Stochastic optimal control and time changed Lévy noises

Steffen Sjursen

January 9, 2014

© Steffen Sjursen, 2014

*Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1472*

ISSN 1501-7710

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinssen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika Publishing.
The thesis is produced by Akademika Publishing merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

Acknowledgements

Doing a PhD has been very fun. A great thanks to my supervisor Giulia Di Nunno for her clever advice, patience and the interesting discussions. I would also like to thank Bernt Øksendal for his very friendly advice and the financial support from the INNOSTOCH program. The CMA and the Department of mathematics has created for a very welcoming and productive atmosphere that I'm happy to have been a part of. I have thoroughly enjoyed all the discussions with my colleagues both in the department and those I met on various conferences and workshops.

Contents

1	Introduction	7
1.1	Some mathematical preliminaries	8
1.1.1	The martingale random field	9
1.1.2	The non-anticipating stochastic derivative	10
1.1.3	Processes with conditionally independent increments	10
1.1.4	Time-changed Lévy processes	12
1.2	On chaos representation and orthogonal polynomials for the doubly stochastic Poisson process	13
1.3	BSDEs for time-changed Lévy processes and application to optimal control . .	14
1.4	Maximum principles for martingale random fields	15
1.5	Optimal portfolio problems in the presence of default and the value of information	16
2	Chaos representation and orthogonal polynomials	19
2.1	Introduction	19
2.2	The doubly stochastic Poisson process	21
2.3	Multilinear forms, polynomials, and chaos expansions	25
2.4	Non-anticipating stochastic integration	31
2.5	Non-anticipating stochastic derivative and representation theorem	35
2.6	Iterated integrals and chaos expansions	39
2.7	Anticipative stochastic derivatives and integral representations	42
3	BSDEs for time-changed Lévy processes and application to optimal control	51
3.1	Introduction	51
3.2	The framework	54
3.2.1	The random measures and their properties	54
3.2.2	Stochastic non-anticipating integration	57
3.3	Integral and martingale representation theorems	58
3.4	BSDE: Existence and uniqueness of the solution	62
3.5	Linear BSDEs and a comparison theorem	69
3.6	Sufficient stochastic maximum principle	73
3.7	Necessary maximum principle	78
3.8	Optimal portfolio problems	83
3.8.1	The mean-variance portfolio selection	85
3.8.2	Utility maximization of final wealth	88

4	Maximum principles for martingale random fields via non-anticipating stochastic derivatives	91
4.1	Introduction	91
4.2	The martingale random field	93
4.3	Duality relations with non-anticipating stochastic derivatives	95
4.4	Optimization problem	96
4.5	Maximum principle	98
4.6	A remark on the technique used	105
4.7	Lévy Processes	106
4.8	Application to default risk	107
5	Information and optimal investment in defaultable assets	111
5.1	Introduction: The model, the optimization problem, the streams of information	111
5.2	Mathematical framework: Forward Integrals	115
5.2.1	The Itô formula for forward integrals	118
5.3	Local maximums	121
5.3.1	Some examples with logarithmic utility	127
5.4	On the driving processes as semi-martingales	132
	References	136

Chapter 1

Introduction

This thesis consist of this introductory chapter and a collection of four papers on the stochastic calculus and optimization problems related to the doubly stochastic Poisson process:

1. *On chaos representation and orthogonal polynomials for the doubly stochastic Poisson process*, together with Giulia Di Nunno. Published in *Seminar on stochastic analysis, random fields and applications VII*, R. Daland, M. Dozzi and F. Russo (eds), vol 67 of Progress in Probability, Springer Basel 2013.
2. *BSDEs for time-changed Lévy processes and application to optimal control*, together with Giulia Di Nunno. To appear in *Stochastic Processes and their Applications*.
3. *Maximum principles for martingale random fields via non-anticipating stochastic derivatives*. Submitted.
4. *Information and optimal investment in defaultable assets*, together with Giulia Di Nunno. Submitted.

The introduction will outline some of the key concepts explored in the thesis and gives a brief motivation for each individual paper.

The doubly stochastic Poisson process (DSPP) is an extension of the Poisson process or Poisson random field with a randomization of the intensity parameter. For a detailed exposition on the DSPP we refer to [50], and a brief description can be found in Chapter 2. To put it informally, for two random measures H and α on a space $\mathcal{X}$, H is DSPP with intensity α if it satisfies

$$\mathbb{P}(H(\Delta) = k | \alpha(\Delta)) = \frac{\alpha(\Delta)^k e^{-\alpha(\Delta)}}{k!}, \quad \Delta \subset \mathcal{X}, \quad k \in \mathbb{N}.$$

Meaning that conditionally on α , H is Poisson distributed. The DSPP is an extension of the Poisson process but the extra “parameter” α allows for a dependence between the increments. This dependence is highly useful for credit risk and asset models with stochastic volatility (see, e.g. [74] and [25] respectively). There is however a computational downside of the dependence, the DSPP is less tractable than a Poisson process and is in general not Markovian.

