominated convergence and Lemma 3.6 that

$$\begin{aligned}
 &\lim_{\epsilon \downarrow 0} \mathbb{E}\left(\sum_{i=1}^n \int_{0+}^{\tau_m} \int_E \left(1_{A_i} \frac{k_s(\cdot, z, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon}\right) \pi^{\mathbb{F}}(dz, ds)\right) \\
 &= \mathbb{E}\left(\int_{0+}^{\tau_m} \int_E \delta_s(z, \omega) \pi^{\mathbb{F}}(dz, ds)\right) \\
 &= 0.
 \end{aligned}$$

This yields

$$\begin{aligned}
 &I(\sigma(\Delta) \|\mathcal{F}_{\tau_m}) \\
 &= \lim_{\epsilon \downarrow 0} \mathbb{E}\left(\sum_{i=1}^n \int_{0+}^{\tau_m} \int_E 1_{A_i} \log\left(1 + \frac{k_s(\cdot, z, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon}\right) (1 + \delta_s(z, \omega)) \pi^{\mathbb{F}}(dz, ds)\right).
 \end{aligned} \tag{4.2}$$

Now suppose that  $1_{[0, \tau_m]}(1 + \delta) \log(1 + \delta)$  is integrable with respect to  $\pi^{\mathbb{F}} \times \mathbb{P}$ . It is easy to verify that

$$\left| (1 + \delta) \log\left(1 + \frac{k}{P + \epsilon}\right) \right| \leq |(1 + \delta) \log(1 + \delta)|,$$

and hence, by dominated convergence we may interchange integration and limits in (4.2), which yields

$$I(\sigma(\Delta) \|\mathcal{F}_{\tau_m}) = \mathbb{E} \int_0^{\tau_m} \int_E \log(1 + \delta(s, z)) \pi^{\mathbb{G}}(ds, dz).$$

Finally, suppose that  $1_{[0, \tau_m]}(1 + \delta) \log(1 + \delta)$  is not integrable. In this case, Fatou’s lemma implies

$$\begin{aligned}
 &\liminf_{\epsilon \downarrow 0} \mathbb{E}\left(\sum_{i=1}^n \int_{0+}^{\tau_m} \int_E 1_{A_i} \log\left(1 + \frac{k_s(\cdot, z, A_i)}{\mathbb{P}_{s-}(\cdot, A_i) + \epsilon}\right) (1 + \delta_s(z, \omega, \omega)) \pi^{\mathbb{F}}(dz, ds)\right) \\
 &\geq \mathbb{E} \int_0^{\tau_m} \int_E \log(1 + \delta(s, z)) \pi^{\mathbb{G}}(ds, dz) = \infty.
 \end{aligned}$$

To sum up, we have shown that for all  $m \geq 1$ ,

$$I(\sigma(\Delta) \| \mathcal{F}_{\tau_m}) = \mathbb{E} \int_0^{\tau_m} \int_E \log \left( \frac{d(\pi^{\mathbb{G}} \times \mathbb{P})}{d(\pi^{\mathbb{F}} \times \mathbb{P})} \Big|_{\mathcal{P}_T(\mathbb{G}) \otimes \mathcal{E}} \right) \pi^{\mathbb{G}}(ds, dz).$$

By letting  $m \rightarrow \infty$  and using Lemma 4.5, we obtain the result. □

**Corollary 4.9** *Let  $G$  be an  $\mathcal{F}_T$ -measurable random value taking only finite number of values  $g_1, \dots, g_n$ . Consider the initial enlargement of the form  $\mathcal{G}_t^0 = \mathcal{F}_t^0 \vee \sigma(G)$ . Then*

$$\mathcal{H}_{\mathcal{P}_T(\mathbb{G}) \times \mathcal{E}}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}) = - \sum_{i=1}^n \mathbb{P}(G = g_i) \log(\mathbb{P}(G = g_i)).$$

*Proof* It suffices to observe that the initial enlargement by  $G$  is given as an enlargement by partition of the form  $A_i = \{G = g_i\}$ . Then, Theorem 4.8 completes the proof. □

For the proof of Theorem 4.7, we need some further auxiliary results. We start with the following one.

**Lemma 4.10** *Assume that  $\nu(E) < \infty$ ,  $T \in \mathbb{R}_+$ , and let  $\mathbb{H} = (\mathcal{H}_t)$  and  $\mathbb{G} = (\mathcal{G}_t)$  be two filtrations such that  $\mathcal{F}_t \subset \mathcal{H}_t \subset \mathcal{G}_t$  for all  $t \in [0, T]$ . Then  $\pi^{\mathbb{G}} \times \mathbb{P} = \pi^{\mathbb{H}} \times \mathbb{P}$  on the  $\sigma$ -algebra  $\mathcal{P}(\mathbb{H}) \otimes \mathcal{E}$ , and*

$$\mathcal{H}_{\mathcal{P}_T(\mathbb{H}) \times \mathcal{E}}(\pi^{\mathbb{H}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}) \leq \mathcal{H}_{\mathcal{P}_T(\mathbb{G}) \times \mathcal{E}}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}).$$

