OPTIMAL REINSURANCE VIA BSDES IN A PARTIALLY OBSERVABLE CONTAGION MODEL WITH JUMP CLUSTERS

MATTEO BRACHETTA, GIORGIA CALLEGARO, CLAUDIA CECI, AND CARLO SGARRA

Abstract. We investigate the optimal reinsurance problem when the loss process exhibits jump clustering features and the insurance company has restricted information about the loss process. We maximize expected exponential utility of terminal wealth and show that an optimal solution exists. By exploiting both the Kushner-Stratonovich and Zakai approaches, we provide the equation governing the dynamics of the (infinite-dimensional) filter and characterize the solution of the stochastic optimization problem in terms of a BSDE, for which we prove existence and uniqueness of solution. After discussing the optimal strategy for a general reinsurance premium, we provide more explicit results for proportional reinsurance under the expected value premium principle.

Keywords: Optimal reinsurance; Partial information; Hawkes processes; Cox processes with shot noise; BSDEs; Proportional Reinsurance Premium.

JEL Classification: G11, G22, C61.

AMS Classification: 60G55, 60J60, 91G05, 91G10, 93E20.

1. Introduction

Optimal reinsurance problems have attracted special attention during the past few years and they have been investigated in many different model settings. Insurance companies can hardly deal with all the different sources of risk in the real world, so they hedge against at least part of them, by re-insuring with other institutions. A reinsurance agreement allows the primary insurer to transfer part of the risk to another company and it is well known that this is an effective tool in risk management. Moreover, the subscription of such contracts is required by some financial regulators, see e.g. the Directive Solvency II in the European Union. Large part of the existing literature focuses mainly on classical reinsurance contracts such as the proportional and the excess-of-loss, which were extensively investigated under a variety of optimization criteria, e.g. ruin probability minimization, dividend optimization and expected utility maximization. Here we are interested in

Matteo Brachetta, Department of Mathematics, Politecnico di Milano
Giorgia Callegaro, Department of Mathematics, University of Padova
Claudia Ceci, Department of Economic Studies, University “G. D’Annunzio” of Chieti-Pescara, Viale Pindaro, 42, I-65127 Pescara, Italy.
Carlo Sgarra, Corresponding Author, Department of Mathematics, Politecnico di Milano
E-mail addresses: matteo.brachetta@polimi.it, gcallega@math.unipd.it, c.ceci@unich.it, carlo.sgarra@polimi.it.
Date: July 13, 2022.
the latter approach (see Irgens and Paulsen [21], Mania and Santacroce [28], Brachetta and Ceci [4] and references therein). Some of the classical papers devoted to the subject assume a diffusion type dynamics for the surplus process, while the more recent literature considers surplus processes including jumps.

The pioneering risk model with jumps in nonlife insurance is the classical Cramér-Lundberg model, where the claims arrival process is a Poisson process with constant intensity. This assumption implies that the instantaneous probability that an accident occurs is always constant, which is in a way too restrictive in the real world, as already motivated by Grandell [19]. In recent years, many authors made a great effort to go beyond the classical model formulation. For example, Cox processes were employed to introduce a stochastic intensity for the claims arrival process, see e.g. Albrecher and Asmussen [1], Bjork and Grandell [3], Embrechts et al. [18]. Moreover, other authors introduced Hawkes processes in order to capture the self-exciting property of the insurance risk model in presence of catastrophic events. Hawkes processes were introduced by Hawkes [20] to the end of describing geological phenomena with clustering features like earthquakes. Hawkes processes with general kernels are not Markov processes: they can eventually include long-range dependence, while Hawkes processes with exponential kernel exhibit the appealing property that the couple process-intensity is Markovian; moreover they are affine processes according to the definition provided by Duffie, Filipovic and Schachermayer [17]. For the latter literature strand here we mention Stabile and Torrisi [35] and Swishchuk et al. [37].

Dassios and Zhao [13] proposed a model which combines the two approaches by introducing a Cox process with shot noise intensity and a Hawkes process with exponential kernel for describing the claim arrival dynamics. Recently Cao, Landriault and Li [8] investigated the optimal reinsurance-investment problem in the model setting proposed by Dassios and Zhao [13] with a reward function of mean-variance type. This choice of optimization criterion raises the problem of time-inconsistency and the failing of the dynamic programming principle, so the authors attack the optimization problem by describing the decision-making process as a non-cooperative game against all strategies adopted by future players.

A different line of research related to the optimal-reinsurance investment problem focuses on the possibility that the insurer does not have access to all the information when choosing the reinsurance strategy. As a matter of fact, only the claims arrival and the corresponding disbursements are observable. In this case we need to solve a stochastic optimization problem under partial information. Liang and Bayraktar [25] were the first to introduce a partial information framework in optimal reinsurance problems. They consider the optimal reinsurance and investment problem in an unobservable Markov-modulated compound Poisson risk model, where the intensity and jump size distribution are not known, but have to be inferred from the observations of claim arrivals. Ceci, Colaneri and Cretarola [11] derive risk-minimizing investment strategies when information available to investors is restricted and they provide optimal hedging strategies for unit-linked life insurance contracts. Jang, Kim and Lee [23] present a systematic comparison between optimal reinsurance strategies in complete and partial information framework and quantify the information value in a diffusion setting.
More recently, Brachetta and Ceci [5] investigate the optimal reinsurance problem under the criterion of maximizing the expected exponential utility of terminal wealth when the insurance company has restricted information on the loss process in a model with claim arrival intensity and claim sizes distribution affected by an unobservable environmental stochastic factor. By filtering techniques (with marked point process observations), they reduce the original problem to an equivalent stochastic control problem under full information. Since the classical Hamilton-Jacobi-Bellman approach does not apply in this setting, due to the infinite dimensionality of the filter, they choose an alternative approach based on Backward Stochastic Differential Equations (henceforth BSDEs) and they characterize the value process and the optimal reinsurance strategy in terms of the unique solution to a BSDE driven by a marked point process.

In the present paper we investigate the optimal reinsurance strategy for a risk model with jump clustering properties in a partial information setting. The risk model is similar to that proposed by Dassios and Zhao [13] and it includes two different jump processes driving the claims arrivals: one process with constant intensity describing the exogenous jumps and one process with stochastic intensity describing the endogenous jumps, that exhibits self-exciting features. The externally-excited component represents catastrophic events, which generate claims clustering increasing the claim arrival intensity. The endogenous part allows us to capture the clustering effect due to self-exciting features. That is, when an accident occurs, it increases the likelihood of such events. The insurance company has only partial information at disposal, more precisely the insurer can only observe the cumulative claims process. The externally-excited component of the intensity is not observable and the insurer needs to estimate the stochastic intensity by solving a filtering problem. Our approach is substantially different from that of Cao et Al. [8] in several respects: firstly, we work in a partial information setting, secondly, the intensity of the self-excited claims arrival exhibits a slight more general dependence on the claims severity; finally, we maximize an exponential utility function instead of following a mean-variance criterion. In a partially observable framework, our goal is to characterize the value process and the optimal strategy. The optimal stochastic control problem in our case turns out to be infinite dimensional and the characterization of the optimal strategy cannot be performed by solving a Hamilton-Jacobi-Bellman equation, but via a BSDE.

A difficulty naturally arises when dealing with Hawkes processes: the intensity of the jumps is not bounded a priori, although a non-explosive condition holds. Hence we are not able to exploit some relevant bounds, which are usually required to prove a verification theorem and results on existence and uniqueness of the solution for the related BSDE. Nevertheless, we are going to show that the optimal stochastic control problem has a solution, which admits a characterization in terms of a unique solution to a suitable BSDE.

Our paper aims to contribute in different directions of the literature on optimal reinsurance problems: First, we provide a rigorous and formal construction of the dynamic contagion model. Second, we study the filtering problem associated to our problem, providing a characterization of the filter process in terms of the Kushner-Stratonovich equation and the Zakai equation as well. To the best of our knowledge, this problem has not been addressed in the existing literature. We can just
refer to Dassios and Jang [14] for a similar problem without the self-exciting component. Third, we solve the optimal reinsurance problem under the expected utility criterion.

The paper is organized as follows. In Section 2 we are going to introduce the risk model and to specify what information is available to the insurer. A rigorous mathematical construction is provided, based on a measure change approach, necessary to develop the following analysis in full details. In Section 3 the filtering problem is investigated in order to reduce the optimal stochastic control problem to a complete information setting. The stochastic differential equation satisfied by the filter is obtained, by exploiting both the Kushner-Stratonovich and the Zakai approaches. To our knowledge, this is the first equation for a filter of this kind obtained in the literature, and this is another new contribution of this paper. In Section 4 the optimal stochastic control problem is formulated, while in Section 5 a characterization of the value process associated with the optimal stochastic control problem is illustrated. Due to infinite dimension of the filter, the approach based on the Hamilton-Jacobi-Bellman equation cannot be exploited, so in Section 5 the optimal solution to the problem considered is characterized as the solution of a BSDE. In Section 6 we prove existence and uniqueness of the solution for the BSDE obtained. In Section 7 the optimal reinsurance strategy is characterized under general assumptions and the special case of proportional reinsurance under the expected value principle for the reinsurance premium is discussed. Some useful proofs and computations are collected in Appendices A and B.

2. The mathematical model

Let \((\Omega, \mathcal{F}, P; \mathbb{F})\) be a filtered probability space and assume that the filtration \(\mathbb{F} = \{\mathcal{F}_t, t \in [0, T]\}\) satisfies the usual hypotheses. The time \(T > 0\) is a finite time horizon that represents the maturity of a reinsurance contract. Here we start by giving an overview of the optimal reinsurance problem from the primary insurer’s point of view, then, in Section 2.1 we provide a rigorous construction of our model setting.

Our aim is to introduce a dynamic contagion process which generalizes the Hawkes and Cox processes with shot noise intensity introduced e.g. by Dassios and Zhao [13]. More precisely, the claims counting process \(N^{(1)}\) has the following \((P, \mathbb{F})\)-stochastic intensity, for \(t \in [0, T]\):

\[
\lambda_t = \beta + (\lambda_0 - \beta) e^{-\alpha t} + \sum_{j=1}^{N^{(1)}_t} e^{-\alpha (t - T^{(1)}_j)} \ell(Z^{(1)}_j) + \sum_{j=1}^{N^{(2)}_t} e^{-\alpha (t - T^{(2)}_j)} Z^{(2)}_j,
\]

(2.1)

where

- \(\beta > 0\) is the constant reversion level;
- \(\lambda_0 > 0\) is the initial value;
- \(\alpha > 0\) is the constant rate of exponential decay;
- \(N^{(2)}\) is a Poisson process with constant intensity \(\rho > 0\);
- \(\{T^{(1)}_n\}_{n \geq 1}\) are the jump times of \(N^{(1)}\), i.e., the time instants when claims are reported;
• \( \{T_n^{(2)}\}_{n \geq 1} \) are the jump times of \( N^{(2)} \), i.e., when exogenous/external factors make intensity jump;
• \( \{Z_n^{(1)}\}_{n \geq 1} \) represent the claim size and they are modeled as a sequence of i.i.d. \( \mathbb{R}^+ \)-valued random variables with distribution function \( F^{(1)} : (0, +\infty) \to [0, 1] \) such that \( \mathbb{E}[Z^{(1)}] < +\infty \);
• \( \ell : [0, +\infty) \to [0, +\infty) \) is a measurable function (for instance we could take \( \ell(z) = az \), \( a > 0 \), and the self-exciting jumps would be proportional to claims sizes) such that \( \mathbb{E}[\ell(Z^{(1)})] < +\infty \);
• \( \{Z_n^{(2)}\}_{n \geq 1} \) are the externally-excited jumps and they are modeled as a sequence of i.i.d. \( \mathbb{R}^+ \)-valued random variables with distribution function \( F^{(2)} : (0, +\infty) \to [0, 1] \), such that \( \mathbb{E}[Z^{(2)}] < +\infty \).

The following assumption will hold from now on:

**Assumption 2.1.** We assume \( N^{(2)} \), \( \{Z_n^{(1)}\}_{n \geq 1} \) and \( \{Z_n^{(2)}\}_{n \geq 1} \) to be independent of each other.

We define the cumulative claim process \( C_t \), \( t \in [0, T] \) at time \( t \) as

\[
C_t = \sum_{j=1}^{N_t^{(1)}} Z_j^{(1)}, \quad t \in [0, T].
\] (2.2)

**Remark 2.2.** Our model includes many meaningful properties of risk models. The claim arrival process has stochastic intensity, reflecting random changes in the instantaneous probability that accidents occur. Most importantly, our framework captures both self-exciting (endogenous) and externally-exciting (exogenous) factors, via, respectively, the claim arrival times and sizes \( \{T_n^{(1)}, Z_n^{(1)}\}_{n \geq 1} \) and \( \{T_n^{(2)}, Z_n^{(2)}\}_{n \geq 1} \). For this reason, it is well suited to describe, for instance, catastrophic events, see Cao, Landriault and Li [8]. Differently from Cao, Landriault and Li [8], where self-exciting jump sizes are independent on claims severity, in our model they depend on claim sizes: \( l(Z_j^{(1)}) \). Moreover, the decay coefficient is considered, because the catastrophic events typically exhibit this behavior.

The insurance company is allowed to subscribe a reinsurance contract with a dynamic retention level \( u_t \in [0, 1] \) \( \forall t \in [0, T] \) (the control) and for this service a reinsurance premium \( q_t^u \), \( t \in [0, T] \) must be paid. The reinsurance policy is described by a retention function \( \Phi(z, u) \), that is, under a dynamic retention level \( u_t \in [0, 1] \) the aggregate losses covered by the insurer, denoted by \( C^u_t \), read

\[
C^u_t = \sum_{j=1}^{N_t^{(1)}} \Phi(Z_j^{(1)}, u_{T_j^{(1)}}), \quad t \in [0, T]
\]

so that the remaining losses \( (C - C^u) \) will be undertaken by the reinsurer. Hence the primary insurer receives the insurance premium \( c \), pays the reinsurance premium \( q^u \) and bears the aggregate
losses $C_u$, so that the surplus process, $R_u$, follows the SDE:
\[ dR_t^u = (c_t - q_t^u) dt - dC_t^u, \quad R_0^u = R_0 \in \mathbb{R}^+, \]
where $R_0$ denotes the initial capital. Investing the surplus in a risk-free asset with interest rate $r > 0$, the total wealth $X^u$ of the primary insurer is
\[ dX_t^u = dR_t^u + rX_t^u dt, \quad X_0^u = R_0 \in \mathbb{R}^+. \]

We assume that the information at disposal is limited: the insurer only observes the cumulative claims process $C$ in Equation (2.2). Let us denote by
\[ \mathbb{H} = \mathbb{F}^C = \{ \mathcal{F}_t^C, \ t \in [0, T] \}, \mathcal{F}_t^C = \sigma \{ C_s, 0 \leq s \leq t \} \] (2.3)
the natural filtration generated by $C$. We assume that the insurer and the reinsurer have the same information represented by $\mathbb{H}$. Therefore, the insurance and the reinsurance premium have to be $\mathbb{H}$-predictable. The same applies to the insurer’s control, which is the retention level $u$. The insurer aims at maximizing the expected exponential utility of terminal wealth over a suitable class of $\mathbb{H}$-predictable strategies $\mathcal{U}$ (which will be made precise later in Definition 4.4):
\[ \sup_{u \in \mathcal{U}} \mathbb{E} \left[ 1 - e^{-\eta X_T^u} \right], \]
where $\eta > 0$ denotes the insurer’s risk aversion. More mathematical details on the control problem to be solved will be given in Section 4.

This setting leads to investigate a stochastic control problem under partial information. Due to the presence of the externally-excited component, the claim arrival intensity in Equation (2.1) is $\mathbb{F}$-adapted rather than $\mathbb{H}$-adapted, hence it is not observable by the insurance and reinsurance companies. We can reduce the original problem to a stochastic control problem under complete information by solving a filtering problem (see Section 3). The knowledge of the filter process allows to compute the $\mathbb{H}$-adapted (predictable) intensity of the claim arrival process $N^{(1)}$, which represents the best estimate of the stochastic intensity $\lambda$ based on the available information.

The next subsection provides a formal and rigorous construction of our model.

2.1. Model construction. We are going to introduce the dynamic contagion model by a suitable measure change, starting from two Poisson processes with constant intensity on a given probability space $(\Omega, \mathcal{F}, \mathbb{Q})$: $N^{(1)}$ is standard and $N^{(2)}$ has constant intensity $\rho > 0$. Moreover, we take two sequences $\{Z^{(1)}_n\}_{n \geq 1}$ and $\{Z^{(2)}_n\}_{n \geq 1}$ of i.i.d. positive random variables with distribution functions $F^{(1)}$ and $F^{(2)}$, respectively, and such that $\mathbb{E}^Q[L(Z^{(1)})] < +\infty$ and $\mathbb{E}^Q[Z^{(2)}] < +\infty$. We assume $N^{(1)}$, $N^{(2)}$, $\{Z^{(1)}_n\}_{n \geq 1}$ and $\{Z^{(2)}_n\}_{n \geq 1}$ to be independent of each other under $\mathbb{Q}$.

The key idea behind our construction is to introduce a new measure $\mathbb{P}$, equivalent to $\mathbb{Q}$ on $(\Omega, \mathcal{F}; \mathbb{F})$, such that, under $\mathbb{P}$, $N^{(1)}$ is a counting process with stochastic intensity $\lambda$ given by Equation (2.1).
Remark 2.3. The intensity of $N^{(2)}$ and the distributions of $\{Z^{(1)}_n\}_{n \geq 1}$ and $\{Z^{(2)}_n\}_{n \geq 1}$ do not change. Notice that $N^{(1)}$, $N^{(2)}$, $\{Z^{(1)}_n\}_{n \geq 1}$ and $\{Z^{(2)}_n\}_{n \geq 1}$ independent under $Q$ and not under $P$, while $N^{(2)}$, $\{Z^{(1)}_n\}_{n \geq 1}$ and $\{Z^{(2)}_n\}_{n \geq 1}$ are independent both under $P$ and under $Q$.