Of particular interest in our study are optimization problems of type

$$\sup_{u \in \mathcal{A}} J(u) = \sup_{u \in \mathcal{A}} \mathbb{E} \left[\int_0^T f_s(u_s, X_s(u)) ds + l(X_T(u)) \right] \quad (1.0.1)$$

for controls u in a set of admissible controls $\mathcal{A}$. The state process X is driven by a DSPP or a more general type of noise (like a process with conditionally independent increments or a martingale random field). Problems of type (1.0.1) include utility and consumption problems, mean variance portfolio selection, mean variance hedging and more. Thus solving (1.0.1) is relevant for asset models with the random noises considered here, like the DSPP or other types of time changed Lévy noises.

In the literature there are two main approaches to optimization problems of type (1.0.1), dynamic programming and the maximum principle. The non-Markovian structure has an important implication for optimization problems with DSPPs, namely that we cannot apply dynamic programming techniques.

A maximum principle with BSDEs requires the existence and uniqueness of the relevant BSDE. Such BSDEs had not been studied for the DSPP except for the very special case considered by [77].

Another option to solving (1.0.1) in the Lévy case is maximum principles via the duality relation of the Malliavin derivative [82]. However, this would depend on a Malliavin type calculus, which for the DSPP has only been studied in [107]. Chaos expansions, integral representation, stochastic differentiation and Malliavin calculus also have independent interest both in the theory of stochastic analysis and in applications.

The three first papers consider the above issues directly. The first paper studies chaos expansions in filtered spaces generated by the DSPP, the non-anticipative stochastic derivative, integral representation and their relations to the Malliavin calculus for the DSPP. The second paper shows existence and uniqueness of BSDEs for filtrations generated by time-changed Lévy processes (a class of processes related to the DSPP). Using these BSDEs we find sufficient and necessary maximum principles for problems of type (1.0.1). The third paper considers necessary conditions for optimal control for (1.0.1) for general martingale random fields using the non-anticipative stochastic derivative. There we also consider the DSPP as a special case.

The fourth paper discuss utility maximization with a credit risky asset and the role of the filtration. The DSPP has been used to model credit risk events in enlargement of filtration models. Here we explore the utility maximization with minimal assumptions on the credit events. We find sufficient and necessary conditions for an optimal strategy and study the implications of the existence of such a strategy.

The papers will be given in full in the next chapters, presented in the same form as they were submitted or published.

1.1 Some mathematical preliminaries

We introduce some mathematical concepts frequently used in this thesis, the *martingale random field*, the *non-anticipating stochastic derivative* and *processes with conditionally independent increments* before discussing the four different papers.

1.1.1 The martingale random field

We define integration and the non-anticipating stochastic derivative over a martingale random field μ . We refer to [38] for a detailed discussion on these concepts.

Let $(\Omega, \mathcal{G}, \mathbb{P})$ be a complete probability space equipped with a right-continuous filtration $\mathbb{G} := \{\mathcal{G}_t, t \in [0, T]\}$. Let $\mathcal{Z}$ be a separable Hausdorff topological space. We denote $\mathcal{B}_{\mathcal{Z}}$ as the Borel σ -algebra on $\mathcal{Z}$ and $\mathcal{B}_{[0, T] \times \mathcal{Z}}$ as the Borel σ -algebra on the product space $[0, T] \times \mathcal{Z}$.

We define the martingale random field as in [38, Definition 2.1]:

Definition 1.1.1. *We say that the stochastic set function $\mu(\Delta)$, $\Delta \in \mathcal{B}_{[0, T] \times \mathcal{Z}}$ is a martingale random field on $[0, T] \times \mathcal{Z}$ (with square integrable values) with respect to $\mathbb{G}$ if it satisfies the following properties*

- i) μ has a tight, σ -finite variance measure $V(\Delta) = E[\mu(\Delta)^2]$, $\Delta \in \mathcal{B}_{[0, T] \times \mathcal{Z}}$, which satisfies $V(\{0\} \times \mathcal{Z}) = 0$.
- ii) μ is additive, i.e. for pairwise disjoint sets $\Delta_1, \dots, \Delta_K$: $V(\Delta_k) < \infty$,
 $\mu(\bigcup_{k=1}^K \Delta_k) = \sum_{k=1}^K \mu(\Delta_k)$ and σ -additive in $L_2(\Omega)$.
- iii) μ is $\mathbb{G}$ -adapted.
- iv) μ has the martingale property. For any $\Delta \subseteq (t, T] \times \mathcal{Z}$ we have: $\mathbb{E}[\mu(\Delta) \mid \mathcal{G}_t] = 0$.