*Proof* Let  $f$  be a nonnegative  $\mathcal{P}(\mathbb{H}) \otimes \mathcal{E}$ -measurable process. Then

$$\mathbb{E}(f * \pi^{\mathbb{H}})_T = \mathbb{E}(f * \mu)_T = \mathbb{E}(f * \pi^{\mathbb{G}})_T,$$

which implies that  $\pi^{\mathbb{H}} \times \mathbb{P}$  coincides with  $\pi^{\mathbb{G}} \times \mathbb{P}$  on  $\mathcal{P}(\mathbb{H}) \otimes \mathcal{E}$ . Moreover, by using Lemma 4.4, we get

$$\begin{aligned} \mathcal{H}_{\mathcal{P}_T(\mathbb{H}) \otimes \mathcal{E}}(\pi^{\mathbb{H}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}) &= \mathcal{H}_{\mathcal{P}_T(\mathbb{H}) \otimes \mathcal{E}}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}) \\ &\leq \mathcal{H}_{\mathcal{P}_T(\mathbb{G}) \otimes \mathcal{E}}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}), \end{aligned}$$

and hence the result. □

We proceed by showing that if  $E_1, E_2, \dots$  is a sequence with  $E_n \uparrow E$  and  $\nu(E_n) < \infty$ , then  $\mathcal{H}_{\mathcal{P}_T(\mathbb{G}) \times (\mathcal{E} \cap E_n)}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{G}} \times \mathbb{P})$  is an increasing sequence with a limit not depending on the choice of the approximating sequence  $E_n$ . To simplify the notation, we will use the abbreviation  $\mathcal{H}_n(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{G}} \times \mathbb{P}) = \mathcal{H}_{\mathcal{P}_T(\mathbb{G}) \times (\mathcal{E} \cap E_n)}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{G}} \times \mathbb{P})$ .

**Lemma 4.11** *Let  $T \in \mathbb{R}_+$ , and let  $E_1, E_2, \dots$  be a sequence of measurable sets with  $E_n \uparrow E$  and  $\nu(E_n) < \infty$ . Then  $\mathcal{H}_n(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{G}} \times \mathbb{P})$  is an increasing sequence. Moreover, its supremum does not depend on the choice of the approximating sequence  $E_n$ .*



*Proof* Let  $A \in \mathcal{E}$  with  $\nu(A) < \infty$ , and let  $c = \pi^{\mathbb{F}}([0, T] \times A)$ . Note that also  $\pi^{\mathbb{G}}([0, T] \times A) = c$ . By Jensen’s inequality we have

$$c \int_{\Omega \times [0, T] \times A} \frac{d(\pi^{\mathbb{G}} \times \mathbb{P})}{d(\pi^{\mathbb{F}} \times \mathbb{P})} \log \left( \frac{d(\pi^{\mathbb{G}} \times \mathbb{P})}{d(\pi^{\mathbb{F}} \times \mathbb{P})} \right) \frac{d(\pi^{\mathbb{F}} \times \mathbb{P})}{c} \geq 0,$$

which implies that  $\mathcal{H}_n(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{G}} \times \mathbb{P})$  is an increasing sequence.

Let  $(E_n)$  and  $(D_m)$  be two increasing sequences of measurable sets with  $\bigcup_n E_n = \bigcup_m D_m = E$ ,  $\nu(E_n) < \infty$ , and  $\nu(D_m) < \infty$  for all  $n, m \geq 1$ . Then, by dominated convergence,

$$\begin{aligned} & \sup_n \int 1_{E_n} \log \left( \frac{d(\pi^{\mathbb{G}} \times \mathbb{P})}{d(\pi^{\mathbb{F}} \times \mathbb{P})} \right) d(\pi^{\mathbb{F}} \times \mathbb{P}) \\ &= \sup_n \sup_m \int 1_{E_n \cap D_m} \log \left( \frac{d(\pi^{\mathbb{G}} \times \mathbb{P})}{d(\pi^{\mathbb{F}} \times \mathbb{P})} \right) d(\pi^{\mathbb{F}} \times \mathbb{P}) \\ &= \sup_m \sup_n \int 1_{E_n \cap D_m} \log \left( \frac{d(\pi^{\mathbb{G}} \times \mathbb{P})}{d(\pi^{\mathbb{F}} \times \mathbb{P})} \right) d(\pi^{\mathbb{G}} \times \mathbb{P}) \\ &= \sup_m \int 1_{D_m} \log \left( \frac{d(\pi^{\mathbb{G}} \times \mathbb{P})}{d(\pi^{\mathbb{F}} \times \mathbb{P})} \right) d(\pi^{\mathbb{G}} \times \mathbb{P}), \end{aligned}$$