Let us introduce the integer valued random measures $m^{(i)}(dt, dz)$, $i = 1, 2$

$$m^{(i)}(dt, dz) = \sum_{n \geq 1} \delta_{(T^{(i)}_n, Z^{(i)}_n)}(dt, dz)\mathbb{I}_{\{T^{(i)}_n < +\infty\}},$$

where $\delta_{(t,z)}$ denotes the Dirac measure in $(t, z)$. Under $Q$, $m^{(i)}(dt, dz)$, $i = 1, 2$, are independent Poisson measures with compensator measures given respectively by

$$\nu^{(1)}(dt, dz) = F^{(1)}(dz)dt, \quad \nu^{(2)}(dt, dz) = \rho F^{(2)}(dz)dt.$$

The measure change from $(Q, \mathcal{F})$ to $(P, \mathcal{F})$ will be performed via the stochastic process $L_t$ defined as follows, for $t \in [0, T]$:

$$L_t = \mathcal{E} \left( \int_0^t \int_0^\infty (\lambda_{s^-} - 1) (m^{(1)}(ds, dz) - F^{(1)}(dz)ds) \right) = \mathcal{E} \left( \int_0^t (\lambda_{s^-} - 1)(dN^{(1)}_s - ds) \right),$$

where $\mathcal{E}(M_t)$ denotes the Doléans-Dade exponential of a martingale $M$. This process will be proved to be a $(Q, \mathcal{F})$-martingale under the following:

Assumption 2.4. We assume that there exists $\varepsilon > 0$ such that

$$\mathbb{E}^Q \left[ e^{\varepsilon t(Z^{(1)})} \right] < +\infty, \quad \mathbb{E}^Q \left[ e^{\varepsilon t(Z^{(2)})} \right] < +\infty.$$

Before proving the martingale property, we notice the following:

Remark 2.5. Let us observe that $\{\int_0^t (\lambda_{s^-} - 1)(dN^{(1)}_s - ds), t \in [0, T]\}$ is a $(Q, \mathcal{F})$-martingale, since $\mathbb{E}^Q[\int_0^t \lambda_s ds] < +\infty, \forall t \in [0, T]$. In fact, by Equation (2.1)

$$\lambda_t \leq \max \{\lambda_0, \beta\} + \sum_{j=1}^{N^{(1)}_t} \ell(Z^{(1)}_j) + \sum_{j=1}^{N^{(2)}_t} Z^{(2)}_j$$

and

$$\mathbb{E}^Q[\lambda_t] \leq \max \{\lambda_0, \beta\} + \mathbb{E}^Q[N^{(1)}_t]E^Q[\ell(Z^{(1)})] + \mathbb{E}^Q[N^{(2)}_t]E^Q[Z^{(2)}]$$

$$\leq \max \{\lambda_0, \beta\} + (E^Q[\ell(Z^{(1)})] + \rho E^Q[Z^{(2)}])t.$$
Proposition 2.6. Under Assumption 2.4 the Radon-Nikodym density process \( L \) given in Equation (2.6) is a \((Q,F)\)-martingale.

Proof. This proof is based on Sokol and Hansen [34, Corollary 2.5]. We observe that \( \lambda \) in Equation (2.1) is nonnegative, predictable and locally bounded. Hence Sokol and Hansen [34, Corollary 2.5] can be straightforwardly applied after we prove that condition (2.6) therein holds: there exists \( \varepsilon > 0 \) such that whenever \( 0 \leq u \leq t, t - u \leq \varepsilon \)

\[
\mathbb{E}^Q \left[ e^{\int_u^t \log_+ (\lambda_s - ) \, dN_s^{(1)}} \right] < +\infty,
\]

where \( \log_+ (x) := \max\{0, \log x\} \).

Applying Lemma A.1 under the measure \( Q \), we obtain that

\[
\mathbb{E}^Q \left[ e^{\int_u^t \log_+ (\lambda_s - ) \, dN_s^{(1)}} \right] \leq \mathbb{E}^Q \left[ e^{\int_u^t (\lambda_s - 1) \, ds} \right].
\]

Hence condition (2.7) is fulfilled if the expectation \( \mathbb{E}^Q \left[ e^{\int_u^t \lambda_s - \, ds} \right] \) is finite. Notice that by definition of the intensity in Equation (2.1)

\[
\lambda_t \leq \max\{\lambda_0, \beta\} + \sum_{j=1}^{N^{(1)}_t} \ell(Z^{(1)}_j) + \sum_{j=1}^{N^{(2)}_t} Z^{(2)}_j
\]

and so we have

\[
\mathbb{E}^Q \left[ e^{\int_u^t \lambda_s - \, ds} \right] \leq e^{\varepsilon \lambda_0 \vee \beta} \mathbb{E}^Q \left[ e^{\varepsilon \sum_{j=1}^{N^{(1)}_t} \ell(Z^{(1)}_j)} \cdot e^{\varepsilon \sum_{j=1}^{N^{(2)}_t} Z^{(2)}_j} \right]
\]

\[
\leq e^{\varepsilon \lambda_0 \vee \beta} \mathbb{E}^Q \left[ e^{\varepsilon \sum_{j=1}^{N^{(1)}_t} \ell(Z^{(1)}_j)} \right] \mathbb{E}^Q \left[ e^{\varepsilon \sum_{j=1}^{N^{(2)}_t} Z^{(2)}_j} \right],
\]

where we used the mutual independence of \( N^{(1)}, N^{(2)}, \{Z^{(1)}_n\}_{n \geq 1} \) and \( \{Z^{(2)}_n\}_{n \geq 1} \) which holds by construction under \( Q \). To conclude the proof, we explicitly compute the expectations above, starting from the first one (notice that the sum below starts from one, instead of zero and we implicitly use \( \ell(Z^{(1)}_0) = 0 \):
Using similar arguments, one shows that

\[
\mathbb{E}^{Q} \left[ e^{\sum_{j=1}^{N^{(1)}_{t}} \ell(Z^{(1)}_{j})} \right] = \sum_{n=1}^{+\infty} \mathbb{E}^{Q} \left[ e^{\sum_{j=1}^{n} \ell(Z^{(1)}_{j})} \right] Q(N^{(1)}_{t} = n) + Q(N^{(1)}_{t} = 0)
\]

\[
= \sum_{n=1}^{+\infty} \mathbb{E}^{Q} \left[ \prod_{j=1}^{n} e^{\ell(Z^{(1)}_{j})} \right] Q(N^{(1)}_{t} = n) + Q(N^{(1)}_{t} = 0)
\]

\[
= \sum_{n=1}^{+\infty} n \mathbb{E}^{Q} \left[ e^{\ell(Z^{(1)}_{j})} \right] Q(N^{(1)}_{t} = n) + Q(N^{(1)}_{t} = 0)
\]

\[
= e^{-t} \sum_{n=0}^{+\infty} \mathbb{E}^{Q} \left[ e^{\ell(Z^{(1)}_{j})} \right] t^{n} \frac{n!}{n!}
\]

\[
= e^{t \left( \mathbb{E}^{Q}[e^{\ell(Z^{(1)}_{j})}] - 1 \right)} < +\infty.
\]

Using similar arguments, one shows that

\[
\mathbb{E}^{Q} \left[ e^{\sum_{j=1}^{N^{(2)}_{t}} \ell(Z^{(2)}_{j})} \right] = e^{\rho \left( \mathbb{E}^{Q}[e^{\ell(Z^{(2)}_{j})}] - 1 \right)} < +\infty.
\]

Now that the change of measure has been rigorously introduced, we can safely introduce the \((P, F)\) compensator measures of \(m^{(i)}(dt, dz), i = 1, 2\).

**Remark 2.7.** By the Girsanov Theorem the \((P, F)\)-predictable projections measures (the so-called compensator measures) of \(m^{(1)}(dt, dz)\) and \(m^{(2)}(dt, dz)\) (see Equation (2.4)) are given respectively by

\[
\nu^{(1)}(dt, dz) = \lambda_{1} F^{(1)}(dz) dt, \quad \nu^{(2)}(dt, dz) = \rho F^{(2)}(dz) dt.
\]

In particular, \(N^{(1)}\) is a point process with \((P, F)\)-predictable intensity \(\{\lambda_{s}\}_{s \in [0, T]}\), while \(N^{(2)}\) remains a point process with constant \((P, F)\) intensity \(\rho > 0\).

It turns out that for any \(F\)-predictable random field \(\{H(t, z), t \in [0, T], z \in [0, +\infty)\}\) and \(i = 1, 2\)

\[
\mathbb{E} \left[ \int_{0}^{T} \int_{0}^{+\infty} H(s, z) m^{(i)}(ds, dz) \right] = \mathbb{E} \left[ \int_{0}^{t} \int_{0}^{+\infty} H(s, z) \nu^{(i)}(ds, dz) \right], \quad \forall t \in [0, T],
\]

where \(\nu^{(i)}(ds, dz), i = 1, 2\), are defined in Equation (2.8). Moreover, under the condition

\[
\mathbb{E} \left[ \int_{0}^{T} \int_{0}^{+\infty} |H(s, z)| \nu^{(i)}(ds, dz) \right] < +\infty,
\]

the process

\[
\int_{0}^{t} \int_{0}^{+\infty} H(s, z) \left( m^{(i)}(ds, dz) - \nu^{(i)}(ds, dz) \right), \quad t \in [0, T],
\]

is a \((P, F)\)-martingale.
2.2. **Markov property.** In this subsection we discuss and characterize the Markov structure of the class of processes introduced. From now on we work on the filtered probability space \((\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})\). Equation (2.1) reads as

\[
\frac{d\lambda_t}{\lambda_t} = \alpha(\beta - \lambda_t)dt + d\left(\sum_{j=1}^{N_1(t)} \ell(Z_j^{(1)})\right) + d\left(\sum_{j=1}^{N_2(t)} Z_j^{(2)}\right)
\]

\[
\begin{align*}
&= \alpha(\beta - \lambda_t)dt + \int_0^{+\infty} \ell(z)m^{(1)}(dt, dz) + \int_0^{+\infty} zm^{(2)}(dt, dz).
\end{align*}
\] (2.9)

Now we can characterize the generator of the intensity process \(\lambda\).

**Definition 2.8.** The set \(\mathcal{D}(\mathcal{L})\) denotes the class of functions \(f \in C^1(0, +\infty)\) such that

\[
\begin{align*}
\mathbb{E}\left[\int_0^t \int_0^{+\infty} |f(\lambda_s + \ell(z)) - f(\lambda_s)|\lambda_s F^{(1)}(dz)ds\right] &< +\infty, \\
\mathbb{E}\left[\int_0^t \int_0^{+\infty} |f(\lambda_s + z) - f(\lambda_s)|F^{(2)}(dz)ds\right] &< +\infty,
\end{align*}
\] (2.10)

and

\[
\mathbb{E}\left[\int_0^t \lambda_s |f'(\lambda_s)|ds\right] < +\infty.
\] (2.11)

**Proposition 2.9.** The process \(\lambda\) is a \((\mathbb{P}, \mathbb{F})\)-Markov process with generator

\[
\mathcal{L}f(\lambda) = \alpha(\beta - \lambda)f'(\lambda) + \int_0^{+\infty} [f(\lambda + \ell(z)) - f(\lambda)]\lambda F^{(1)}(dz) + \int_0^{+\infty} [f(\lambda + z) - f(\lambda)]\rho F^{(2)}(dz).
\] (2.12)

The domain of the generator \(\mathcal{L}f(\lambda)\) is \(\mathcal{D}(\mathcal{L})\).

**Proof.** By Remark 2.7 and Itô formula we get that for any \(f \in C^1(0, +\infty)\)

\[
f(\lambda_t) = f(\lambda_0) + \int_0^t \mathcal{L}f(\lambda_s)ds + m^f_t,
\] (2.13)

where \(m^f\) is the \((\mathbb{P}, \mathbb{F})\)-local martingale given by

\[
m^f_t = \int_0^t \int_0^{+\infty} [f(\lambda_{s-} + \ell(z)) - f(\lambda_{s-})](m^{(1)}(ds, dz) - \lambda^{(1)}F^{(1)}(dz))
\]

\[
+ \int_0^t \int_0^{+\infty} [f(\lambda_{s-} + z) - f(\lambda_{s-})](m^{(2)}(ds, dz) - \rho F^{(2)}(dz)ds).
\] (2.14)

Finally, \(\forall f \in \mathcal{D}(\mathcal{L})\), the stochastic process \(m^f\) in Equation (2.14) is a \((\mathbb{P}, \mathbb{F})\)-martingale and \(\mathbb{E}[\int_0^t |\mathcal{L}f(\lambda_s)|ds] < +\infty. \)

In what follows we will need the following, which will be crucial to prove Proposition 2.11.
Assumption 2.10.
\[ \mathbb{E}[(\ell(Z^{(1)})^k] < +\infty, \quad \mathbb{E}[(Z^{(2)})^k] < +\infty, \quad \forall k = 1, 2, \ldots. \]

Proposition 2.11. Under Assumption 2.10, for any \( t > 0 \)
\[ \mathbb{E} \left[ \int_0^t \lambda_s^k ds \right] < +\infty, \quad \forall k = 1, 2, \ldots. \]

Proof. We proceed by induction on \( k \). We first prove that \( \mathbb{E}[\lambda_t] \leq h_1(t), \ t \geq 0, \) with \( h_1 \) a measurable, nonnegative function such that \( \int_0^T h_1(t)dt < +\infty \). Let us observe that Equation (2.16) reads as
\[ \lambda_t = \beta + (\lambda_0 - \beta) e^{-at} + \int_0^t \int_0^{+\infty} e^{-a(t-s)}(s) m^{(1)}(ds, dz) + \int_0^t \int_0^{+\infty} e^{-a(t-s)} zm^{(2)}(ds, dz), \quad (2.15) \]
hence by Remark 2.7
\[ \mathbb{E} [\lambda_t] = \beta + \left( \lambda_0 - \beta - \frac{\rho \mathbb{E}[Z^{(2)}]}{\alpha} \right) e^{-at} + \frac{1}{\alpha} \rho \mathbb{E}[Z^{(2)}] + \mathbb{E}[\ell(Z^{(1)})] \int_0^t e^{-a(t-s)} \mathbb{E}[\lambda_s] ds. \quad (2.16) \]
By applying Gronwall’s Lemma we obtain
\[ \mathbb{E}[\lambda_t] \leq \left( \beta + (\lambda_0 - \beta - \frac{\rho \mathbb{E}[Z^{(2)}]}{\alpha}) e^{-at} + \frac{1}{\alpha} \rho \mathbb{E}[Z^{(2)}] \right) e^{\mathbb{E}[\ell(Z^{(1)})] \frac{1-e^{-at}}{-a}} = h_1(t), \quad t \geq 0. \quad (2.17) \]

It is immediate to verify \( h_1(t) \geq 0 \) and that \( \int_0^T h_1(t)dt < +\infty \). Let us assume that \( \mathbb{E}[\lambda_t] \leq h_i(t), \) with \( h_i \) a measurable, nonnegative function on \([0, +\infty) \) such that \( \int_0^T h_i(t)dt < +\infty \) for any \( i = 1, 2, \ldots k - 1 \). By Itô formula we get
\[
\lambda_t^k = \lambda_0^k + \int_0^t \alpha (\lambda - \lambda_s) k \lambda_s^{k-1} ds + \int_0^t \int_0^{+\infty} [(\lambda_s^- + \ell(z))^{k} - (\lambda_s^-)^k] m^{(1)}(ds, dz) \\
+ \int_0^t \int_0^{+\infty} [(\lambda_s^- + z)^k - (\lambda_s^-)^k] m^{(2)}(ds, dz) \\
= \lambda_0^k + \int_0^t \alpha (\lambda - \lambda_s) k \lambda_s^{k-1} ds + \int_0^t \int_0^{+\infty} \sum_{i=0}^{k-1} \binom{k}{i} (\lambda_s^-)^{i} z^{k-i} m^{(1)}(ds, dz) \\
+ \int_0^t \int_0^{+\infty} \sum_{i=0}^{k-1} \binom{k}{i} (\lambda_s^-)^{i} z^{k-i} m^{(2)}(ds, dz).
\]
Then there exist $c_i > 0, i = 1, 2, \ldots k$ such that

$$
\mathbb{E}[\lambda^k_t] = \lambda^k_0 + \int_0^t \alpha k(\beta \mathbb{E}[\lambda^{k-1}_s] - \mathbb{E}[\lambda^k_s]) ds + \int_0^t \sum_{i=0}^{k-1} \binom{k}{i} \mathbb{E}[\lambda^{i+1}_s] \mathbb{E}[\ell(Z^{(1)})^{k-i}] ds \\
+ \int_0^t \sum_{i=0}^{k-1} \binom{k}{i} \mathbb{E}[\lambda^{i}_s] \rho \mathbb{E}[(Z^{(2)})^{k-i}] ds \\
\leq \lambda^k_0 + \int_0^t \sum_{i=0}^{k-1} c_i h_i(s) ds + \int_0^t c_k \mathbb{E}[\lambda^k_s] ds,
$$

and again by Gronwall’s Lemma it follows that $\mathbb{E}[\lambda^k_t] \leq h_k(t)$, with $h_k$ a measurable, integrable and nonnegative function on $[0, T]$, and this concludes the proof. □

**Proposition 2.12.** The functions $f_k(\lambda) := \lambda^k, k = 1, 2, \ldots$ belong to $\mathcal{D}(\mathcal{L})$.

**Proof.** Under Assumption 2.10, by computations similar to those performed in the proof of Proposition 2.11 we get the claim. □

## 3. The Filtering Problem

We assume that the insurance company has a partial information because the externally-exciting component in the intensity process $\lambda$ introduced in Equation (2.1) of the claims counting process $N^{(1)}$ is not observable. For filtering of Cox processes with shot noise intensity, that is without the self-exciting component in Equation (2.1), we refer to Dassios and Jang [14], where the estimation of the intensity $\lambda$ given the observations of the claim arrival process $N^{(1)}$ reduces to the use of the classical Kalman-Bucy filter after a Gaussian approximation of the intensity is performed. This result applies in the case where the intensity $\rho$ of the externally-exciting component is sufficiently large. Their working setting can be seen as a particular case of our contagion model and their results can then be obtained as special cases, with no assumption on $\rho$ needed (see also Remark 3.7).