The martingale random field μ has conditionally orthogonal values if

- v) for any $\Delta_1, \Delta_2 \subseteq (t, T] \times \mathcal{Z}$ such that $\Delta_1 \cap \Delta_2 = \emptyset$ we have: $\mathbb{E}[\mu(\Delta_1)\mu(\Delta_2) \mid \mathcal{G}_t] = 0$.

Here we will always assume that the martingale random field has orthogonal values. In particular, any finite sums of orthogonal, square integrable martingales would be a martingale random field in the sense of i)-ii)-iii)-iv)-v) above. The random field μ has a σ -finite conditional random variance measure [38, Theorem 2.1], which we denote by Λ . For martingale processes the conditional variance measure is the $\mathbb{G}$ -predictable compensator.

We denote $\mathcal{I}$ as the set of $\mathbb{G}$ -predictable random fields $\phi : \Omega \times [0, T] \times \mathcal{Z} \rightarrow \mathbb{R}$ satisfying

$$\|\phi\|_{\mathcal{I}} := \mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} \phi(s, z)^2 \Lambda(ds, dz) \right]^{\frac{1}{2}} < \infty.$$

We say that $\phi \in \mathcal{I}$ is a simple random field if can be expressed as a finite sum of type

$$\phi(s, z, \omega) = \sum_{i=1}^N \phi_i(\omega) \mathbf{1}_{\Delta_i}(s, z), \quad \Delta_i \in \mathcal{B}_{[0, T] \times \mathcal{Z}}, \quad (1.1.1)$$

where ϕ_i are bounded random variables (for $i = 1, \dots, N < \infty$). Simple, $\mathbb{G}$ -predictable random fields are dense in $\mathcal{I}$ by the usual Itô integration type arguments and we have, for every $\phi \in \mathcal{I}$:

$$\mathbb{E} \left[\left(\int_0^T \int_{\mathcal{Z}} \phi(s, z) \mu(ds, dz) \right)^2 \right] = \mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} \phi(s, z)^2 \Lambda(ds, dz) \right].$$

1.1.2 The non-anticipating stochastic derivative

The *non-anticipating stochastic derivative* is a characterization of the integrand in the Kunita-Watanabe decomposition, developed in [34, 35, 44, 37, 38]. It is the adjoint (linear) operator $\mathcal{D} = I^*$ of the stochastic integral:

$$\mathcal{D} : L_2(\Omega, \mathcal{G}, \mathbb{P}) \Longrightarrow \mathcal{I}.$$

The non-anticipating stochastic derivative can be explicitly represented as the limit [38, Theorem 3.1]

$$\mathcal{D}\xi = \lim_{n \rightarrow \infty} \phi_n \quad (1.1.2)$$

with convergence in $\mathcal{I}$ of the stochastic functions of type (1.1.1) given by

$$\phi_n(t, z) := \sum_{k=1}^{K_n} \mathbb{E} \left[\xi \frac{\mu(\Delta_{n,k})}{\Lambda(\Delta_{n,k})} \middle| \mathcal{G}_{s_{n,k}} \right] \mathbf{1}_{\Delta_{n,k}}(t, z) \quad (1.1.3)$$

where $\Delta_{n,k} = (s_{n,k}, t_{n,k}] \times Z_{n,k}$ refers to decreasing partions of $A_n \subseteq [0, T] \times \mathcal{Z}$, where A_n is an increasing series of sets. Remark that the series of sets are chosen as elements of a dissecting system, for details see Definition 2.2.6 and Section 2.5. We have the following result [38, Theorem 3.1]:

Theorem 1.1.2. *All $\xi \in L_2(\Omega, \mathcal{G}, \mathbb{P})$ have representation*

$$\xi = \xi_0 + \int_0^T \int_{\mathcal{Z}} \mathcal{D}_{t,z} \xi \mu(dt, dz). \quad (1.1.4)$$

Moreover $\mathcal{D}\xi_0 = 0$ and $\xi_0 \in L_2(\Omega, \mathcal{G}, \mathbb{P})$ is orthogonal to space generated by $\{I(\phi), \phi \in \mathcal{I}\}$.