which completes the proof. □

*Proof of Theorem 4.7* First assume that  $T \in \mathbb{R}_+$  and  $\nu(E) < \infty$ . Let  $\mathcal{P}_n$  be a sequence of finite partitions of  $\Omega$  such that  $\mathcal{P}_n \subset \mathcal{P}_{n+1}$  and  $\mathcal{A} = \bigvee_{n \geq 1} \sigma(\mathcal{P}_n)$ . Let  $\mathcal{G}_t^n = \bigcap_{s > t} (\mathcal{F}_s \vee \sigma(\mathcal{P}_n))$  and set  $\mathbb{G}^n = (\mathcal{G}_t^n)$ . Then, by Lemma 4.8,  $\mathcal{H}_{\mathcal{P}_T(\mathbb{G}^n) \otimes \mathcal{E}}(\pi^{\mathbb{G}^n} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}) = I(\sigma(\mathcal{P}_n) \| \mathcal{F}_T)$ . Moreover, Lemmas 4.4 and 4.10 yield

$$\begin{aligned} I(\mathcal{A} \| \mathcal{F}_T) &= \sup_n I(\sigma(\mathcal{P}_n) \| \mathcal{F}_T) = \sup_n \mathcal{H}_{\mathcal{P}_T(\mathbb{G}^n) \otimes \mathcal{E}}(\pi^{\mathbb{G}^n} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}) \\ &= \sup_n \mathcal{H}_{\mathcal{P}_T(\mathbb{G}^n) \otimes \mathcal{E}}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}) \\ &= \mathcal{H}_{\mathcal{P}_T(\mathbb{G}) \otimes \mathcal{E}}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}), \end{aligned}$$

which proves statement (i) of Theorem 4.7.

Now suppose that  $\nu(E) = \infty$ , and let  $E_1, E_2, \dots$  be a sequence of sets in  $\mathcal{E}$  with  $E_n \uparrow E$  and  $\nu(E_n) < \infty$ . Let  $\mathbb{F}^n$  be the associated approximating sequence of filtrations as defined in (2.2), and let  $\mathbb{G}^n = (\mathcal{G}_t^n)$  be defined by  $\mathcal{G}_t^n = \bigcap_{s > t} (\mathcal{F}_s^n \vee \sigma(G))$ . Then

$$I(\sigma(G) \| \mathcal{F}_T^n) = \mathcal{H}_{\mathcal{P}_T(\mathbb{G}^n) \otimes (\mathcal{E} \cap E_n)}(\pi^{\mathbb{G}^n} \times \mathbb{P} \| \pi^{\mathbb{F}^n} \times \mathbb{P}),$$

which, due to (2.3) and Lemma 4.10, simplifies to

$$I(\sigma(G) \| \mathcal{F}_T^n) = \mathcal{H}_{\mathcal{P}_T(\mathbb{G}^n) \otimes (\mathcal{E} \cap E_n)}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}).$$

If  $T_n \in \mathbb{R}_+$  are such that  $T_n \uparrow \infty$ , then Lemma 4.4 implies

$$\begin{aligned} I(\sigma(G) \| \mathcal{F}_\infty) &= \sup_n \mathcal{H}_{\mathcal{P}_{T_n}(\mathbb{G}^n) \otimes (\mathcal{E} \cap E_n)}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}) \\ &= \sup_n \sup_m \mathcal{H}_{\mathcal{P}_{T_n}(\mathbb{G}^m) \otimes (\mathcal{E} \cap E_n)}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}) \\ &= \sup_n \mathcal{H}_{\mathcal{P}_{T_n}(\mathbb{G}) \times (\mathcal{E} \cap E_n)}(\pi^{\mathbb{G}} \times \mathbb{P} \| \pi^{\mathbb{F}} \times \mathbb{P}), \end{aligned}$$

and hence statement (ii) of Theorem 4.7. □

### 5 Embeddings

In this section we will analyse some standard norms on the set of semimartingales and see how they depend on the filtration relative to which they are defined. We shall see how the results from the preceding section can be used in order to give some general estimates of changes in the semimartingale norms when enlarging  $\mathbb{F}$  to a filtration  $\mathbb{G}$ . In particular we show that if  $\mathbb{G}$  is an initial enlargement of  $\mathbb{F}$  by some random variable  $G$  with *finite entropy*, then every square integrable  $(\mathbb{F}, \mathbb{P})$ -martingale is a  $(\mathbb{G}, \mathbb{P})$ -semimartingale and possesses an information drift which is  $L^1(\lambda \times \mathbb{P})$ -integrable. We will start with an easy observation concerning the preservation of the semimartingale property.