The insurance company aims at estimating the intensity $\lambda$ by observing the cumulative catastrophic claim process $C$ defined in Equation (2.2), that is, by observing the double sequence $\{(T_n^{(1)}, Z_n^{(1)})\}_{n \geq 1}$ of arrival times and claim sizes. This leads to a filtering problem with marked point processes observations.

Let us recall that $H = \mathbb{F} \subseteq \mathbb{C}$, defined in Equation (2.3), is the observation flow, representing the information at disposal by the insurance company. So, the estimate of the intensity $\lambda$ can be described through the filter process $\pi = \{\pi_t, t \geq 0\}$ which provides the conditional distribution of $\lambda_t$ given $\mathcal{H}_t$, for any time $t \geq 0$. More in details, the filter is the $\mathbb{H}$-càdlàg process (right-continuous with left limits) process taking values in the space of probability measures on $[0, +\infty)$ such that

$$
\pi_t(f) = \mathbb{E}[f(\lambda_t)|\mathcal{H}_t],
$$

where
for any function $f$ satisfying $\mathbb{E}[\int_0^t |f(\lambda_s)|\,ds] < +\infty$, $\forall t \geq 0$. It is easy to verify that $\{\pi_t(\lambda), t \geq 0\}$, where $\pi_t(\lambda) = \mathbb{E}[\lambda | \mathcal{H}_t]$ and $\pi_t^{-}(\lambda) = \lim_{s \to t^-} \pi_t(\lambda)$ provides the $\mathbb{H}$-predictable intensity of $N^{(1)}$.

**Remark 3.1.** For any function $f$ satisfying $\mathbb{E}[\int_0^t |f(\lambda_s)|\,ds] < +\infty$, $t \geq 0$, we have that $\mathbb{E}[\int_0^t \pi_s(f)\,ds] < +\infty$ and

$$\mathbb{E}\left[\int_0^t \pi_s(f)\,ds\right] = \mathbb{E}\left[\int_0^t f(\lambda_s)\,ds\right], t \geq 0,$$

because $\mathbb{E}[\pi_s(f)] = \mathbb{E}[f(\lambda_s)]$ by definition of the filter and Jensen’s inequality implies $|\pi_s(f)| \leq \pi_s(|f|)$.

By applying the innovation method (see for instance Brémaud [4, Chapter IV]) we will characterize the filter in terms of the so called Kushner-Stratonovich (KS henceforth) equation.

**Theorem 3.2** (Kushner-Stratonovich equation). For any $f \in \mathcal{D}(\mathcal{L})$, the filter is the unique strong solution to the filtering equation

$$\pi_t(f) = f(\lambda_0) + \int_0^t \pi_s(\mathcal{L}f)\,ds$$

$$+ \int_0^t \int_0^{+\infty} \left(\frac{\pi_{s-}(f(\lambda + \ell(z))\lambda)}{\pi_{s-}(\lambda)} - \pi_{s-}(f)\right) \left(m^{(1)}(ds, dz) - \pi_{s-}(\lambda)F^{(1)}(dz)ds\right).$$

(3.1)

**Proof.** We denote by $\hat{R}$ the $(\mathbb{P}, \mathbb{H})$-optional projection of an $\mathbb{F}$-progressively measurable process $R$ such that $\mathbb{E}[|R_t|] < +\infty$ $\forall t \geq 0$. We will use the the two-well known facts:

- For every $(\mathbb{P}, \mathbb{F})$-martingale $m$, the $(\mathbb{P}, \mathbb{H})$-optional projection $\hat{m}$ is a $(\mathbb{P}, \mathbb{H})$-martingale;
- For any $\mathbb{F}$-progressively measurable process $\Psi$ we have that

$$\int_0^t \hat{\Psi}_s\,ds - \int_0^t \hat{\Psi}_s\,ds$$

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

Taking the $(\mathbb{P}, \mathbb{H})$-optional projection in Equation (2.13), for any $f \in \mathcal{D}(\mathcal{L})$ we get

$$\hat{f}(\lambda_t) = f(\lambda_0) + \int_0^t \hat{\mathcal{L}f}(\lambda_s)\,ds + M^f_t$$

(3.3)

where $M^f$ is a $(\mathbb{P}, \mathbb{H})$-martingale. By the martingale representation theorem there exists an $\mathbb{H}$-predictable random field, $h^f = \{h^f_t(z), t \geq 0, z \in [0, +\infty)\}$, such that

$$M^f_t = \int_0^t \int_0^{+\infty} h^f_s(z) \left(m^{(1)}(ds, dz) - \pi_{s-}(\lambda)F^{(1)}(dz)ds\right)$$

(3.4)

and $\mathbb{E}\left[\int_0^t \int_0^{+\infty} |h^f_s(z)|\pi_{s-}(\lambda)F^{(1)}(dz)ds\right] < +\infty$. To derive the expression of $h^f$, we consider an $\mathbb{H}$-adapted and bounded process

$$\Gamma_t = \int_0^t \int_0^{+\infty} u(z)m^{(1)}(ds, dz)$$
with $U$ an $\mathbb{H}$-predictable bounded random field. Since $\Gamma$ is $\mathbb{H}$-adapted the following equality holds
\[
\tilde{\Gamma}_t f(\lambda_t) = \Gamma_t f(\lambda_t). \tag{3.5}
\]
By applying the product rule we get
\[
d(\Gamma_t f(\lambda_t)) = \Gamma_t - df(\lambda_t) + f(\lambda_t -)d\Gamma_t + d([\Gamma_t, f(\lambda_t)])
\]
\[
= \Gamma_t - \mathcal{L}f(\lambda_t)dt + \Gamma_t - dm_t^f + \int_0^{+\infty} f(\lambda_t -)U_t(z)m^{(1)}(dt, dz)
\]
\[
+ \int_0^{+\infty} U_t(z)[f(\lambda_t - + \ell(z)) - f(\lambda_t -)]m^{(1)}(dt, dz)
\]
\[
= \Gamma_t - \mathcal{L}f(\lambda_t)dt + \int_0^{+\infty} U_t(z)f(\lambda_t - + \ell(z))\lambda_tF^{(1)}(dz)dt + \overline{m}_t^f,
\]
where $\overline{m}$ is a $(\mathbf{P}, \mathbb{F})$-martingale. Taking the $(\mathbf{P}, \mathbb{H})$-optional projection we obtain that
\[
d(\Gamma_t f(\lambda_t)) = \left[\Gamma_t - \mathcal{L}\tilde{f}(\lambda_t) + \int_0^{+\infty} U_t(z)(\lambda_t\tilde{f}(z))F^{(1)}(dz)\right]dt + \mathcal{M}_t^f, \tag{3.6}
\]
where $\mathcal{M}$ is a $(\mathbf{P}, \mathbb{H})$-martingale. On the other hand we have that
\[
d(\Gamma_t \tilde{f}(\lambda_t)) = \Gamma_t - df(\lambda_t) + \tilde{f}(\lambda_t -)d\Gamma_t + d([\Gamma_t, \tilde{f}(\lambda_t)]) =
\]
\[
= \left[\Gamma_t - \mathcal{L}\tilde{f}(\lambda_t) + \int_0^{+\infty} U_t(z)(h_t^f(z) + \tilde{f}(\lambda_t))\lambda_tF^{(1)}(dz)\right]dt + \overline{\mathcal{M}}_t^f, \tag{3.7}
\]
where $\overline{\mathcal{M}}$ is a $(\mathbf{P}, \mathbb{H})$-martingale. By (3.5) we have that the finite variation parts in Equations (3.6) and (3.7) have to coincide
\[
\int_0^t \int_0^{+\infty} U_s(z)h_s^f(z)\lambda_sF^{(1)}(dz) = \int_0^t \int_0^{+\infty} U_s(z)[(\lambda_s\tilde{f}(z))\lambda_s - \tilde{f}(\lambda_t)\lambda_s]F^{(1)}(dz), \quad \forall t \geq 0. \tag{3.8}
\]
We select $U_t(z)$ of the form $U_t(z) = U_t \mathbb{I}_A(z) \mathbb{I}_{\{t \leq T_n^1\}}$ with $U = \{U_t, t \geq 0\}$ any bounded $\mathbb{H}$-predictable, positive process and $A \in \mathcal{B}([0, +\infty))$. With this choice we get that $\Gamma$ is bounded and $\forall A \in \mathcal{B}([0, +\infty))$ and $t \leq T_n^1$
\[
\int_A h_t^f(z)\lambda_tF^{(1)}(dz) = \int_A [(\lambda_t\tilde{f}(z))\lambda_t - \tilde{f}(\lambda_t)\lambda_t]F^{(1)}(dz)
\]
and recalling that $\lambda_t > 0, \forall t \geq 0$ (which implies $\lambda_t \geq \pi_t(\lambda) > 0 \forall t \geq 0$), we obtain that
\[
h_t^f(z) = \frac{\pi_t(\lambda + \ell(z))\lambda - \pi_t(\lambda)\lambda_t}{\pi_t(\lambda)}, \quad t \leq T_n^1. \tag{3.9}
\]
Finally since the counting process $N^{(1)}$ is not explosive we have that $T^{(1)}_n \to +\infty$ as $n \to +\infty$ and by Equations (3.3) and (3.4) we obtain that the filter is solution to the KS Equation (3.1).

It remains to prove uniqueness for this equation. As in Theorem 3.3 in Ceci and Colaner [9] we have that strong uniqueness of the solution to the KS Equation follows by uniqueness of the Filtered Martingale Problem (FMP($\tilde{\mathcal{L}}, \lambda_0, C_0$)) associated to the generator $\tilde{\mathcal{L}}$ of the pair $\{((\lambda_t, C_t))_{t \geq 0}$ for any initial condition $(\lambda_0, C_0) \in (0, +\infty) \times [0, +\infty)$. For details on Filtered Martingale Problem we refer to Kurtz and Ocone [24].

For any function $f(\lambda, C), C^1$ w.r.t. $\lambda \in (0, +\infty)$ and such that $\forall t \geq 0$

\[
\mathbb{E} \left[ \int_0^t \int_0^{+\infty} |f(\lambda_s + \ell(z), C_s + z) - f(\lambda_s, C_s)| \lambda_s F^{(1)}(dz) ds \right] < +\infty,
\]

(3.10)

\[
\mathbb{E} \left[ \int_0^t \int_0^{+\infty} |f(\lambda_s + z, C_s)) - f(\lambda_s, C_s)| F^{(2)}(dz) ds \right] < +\infty,
\]

and

\[
\mathbb{E} \left[ \int_0^t \lambda_s \left| \frac{\partial f}{\partial \lambda}(\lambda_s, C_s) \right| ds \right] < +\infty,
\]

(3.11)

the operator $\tilde{\mathcal{L}}$ is given by

\[
\tilde{\mathcal{L}}f(\lambda, C) = \alpha(\beta - \lambda) \frac{\partial f}{\partial \lambda}(\lambda, C) + \int_0^{+\infty} [f(\lambda + \ell(z), C + z) - f(\lambda, C)] \lambda F^{(1)}(dz) \]

\[+ \int_0^{+\infty} [f(\lambda + z, C) - f(\lambda, C)] \rho F^{(2)}(dz). \]

(3.12)

Next, to prove that the FMP($\tilde{\mathcal{L}}, \lambda_0, C_0$) has a unique solution we apply Theorem 3.3 in Kurtz and Ocone [24] after checking that the required hypotheses are fulfilled. First, let us observe that the martingale problem for the operator $\tilde{\mathcal{L}}$ is well posed on the space of cadlag $(0, +\infty) \times [0, +\infty)$. Furthermore, we can choose a domain $\mathcal{D}(\tilde{\mathcal{L}})$ such that for any $f \in \mathcal{D}(\tilde{\mathcal{L}})$ then $\tilde{\mathcal{L}}f \in C_h((0, +\infty) \times [0, +\infty))$. Let $\mathcal{D}(\tilde{\mathcal{L}})$ the set of functions $f \in C((0, +\infty) \times [0, +\infty))$ having compact support and $C^1$ w.r.t. $\lambda \in (0, +\infty)$. Then for any $f \in \mathcal{D}(\tilde{\mathcal{L}})$ there exists $R_f > 0$ such that

\[
|\tilde{\mathcal{L}}f(\lambda, C)| \leq \left\{ \alpha(\beta + \lambda) \left| \frac{\partial f}{\partial \lambda}(\lambda, C) \right| + 2|f| (\lambda + \rho) \right\} 1_{\{||f|| \leq R_f \}} \leq K
\]

with $K$ positive constant. Moreover, it is easy to verify that $\tilde{\mathcal{L}}f(\lambda, C)$ is a continuous function of their arguments. Finally, $\mathcal{D}(\tilde{\mathcal{L}})$ is dense in the space of continuous functions which vanish at infinity and so all hypotheses of Theorem 3.3 in Kurtz and Ocone [24] are satisfied and this concludes the proof.
The filtering Equation \((3.1)\) has a natural recursive structure in terms of the sequence \(\{T_n^{(1)}\}_{n \geq 1}\). Indeed, between two consecutive jump times, for \(t \in [T_n^{(1)}, T_{n+1}^{(1)}]\) Equation \((3.1)\) reads as

\[
\begin{align*}
\text{d}\pi_t(f) &= \pi_t(\mathcal{L}f)dt - [\pi_t(f(\lambda + \ell(z)))\lambda - \pi_t(f)f(\lambda)]F^{(1)}(dz) \\
&= \pi_t(\tilde{\mathcal{L}}f)dt - [\pi_t(\lambda f) - \pi_t(\lambda)f(\lambda)]dt,
\end{align*}
\]  

(3.13)

(3.14)

where

\[
\tilde{\mathcal{L}}f(\lambda) = \alpha(\beta - \lambda)f'(\lambda) + \int_0^{+\infty} [f(\lambda + z) - f(\lambda)]\rho F^{(2)}(dz).
\]  

(3.15)

At a jump time \(T_n^{(1)}\), we have that the value of the filter is completely determined by the knowledge of the filter \(\pi_t\), with \(t \in (T_{n-1}^{(1)}, T_n^{(1)})\) and the observed data \((T_n^{(1)}, Z_n^{(1)})\), precisely

\[
\pi_{T_n^{(1)}}(f) = \frac{\pi_{T_n^{(1)}}(\lambda f + \ell(Z_n^{(1)}))}{\pi_{T_n^{(1)}}(\lambda)}.
\]  

(3.16)

Notice that \(\tilde{\mathcal{L}}\) is the Markov generator of a shot noise Cox process, obtained taking \(\ell(z) = 0\) in Equation \((2.1)\).

**Remark 3.3.** Let us consider \(f_k(\lambda) = \lambda^k\), \(k = 1, 2, \ldots\) since

\[
\tilde{\mathcal{L}}f_k(\lambda) = \alpha(\beta - \lambda)kf_{k-1}(\lambda) + \int_0^{+\infty} [(\lambda + z)^k - \lambda^k]\rho F^{(2)}(dz)
\]

we get by Equations \((3.13)\) and \((3.16)\), that, for any \(k = 1, 2, \ldots\), between two consecutive jump times

\[
\begin{align*}
\text{d}\pi_t(f_k) &= \pi_t(\tilde{\mathcal{L}}f_k)dt - [\pi_t(f_{k+1}) - \pi_t(f_1)\pi_t(f_k)]dt \\
&= \alpha(\beta\pi_t(f_{k-1}) - \pi_t(f_k))kt + \sum_{i=0}^{k-1} \binom{k}{i} \pi_t(f_i)\rho E[(Z_n^{(2)})^{k-i}]dt - [\pi_t(f_{k+1}) - \pi_t(f_1)\pi_t(f_k)]dt
\end{align*}
\]  

(3.17)

(3.18)

and at a jump time \(T_n^{(1)}\)

\[
\pi_{T_n^{(1)}}(f_k) = \frac{\pi_{T_n^{(1)}}(\lambda(\lambda + \ell(Z_n^{(1)})))^k}{\pi_{T_n^{(1)}}(f_1)} = \sum_{i=0}^{k} \binom{k}{i} \frac{\pi_{T_n^{(1)}}(f_{i+1})\ell(Z_n^{(1)})^{k-i}}{\pi_{T_n^{(1)}}(f_1)}.
\]  

(3.19)

In particular for \(k = 1\) we have that \(\pi_t(f_1) = \pi_t(\lambda)\) provides the \((\mathbb{P}, \mathbb{H})\)-intensity of \(N^{(1)}\), and the KS equation reads as
\[ \pi_t(\lambda) = \lambda_0 + \int_0^t \pi_s(\mathcal{L}_f) ds \\
+ \int_0^t \int_0^{+\infty} \left( \frac{\pi_s((\lambda + \ell(z))\lambda)}{\pi_s(\lambda)} - \pi_s(\lambda) \right) (m^{(1)}(ds, dz) - \pi_s^{-1}(\lambda) F^{(1)}(dz) ds) \\
= \lambda_0 + \int_0^t \left[ \alpha(\beta - \pi_s(\lambda)) + \rho \mathbb{E}[Z^{(2)}] - (\pi_s(\lambda^2) - \pi_s(\lambda)^2) \right] ds \\
+ \int_0^t \int_0^{+\infty} \left[ \ell(z) + \frac{\pi_s(\lambda^2) - \pi_s^{-1}(\lambda)^2}{\pi_s(\lambda)} \right] m^{(1)}(ds, dz), \]
that is
\[ d\pi_t(\lambda) = \alpha \left( \beta + \rho \mathbb{E}[Z^{(2)}] \alpha - \pi_t(\lambda) \right) dt - (\pi_t(\lambda^2) - \pi_t(\lambda)^2) dt \\
+ \int_0^{+\infty} \ell(z) m^{(1)}(ds, dz) + \frac{\pi_t(\lambda^2) - \pi_t^{-1}(\lambda)^2}{\pi_t(\lambda)} dN_t^{(1)}. \tag{3.20} \]

Notice that the equations for \( \pi_t(f_k) \) depend on \( \pi_t(f_1), \ldots, \pi_t(f_{k+1}) \), for any \( k = 1, 2, \ldots \). Thus the \( (\mathbb{P}, \mathbb{H}) \)-predictable intensity of \( N^{(1)} \), \( \pi_t^{-1}(\lambda) = \pi_t^{-1}(f_1) \), is completely characterized by a countable systems of equations given in (3.17) and (3.19). Moreover, Equation (3.20) involves the process \( \pi_t(\lambda) \) and \( \pi_t(\lambda^2) - \pi_t(\lambda)^2 = \mathbb{E}[ (\lambda_t - \pi_t(\lambda))^2 | \mathcal{H}_t] = \text{Var}(\lambda_t | \mathcal{H}_t). \)

**Remark 3.4.** By Jensen’s inequality, since \( \pi_t(\lambda^2) \geq \pi_t(\lambda)^2 \), we get by Equation (3.20) and a comparison result that
\[ \pi_t(\lambda) \leq Y_t, \quad \mathbb{P} - \text{a.s. } \forall t \geq 0, \]
where the process \( Y \) has the same jumps of \( \pi(\lambda) \) and solves between two consecutive jumps the following SDE
\[ dY_t = \alpha (\tilde{\beta} - Y_t) dt, \]
where \( \tilde{\beta} = \beta + \frac{\rho \mathbb{E}[Z^{(2)}]}{\alpha} \). More precisely, for \( t \in [T_n^{(1)}, T_{n+1}^{(1)}) \)
\[ Y_t = \tilde{\beta} + (\pi_{T_n^{(1)}}(\lambda) - \tilde{\beta}) e^{-\alpha(t - T_n^{(1)})}. \]
Hence the filter is dominated by a process with exponential decay behavior between consecutive jump times.

Thanks to Theorem 3.2 we have characterized the filter in terms of a nonlinear stochastic equation. In our framework it is possible to describe the filter also in terms of the unnormalized filter as solution of the so-called Zakai equation which has the advantage of being linear.

By the Kallianpur-Striebel formula we get that
\[ \pi_t(f) = \frac{\mathbb{E}^Q[L_t f(\lambda_t)| \mathcal{H}_t]}{\mathbb{E}^Q[L_t | \mathcal{H}_t]} = \frac{\sigma(f)}{\sigma(t)} \tag{3.21} \]
where $Q$ is the equivalent probability measure introduced in Section 2.1. $L$ is given in Equation (2.6). The process $\sigma_t(f) = \mathbb{E}^Q[L_t f(\lambda_t) | \mathcal{H}_t]$ denotes the unnormalized filter and is a finite measure-valued $\mathbb{H}$- càdlàg process.

**Proposition 3.5. (Zakai equation)** For any $f \in \mathcal{D}(\mathcal{L})$, the unnormalized filter is the unique strong solution to the Zakai equation

$$
\sigma_t(f) = f(\lambda_0) + \int_0^t \sigma_s(\mathcal{L}f) ds + \int_0^t \int_0^{+\infty} \left( \sigma_s^{-}(\lambda f(\lambda + \ell(z))) - \sigma_s^{-}(f) \right) (m^{(1)}(ds, dz) - F^{(1)}(dz)ds).
$$

(3.22)

**Proof.** First let us observe that

$$
\sigma_t(1) = \mathbb{E}^Q[L_t | \mathcal{H}_t] = \left. \frac{dP}{dQ} \right|_{\mathcal{H}_t}.
$$

Thus the dynamics of $\sigma(1)$ can be easily obtained by considering the effect of the Girsanov change measure, that is $\sigma(1)$ is the Doléans-Dade exponential of the $(Q, \mathbb{H})$-martingale $\int_0^t (\pi^{s-}(\lambda) - 1)(dN_s^{(1)} - ds)$

$$
\sigma_t(1) = \mathcal{E} \left( \int_0^t (\pi^{s-}(\lambda) - 1)(dN_s^{(1)} - ds) \right).
$$

Hence it solves

$$
d\sigma_t(1) = \sigma_t^{-}(1)(\pi_t^{-}(\lambda) - 1)(dN_t^{(1)} - dt). \quad (3.23)
$$

By Itô’s formula we get that

$$
d\sigma_t(f) = \pi_t^{-}(f) d\sigma_t(1) + \sigma_t^{-}(1) d\pi_t(f) + d \left( \sum_{s \leq t} \Delta \pi_s(f) \Delta \sigma_s(1) \right).
$$

Taking into account Equations (3.1) and (3.25) and that

$$
d \left( \sum_{s \leq t} \Delta \pi_s(f) \Delta \sigma_s(1) \right) = \int_0^{+\infty} \sigma_t^{-}(1)(\pi_t^{-}(\lambda) - 1)\left( \frac{\pi_t^{-}(\lambda f(\lambda + \ell(z)))}{\pi_t^{-}(\lambda)} - \pi_t^{-}(f) \right) m^{(1)}(dt, dz)
$$

we get Equation (3.22). Finally as in Theorem 4.7 in Ceci and Colaneri [10] we can prove strong uniqueness for the Zakai equation from strong uniqueness of the KS-equation.

□

The Zakai equation can be written also as

$$
d\sigma_t(f) = [\sigma_t(\mathcal{L}f) - \sigma_t((\lambda - 1)f)] dt + \int_0^{+\infty} \left( \sigma_t^{-}(\lambda f(\lambda + \ell(z))) - \sigma_t^{-}(f) \right) m^{(1)}(ds, dz), \quad (3.24)
$$
where the operator $\tilde{L}$ is defined in Equation (3.15) and as the KS-equation it has a natural recursive structure in terms of the sequence $\{T_n^{(1)}\}_{n\geq 1}$. Indeed, between two consecutive jump times, for $t \in [T_n^{(1)}, T_{n+1}^{(1)}]$ it reads as
\[
d\sigma_t(f) = [\sigma_t(\tilde{L} f) - \sigma_t((\lambda - 1)f)]dt
\] (3.25)
and at a jump time $T_n^{(1)}$
\[
\sigma_{T_n^{(1)}}(f) = \sigma_{T_n^{(1)}}(\lambda f + \ell(Z_n^{(1)})).
\] (3.26)
By the linear structure of the Zakai between consecutive jumps we get a convenient expression of the filter.

**Proposition 3.6.** The following representation holds, for any $f \in \mathcal{D}(\mathcal{L})$ and $\forall n = 1, 2, \ldots$
\[
\pi_t(f) = \frac{\mathbb{E}[f(\tilde{\lambda}^n_t)e^{-\int_t^t s \tilde{\lambda}^n_u du}]_{s=T_n^{(1)}-1}}{\mathbb{E}[e^{-\int_t^t s \tilde{\lambda}^n_u du}]_{s=T_n^{(1)}-1}}, \quad t \in (T_{n-1}^{(1)}, T_n^{(1)})
\] (3.27)
where $\tilde{\lambda}^n$ is the shot noise Cox process, solution $\forall t \in (T_{n-1}^{(1)}, T_n^{(1)})$, of the SDE
\[
d\tilde{\lambda}_t^n = \alpha(\beta - \tilde{\lambda}_t^n)dt + \int_0^{+\infty} zm^{(2)}(dt, dz),
\] (3.28)
with initial law $\pi_{T_{n-1}^{(1)}}$.

**Proof.** Let $\tilde{\lambda}_t^{s,x}$ denotes the solution to Equation (3.28) with initial condition $(s, x) \in [0, +\infty) \times (0, +\infty)$. By Itô’s formula $\forall t > s$
\[
f(\tilde{\lambda}_t^{s,x}) = f(x) + \int_s^t \tilde{L} f(\tilde{\lambda}_u^{s,x}) du + M_t - M_s,
\]
with $M$ a $(\mathbb{P}, \mathbb{F})$-martingale. Setting $\gamma_t = e^{-\int_t^t s \tilde{\lambda}_u^{s,x} - 1}du$ by the product rule we obtain
\[
f(\tilde{\lambda}_t^{s,x})\gamma_t = f(x) + \int_s^t \tilde{L} f(\tilde{\lambda}_u^{s,x})\gamma_u du - \int_s^t f(\tilde{\lambda}_u^{s,x})(\tilde{\lambda}_u^{s,x} - 1)\gamma_u du + \int_s^t \gamma_u dM_u
\]
and taking the expectation we obtain
\[
\mathbb{E}[f(\tilde{\lambda}_t^{s,x})\gamma_t] = f(x) + \int_s^t \mathbb{E}[\tilde{L} f(\tilde{\lambda}_u^{s,x})\gamma_u] du - \int_s^t \mathbb{E}[f(\tilde{\lambda}_u^{s,x})(\tilde{\lambda}_u^{s,x} - 1)\gamma_u] du.
\]
Thus for any $f \in \mathcal{D}(\mathcal{L})$, $\Psi_t(s, x)(f) := \mathbb{E}[f(\tilde{\lambda}_t^{s,x})\gamma_t]$ solves Equation (3.25) and as a consequence $\Psi_t(s, x)(f)/(\Psi_t(s, x)(1))$ solves the KS-equation between two consecutive jump times given in Equation (3.13).
Finally the thesis follows by uniqueness of the KS-equation observing that
\[
\int_{0}^{+\infty} \Psi_{t}(T_{n-1}^{(1)}(x))(f)\pi_{n-1}^{(1)}(dx)
\]
coincides with the filter at jump time \(T_{n-1}^{(1)}\).

\[\Box\]

**Remark 3.7.** [Filtering of a shot noise Cox process] Taking \(\beta = 0\) and \(\ell(z) = 0\) in Equation (2.1) the claim arrival process \(N^{(1)}\) reduces to the Cox process with shot noise intensity considered in Dassios and Jang [15]. Denoting by \(L^{SN}\) the Markov generator given by
\[
L^{SN} f(\lambda) = -\alpha \lambda f'(\lambda) + \int_{0}^{+\infty} \left[ f(\lambda + z) - f(\lambda) \right] \rho F^{(2)}(dz),
\]
in this special case the KS and the Zakai equations are driven by \(N^{(1)}\) and are given by
\[
d\pi_{t}(f) = \pi_{t}(L^{SN} f)ds + \int_{0}^{+\infty} \left( \frac{\pi_{t-}(\lambda f)}{\pi_{t-}(\lambda)} - \pi_{t-}(f) \right) \left( dN_{t}^{(1)} - \pi_{t-}(\lambda)dt \right),
\]
and
\[
d\sigma_{t}(f) = \sigma_{t}(L^{SN} f)dt + \left( \sigma_{t-}(\lambda f) - \sigma_{t-}(f) \right) \left( dN_{t}^{(1)} - dt \right),
\]
respectively. In particular, the KS-equation between two consecutive jump times coincides with that in the general case in Equation (3.13) (with \(\tilde{L}\) replaced by \(L^{SN}\)) while the update at a jump time \(T_{n}^{(1)}\) (see Equation (3.16)) is given by
\[
\pi_{T_{n}^{(1)}}(f) = \frac{\pi_{T_{n}^{(1)}(\lambda f)}}{\pi_{T_{n-1}^{(1)}(\lambda)}}.
\]

Analogously, the Zakai-equation between two consecutive jump times coincides with that in the general case in Equation (3.25) (with \(\tilde{L}\) replaced by \(L^{SN}\)), while the update at a jump time \(T_{n}^{(1)}\) (see Equation (3.26)) is given by
\[
\sigma_{T_{n}^{(1)}}(f) = \sigma_{T_{n-1}^{(1)}(\lambda f)}.
\]

**4. The Reduced Optimal Control Problem Under Complete Information**

By the filtering techniques developed in Section 3, the original problem under partial information is now reduced to a complete observation stochastic control problem, which involves only processes adapted or predictable w.r.t. the filtration \(\mathbb{H}\), under \(P\). The \((\mathbb{P}, \mathbb{H})\)-predictable projection measure of \(m^{(1)}(dt, dz)\) (see Equation (2.4)) associated with the loss process \(C\) can be written in terms of the filter \(\pi\):
\[
\pi_{t-}(\lambda)F^{(1)}(dz)dt.
\]

In the sequel we shall denote by \(\tilde{m}^{(1)}(dt, dz)\) the \((\mathbb{P}, \mathbb{H})\)-compensated jump-measure
\[
\tilde{m}^{(1)}(dt, dz) = m^{(1)}(dt, dz) - \pi_{t-}(\lambda)F^{(1)}(dz)dt.
\]
We are now ready to state the analogous of Remark 2.7 in \((P, \mathbb{H})\):

**Remark 4.1.** For any \(\mathbb{H}\)-predictable random field \(\{H(t, z), t \in [0, T], z \in [0, +\infty)\}\) and for \(i = 1, 2\) the following equation holds:

\[
\mathbb{E} \left[ \int_0^t \int_0^{+\infty} H(s, z) m^{(1)}(ds, dz) \right] = \mathbb{E} \left[ \int_0^t \int_0^{+\infty} H(s, z) \pi_s - (\lambda) F^{(1)}(dz)ds \right], \quad t \in [0, T].
\]

Moreover, under the condition

\[
\mathbb{E} \left[ \int_0^T \int_0^{+\infty} |H(s, z)| \pi_s - (\lambda) F^{(1)}(dz)ds \right] < +\infty,
\]

the process

\[
\int_0^t \int_0^{+\infty} H(s, z) \tilde{m}^{(1)}(ds, dz), \quad t \in [0, T]
\]

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

The primary insurer wishes to subscribe a reinsurance contract to optimally control her wealth. The surplus process without reinsurance evolves according to the following equation:

\[
dR_t = c_t dt - \int_0^{+\infty} z m^{(1)}(dt, dz), \quad R_0 = R_0 \in \mathbb{R}^+,
\]

where \(\{c_t\}_{t \in [0, T]}\) denotes the insurance premium and \(R_0\) is the initial capital. We assume that the primary insurer subscribes a generic reinsurance contract so that, under an admissible strategy \(u \in U\) (the definition of admissibility set \(U\) will be given in Definition 4.4), she retains the amount \(\Phi(Z_j, u_T(j))\) of the \(j\)-th claim, while the remaining \(Z_j - \Phi(Z_j, u_T(j))\) is paid by the reinsurer. The function \(\Phi(z, u) : [0, +\infty) \times [0, I] \to [0, +\infty]\), for \(I > 0\), characterizes the reinsurance agreement and it is called self-insurance function. We summarize its properties in the next remark.

**Remark 4.2.** According to the classical risk theory\(^1\), here are some useful properties of the self-insurance function:

- \(\Phi\) is increasing in both the variables \(z, u\); moreover, it is continuous in \(u \in [0, I]\);
- \(\Phi(z, u) \leq z \forall u \in [0, I]\), because the retained loss is always less than or equal to the claim amount;
- \(\Phi(z, 0) = 0 \forall z \in [0, +\infty)\), because \(u = 0\) is the full reinsurance;
- \(\Phi(z, I) = z \forall z \in [0, +\infty)\), because \(u = I\) is the null reinsurance.