**Theorem 5.1** *Let  $v(E) < \infty$ , and let  $\mathbb{G}$  be an arbitrary (not necessarily initial) enlargement of  $\mathbb{F}$ . Then, every  $(\mathbb{F}, \mathbb{P})$ -semimartingale is also a  $(\mathbb{G}, \mathbb{P})$ -semimartingale.*

*Proof* Let  $X$  be an  $(\mathbb{F}, \mathbb{P})$ -semimartingale with Doob–Meyer decomposition  $X_t = M_t + A_t$ . From the predictable representation property of  $\mu$  we deduce the existence of an  $\mathbb{F}$ -predictable process  $\psi$  such that

$$M_t = M_0 + \int_0^t \int_E \psi(s, z) \mu^{\mathbb{F}}(dz, ds).$$

If  $v(E) < \infty$ , then one can show that  $\psi \in \mathbb{L}^1(\pi^{\mathbb{F}} \times \mathbb{P})$  (see Theorem 4.37 in Chap. III of [10]). This means that  $X$  is a finite-variation process and hence a semimartingale with respect to  $\mathbb{G}$ . □

**Definition 5.2** Let  $X$  be an  $(\mathbb{F}, \mathbb{P})$ -semimartingale. Given the semimartingale decomposition  $X = M + A$ , for all  $1 \leq p < \infty$  and  $T \in [0, \infty]$ , we define

$$j_{p,T}(M, A) = \left\| [M, M]_T^{\frac{1}{2}} + \int_0^T |dA_s| \right\|_{\mathbb{L}^p}.$$

Further, we shall simply denote  $j_{p,\infty}$  by  $j_p$ .

**Definition 5.3** For given  $1 \leq p < \infty$ , the space of all  $(\mathbb{F}, \mathbb{P})$ -semimartingales  $X$  such that

$$\inf_{X=M+A} j_{p,T}(M, A) < \infty,$$

where the infimum is taken over all semimartingale decompositions, is denoted by  $\mathcal{S}_T^p$ . We will adopt the notation  $\mathcal{S}^p := \mathcal{S}_\infty^p$ .

Let  $1 \leq p < \infty$ . One can show that every semimartingale in  $\mathcal{S}_T^p$  is special, which means that there exists a decomposition  $X = M' + A'$  such that  $A'$  is  $\mathbb{F}$ -predictable and  $A'_0 = 0$ .

**Definition 5.4** We define a norm on  $\mathcal{S}_T^p$  by

$$\|X\|_{\mathcal{S}^p} = j_{p,T}(M', A').$$

**Lemma 5.5** *There exists a constant  $C = C(p)$ , such that*

$$C\|X\|_{\mathcal{S}^p} \leq \inf_{X=M+A} j_p(M, A) \leq \|X\|_{\mathcal{S}^p}.$$

*Proof* See Remark (c) of Sect. 98 in Chap. VII of [5] for the case  $T = \infty$ . The proof in the case  $T < \infty$  is analogous. □

**Definition 5.6** The subspace of  $\mathcal{S}_T^p$  consisting of martingales will be denoted as  $\mathcal{H}_T^p$ .

**Definition 5.7** Let  $f : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$  be a convex function on the real line. The function  $f^* : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$  defined by

$$f^*(y) = \sup_{x \in \mathbb{R}} (xy - f(x))$$

is called the *Legendre transform* of  $f$ .

We recall here the well-known *Young inequality*, which uses Legendre transforms. It states that for convex  $f$  and its Legendre transform  $f^*$ , the following inequality holds:

$$xy \leq f(x) + f^*(y).$$

From Young’s inequality we easily obtain the following lemma.

**Lemma 5.8** *For all  $x \in \mathbb{R}$  and  $y \geq -1$ , we have*

$$xy \leq e^x - x - 1 + (y + 1) \log(y + 1) - y.$$

*Proof* It suffices to apply Young’s inequality to the function  $f(x) = e^x - x - 1$ , and a straightforward calculation yields the result. □

We now use the preceding results to prove the following embedding theorems. First, let us begin with the following easy lemma.

**Lemma 5.9** *Consider an arbitrary enlargement  $\mathbb{G}$  of  $\mathbb{F}$  with finite time horizon  $T < \infty$ , and assume that  $\nu(E) < \infty$ . Then  $\mathcal{H}^2(\mathbb{F}) \subset \mathcal{S}^1(\mathbb{G})$ . Moreover, the inclusion mapping  $X \mapsto X$  is continuous.*

*Proof* First, observe that if  $X \in \mathcal{H}^2_T(\mathbb{F})$ , then  $X$  is a finite-variation process. Indeed, by Jensen’s inequality we obtain

$$\begin{aligned} \mathbb{E} \int_0^T |dX_s| &\leq 2\mathbb{E} \left( \int_0^T \int_E |\psi(s, z)| \pi^{\mathbb{F}}(dz, ds) \right) \\ &\leq 2\sqrt{T\nu(E)} \left( \int_0^T \int_E |\psi(s, z)|^2 \pi^{\mathbb{F}} \right)^{\frac{1}{2}} \\ &= 2\sqrt{T\nu(E)} \|X\|_{\mathcal{H}^2_T(\mathbb{F})}. \end{aligned}$$