**Example 4.3.** The most relevant reinsurance agreements are the proportional and the excess-of-loss:

1. Under a proportional reinsurance the insurer transfers a percentage \((1 - u)\) of any future loss to the reinsurer, so \(I = 1\) and

\[
\Phi(z, u) = uz, \quad u \in [0, 1].
\]

\(^1\)See e.g. Schmidli \[32, Chapter 4\] or Schmidli \[33\].
(2) Under an excess-of-loss reinsurance policy the reinsurer covers all the losses exceeding a threshold $u$, hence $I = +\infty$ and

$$\Phi(z, u) = u \land z, \quad u \in [0, +\infty).$$

Clearly the insurer will have to pay a reinsurance premium $\{q^u_t\}_{t \in [0, T]}$, which depends on the retention level $u$. We assume that the reinsurance premium admits the following representation:

$$q^u_t(\omega) = q(t, \omega, u) \quad \forall (t, \omega, u) \in [0, T] \times \Omega \times [0, I], \quad (4.4)$$

for a given function $q(t, \omega, u)$: $[0, T] \times \Omega \times [0, I] \rightarrow [0, +\infty)$ continuous and decreasing in $u$, with partial derivative $\frac{\partial q(t, \omega, u)}{\partial u}$ continuous in $u$, where $\frac{\partial q(t, \omega, 0)}{\partial u}$ and $\frac{\partial q(t, \omega, I)}{\partial u}$ are interpreted as right and left derivatives, respectively. We assume that

$$q(t, \omega, I) = 0 \quad \forall (t, \omega) \in [0, T] \times \Omega,$$

since a null protection cannot be expensive, and

$$q(t, \omega, 0) > c_t \quad \forall (t, \omega) \in [0, T] \times \Omega,$$

in order to prevent the insurer from gaining a risk-free profit. In the following $\{q^u_t\}_{t \in [0, T]}$ will denote the reinsurance premium associated with the dynamic reinsurance strategy $\{u_t\}_{t \in [0, T]}$. Both insurance and reinsurance premia are assumed to be $\mathbb{H}$-predictable, since insurer and reinsurer have the same information. Finally, we require the following integrability condition:

$$\mathbb{E}\left[ \int_0^T q^0_t dt \right] < +\infty,$$

which ensures that both the premium $c$ and $q^u_t, \forall u \in \mathcal{U}$, satisfy

$$\mathbb{E}\left[ \int_0^T c_s ds \right] < +\infty, \quad \mathbb{E}\left[ \int_0^T q^u_s ds \right] < +\infty.$$

Summarizing, the surplus process with reinsurance evolves according to

$$dR^u_t = (c_t - q^u_t) dt - \int_0^{+\infty} \Phi(z, u_t) m^{(1)}(dt, dz), \quad R^u_0 = R_0 \in \mathbb{R}^+. \quad (4.5)$$

Let us observe that

$$\int_0^t \int_0^{+\infty} \Phi(z, u_s) \tilde{m}^{(1)}(ds, dz), \quad t \in [0, T]$$

turns out to be a $(\mathbf{P}, \mathbb{H})$-martingale, because

$$\mathbb{E}\left[ \int_0^T \int_0^{+\infty} \Phi(z, u_s) \pi_{s-}(\lambda) F^{(1)}(dz) ds \right] \leq \mathbb{E}\left[ \int_0^T \int_0^{+\infty} z \pi_{s-}(\lambda) F^{(1)}(dz) ds \right]$$

$$= \mathbb{E}\left[ Z^{(1)} \right] \mathbb{E}\left[ \int_0^T \lambda_s ds \right] < +\infty,$$

which is finite since Proposition 2.11 holds, and Remarks 3.1, 4.1 apply.
The insurance company invests its surplus in a risk-free asset with constant interest rate \( r > 0 \), so that for any reinsurance strategy \( u \in \mathcal{U} \) the wealth dynamics is

\[
\frac{dX^u_t}{dt} = dR^u_t + rX^u_t dt, \quad X^u_0 = R_0 \in \mathbb{R}^+,
\]

whose solution is given by

\[
X^u_t = R_0 e^{rt} + \int_0^t e^{r(t-s)} (c_s - q^u_s) \, ds - \int_0^t \int_0^{+\infty} e^{r(t-s)} \Phi(z, u_s) m^{(1)}(ds, dz).
\]

As announced before, the insurer aims at optimally controlling her wealth using reinsurance. More formally, she aims at maximizing the expected exponential utility of terminal wealth, that is:

\[
\sup_{u \in \mathcal{U}} \mathbb{E}[1 - e^{-\eta X^u_T}],
\]

which turns out trivially to be equivalent to the minimization problem:

\[
\inf_{u \in \mathcal{U}} \mathbb{E}[e^{-\eta X^u_T}],
\]

where \( \eta > 0 \) denotes the insurer’s risk aversion.

**Definition 4.4.** We define by \( \mathcal{U} \) the class of admissible strategies, which are all the \([0, 1]-valued\) and \( \mathbb{H} \)-predictable processes such that

\[
\mathbb{E}[e^{-\eta X^u_T}] < +\infty.
\]

Given \( t \in [0, T] \), we will denote by \( \mathcal{U}_t \) the class \( \mathcal{U} \) restricted to the time interval \([t, T]\).

Clearly, the admissible strategies must be \( \mathbb{H} \)-predictable, since they are based on the information at disposal. The next assumptions are required in the sequel.

**Assumption 4.5.** We assume that for every \( a > 0 \)

\begin{enumerate}
  \item \( \mathbb{E}[e^{a\ell(Z^{(1)})}] < +\infty \) and \( \mathbb{E}[e^{aZ^{(2)}]} < +\infty \).
  \item \( \mathbb{E}[e^{a \int_0^t q^{(i)} dt}] < +\infty \).
\end{enumerate}

**Lemma 4.6.** Under Assumption 4.5 i) for every \( a > 0 \) we have that

\[
\mathbb{E}[e^{aC_T}] < +\infty.
\]

**Proof.** See Appendix B. \( \square \)

The class of admissible strategies is non empty, as shown by the next result.

**Proposition 4.7.** Under Assumption 4.5, every \( \mathbb{H} \)-predictable process \( \{u_t\}_{t \in [0, T]} \) with values in \([0, 1]\) is admissible.

**Proof.** Thanks to Lemma 4.6 the proof is basically the same as in Bracchetta and Ceci [5, Prop. 2.2, pag. 4]. \( \square \)
5. The value process

In this section we study the value process associated to the problem in Equation (4.8). Let us introduce the Snell envelope for any \( u \in \mathcal{U} \):

\[
W_t^u \doteq \text{ess inf}_{\bar{u} \in \mathcal{U}(t,u)} \mathbb{E} \left[ e^{-\eta X_t^u} \mid \mathcal{H}_t \right], \forall t \in [0, T]
\]  

(5.1)

with \( \mathcal{U}(t, u) \) defined, for an arbitrary control \( u \in \mathcal{U} \), as the restricted class of controls almost surely equal to \( u \) over \([0, t]\)

\[
\mathcal{U}(t, u) := \left\{ \bar{u} \in \mathcal{U} : \bar{u}_s = u_s \text{ a.s. for all } s \leq t \leq T \right\}.
\]

By denoting by \( \bar{X}_t^u = e^{-rt}X_t^u \) the discounted wealth:

\[
\bar{X}_t^u = R_0 + \int_0^t e^{-rs}(c_s - q_s^u) \, ds - \int_0^t \int_0^{+\infty} e^{-rs} \Phi(z, u_s) \mu^{(1)}(ds, dz),
\]  

(5.2)

and introducing the value process as follows,

\[
V_t \doteq \text{ess inf}_{\bar{u} \in \mathcal{U}_t} \mathbb{E} \left[ e^{-\eta \int_T^t e^{r(T-s)}c_s \, ds} \mid \mathcal{H}_t \right], \forall t \in [0, T]
\]  

(5.3)

we can show that \( \forall u \in \mathcal{U} \)

\[
W_t^u = e^{-\eta X_t^u} e^{rT} V_t,
\]  

(5.4)

and, in turn,

\[
V_t = e^{\eta \bar{X}_t^u e^{rT}} W_t^1, \forall t \in [0, T].
\]  

(5.5)

Our aim is to develop a BSDE characterization for the process \( \{W_t^I\}_{t \in [0, T]} \) (the Snell envelope in Equation (5.1) associated to a null reinsurance) which also provides a complete description of the value process \( \{V_t\}_{t \in [0, T]} \) in Equation (5.3).

We first give some preliminary results.

**Proposition 5.1.** Under Assumption [4.5] i), we have that

\[
0 < M_t^{(1)} \leq W_t^I \leq M_t^{(2)}, \quad t \in [0, T],
\]  

(5.6)

where \( M^{(i)}, i = 1, 2 \), are the following \((\mathbb{P}, \mathbb{H})\)-martingales

\[
M_t^{(1)} = e^{-\eta R_0 e^{rT}} \mathbb{E} \left[ e^{-\eta \int_0^T e^{r(T-s)}c_s \, ds} \mid \mathcal{H}_t \right], \quad M_t^{(2)} = \mathbb{E} \left[ e^{\eta R_0 e^{rT}} \mid \mathcal{H}_t \right], \quad t \in [0, T].
\]  

(5.7)

Moreover,

\[
\mathbb{E} \left[ (\sup_{t \in [0, T]} W_t^I)^2 \right] < +\infty.
\]  

(5.8)
Proof. The discounted wealth in Equation (5.2) for \( u = I \) becomes

\[
X_t^I = R_0 + \int_0^t e^{-rs} c_s \, ds - \int_0^t \int_0^{\infty} e^{-rs} m^{(1)}(ds, dz),
\]

hence Equation (5.3) implies that

\[
0 \leq V_t \leq \mathbb{E} \left[ e^{-\eta X_T^I} \left| \mathcal{H}_t \right. \right] \leq \mathbb{E} \left[ e^{\eta e^{-rT}(C_T - C_t)} \left| \mathcal{H}_t \right. \right] \quad \mathbb{P} - \text{a.s.} \quad \forall t \in [0, T].
\]

By Equation (5.5) we get that

\[
W_t^I \leq e^{-\eta X_t^I} e^{rT} \mathbb{E} \left[ e^{\eta e^{-rT}(C_T - C_t)} \left| \mathcal{H}_t \right. \right] \leq \mathbb{E} \left[ e^{\eta e^{-rT}C_T} \left| \mathcal{H}_t \right. \right] \equiv M_t^{(2)} \quad \mathbb{P} - \text{a.s.} \quad \forall t \in [0, T],
\]

where \( M^{(2)} \) is a \((\mathbb{P}, \mathbb{H})\)-martingale. Moreover, we have that

\[
W_t^I = \text{ess inf}_{u \in \mathcal{U}(t, T)} \mathbb{E} \left[ e^{-\eta X_t^I} \left| \mathcal{H}_t \right. \right] \geq \mathbb{E} \left[ e^{-\eta X_t^I} \left| \mathcal{H}_t \right. \right] \geq e^{-\eta R_0 e^{rT}} \mathbb{E} \left[ e^{-\eta T^{(1)} e^{r(T-s)} c_s} ds \left| \mathcal{H}_t \right. \right] \equiv M_t^{(1)} > 0 \quad \forall t \in [0, T].
\]

To complete the proof, we observe that Doob's martingale inequality implies that

\[
\mathbb{E} \left[ (\sup_{t \in [0, T]} W_t^I)^2 \right] \leq \mathbb{E} \left[ (\sup_{t \in [0, T]} M_t^{(1)})^2 \right] \leq 4 \mathbb{E} [M_T^2] = 4 \mathbb{E} \left[ e^{2\eta e^{-rT}C_T} \right] < +\infty,
\]

which is finite according to Lemma 4.6. \( \square \)

**Proposition 5.2** (Bellman’s Optimality Principle). **Under Assumption 4.5**

i) \( \{W_t^u\}_{t \in [0, T]} \) in a \((\mathbb{P}, \mathbb{H})\)-submartingale \( \forall u \in \mathcal{U} \);

ii) \( \{W_t^{u^*}\}_{t \in [0, T]} \) in a \((\mathbb{P}, \mathbb{H})\)-martingale if and only if \( u^* \in \mathcal{U} \) is an optimal control.

**Proof.** The proof follows the same lines of Brachetta and Ceci [5, Proposition 3.2]. \( \square \)

The following definitions will play a key role for our BSDE characterization and its solution.

**Definition 5.3.** We define these classes of stochastic processes:

- \( \mathcal{S}^2 \) denotes the space of càdlàg \( \mathbb{H} \)-adapted processes \( Y \) such that:
  \[
  \mathbb{E} \left[ (\sup_{t \in [0, T]} |Y_t|)^2 \right] < +\infty.
  \]

- \( \mathcal{L}^2 \) denotes the space of càdlàg \( \mathbb{H} \)-adapted processes \( Y \) such that:
  \[
  \mathbb{E} \left[ \int_0^T |Y_t|^2 \, dt \right] < +\infty.
  \]
\( \hat{L}^2 \) denotes the space of \([0, +\infty)\)-indexed \( \mathbb{H} \)-predictable random fields \( \Theta = \{ \Theta_t(z), t \geq 0, z \in [0, +\infty) \} \) such that:

\[
\mathbb{E} \left[ \int_0^T \int_0^{+\infty} \Theta_t(z) \pi_{t-}(\lambda) F^{(1)}(dz) \, dt \right] < +\infty.
\]

**Definition 5.4.** We define

\[
\mathcal{M} = \{(t, \omega, y, \theta(\cdot)) : (t, \omega, y) \in [0, T] \times \Omega \times [0, +\infty) \text{ and } \theta : [0, +\infty) \to \mathbb{R} \text{ measurable}\}.
\]

and, similarly, we denote by \( \mathcal{M}^u \) the same set augmented with the variable \( u \in [0, 1] \), i.e.

\[
\mathcal{M}^u = \{(t, \omega, y, \theta(\cdot), u) : (t, \omega, y, u) \in [0, T] \times \Omega \times [0, +\infty) \times [0, 1] \text{ and } \theta : [0, +\infty) \to \mathbb{R} \text{ measurable}\}.
\]

**Definition 5.5.** Let \( \xi \) be an \( \mathcal{H}_T \)-measurable random variable. A solution to a BSDE driven by the compensated random measure \( \tilde{m}^{(1)}(dt, dz) \) given in Equation (1.2) and generator \( g \) is a pair \((Y, \Theta Y) \in \mathcal{L}^2 \times \hat{L}^2 \) such that

\[
Y_t = \xi + \int_t^T g(s, Y_s, \Theta_s Y(\cdot))ds - \int_t^T \int_0^{+\infty} \Theta_s Y(z) \tilde{m}^{(1)}(ds, dz), \quad t \in [0, T], \ P \text{-a.s.},
\]

where \( g(t, \omega, y, \theta(\cdot)) \) is a real-valued function on \( \mathcal{M} \) which is \( \mathbb{H} \)-predictable w.r.t. \((t, \omega) \in [0, T] \times \Omega\).

Now we are ready to prove the main result of this section.

**Theorem 5.6.** Let \( u^* \in \mathcal{U} \) be an optimal control. Under Assumption 4.3, \((W^I, \Theta^{W^I}) \in S^2 \times \hat{L}^2 \) satisfies the following BSDE:

\[
W_t^I = \xi - \int_t^T \int_0^{+\infty} \Theta_t^{W^I}(z) \tilde{m}^{(1)}(ds, dz) - \int_t^T \text{ess sup}_{u \in \mathcal{U}} \tilde{f}(s, W_s^I, \Theta_s^{W^I}(z), u_s) \, ds, \quad (5.9)
\]

with terminal condition \( \xi = e^{-nX_I^I} \), i.e., in differential form

\[
dW_t^I = \int_0^{+\infty} \Theta_t^{W^I}(z) \tilde{m}^{(1)}(dt, dz) + \text{ess sup}_{u \in \mathcal{U}} \tilde{f}(t, W_t^I, \Theta_t^{W^I}(z), u_t) \, dt \quad (5.10)
\]

where

\[
\tilde{f}(t, W_t^I, \Theta_t^{W^I}(\cdot), u_t) = -W_t^I \eta e^{r(t-T)} q_t^u - \int_0^{+\infty} [W_t^{I \perp} + \Theta_t^{W^I}(z)] \left[ e^{-n\lambda(T-t)} (z, \Phi(z, u_t)) - 1 \right] \pi_{t-}(\lambda) F^{(1)}(dz). \quad (5.11)
\]

Moreover, \( u^* \) is such that

\[
\tilde{f}(t, W_t^I, \Theta_t^{W^I}(\cdot), u_t^*) = \text{ess sup}_{u \in \mathcal{U}} \tilde{f}(t, W_t^I, \Theta_t^{W^I}(\cdot), u_t) \quad \forall t \in [0, T]. \quad (5.12)
\]

**Proof.** We proceed as in Brachetta and Ceci \[5\], Lemma 3.1 and Proposition 3.3]. Notice that Brachetta and Ceci \[5\], Lemma 3.1] does not directly apply, because the key hypothesis of boundedness for the claim arrival intensity is not verified in our framework. We split the proof in three steps.
Step 1. By Proposition 5.2 since $u = I \in \mathcal{U}$, $\{W^I_t\}_{t \in [0,T]}$ in a $(\mathbb{P}, \mathbb{H})$-submartingale. As a consequence, by Doob-Meyer decomposition and $\mathbb{H}$-martingale representation theorems, it admits the expression

$$W^I_t = \int_0^t \int_0^\infty \Theta^I_s(z) \tilde{m}^{(1)}(ds,dz) + A_t,$$

where $\Theta^I \in \mathcal{L}^2$ by (5.8) and $\{A_t\}_{t \in [0,T]}$ in an increasing $\mathbb{H}$-predictable process such that $\mathbb{E} \left[ \int_0^T A_t^2 ds \right] < +\infty$.

Step 2. Following the same computations as in Brachetta and Ceci [5, Lemma 3.1], we get the following representation for the Snell envelope

$$dW^u_t = dM^u_t + e^{\eta(X^I_t - X^u_t)}e^{rT} [A_t - \bar{f}(t, W^I_t, \Theta^I_t(\cdot), u_t)] dt,$$

where

$$M^u_t = \int_0^t e^{\eta(X^I_s - X^u_s)}e^{rT} \int_0^\infty \Theta^I_s(z) e^{-\eta e^{r(T-s)}(z-\Phi(z,u_s))} \tilde{m}^{(1)}(ds,dz)$$

$$+ \int_0^t W^I_t e^{\eta(X^I_s - X^u_s)}e^{rT} \int_0^\infty \left( e^{-\eta e^{r(T-s)}(z-\Phi(z,u_s))} - 1 \right) \tilde{m}^{(1)}(ds,dz), \quad t \in [0,T].$$

Step 3. The statement now follows as in Brachetta and Ceci [5, Proposition 3.3] after verifying that for any $u \in \mathcal{U}$ the process $\{M^u_t\}_{t \in [0,T]}$ is a $(\mathbb{P}, \mathbb{H})$-martingale. To this end, it is sufficient to prove that the following two conditions hold

$$\mathbb{E} \left[ \int_0^T e^{\eta(X^I_t - X^u_t)}e^{rT} \int_0^\infty \left| \Theta^I_t(z) e^{-\eta e^{r(T-t)}(z-\Phi(z,u_t))} \pi_t(\lambda) F^{(1)}(dz) dt \right| < +\infty,$$

$$\mathbb{E} \left[ \int_0^T W^I_t e^{\eta(X^I_t - X^u_t)}e^{rT} \int_0^\infty \left| e^{-\eta e^{r(T-t)}(z-\Phi(z,u_t))} - 1 \right| \pi_t(\lambda) F^{(1)}(dz) dt \right] < +\infty.$$

First, recalling Equation (5.2), we notice that

$$\tilde{X}^I_t - \tilde{X}^u_t = \int_0^t e^{-rs} q^u_s ds - \int_0^t \int_0^\infty e^{-rs} (z - \Phi(z,u_s)) \tilde{m}^{(1)}(ds,dz).$$
Using \( \Phi(z, u_t) \leq z \), the well known inequality \( 2ab \leq a^2 + b^2 \) \( \forall a, b \in \mathbb{R} \) and Jensen’s inequality, the first expectation above is dominated by

\[
\mathbb{E} \left[ e^{\eta e^T} \int_0^T e^{-rt} q_t^0 \, dt \int_0^T \int_0^{+\infty} |\Theta_t^{W^I}(z)| \pi_t(\lambda) F^{(1)}(dz) \, dt \right]
\]

\[
\leq \frac{1}{2} \left\{ \mathbb{E} \left[ e^{2\eta e^T} \int_0^T e^{-rt} q_t^0 \, dt \int_0^T \pi_t(\lambda) \, dt \right] + \mathbb{E} \left[ \int_0^T \int_0^{+\infty} |\Theta_t^{W^I}(z)|^2 \pi_t(\lambda) F^{(1)}(dz) \, dt \right] \right\}
\]

\[
\leq \frac{1}{4} \mathbb{E} \left[ e^{4\eta e^T} \int_0^T e^{-rt} q_t^0 \, dt \right] T + \frac{1}{4} \mathbb{E} \left[ \int_0^T \pi_t^2(\lambda) \, dt \right] + \frac{1}{2} \mathbb{E} \left[ \int_0^T \int_0^{+\infty} |\Theta_t^{W^I}(z)|^2 \pi_t(\lambda) F^{(1)}(dz) \, dt \right]
\]

\[
\leq \frac{1}{4} \mathbb{E} \left[ e^{4\eta e^T} \int_0^T e^{-rt} q_t^0 \, dt \right] T + \frac{1}{4} \mathbb{E} \left[ \int_0^T \pi_t^2(\lambda) \, dt \right] + \frac{1}{2} \mathbb{E} \left[ \int_0^T \int_0^{+\infty} |\Theta_t^{W^I}(z)|^2 \pi_t(\lambda) F^{(1)}(dz) \, dt \right]
\]

\[
< +\infty,
\]

which is finite because of Assumption 4.5 ii), Remark 3.1, Proposition 2.11 and recalling that \( \Theta^{W^I} \in \tilde{\mathcal{L}}^2 \) (see Step 1). The second expectation satisfies

\[
\mathbb{E} \left[ e^{\eta e^T} \int_0^T e^{-rt} q_t^0 \, dt \int_0^T W_t^I \pi_t(\lambda) \, dt \right]
\]

\[
\leq \frac{1}{2} \mathbb{E} \left[ \int_0^T |W_t^I|^2 \, dt \right] + \frac{1}{4} \mathbb{E} \left[ e^{4\eta e^T} \int_0^T e^{-rt} q_t^0 \, dt \right] T + \frac{1}{4} \mathbb{E} \left[ \int_0^T \pi_t^4(\lambda) \, dt \right] < +\infty,
\]

where the first term is finite because \( W^I \in \mathcal{S}^2 \subseteq \mathcal{L}^2 \), the second is finite by Assumption 4.5 ii) and the third follows by Remark 3.1 and Proposition 2.11.

\[ \square \]

**Remark 5.7.** Since

\[
X_t^I = R_0 e^{rT} + \int_0^T e^{r(T-t)} c_T \, dt - \int_0^T \int_0^{+\infty} e^{r(T-t)} zm^{(1)}(dt, dz),
\]

we get the inequality

\[
W_t^I = \xi = e^{-nX_t^I} \leq e^{\eta e^T} c_T,
\]

and Lemma 4.6 guarantees that the terminal condition of the BSDE (5.9) is a random variable with finite moments of any order.

**Remark 5.8.** Let us notice that the driver of the BSDE (5.9) is always nonnegative, since recalling Equation (5.11), we easily get that

\[
\text{ess sup}_{u \in \mathcal{U}} \tilde{f}(t, W_t^I, \Theta_t^{W^I}(\cdot), u_t) \geq \tilde{f}(t, W_t^I, \Theta_t^{W^I}(\cdot), I) = 0.
\]

We wish to provide a Verification result in terms of the BSDE (5.9). To this end we recall the following result Brachetta and Ceci [5, Proposition 3.4].

**Proposition 5.9.** Suppose there exists an \( \mathbb{H} \)-adapted process \( D \) such that:
• $D = \{D_t e^{-\eta \bar{X}^I_t e^{-rT}}\}_{t \in [0,T]}$ is an $\mathbb{F}$-sub-martingale for any $u \in \mathcal{U}$ and an $\mathbb{F}$-martingale for some $u^* \in \mathcal{U}$;

• $D_T = 1$.

Then $D_t = V_t$ and $u^*$ is an optimal control.

**Theorem 5.10.** (Verification Theorem) Under Assumption 4.3, let $(Y, \Theta^Y) \in \mathcal{L}^2 \times \hat{\mathcal{L}}^2$ be a solution to the BSDE (5.9) and let $u^* \in \mathcal{U}$ be the process satisfying Equation (5.12). Then $Y$ coincides with $W^I$,

$$V_t = e^{\eta \bar{X}^I_t e^{-rT}}Y_t \quad \forall t \in [0,T],$$

and $u^*$ is an optimal control.

**Proof.** Let $(Y, \Theta^Y) \in \mathcal{L}^2 \times \hat{\mathcal{L}}^2$ be a solution to the BSDE (5.9) and define $D_t = e^{\eta \bar{X}^I_t e^{-rT}}Y_t$. In order to replicate the proof of Brachetta and Ceci [5, Proposition 3.4] we need to verify that for any $u \in \mathcal{U}$ the process $\{\tilde{M}_t^u\}_{t \in [0,T]}$, with

$$\tilde{M}_t^u = \int_0^t e^{\eta (\bar{X}^I_s - \bar{X}^u_s) e^{-rT}} \int_0^{+\infty} \Theta^Y_s(z) e^{-\eta e^{r(T-s)}(z - \Phi(z,u_s))} \tilde{m}^{(1)}(ds, dz)$$

$$+ \int_0^t Y_s e^{\eta (\bar{X}^I_s - \bar{X}^u_s) e^{-rT}} \int_0^{+\infty} \left(e^{-\eta e^{r(T-s)}(z - \Phi(z,u_s))} - 1\right) \tilde{m}^{(1)}(ds, dz), \quad t \in [0,T],$$

is an $\mathbb{H}$-martingale. This can be proved as in Step 2. of Theorem 5.6 because $(Y, \Theta^Y) \in \mathcal{L}^2 \times \hat{\mathcal{L}}^2$.

Then we can mimic the proof of Brachetta and Ceci [5, Theorem 3.1], getting that

$$\{D_t e^{-\eta \bar{X}^I_t e^{-rT}}\}_{t \in [0,T]}$$

is a $(\mathbf{P}, \mathbb{H})$-sub-martingale $\forall u \in \mathcal{U}$ and $\{D_t e^{-\eta \bar{X}^I_t e^{-rT}}\}_{t \in [0,T]}$ is a $(\mathbf{P}, \mathbb{H})$-martingale.

Finally $D_T = e^{\eta \bar{X}^I_T e^{-rT}}Y_T = e^{\eta \bar{X}^I_T e^{-\eta \bar{X}^I_T}} = 1$ and the statement follows by Proposition 5.9. \qed

6. Existence and Uniqueness of the Solution to the BSDE

We proved in Theorem 5.6 that the process $(W^I, \Theta^{W^I})$ is a solution to BSDE (5.9) under the assumption that there exists an optimal control. Next, Verification Theorem 5.10 states that any solution to BSDE (5.9) necessarily coincide with the process $(W^I, \Theta^{W^I})$. Therefore we need a result of existence of solution to BSDE (5.9) in order to completely characterize $(W^I, \Theta^{W^I})$ as the unique solution to this BSDE and so to characterize the value process via Equation (5.5).

So, in this section we apply to our setting the result in [29, Theorem 3.5], which provides a rigorous framework to prove existence and uniqueness for the solution to multidimensional BSDEs driven by a general martingale, under the assumption of a stochastic Lipschitz generator.

**Theorem 6.1.** Under Assumption 4.3, there exists a unique solution $(Y, \Theta^Y) \in \mathcal{L}^2 \times \hat{\mathcal{L}}^2$ to the BSDE (5.9), i.e.,

$$Y_t = \xi - \int_t^T \int_0^{+\infty} \Theta^Y_s(z) \tilde{m}^{(1)}(ds, dz) + \int_t^T f(s, Y_s, \Theta^Y_s(\cdot)) ds,$$
with generator \( f : \mathbb{M} \rightarrow [0, +\infty) \)
\[
\begin{align*}
    f(s, y, \theta(\cdot)) &= - \text{ess sup}_{u \in \mathcal{U}} \tilde{f}(s, y, \theta(\cdot), u_s) \\
    &= \text{ess sup}_{u \in \mathcal{U}} \{-y \eta e^{r(T-s)} q^u_s \\
    &\quad - \int_0^{+\infty} (y + \theta(z)) \left[ e^{-\eta e^{r(T-s)}(z - \Phi(z, u_s))} - 1 \right] \pi_{s+}(\lambda) F^{(1)}(dz)\} ,
\end{align*}
\]
with \( \mathbb{M} \) given in Definition 5.4 and terminal condition \( \xi = e^{-\eta x^f_T} \).

**Proof.** In order to apply Papapantoleon, Possamai and Saplaouras [29, Theorem 3.5] we start by verifying that the BSDE data are standard under \( \hat{\beta} \), i.e., that assumptions (F1) – (F5) therein are satisfied for a \( \hat{\beta} > 0 \). We will show that in our setting any \( \beta > 0 \) works fine (see (F4) below).

(F1) The process \( \{\tilde{C}_t\}_{t \in [0,T]} \), with
\[
\tilde{C}_t = \int_0^t \int_0^{+\infty} z \tilde{m}^{(1)}(ds, dz),
\]
is a \((P, \mathcal{H})\)-martingale because of Remark 4.1. Notice that \( \tilde{C} \) is a pure-jump martingale, since the Brownian part is absent. Moreover,
\[
\mathbb{E}[\tilde{C}_t^2] = \mathbb{E} \left[ \int_0^t \int_0^{+\infty} z^2 \pi_{s-}(\lambda) F^{(1)}(dz) \right] = \mathbb{E} \left[ (Z^{(1)})^2 \right] \mathbb{E} \left[ \int_0^t \pi_{s-}(\lambda) ds \right],
\]
which is finite for every \( t \in [0, T] \) according to Remark 3.1. Hence \( \sup_{t \in [0,T]} \mathbb{E}[\tilde{C}_t^2] < +\infty \) and Papapantoleon, A., Possamai, D. and Saplaouras [29, Assumption 2.10] is satisfied. In particular, the disintegration property is fulfilled according to Equation (4.1), where the transition kernel \( K^\omega \) on \((\Omega \times [0, T], \mathcal{P})\) (here \( \mathcal{P} \) denotes the \( \mathcal{H}\)-predictable sigma-field on \( \Omega \times [0, T] \)) is
\[
K^\omega_t(dz) = \pi_{t-}(\lambda) F^{(1)}(dz).
\]
(F2) Lemma 4.6 guarantees that the terminal condition \( \xi = e^{-\eta x^f_T} \) has finite moments of any order. See also (F4) below for additional details.

(F3) We need to prove that the generator \( f \) satisfies a stochastic Lipschitz condition, i.e., there exist two positive \( \mathcal{H}\)-predictable processes \( \gamma, \bar{\gamma} \) such that on \( \mathbb{M} \)
\[
\left| f(t, \omega, y, \theta(\cdot)) - f(t, \omega, y', \theta'(\cdot)) \right|^2 \leq \gamma_t(\omega) |y - y'|^2 + \bar{\gamma}_t(\omega) (|||\theta(\cdot) - \theta'(\cdot)|||_{L^2}(\omega))^2 ,
\]
where:
\[
(|||\theta(\cdot)|||_{L^2}(\omega))^2 \geq \int_0^{+\infty} \theta^2(z) K^\omega_t(dz) \geq 0.
\]
Exploiting the definition of \( f \) in Equation (6.1), we need first of all to deal with the ess sup:
\[
\left| f(t, \omega, y, \theta(\cdot)) - f(t, \omega, y', \theta'(\cdot)) \right|^2 \leq \left( \text{ess sup}_{u \in [0, T]} |f(t, \omega, y, \theta(\cdot), u) - f(t, \omega, y', \theta'(\cdot), u)| \right)^2 ,
\]
and we preliminarily work on the absolute value difference involving $\tilde{f}$:

$$
\left| \tilde{f}(t, \omega, y, \theta(\cdot), u) - \tilde{f}(t, \omega, y', \theta'(\cdot), u) \right|
= \left| (y - y') \eta e^{r(T-t)} q_t^u(\omega) + \int_0^{\infty} (y - y' + \theta(z) - \theta'(z)) \left( e^{-\eta e^{r(T-t)}(z - \Phi(z,u))} - 1 \right) K_t^\omega(dz) \right|
\leq \left| y - y' \right| \eta e^{r(T-t)} q_t^0(\omega) + \int_0^{\infty} |y - y'| K_t^\omega(dz) + \int_0^{\infty} |\theta(z) - \theta'(z)| K_t^\omega(dz)
$$

where we have used the boundedness of $|e^{-\eta e^{r(T-t)}(z - \Phi(z,u))} - 1|$ and we have taken into account that $q_t^u$ is a nonnegative decreasing function of $u$. Now, since the inequality above does not depend on $u$ we also have that the $\text{ess sup}_{u \in [0,1]}$ satisfies it and we can take its square (we use here the trivial relation $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$), finding:

$$
\left( \text{ess sup}_{u \in [0,1]} \left| \tilde{f}(t, \omega, y, \theta(\cdot), u) - \tilde{f}(t, \omega, y', \theta'(\cdot), u) \right| \right)^2
\leq 3 \left| y - y' \right|^2 \eta^2 e^{2r(T-t)} (q_t^0(\omega))^2 + 3 \left( \int_0^{\infty} |y - y'| K_t^\omega(dz) \right)^2 + 3 \left( \int_0^{\infty} |\theta(z) - \theta'(z)| K_t^\omega(dz) \right)^2.
$$

Recalling now that the transition kernel reads $K_t^\omega(dz) = \pi_{t-}(\lambda) F^{(1)}(dz)$ we use the following, for an integrable function $\vartheta$:

$$
\left( \int_0^{\infty} |\vartheta(\omega, z)| \pi_{t-}(\lambda) F^{(1)}(dz) \right)^2 \leq \int_0^{\infty} |\vartheta(\omega, z)|^2 \pi_{t-}^2(\lambda) F^{(1)}(dz) \cdot \int_0^{\infty} F^{(1)}(dz).
$$

So, we find:

$$
\left( \text{ess sup}_{u \in [0,1]} \left| \tilde{f}(t, \omega, y, \theta(\cdot), u) - \tilde{f}(t, \omega, y', \theta'(\cdot), u) \right| \right)^2
\leq 3 \left| y - y' \right|^2 \eta^2 e^{2r(T-t)} (q_t^0(\omega))^2 + 3 \left| y - y' \right|^2 \pi_{t-}^2(\lambda) + 3 \int_0^{\infty} \left| \theta(z) - \theta'(z) \right|^2 \pi_{t-}(\lambda) K_t^\omega(dz)
= 3 \left| y - y' \right|^2 \left( \eta^2 e^{2r(T-t)} (q_t^0(\omega))^2 + \pi_{t-}^2(\lambda) \right) + 3 \int_0^{\infty} \left| \theta(z) - \theta'(z) \right|^2 \pi_{t-}(\lambda) K_t^\omega(dz).
$$

So, the target, being Equation (6.3), is reached and we have the following values for the stochastic Lipschitz coefficients $\gamma_t$ and $\tilde{\gamma}_t$:

$$
\gamma_t = 3 \eta^2 e^{2r(T-t)} (q_t^0(\omega))^2 + 3 \pi_{t-}^2(\lambda),
\tilde{\gamma}_t = 3 \pi_{t-}(\lambda),
$$

which, as expected, are independent of the control $u$.

**F4** Since by definition $\alpha^2 = \max\{\sqrt{\gamma}, \tilde{\gamma}\}$, here we find:

$$
\alpha^2 = \max \left\{ \sqrt{3 \eta^2 e^{2r(T-t)} (q_t^0(\omega))^2 + 3 \pi_{t-}^2(\lambda)}, 3 \pi_{t-}(\lambda) \right\}
$$
and also $A_t = \int_0^t \alpha_s^2 \, ds$, so that we can easily verify that the inequality $\Delta A_t \leq \Phi, P$-a.s. holds true for any $\Phi > 0$ since $A$ has no jumps. Notice that (F2) requires that the terminal condition $\xi = e^{-\eta X_T}$ belongs to the set of $\mathcal{H}_T$-measurable random variables such that $E\left[e^{\beta A_T} e^{-2\eta X_T} \right] < \infty$, for some $\beta > 0$. This is true for any $\beta > 0$, since $\alpha_s^2 \leq \sqrt{3} \eta e^{r(T-s)} q_0^s + 3\pi_-(\lambda)$ and so

$$E\left[e^{\beta A_T} e^{-2\eta X_T} \right] \leq E\left[e^{\beta \sqrt{3} \eta e^{r(T-s)} q_0^s} \int_0^T e^{-2\eta X_T} \right],$$

which is finite for any $\beta > 0$ thanks to Assumption 4.3 (ii) (see also Lemma B.1).

(F5) Finally, by using the same $\beta > 0$ and $A$ introduced to prove (F4), we find:

$$E \left[ \int_0^T e^{\beta A_t} \frac{|f(t, 0, 0, 0)|^2}{\alpha_t^2} \, dt \right] < \infty, \quad (6.5)$$

since here $f(t, 0, 0, 0) = -\text{ess sup}_{u \in U} \tilde{f}(t, 0, 0, u_t) = 0$.