Therefore,  $X_t = 0 + X_t$  is a semimartingale decomposition with respect to filtration  $\mathbb{G}$ . Hence, by Lemma 5.5 there exists a constant  $c \in \mathbb{R}_+$  such that

$$\|X\|_{\mathcal{S}^1_T(\mathbb{G})} \leq Cj_{1,T}(0, X).$$

Combining this with the previous inequality, we obtain

$$\|X\|_{\mathcal{S}^1_T(\mathbb{G})} \leq 2C\sqrt{T\nu(E)} \|X\|_{\mathcal{H}^2(\mathbb{F})},$$

which ends the proof. □

Although from Theorem 5.1 we know that for  $\nu(E) < \infty$ , if  $X \in \mathcal{H}^2(\mathbb{F})$ , then  $X$  is a  $(\mathbb{G}, \mathbb{P})$ -semimartingale, it is not true that  $\mathcal{H}^2(\mathbb{F}) \subset \mathcal{S}^1(\mathbb{G})$  in general. The next example shows that the assumption  $T < \infty$  cannot be omitted in Lemma 5.9.

*Example 5.10* Let  $\psi$  be a deterministic function given by  $\psi(s, z) = \frac{1}{1+s}$ , and, as usual, let  $X_t = \int_0^t \int_E \psi(s, z) \mu^{\mathbb{F}}(dz, ds)$ . Then

$$\|X\|_{\mathcal{H}^2(\mathbb{F})} = \mathbb{E} \left( \int_0^\infty \int_E \frac{1}{(1+s)^2} \pi^{\mathbb{F}}(dz, ds) \right) = \nu(E) \int_0^\infty \frac{ds}{(1+s)^2} < \infty.$$

On the other hand, if we take  $\mathcal{G}_t = \mathcal{F}_\infty$ , then

$$\|X\|_{\mathcal{S}^1(\mathbb{G})} = \mathbb{E} \int_0^\infty |dX_s| = \int_0^\infty \frac{ds}{1+s} = \infty.$$

Thus, a natural question arises: what happens if we relax the assumption of  $T < \infty$  or  $\nu(E) < \infty$ ? The next theorem gives some answer to this question.

**Theorem 5.11** Consider the initial enlargement of the form  $\mathcal{G}_t = \bigcap_{s>t} \mathcal{F}_s \vee \sigma(G)$  for some random variable  $G$ . If the mutual information between  $\sigma(G)$  and  $\mathcal{F}_\infty$  is finite, i.e.,  $I(\sigma(G)\|\mathcal{F}_\infty) < \infty$ , and  $\nu$  is  $\sigma$ -finite, then

- (i) Every  $X \in \mathcal{H}^2(\mathbb{F})$  possesses an integrable information drift with respect to the enlarged filtration  $\mathbb{G}$ . More precisely, if  $X_t = X_0 + \int_0^t \int_E \psi(s, z) \mu^\mathbb{F}(dz, ds)$ , with  $\psi \in L^2(\pi^\mathbb{F} \times \mathbb{P})$  being  $\mathbb{F}$ -predictable, then  $\eta_s = \int_E \psi(s, z) \delta(s, z) \nu(dz)$  is the associated  $\mathbb{G}$ -information drift.
- (ii)  $\mathcal{H}^2(\mathbb{F}) \subset \mathcal{S}^1(\mathbb{G})$ .

*Proof* Notice first that  $I(\sigma(G)\|\mathcal{F}_\infty) < \infty$ , together with Theorem 4.7, implies condition (3.1), and hence the process  $\delta$  is defined (see Definition 3.2).

Let us begin with part (i). Let  $X \in \mathcal{H}^2(\mathbb{F})$  with  $X_t = X_0 + \int_0^t \int_E \psi(s, z) \mu^\mathbb{F}(dz, ds)$ , and set  $\alpha(s, z) = \psi(s, z) 1_{\{|\psi(s, z)| \leq 1\}}$  and  $\beta(s, z) = \psi(s, z) 1_{\{|\psi(s, z)| > 1\}}$ . Observe first that  $\beta\delta \in L^1(\pi^\mathbb{F} \times \mathbb{P})$ . Indeed, since  $\beta^2(s, z) \geq |\beta(s, z)|$ , we have that  $\beta \in L^1(\pi^\mathbb{F} \times \mathbb{P}) \cap L^2(\pi^\mathbb{F} \times \mathbb{P})$ , and we can apply Lemma 3.3.

Let  $T_n \in \mathbb{R}_+$  with  $T_n \uparrow \infty$ , and let  $E_n \in \mathcal{E}$  be an increasing sequence such that  $E_n \uparrow E$  and  $\nu(E_n) < \infty$ . Then monotone convergence implies

$$\mathbb{E} \left( \int_0^\infty \int_E |\alpha(s, z) \delta(s, z)| \pi^\mathbb{F}(dz, ds) \right) \leq \sup_n \mathbb{E} \left( \int_0^{T_n} \int_{E_n} |\alpha(s, z) \delta(s, z)| \pi^\mathbb{F}(dz, ds) \right).$$

Now, from Lemma 5.8, we get

$$\begin{aligned} & \mathbb{E} \left( \int_0^{T_n} \int_{E_n} |\alpha(s, z) \delta(s, z)| \pi^\mathbb{F}(dz, ds) \right) \\ & \leq \mathbb{E} \left( \int_0^{T_n} \int_{E_n} (e^{|\psi(s, z)| \operatorname{sgn}(\delta(s, z))} - 1 - |\psi(s, z)| \operatorname{sgn}(\delta(s, z))) 1_{\{|\psi| \leq 1\}} \pi^\mathbb{F}(dz, ds) \right) \\ & \quad + \mathbb{E} \left( \int_0^{T_n} \int_{E_n} (\delta(s, z) + 1) \log(\delta(s, z) + 1) - \delta(s, z) \pi^\mathbb{F}(dz, ds) \right). \end{aligned}$$

Observe that

$$\begin{aligned} & \mathcal{H} \mathcal{P}_{T_n}(\mathbb{G}) \otimes \mathcal{E} \cap E_n (\pi^\mathbb{G} \times \mathbb{P} \|\pi^\mathbb{F} \times \mathbb{P}) \\ & = \mathbb{E} \left( \int_0^{T_n} \int_{E_n} (\delta(s, z) + 1) \log(\delta(s, z) + 1) \pi^\mathbb{F}(dz, ds) \right), \end{aligned}$$

which, together with Lemma 3.6 and Theorem 4.7, shows that

$$\mathbb{E} \left( \int_0^{T_n} \int_{E_n} (\delta(s, z) + 1) \log(\delta(s, z) + 1) - \delta(s, z) \pi^\mathbb{F}(dz, ds) \right) \leq I(\sigma(G)\|\mathcal{F}_\infty).$$

Moreover, we have

$$\begin{aligned} & \mathbb{E} \left( \int_0^{T_n} \int_{E_n} (e^{|\psi(s,z)| \operatorname{sgn}(\delta(s,z))} - 1 - |\psi(s,z)| \operatorname{sgn}(\delta(s,z))) 1_{\{|\psi| \leq 1\}} \pi^{\mathbb{F}}(dz, ds) \right) \\ &= \mathbb{E} \left( \int_0^{T_n} \int_{E_n} \sum_{k=2}^{\infty} \frac{(|\psi(s,z)| \operatorname{sgn}(\delta(s,z)))^k}{k!} 1_{\{|\psi(s,z)| \leq 1\}} \pi^{\mathbb{F}}(dz, ds) \right) \\ &\leq \mathbb{E} \left( \int_0^{T_n} \int_{E_n} |\psi(s,z)|^2 \left( \frac{1}{2!} + \frac{1}{3!} + \dots \right) \pi^{\mathbb{F}}(dz, ds) \right) \\ &\leq (e - 2) \mathbb{E} \int_0^{\infty} \int_E |\psi(s,z)|^2 \pi^{\mathbb{F}}(dz, ds) = (e - 2) \|X\|_{\mathcal{H}^2(\mathbb{F})}^2. \end{aligned}$$

Thus,

$$\mathbb{E} \left( \int_0^{\infty} \int_E |\alpha(s,z)\delta(s,z)| \pi^{\mathbb{F}}(dz, ds) \right) \leq I(\sigma(G)\|\mathcal{F}_{\infty}) + (e - 2) \|X\|_{\mathcal{H}^2(\mathbb{F})}^2. \tag{5.1}$$

This further yields that  $\psi\delta = \alpha\delta + \beta\delta$  belongs to  $L^1(\pi^{\mathbb{F}} \times \mathbb{P})$ , and we can use Theorem 3.5 in order to deduce that  $\int_E \psi\delta\nu(dz)$  is the required information drift.

Now, let us proceed with part (ii). We decompose  $X$  into

$$X_t = X_t^{\alpha} + X_t^{\beta},$$

where

$$\begin{aligned} X_t^{\alpha} &= X_0 + \int_0^t \int_E \alpha(s,z) \mu^{\mathbb{F}}(dz, ds), \\ X_t^{\beta} &= \int_0^t \int_E \beta(s,z) \mu^{\mathbb{F}}(dz, ds). \end{aligned}$$

Observe that  $X^{\beta}$  is of finite variation. Indeed, we have

$$\begin{aligned} \mathbb{E} \left( \int_0^{\infty} |dX_s^{\beta}| \right) &\leq 2\mathbb{E} \left( \int_0^{\infty} \int_E |\psi(s,z)| 1_{\{|\psi(s,z)| > 1\}} \mu(dz, ds) \right) \\ &\leq 2\mathbb{E} \left( \int_0^{\infty} \int_E \psi^2(s,z) 1_{\{|\psi(s,z)| > 1\}} \mu(dz, ds) \right) \\ &\leq 2\mathbb{E} \left( \int_0^{\infty} \int_E \psi^2(s,z) \mu(dz, ds) \right) \\ &= 2\mathbb{E}[X, X]_{\infty} = 2\|X\|_{\mathcal{H}^2(\mathbb{F})}^2, \end{aligned}$$

where in the first inequality we used the fact that total variation of a sum (or difference) of two processes is not greater than the sum of total variations of these processes. Thus, by Lemma 5.5 we obtain that there exists a constant  $C \in \mathbb{R}_+$  such that

$$\|X^{\beta}\|_{\mathcal{S}^1(\mathbb{G})} \leq Cj_1(0, X^{\beta}) = C\mathbb{E} \left( \int_0^{\infty} |dX_s^{\beta}| \right).$$