It now remains to prove that the quantity

$$M^\Phi(\beta) = \frac{9}{\beta} + \frac{\Phi^2(2 + 9\beta)}{\sqrt{\beta^2 \Phi^2 + 4} - 2} \exp \left( \frac{\beta \Phi + 2 - \sqrt{\beta^2 \Phi^2 + 4}}{2} \right)$$

with $\Phi > 0$ introduced in (F4) and $\beta > 0$, satisfies $M^\Phi(\beta) < \frac{1}{2}$. Thanks to Papapantoleon, Possamai and Saplaouras [29, Lemma 3.4], for $\beta$ sufficiently large, we know that since $\lim_{\beta \to \infty} M^\Phi(\beta) = 9 e \Phi$ then it suffices to take $\Phi < \frac{1}{18e}$.

The last part of the proof is devoted to showing that $(Y, \Theta Y) \in \mathcal{L}^2 \times \hat{\mathcal{L}}^2$. According to the quoted result Papapantoleon, Possamai and Saplaouras [29, Theorem 3.5], we know that

$$E \left[ \int_0^T e^{\beta A_t} |Y_t|^2 \, dt \right] < +\infty$$

and we also notice that $\alpha_t^2 \geq 3\pi_-(\lambda) \geq 3 \min\{\lambda_0, \beta\}$ and this implies

$$E \left[ \int_0^T e^{\beta A_t} |Y_t|^2 \, dt \right] < +\infty$$

and therefore also

$$E \left[ \int_0^T |Y_t|^2 \, dt \right] < +\infty,$$

which means, by definition, that $Y \in \mathcal{L}^2$. The same argument applies to prove that $\Theta Y \in \hat{\mathcal{L}}^2$.

Finally exploiting the Verification Theorem 5.10 we have that the unique solution to BSDE (5.9) coincides with the process $(W^I, \Theta^W)$. \qed
7. The optimal reinsurance strategy

The aim of this section is to provide more insight into the structure of the optimal reinsurance strategy and investigate the special case of proportional reinsurance under the expected value principle for the reinsurance premium.

By the Verification Theorem [5.11] if \((W^I, \Theta^{W^I}) \in \mathcal{L}^2 \times \hat{\mathcal{L}}^2\) is the solution to the BSDE \((5.9)\), then the optimal strategy is the maximizer satisfying Equation \((5.12)\). Hence, exploiting the expression in Equation \((4.4)\), we look for the maximizer of the function \(\tilde{f}\) w.r.t. to the variable \(w\) in order to obtain some definite results we need to introduce a concavity hypothesis for the function \(f\).

The following general theorem provides a characterization of the optimal reinsurance strategy. In order to obtain some definite results we need to introduce a concavity hypothesis for the function \(\tilde{f}\) w.r.t. to the variable \(u \in [0, I]\).

**Proposition 7.1.** Under Assumption 4.5 let \((W^I, \Theta^{W^I}) \in \mathcal{L}^2 \times \hat{\mathcal{L}}^2\) be the solution to the BSDE \((5.9)\). Suppose moreover that \(\Phi(z, u)\) is differentiable in \(u \in [0, I]\) and \(\tilde{f}\) given in Equation \((7.1)\) is strictly concave in \(u \in [0, I]\). Then the optimal reinsurance strategy is \(u_\ast^t = \{\hat{u}(t, W^I_{t-}, \Theta^{W^I}_t(\cdot))\}_{t \in [0, T]}\), where the function \(\hat{u}\) is given by the following expression:

\[
\hat{u}(t, \omega, w, \theta(\cdot)) = \begin{cases} 
0 & (t, \omega, w, \theta(\cdot)) \in R_0 \\
\hat{u}(t, \omega, w, \theta(\cdot)) & \mathbb{M}\backslash(R_0 \cup R_1) \\
I & (t, \omega, w, \theta(\cdot)) \in R_1,
\end{cases}
\]

where we define the two regions

\[
R_0 = \left\{ (t, \omega, w, \theta(\cdot)) \in \mathbb{M} : \frac{\partial \tilde{f}(t, \omega, w, \theta(\cdot), 0)}{\partial u} < 0 \right\},
\]

\[
R_1 = \left\{ (t, \omega, w, \theta(\cdot)) \in \mathbb{M} : \frac{\partial \tilde{f}(t, \omega, w, \theta(\cdot), I)}{\partial u} > 0 \right\},
\]

and \(\hat{u}\) solves the following equation:

\[
-w \frac{\partial q(t, \omega, u)}{\partial u} = \int_0^\infty [w + \theta(z)] e^{-\eta e^{(T-t)}(z-\Phi(z, u))} \frac{\partial \Phi(z, u)}{\partial u} \pi_{t-}(\lambda)(\omega) F^{(1)}(dz).
\]

**Proof.** We observe that \(\tilde{f}\) given in Equation \((7.1)\) is continuous and strictly concave in \(u \in [0, I]\) by hypothesis. Hence the first order condition, which reads as Equation \((7.3)\), admits a unique solution \(\hat{u}(t, \omega, w, \theta(\cdot))\) measurable function on \(\mathbb{M}\). If we extend the function \(\tilde{f}\) to the whole real line, i.e. \(u \in \mathbb{R}\), it is decreasing for \(u < \hat{u}\) and increasing for \(u > \hat{u}\), hence the maximizer on \([0, I]\) must be given by

\[
\hat{u}(t, \omega, w, \theta(\cdot)) = \max\{0, \min\{\hat{u}(t, \omega, w, \theta(\cdot)), I\}\},
\]

which is equivalent to the Equation \((7.2)\). \(\square\)
Remark 7.2. Let us observe that if $q(t, \omega, u)$ and $\Phi(z, u)$ are linear or convex on $u \in [0, I]$ then $\tilde{f}$ is strictly concave in $u \in [0, I]$ and Proposition 7.1 applies.

7.1. Proportional Reinsurance under the expected value principle. In this subsection we consider the special case of proportional reinsurance, that is $\Phi(z, u) = zu$, $u \in [0, 1]$. Under any admissible reinsurance strategy $u \in \mathcal{U}$ the expected cumulative losses covered by reinsurer in the interval $[0, t]$ are given by

$$E \left[ \int_0^t \int_0^{+\infty} z(1 - u_s) \pi_s - (\lambda) F^{(1)}(dz) ds \right].$$

According to the expected value principle, the premium $q^u$ applied by the reinsurer has to satisfy

$$\forall u \in \mathcal{U}, \forall t \in [0, T] \quad E \left[ \int_0^t q_s^u ds \right] = (1 + \theta_R)E[Z^{(1)}]E \left[ \int_0^t (1 - u_s) \pi_s - (\lambda) ds \right],$$

where $\theta_R > 0$ denotes the safety loading applied by reinsurer. Thus

$$q_t^u = (1 + \theta_R)E[Z^{(1)}] \pi_t - (\lambda)(1 - u_t). \quad (7.4)$$

Let us observe that Assumption 4.5 ii) is automatically satisfied, because for every $a > 0$

$$E \left[ e^{a \int_0^T \pi_t(\lambda) dt} \right] < +\infty,$$

see Appendix [3].

Proposition 7.3. Under Assumption 4.5, let $(W^I, \Theta^{W^I}) \in \mathcal{L}^2 \times \hat{\mathcal{L}}^2$ be the solution to the BSDE (5.9). Under the expected value principle (7.4) and proportional reinsurance (i.e. $\Phi(z, u) = zu$), the optimal control is given by

$$u_t^* = \max \{0, \min \{\bar{u}(t, W^I_t, \Theta^{W^I}_t(\cdot)), I\}\},$$

where $\bar{u}(t, w, \theta(\cdot))$ solves the following equation:

$$(1 + \theta_R)E[Z^{(1)}] = \int_0^{+\infty} \frac{w + \theta(z)}{w} z e^{-\eta e^{(T-t)z(1-u)}} F^{(1)}(dz). \quad (7.5)$$

Proof. Using the assumptions above, function $\tilde{f}$ in Equation (7.1) can be rewritten as

$$\tilde{f}(t, \omega, w, \theta(\cdot), u) = -w e^{(T-t)}(1 + \theta_R)E[Z^{(1)}] \pi_t - (\lambda)(\omega)(1 - u) - \int_0^{+\infty} (w + \theta(z))(e^{-\eta e^{(T-t)z(1-u)}} - 1) \pi_t - (\lambda)(\omega) F^{(1)}(dz).$$

Now we can exploit Proposition 7.1 observing that, in particular, Equation (7.3) translates into Equation (7.5). \qed
Since the right-hand side of Equation (7.3) is increasing in \( u \in [0, 1] \), it is immediate that when the reinsurer’s safety loading \( \theta_R \) increases, then the optimal retention level increases too, so that the insurer reduces the purchase of reinsurance. More precisely, the following result is true.

**Corollary 7.4.** There exist two stochastic thresholds \( \theta_F^t < \theta_N^t \) such that the optimal retention level given in Proposition 7.3 can be rewritten in this way:

\[
u^*_t(\omega) = \begin{cases} 
0 & \text{if } \theta_R < \theta_F^t(\omega) \\
1 & \text{if } \theta_R > \theta_N^t(\omega) \\
\bar{u}(t, \omega, W^I_t(\omega), \Theta^W(\cdot)(\omega)) & \text{otherwise,}
\end{cases}
\]

where

\[
\begin{align*}
\theta_F^t &= \frac{1}{\mathbb{E}[Z^{(1)}]} \int_0^{\infty} \frac{W^I_t + \Theta^W(z)}{W^I_t} \frac{z e^{-\eta e^r(T-t)z} F^{(1)}(dz)}{F^{(1)}(dz)} - 1, \\
\theta_N^t &= \frac{1}{\mathbb{E}[Z^{(1)}]} \int_0^{\infty} \frac{W^I_t + \Theta^W(z)}{W^I_t} z F^{(1)}(dz) - 1.
\end{align*}
\]

**Proof.** We rewrite the result of Proposition 7.3 in view of Equation (7.2):

\[
u^*_t = \begin{cases} 
0 & \text{if } (1 + \theta_R)\mathbb{E}[Z^{(1)}] < \int_0^{\infty} \frac{W^I_t + \Theta^W(z)}{W^I_t} \frac{z e^{-\eta e^r(T-t)z} F^{(1)}(dz)}{F^{(1)}(dz)} \\
1 & \text{if } (1 + \theta_R)\mathbb{E}[Z^{(1)}] > \int_0^{\infty} \frac{W^I_t + \Theta^W(z)}{W^I_t} z F^{(1)}(dz) \\
\bar{u}(t, W^I_t, \Theta^W(\cdot)) & \text{otherwise.}
\end{cases}
\]

Solving the two inequalities with respect to \( \theta_R \), we find \( \theta_F^t \) and \( \theta_N^t \) as above. \( \square \)

Let us briefly comment the previous result. We can distinguish three cases, depending on the stochastic conditions (in particular, depending on the solution of the BSDE (5.9)):

- if the reinsurer’s safety loading \( \theta_R \) is smaller than \( \theta_F^t \), then full reinsurance is optimal;
- if \( \theta_R \) is larger than \( \theta_N^t \), then null reinsurance is optimal and the contract is not subscribed;
- lastly, if \( \theta_F^t < \theta_R < \theta_N^t \), then the optimal retention level takes values in \((0, 1)\), that is, the ceding company transfers to the reinsurance a non null percentage of risk (not the full risk).

In other words, if the reinsurance contract is inexpensive, the full reinsurance is purchased. On the contrary, when the reinsurance cost is excessive, the primary insurer will retain all the risk. In the intermediate case \( \theta_F^t < \theta_R < \theta_N^t \) the retention level takes values in the interval \((0, 1)\). In any case, the concepts of inexpensive and expensive must be related to the underlying risk through the stochastic processes \( W^I \) and \( \Theta^W \), hence the thresholds are stochastic.

**Acknowledgements:** The first and third authors have been partially supported by the Project INdAM-GNAMPA, number: U-UFMBAZ-2020-000791.

The first three authors have been partially supported by the Project INdAM-GNAMPA, number: U-UFMBAZ-2022-000765.
The authors were partially supported by University of Padova Grant BIRD 190200/19.

REFERENCES

[1] Albrecher, H., Asmussen, S. (2006): Ruin probabilities and aggregate claims distributions for shot noise Cox processes, Scand. Actuar. J., 2, 86-110.
[2] Bai, L., Guo, J. (2008): Optimal proportional reinsurance and investment with multiple risky assets and no-shorting constraint, IME, 42(3), 968-975.
[3] Bjork, T., Grandell, J. (1988): Exponential inequalities for ruin probabilities in the Cox case, Scand. Actuar. J., 1-3, 77-111.
[4] Brachetta, M., Ceci, C. (2019): Optimal proportional reinsurance and investment for stochastic factor models. IME, 87, 15-33.
[5] Brachetta, M. and Ceci, C. (2020): A BSDE-based approach for the optimal reinsurance problem under partial information, IME, 95, 1-16.
[6] Brémaud, P. (1981): Point Processes and Queues. Martingale Dynamics. Springer-Verlag.
[7] Brémaud, P. and Massoulié, N. (1996): Stability of nonlinear Hawkes processes, Ann. Probab., 24(3), 1563-1588.
[8] Cao, Y., Landriault, D. and Li, B. (2020): Optimal reinsurance-investment strategy for a dynamic contagion claim model. IME, 93, 206-215.
[9] Ceci, C. and Colaneri, K. (2012): Nonlinear filtering for jump diffusion observations, Adv. Appl. Probab., 44(3), 678-701.
[10] Ceci, C. and Colaneri, K. (2014): The Zakai equation of nonlinear filtering for jump-diffusion observations: existence and uniqueness, Applied Mathematics & Optimization, 69, 47-82.
[11] Ceci, C. and Colaneri, K., Cretarola, A. (2014): A benchmark approach to risk-minimization under partial information, IME, 55, 129-146.
[12] Ceci, C. and Gerardi, A. (2009): Pricing for geometric marked point processes under partial information: entropy approach, International Journal of Theoretical and Applied Finance 12(02), 179-207.
[13] Dassios, A. and Zhao, H. (2011): A dynamic contagion process, Adv. Appl. Prob., 43, 814-846.
[14] Dassios, A. and Jang, J.W. (2003): Pricing a catastrophe reinsurance and derivatives using the Cox process with shot noise intensity, Finance and Stochastics, 7(1), 73-95.
[15] Dassios, A. and Jang, J.W. (2005): Kalman-Bucy filtering for linear systems driven by the Cox process with shot noise intensity and its application to the pricing of reinsurance contracts, J. Appl. Prob., 42, 93-107.
[16] Delong, L. (2013): Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications, Springer, London.
[17] Duffie, D., Filipovic, D., Schachermayer, W. (2003): Affine processes and applications in finance, Annals of Applied Probability, 13(3), 984-1053.
[18] Embrechts, P., Schmidli, H., Grandell, J. (1993): Finite-time Lundberg inequalities in the Cox case, Scand. Actuar. J., 1, 17-41.
[19] Grandell, J. (1991): Aspects of risk theory, Springer-Verlag, New York.
[20] Hawkes, A.G. (1971): Spectra of some self-exciting and mutually exciting point processes, Biometrika, 58, 83-90.
[21] Irgens, C., Paulsen, J. (2004): Optimal control of risk exposure, reinsurance and investments for insurance portfolios, IME, 35(1), 21-51.
[22] Jang, J. and Oh, R. (2021): A review on Poisson, Cox, Hawkes, shot-noise Poisson and dynamic contagion process and their compound processes, Annals of Actuarial Science,
[23] Yang, B.-G., Kim, K.T. and Lee, H.-T. (2022): Optimal Reinsurance and Portfolio Selection: Comparison between Partial and Complete Information Models, European Financial Management, 28(1), 208-232.
[24] Kurtz, T.G. and Ocone, D.L. (1988): Unique characterization of conditional distributions in nonlinear filtering, Annals of Probability, 18, 80-107.
[25] Liang, Z., Bayraktar, E. (2014): Optimal reinsurance and investment with unobservable claim size and intensity, IME, 55, 156-166.
[26] Liu, B, Ma, J. (2009): Optimal reinsurance/investment problems for general insurance models. Ann. Appl. Probab. 19, 1495-1528.
[27] Liu, C.S., Yang, H. (2004): Optimal investment for an insurer to minimize its probability of ruin, N. Am. Acuar., 8(2), 11-31.
[28] Mania, M. and Santacroce, M. (2010): Exponential utility maximization under partial information, Finance and Stochastics, 14(3), 419-448.
[29] Papapantoleon, A., Possamai, D. and Saplaouras, A. (2018): Existence and uniqueness results for BSDE with jumps: the whole nine yards, Electron. J. Probab., 23(121), 1-68.
[30] Rosazza Gianin, E. (2006): Risk measures via $g$-expectations, IME, 39(1), 19-34.
[31] Schmidli, H. (2002): On minimizing the ruin probability by investment and reinsurance, Annals of Applied Probability, 12(3), 415-429.
[32] Schmidli, H. (2008): Stochastic Control in Insurance, Springer-Verlag.
[33] Schmidli, H. (2018): Risk Theory, Springer Actuarial. Springer International Publishing.
[34] Sokol, A., Hansen, N.R. (2015): Exponential Martingales and Changes of Measure for Counting Processes, Stochastic Analysis and Applications, 33(5), 823-843.
[35] Stabile, G., Torrisi, G.L. (2010): Risk processes with non-stationary claims arrivals, Methodology Comput. Appl. Prob., 12, 415-429.
[36] Swishchuk, A., Zagst, R. Zeller, G. (2020): Hawkes processes in insurance: Risk model, application to empirical data and optimal investment, IME, 101(A), 107-124.
[37] Swishchuk, A. (2021): Merton Investment Problems in Finance and Insurance for Hawkes-based Models, Preprint. Available at https://arXiv:2014.02694.
[38] Zeng, Y., Li, Z. (2011): Optimal time-consistent investment and reinsurance policies for mean-variance insurers, IME, 49(1), 145-154.
Lemma A.1. Let \((\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})\) be a filtered probability space and assume that the filtration \(\mathbb{F} = \{\mathcal{F}_t, \ t \in [0, T]\}\) satisfies the usual hypotheses. Let \(N\) be a standard Poisson process with \(\mathbb{F}\)-intensity \(\lambda > 0\) and let \(\{b_t\}_{t \geq 0}\) an \(\mathbb{F}\)-predictable process. Then
\[
\mathbb{E}\left[ e^{\int_0^T b_t \, dN_t} \right] = \mathbb{E}\left[ e^{\int_0^T (e^{bt} - 1) \lambda \, dt} \right],
\]
provided that the last expectation is finite.