This implies that

$$\|X^\beta\|_{\mathcal{S}^1(\mathbb{G})} \leq 2C\|X\|_{\mathcal{H}^2(\mathbb{F})}^2.$$

On the other hand, we know that the norm  $\|X^\alpha\|_{\mathcal{S}^1(\mathbb{G})}$  can be estimated by

$$\|X^\alpha\|_{\mathcal{S}^1(\mathbb{G})} \leq \mathbb{E}[X^\alpha, X^\alpha]_\infty^{\frac{1}{2}} + 2\mathbb{E}\left(\int_0^\infty \int_E |\alpha(s, z)\delta(s, z)|\pi^\mathbb{F}(dz, ds)\right).$$

By combining this with inequality (5.1) we get

$$\begin{aligned} \|X^\alpha\|_{\mathcal{S}^1(\mathbb{G})} &\leq \mathbb{E}[X, X]_\infty^{\frac{1}{2}} + 2(e - 2)\|X\|_{\mathcal{H}^2(\mathbb{F})}^2 + 2I(\sigma(G)\|\mathcal{F}_\infty) \\ &\leq \sqrt{\mathbb{E}[X, X]_\infty} + 2(e - 2)\|X\|_{\mathcal{H}^2(\mathbb{F})}^2 + 2I(\sigma(G)\|\mathcal{F}_\infty) \\ &= \|X\|_{\mathcal{H}^2(\mathbb{F})} + 2(e - 2)\|X\|_{\mathcal{H}^2(\mathbb{F})}^2 + 2I(\sigma(G)\|\mathcal{F}_\infty), \end{aligned}$$

where the second inequality follows from Jensen’s inequality. Finally, we obtain

$$\begin{aligned} \|X\|_{\mathcal{S}^1(\mathbb{G})} &\leq \|X^\alpha\|_{\mathcal{S}^1(\mathbb{G})} + \|X^\beta\|_{\mathcal{S}^1(\mathbb{G})} \\ &\leq \|X\|_{\mathcal{H}^2(\mathbb{F})} + 2(e - 2 + C)\|X\|_{\mathcal{H}^2(\mathbb{F})}^2 + 2I(\sigma(G)\|\mathcal{F}_\infty), \end{aligned}$$

which completes the proof. □

We remark that if  $I(\sigma(G)\|\mathcal{F}_\infty)$  is finite, then a condition introduced in [9], and often referred to as *Jacod’s condition*, is satisfied. Let  $P_t(\omega, \cdot)$  denote a version of the regular conditional distribution of  $G$  with respect to  $\mathcal{F}_t$ . The condition introduced in [9] states that for all  $t$ , there exists a positive  $\sigma$ -finite measure  $\eta_t$  on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  such that

$$P_t(\omega, \cdot) \ll \eta_t(\cdot) \tag{5.2}$$

for  $\mathbb{P}$ -a.a.  $\omega$ . It is shown in [9] that (5.2) implies that every  $\mathbb{F}$ -semimartingale also is a  $\mathbb{G}$ -semimartingale.

In Theorem 5.11, Jacod’s condition is replaced with the stronger assumption  $I(\sigma(G)\|\mathcal{F}_\infty) < \infty$ , thus allowing us to obtain some estimates of the FV part in the  $\mathbb{G}$ -Doob–Meyer decompositions in terms of the mutual information.

We close this section with some examples, verifying whether the assumption  $I(\sigma(G)\|\mathcal{F}_\infty) < \infty$  is satisfied.

*Example 5.12* Let  $\psi \in L^1(\pi^\mathbb{F})$  be a deterministic positive function and define  $G = \int_0^\infty \int_E \psi(s, z)\mu(ds, dz)$ . Then  $G$  is Poisson distributed with parameter  $\lambda = \int_0^\infty \int_E \psi(s, z)\pi^\mathbb{F}(ds, dz) \in \mathbb{R}_+$ . The mutual information between  $G$  and  $\mathcal{F}_\infty$  is given by

$$I(\sigma(G)\|\mathcal{F}_\infty) = - \sum_{i=0}^\infty e^{-\lambda} \frac{\lambda^i}{i!} \log\left(e^{-\lambda} \frac{\lambda^i}{i!}\right),$$

which is easily shown to be finite. Consequently, by Theorem 5.11, every square-integrable  $\mathbb{F}$ -martingale possesses a  $\mathbb{G}$ -information drift and belongs to the class of semimartingales  $\mathcal{S}^1(\mathbb{G})$ , where the filtration  $\mathbb{G} = (\mathcal{G}_t)$  is defined by

$$\mathcal{G}_t = \bigcap_{s>t} \mathcal{F}_s \vee \sigma(G).$$

*Example 5.13* Let  $\tau$  be the first jump time of a standard Poisson process  $N = (N_t)_{t \geq 0}$  defined on our probability space. The mutual information between  $\tau$  and  $\mathcal{F}_\infty$  is infinite, since the distribution of  $\tau$  is absolutely continuous with respect to the Lebesgue measure. By adding some noise, however, the mutual information may become finite again. More precisely, let  $X$  be centered Gaussian random variable, independent of  $\mathcal{F}_\infty$ , with distribution that is absolutely continuous with respect to the Lebesgue measure. Let  $\hat{\tau} = \tau + X$ , and consider the enlargement

$$\mathcal{G}_t = \bigcap_{s>t} \mathcal{F}_s \vee \sigma(\hat{\tau}).$$

According to Lemma 6.2 in [2], the mutual information satisfies

$$I(\sigma(\hat{\tau}) \| \mathcal{F}_\infty) \leq \frac{1}{2} \log \left( \frac{\text{Var}(\tau) + \text{Var}(X)}{\text{Var}(X)} \right) < \infty.$$

Notice that the variance  $\text{Var}(\tau)$  is defined since  $\tau$  is exponentially distributed.

**Acknowledgements** The authors would like to thank an anonymous referee for valuable suggestions and corrections which improved the style of this paper.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Ankirchner, S.: On filtration enlargements and purely discontinuous martingales. *Stoch. Process. Appl.* **118**(9), 1662–1678 (2008)
2. Ankirchner, S., Dereich, S., Imkeller, P.: The Shannon information of filtrations and the additional logarithmic utility of insiders. *Ann. Probab.* **34**(2), 743–778 (2006)
3. Ankirchner, S., Dereich, S., Imkeller, P.: Enlargement of filtrations and Girsanov-type embeddings. In: *Séminaire de Probabilités XL. Lecture Note in Mathematics*, vol. 1899, pp. 389–410. Springer, Berlin (2007)
4. Bogachev, V.I.: *Measure Theory*, vols. I, II. Springer, Berlin (2007)
5. Dellacherie, C., Meyer, P.-A.: *Probabilities and Potential*, B. North-Holland Mathematics Studies, vol. 72. North-Holland, Amsterdam (1982). *Theory of martingales*, translated from the French by J.P. Wilson
6. Gasbarra, D., Valkeila, E.: Initial enlargement: a Bayesian approach. *Theory Stoch. Process.* **9**(3–4), 26–37 (2003)
7. Gasbarra, D., Valkeila, E., Vostrikova, L.: Enlargement of filtration and additional information in pricing models: Bayesian approach. In: *From Stochastic Calculus to Mathematical Finance*, pp. 257–285. Springer, Berlin (2006)



8. Imkeller, P.: Malliavin's calculus in insider models: additional utility and free lunches. *Math. Finance* **13**(1), 153–169 (2003). Conference on Applications of Malliavin Calculus in Finance (Rocquencourt, 2001)
9. Jacod, J.: Grossissement initial, hypothèse  $(H')$ , et théorème de Girsanov. In: Jeulin, T., Yor, M. (eds.) *Grossissements de Filtrations: Exemples et Applications*, pp. 15–35. Springer, Berlin (1985)
10. Jacod, J., Shiryaev, A.N.: *Limit Theorems for Stochastic Processes*, 2nd edn. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 288. Springer, Berlin (2003)
11. Jeulin, T.: *Semi-Martingales et Grossissement d'une Filtration*. *Lecture Notes in Mathematics*, vol. 833. Springer, Berlin (1980)
12. Jeulin, T., Yor, M. (eds.): *Grossissements de Filtrations: Exemples et Applications*. *Lecture Notes in Mathematics*, vol. 1118. Springer, Berlin (1985). Papers from the seminar on stochastic calculus held at the Université de Paris VI, Paris, 1982/1983
13. Liese, F., Vajda, I.: *Convex Statistical Distances*. *Teubner-Texte zur Mathematik [Teubner Texts in Mathematics]*, vol. 95. BSB B.G. Teubner Verlagsgesellschaft, Leipzig (1987). With German, French and Russian summaries
14. Mansuy, R., Yor, M.: *Random Times and Enlargements of Filtrations in a Brownian Setting*. *Lecture Notes in Mathematics*, vol. 1873. Springer, Berlin (2006)
15. Protter, P.E.: *Stochastic Integration and Differential Equations*, 2nd edn. *Stochastic Modelling and Applied Probability*, vol. 21. Springer, Berlin (2005). Version 2.1, Corrected third printing
16. Yor, M.: Entropie d'une partition, et grossissement initial d'une filtration. In: Jeulin, T., Yor, M. (eds.) *Grossissements de Filtrations: Exemples et Applications*, pp. 45–58. Springer, Berlin (1985)