Proof. In order to show that the statement is valid for any bounded \(\mathbb{F}\)-predictable process, see Brémaud \cite{T4, Theorem, Appendix A1}, it is sufficient to prove our result for any arbitrary process
\[
b_t = 1_{(t_1, t_2]}(t)1_A, \quad 0 \leq t_1 < t_2 \leq T, \quad A \in \mathcal{F}_{t_1},
\]
Let \(0 \leq t_1 < t_2 \leq T, A \in \mathcal{F}_{t_1}\) and denote by \(A^C\) the complementary set of \(A\). Then we have that
\[
\mathbb{E}\left[ e^{\int_0^T b_t \, dN_t} \right] = \mathbb{E}\left[ e^{\int_0^{t_2} 1_A \, dN_t} \right] = \mathbb{E}\left[ e^{(N_{t_2} - N_{t_1})1_A} (1 + 1_{A^C}) \right]
= \mathbb{E}\left[ \mathbb{E}[e^{(N_{t_2} - N_{t_1})} | \mathcal{F}_{t_1}]1_A + 1_{A^C} \right]
= \mathbb{E}\left[ \mathbb{E}[e^{(N_{t_2} - N_{t_1})}]1_A + 1_{A^C} \right].
\]
Now the inner expectation corresponds to the Laplace transform of a Poisson random variable, since \((N_{t_2} - N_{t_1}) \sim \text{Po}(\lambda(t_2 - t_1))\), namely \(\mathbb{E}[e^{(N_{t_2} - N_{t_1})}] = e^{(e-1)(t_2-t_1)}\). Substituting and rearranging the terms we then get that
\[
\mathbb{E}\left[ e^{\int_0^T b_t \, dN_t} \right] = \mathbb{E}\left[ e^{(e-1)(t_2-t_1)}1_A \right].
\]
On the other hand, we notice that
\[
e^{bs} - 1 = e^{\sum_{t_1 \leq \tau \leq t_2} (t)1_A} - 1 = e \cdot 1_{(t_1, t_2]}(t)1_A - 1_{(t_1, t_2]}(t)1_A = (e - 1)1_{(t_1, t_2]}(t)1_A
\]
and so
\[
\mathbb{E}\left[ e^{\int_0^T (e^{bt} - 1) \lambda \, dt} \right] = \mathbb{E}\left[ e^{\int_0^T (e^{bt} - 1) \lambda \, dt} \right] = \mathbb{E}\left[ e^{(e-1)(t_2-t_1)}1_A \right],
\]
which proves the statement for any bounded \(\mathbb{F}\)-predictable process. To complete the proof, we extend this result to unbounded processes.

Assume that \(\{b_t\}_{t \geq 0}\) is an arbitrary \(\mathbb{F}\)-predictable process and define a sequence of \(\mathbb{F}\)-stopping times
\[
\tau_n = \inf\{t \geq 0 : b_t > n\}, \quad n \geq 1.
\]
Clearly, \(\tau_n \to +\infty\) as \(n \to +\infty\). By the first part of the proof, we know that
\[
\mathbb{E}\left[ e^{\int_0^{T\wedge \tau_n} b_t \, dN_t} \right] = \mathbb{E}\left[ e^{\int_0^{T\wedge \tau_n} (e^{bt} - 1) \lambda \, dt} \right],
\]
so that, to complete the proof, it remains to pass to the limit \(n \to +\infty\) and to apply the monotone convergence theorem to the family of random variables \(X_n := e^{\int_0^{T\wedge \tau_n} b_t \, dN_t}, n \geq 1\), in the case when \(b\) is positive, or to \(\overline{X}_n := \frac{e^{\int_0^{T\wedge \tau_n} b_t^2 \, dN_t}}{e^{\int_0^{T\wedge \tau_n} b_t \, dN_t}}, n \geq 1\) for a general \(b\). \(\square\)
Lemma A.2. Let $(\Omega, \mathcal{F}, P; \mathbb{F})$ be a filtered probability space and assume that the filtration $\mathbb{F} = \{\mathcal{F}_t, \ t \in [0,T]\}$ satisfies the usual hypotheses. Let $N(dt, dz)$ be a Poisson random measure on $[0,T] \times [0, +\infty)$ with $\mathbb{F}$-intensity kernel $\lambda F(dz) dt$. Then for any $\mathbb{F}$-predictable and $[0, +\infty)$-indexed process $\{H_t(z)\}_{t \geq 0}$ we have that

$$\mathbb{E} \left[ e^{\int_0^T \int_0^\infty H_t(z) N(dt,dz)} \right] = \mathbb{E} \left[ e^{\int_0^T \int_0^\infty (e^{H_t(z)} - 1) \lambda F(dz) dt} \right],$$

provided that the last expectation is finite.

Proof. It is sufficient to prove the result for any process $\{H_t(z)\}_{t \geq 0}$ of this form:

$$H_t(z) = b_t \mathbb{1}_A, \ \forall t \geq 0, A \in \mathcal{B}([0, +\infty)),$$

where $b_t$ is $\mathbb{F}$-predictable and $\mathcal{B}([0, +\infty))$ denotes the Borel $\sigma$-algebra of subsets of $[0, +\infty)$. By Lemma A.1 we readily obtain that

$$\mathbb{E} \left[ e^{\int_0^T H_t(z) N(dt,dz)} \right] = \mathbb{E} \left[ e^{\int_0^T b_t N(dt,A)} \right]
\begin{align*}
&= \mathbb{E} \left[ e^{\int_0^T (e^{b_t} - 1) \int_A F(dz) \lambda dt} \right] \\
&= \mathbb{E} \left[ e^{\int_0^T \int_0^\infty (e^{b_t} - 1) \lambda dt} A(z) F(dz) \lambda dt \right] \\
&= \mathbb{E} \left[ e^{\int_0^T \int_0^\infty (e^{b_t A(z)} - 1) F(dz) \lambda dt} \right] \\
&= \mathbb{E} \left[ e^{\int_0^T \int_0^\infty (e^{H_t(z)} - 1) F(dz) \lambda dt} \right],
\end{align*}

where we have used that $N((0,t] \times A)$ is a Poisson process with intensity $\int_A F(dz) \lambda$. \hfill \Box

Appendix B. Proof of Key Lemmas

We focus here on the finiteness of $\mathbb{E} \left[ e^{a N^{(1)}_T} \right]$, $\mathbb{E} \left[ e^{a \int_0^T \lambda ds} \right]$, $\mathbb{E} \left[ e^{a \int_0^T \pi_s \lambda ds} \right]$ and $\mathbb{E}[e^{a C_T}]$, which are computed under $P$ for an arbitrary real constant $a > 0$. Here $N^{(1)}$ is a standard Poisson process under $(Q, \mathbb{F})$ and a counting process with intensity $\lambda$ (given in Equation (2.1)) under $(P, \mathbb{F})$. We will exploit the measure change introduced in detail in Section 2 and we will work under Assumption 4.5 i).

We prove the following:

Lemma B.1. Under Assumption 4.5 i)

$$\mathbb{E} \left[ e^{a N^{(1)}_T} \right] < +\infty, \ \mathbb{E} \left[ e^{a \int_0^T \lambda ds} \right] < +\infty, \ \text{and} \ \mathbb{E} \left[ e^{a \int_0^T \pi_s \lambda ds} \right] < +\infty.$$

Proof. First of all, we show that under Assumption 4.5 i) we have

$$\mathbb{E} Q \left[ e^{a \int_0^T \lambda ds} \right] < +\infty. \quad (B.1)$$
Recalling that for every \( s \in [0, T] \) the intensity satisfies

\[
\lambda_s \leq \max\{\lambda_0, \beta\} + \sum_{j=1}^{N_T^{(1)}} \ell(Z_j^{(1)}) + \sum_{j=1}^{N_T^{(2)}} Z_j^{(2)},
\]

for a suitable \( c_1 > 0 \) and for \( c_2 = aT \) we find that

\[
\mathbb{E}^Q \left[ e^{aT} \int_0^T \lambda_s \, ds \right] \leq \mathbb{E}^Q \left[ e^{aT} \left( \max\{\lambda_0, \beta\} + \sum_{j=1}^{N_T^{(1)}} \ell(Z_j^{(1)}) + \sum_{j=1}^{N_T^{(2)}} Z_j^{(2)} \right) \right]
\]

\[
\leq c_1 \mathbb{E}^Q \left[ e^{c_2 \sum_{j=1}^{N_T^{(1)}} \ell(Z_j^{(1)})} \sum_{j=1}^{N_T^{(2)}} Z_j^{(2)} \right]
\]

\[
= c_1 e^{c_2 \sum_{j=1}^{N_T^{(1)}} \ell(Z_j^{(1)})} \mathbb{E}^Q \left[ e^{c_2 \sum_{j=1}^{N_T^{(2)}} Z_j^{(2)}} \right]
\]

\[
= c_1 e^{T \left( \mathbb{E}^Q [e^{c_2 \ell(Z_1^{(1))}}] - 1 \right)} e^{T \left( \mathbb{E}^Q [e^{c_2 Z_2^{(2)}]} - 1 \right)} < +\infty,
\]

where we used the mutual independence of \( N^{(1)}, N^{(2)}, \{Z_n^{(1)}\}_{n \geq 1}, \{Z_n^{(2)}\}_{n \geq 1} \) under \( Q \) (see Remark 2.3) and, in the last equality, we followed the path traced in the proof of Proposition 2.6. Finally Assumption 4.5(i) gives the finiteness of the expectation under \( Q \).

To prove that \( \mathbb{E} \left[ e^{a N_T^{(1)}} \right] \) is finite we exploit the change of measure from \( P \) to \( Q \) via \( \frac{dP}{dQ} \big|_{\mathcal{F}_T} = L_T \), with \( L_T \) given in Equation (2.6), so that

\[
\mathbb{E} \left[ e^{a N_T^{(1)}} \right] = \mathbb{E}^Q \left[ L_T e^{a N_T^{(1)}} \right] = \mathbb{E}^Q \left[ e^{- \int_0^T (\lambda_s - 1) \, ds + \int_0^T (\ln(\lambda_s) + a) \, dN_s^{(1)}} \right]
\]

\[
\leq C \mathbb{E}^Q \left[ e^{\int_0^T (\ln(\lambda_s) + a) \, dN_s^{(1)}} \right]
\]

for a suitable constant \( C > 0 \).

Now we recall that under \( Q \) the Poisson process \( N^{(1)} \) has unitary intensity and for any predictable process \( b \) we have that \( \mathbb{E}^Q \left[ e^{\int_0^T b_s dN_s^{(1)}} \right] = \mathbb{E}^Q \left[ e^{\int_0^T (e^{b_s - 1}) \, ds} \right] \), according to Lemma A.1. Hence, taking \( b_s = (\ln(\lambda_s) + a) \), we obtain

\[
\mathbb{E} \left[ e^{a N_T^{(1)}} \right] \leq C \mathbb{E}^Q \left[ e^{\int_0^T (\ln(\lambda_s) + a) \, dN_s^{(1)}} \right] = C \mathbb{E}^Q \left[ e^{\int_0^T (-\lambda_s - 1) \, ds} \right] < +\infty,
\]

which is finite because of Equation (B.1).
We show now that $\mathbb{E} \left[ e^{x_0^T \lambda_s ds} \right] < +\infty \ \forall x > 0$. We proceed as above: passing under $Q$ via $L_T$, recalling Equation B.2 and introducing the integer valued random measure $m^{(1)}(dt, dz)$, we find

$$
\mathbb{E} \left[ e^{x_0^T \lambda_s ds} \right] = \mathbb{E}^Q \left[ L_T e^{x_0^T \lambda_s ds} \right] = \mathbb{E}^Q \left[ e^{x_0^T \lambda_s ds} \right] 
\leq C_1 \mathbb{E}^Q \left[ e^{C_2 \sum_{j=1}^{N_T^{(1)}} Z_j^{(2)}} \right] \mathbb{E}^Q \left[ e^{C_2 \sum_{j=1}^{N_T^{(1)}} Z_j^{(2)}} e^{T_T \ln(\lambda_s) dN_s^{(1)}} \right] 
= C_1 \mathbb{E}^Q \left[ e^{C_2 \sum_{j=1}^{N_T^{(1)}} Z_j^{(2)}} \right] \mathbb{E}^Q \left[ e^{C_2 \sum_{j=1}^{N_T^{(1)}} Z_j^{(2)}} e^{T_T \ln(\lambda_s) dN_s^{(1)}} \right] 
= C_1 \mathbb{E}^Q \left[ e^{C_2 \sum_{j=1}^{N_T^{(1)}} Z_j^{(2)}} \right] \mathbb{E}^Q \left[ e^{T_T \int_0^T [C_2 \ell(z) + \ln(\lambda_s)] m^{(1)}(dt, dz)} \right] 
$$

for a suitable constant $C_1 > 0$. We now apply Lemma A.2 under $Q$ and for $H_T = [C_2 \ell(z) + \ln(\lambda_s)]$ and with $\nu^{(1),Q}(dt, dz) = F^{(1)}(dz) dt$ and we get:

$$
\mathbb{E} \left[ e^{x_0^T \lambda_s ds} \right] \leq C_1 \mathbb{E}^Q \left[ e^{C_2 \sum_{j=1}^{N_T^{(1)}} Z_j^{(2)}} \right] \mathbb{E}^Q \left[ e^{T_T \int_0^T [C_2 \ell(z) + \ln(\lambda_s)] m^{(1)}(dt, dz)} \right] 
= C_1 \mathbb{E}^Q \left[ e^{C_2 \sum_{j=1}^{N_T^{(1)}} Z_j^{(2)}} \right] \mathbb{E}^Q \left[ e^{T_T \int_0^T \left( \lambda_s \nu^{(1),Q} \right) \ell^{(1)}(z) - 1\right] ds ] 
$$

which is finite under Assumption 4.5 (i).

It remains to prove that $\mathbb{E} \left[ e^{x_0^T \pi_{T^{(1)}}(\lambda) ds} \right] < +\infty \ \forall x > 0$. The structure of the filtering equation implies that over $[0, T]$ the filter attains its maximum value at a jump time. More precisely, we showed in Remark 3.4 that the filter is dominated by a process with exponential decay behavior between two consecutive jumps, hence the maximum over $[0, T]$ is attained at a jump time $\tau \leq T$ such that

$$
\pi_{T^{(1)}}(\lambda) = \max \left\{ \pi_{T^{(1)}}(\lambda), \ldots, \pi_{T^{(1)}}(\lambda) \right\}.
$$

Notice that the maximum is taken over a finite number of elements, because the jump process $N^{(1)}$ is non explosive. Then, using Jensen’s inequality we have that

$$
\mathbb{E} \left[ e^{x_0^T \pi_{T^{(1)}}(\lambda) ds} \right] \leq \mathbb{E} \left[ e^{xT \pi_{T^{(1)}}(\lambda)} \right] \leq \mathbb{E} \left[ e^{xT \pi_{T^{(1)}}(\lambda)} \right] = \mathbb{E} \left[ e^{xT \pi_{T^{(1)}}(\lambda)} \right] < +\infty.
$$
The last inequality is implied by the fact that $\tau \leq T$ and so the following inequalities hold
\[
\mathbb{E} \left[ e^{aT\lambda_T} \right] = \mathbb{E}^Q \left[ L_T e^{aT\lambda_T} \right] \\
\leq C_1 \mathbb{E}^Q \left[ e^{aT\lambda_T} \lambda_T \int_0^T \ln(\lambda_T - s) dN_t^{(1)} \right] \\
\leq C_1 \mathbb{E}^Q \left[ \frac{C_2}{e} \left( \sum_{j=1}^{N_T^{(1)}} e^{(Z_j^{(1)} + \sum_{j=1}^{N_T^{(2)}} Z_j^{(2)})} \right) e^{\int_0^T \ln(\lambda_T - s) dN_t^{(1)}} \right]
\]
for suitable constants $C_i > 0$, $i = 1, 2$, and we can prove the finiteness by doing the same computations to prove that $\mathbb{E} \left[ e^{a\int_0^T \lambda_t ds} \right] < +\infty$.

\[\square\]

Based on the previous Lemma, we conclude this section proving the useful result given in Lemma 4.6, i.e. for every $a > 0$
\[
\mathbb{E}[e^{aC_T}] < +\infty.
\]

**Proof of Lemma 4.6.** We have that for a suitable constant $\kappa > 0$, passing under $Q$ via the Radon-Nikodym derivative $L_T$ given in Equation 2.6 and using Lemma A.2,
\[
\mathbb{E}[e^{aC_T}] = \mathbb{E}^Q \left[ e^{-\int_0^T (\lambda_T - 1) dt + \int_0^T \log \lambda_T - dN_t^{(1)} e^{\int_0^T +\infty} az m^{(1)}(dt,dz) } \right] \\
= \kappa \mathbb{E}^Q \left[ e^{\int_0^T +\infty} (\log \lambda_T + az) m^{(1)}(dt,dz) } \right] \\
= \kappa \mathbb{E}^Q \left[ e^{\int_0^T +\infty} (e^{\log \lambda_T + az} - 1) F^{(1)}(dz) dt \right] \\
= \kappa \mathbb{E}^Q \left[ e^{\int_0^T \lambda_T - (e^{aZ^{(1)}} - 1) dt } \right] < +\infty,
\]
where the finiteness comes from Equation (B.1) and Assumption 4.5 i).

\[\square\]