By the orthogonality of the sum in (1.1.4), one can see that the following duality rule is verified: Let $\xi \in L_2(\Omega, \mathcal{G}, \mathbb{P})$ and $\kappa \in \mathcal{I}$, then

$$\mathbb{E} \left[\xi \int_0^T \int_{\mathcal{Z}} \kappa(s, z) \mu(ds, dz) \right] = \mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} (\mathcal{D}_{s,z} \xi) \kappa(s, z) \Lambda(ds, dz) \right]. \quad (1.1.5)$$

1.1.3 Processes with conditionally independent increments

Next, we specify the martingale random field μ to have conditionally independent increments (CII). A stochastic process or martingale random field μ has CII with respect to a σ -algebra $\mathcal{F}^\Lambda$ if

$$\mathbb{P} \left(\prod_{i=1}^N \mu(\Delta_i) < x_i \middle| \mathcal{F}^\Lambda \right) = \prod_{i=1}^N \mathbb{P} \left(\mu(\Delta_i) < x_i \middle| \mathcal{F}^\Lambda \right),$$

for any $N < \infty$, $x, \dots, x_N \in \mathbb{R}^N$ and pairwise disjoint $\Delta_1, \dots, \Delta_N \subseteq [0, T] \times \mathcal{Z}$. Processes with CII are intrinsically linked to the normal and Poisson distributions, μ has CII if and only if the conditional probability distributions are of normal and/or Poisson type [100, 51]. This also

implies a link between time-changed Lévy processes and the processes with CII, which we will discuss below.

Hence we can define μ via the conditional probability distributions. Set $\mathcal{Z} = \mathcal{Z}_0 \times \{0\}$, where $\mathcal{Z}$ is a complete metric space. Let Λ^B be a positive random measure on $[0, T] \times \{0\}$ and Λ^H be a positive random measure on $[0, T] \times \mathcal{Z}_0$. Define the random measure Λ on $[0, T] \times \mathcal{Z}$ by

$$\Lambda(\Delta) := \int_0^T \mathbf{1}_{\{(t,0) \in \Delta\}}(t) \Lambda^B(dt) + \int_0^T \int_{\mathcal{Z}_0} \mathbf{1}_{\Delta}(t, z) \Lambda^H(dt, dz) \quad (1.1.6)$$

as the mixture of measures on disjoint sets. We assume that Λ is non-atomic. We denote $\mathcal{F}^\Lambda$ as the σ -algebra generated by Λ .

Definition 1.1.3. *B is a signed random measure on the Borel sets of $[0, T] \times \{0\}$ satisfying,*

$$A1) \quad \mathbb{P}\left(B(\Delta) \leq x \mid \mathcal{F}^\Lambda\right) = \mathbb{P}\left(B(\Delta) \leq x \mid \Lambda^B(\Delta)\right) = \Phi\left(\frac{x}{\sqrt{\Lambda^B(\Delta)}}\right), \quad x \in \mathbb{R}, \Delta \subseteq [0, T] \times \{0\},$$

A2) $B(\Delta_1)$ and $B(\Delta_2)$ are conditionally independent given $\mathcal{F}^\Lambda$ whenever Δ_1 and Δ_2 are disjoint sets.

Here Φ stands for the cumulative probability distribution function of a standard normal random variable.

H is a random measure on the Borel sets of $[0, T] \times \mathcal{Z}_0$ satisfying

$$A3) \quad \mathbb{P}\left(H(\Delta) = k \mid \mathcal{F}^\Lambda\right) = \mathbb{P}\left(H(\Delta) = k \mid \Lambda^H(\Delta)\right) = \frac{\Lambda^H(\Delta)^k}{k!} e^{-\Lambda^H(\Delta)}, \quad k \in \mathbb{N}, \Delta \subseteq [0, T] \times \mathcal{Z}_0,$$

A4) $H(\Delta_1)$ and $H(\Delta_2)$ are conditionally independent given $\mathcal{F}^\Lambda$ whenever Δ_1 and Δ_2 are disjoint sets.

Furthermore we assume that

A5) B and H are conditionally independent given $\mathcal{F}^\Lambda$.

Conditions A1) and A3) mean that conditional on Λ , B is a Gaussian random measure and H is Poisson a random measure. Let $\tilde{H} := H - \Lambda^H$ be the signed random measure given by

$$\tilde{H}(\Delta) = H(\Delta) - \Lambda^H(\Delta), \quad \Delta \subset [0, T] \times \mathcal{Z}_0. \quad (1.1.7)$$

Set μ to be given by

$$\mu(\Delta) := B\left(\Delta \cap [0, T] \times \{0\}\right) + \tilde{H}\left(\Delta \cap [0, T] \times \mathbb{R}_0\right), \quad \Delta \subseteq [0, T] \times \mathcal{Z}, \quad (1.1.8)$$

and

$$\mathbb{F} := \{\mathcal{F}_t\} \text{ as the smallest right-continuous filtration to which } \mu \text{ is adapted.} \quad (1.1.9)$$

It is possible to show that B , H and Λ are all $\mathbb{F}$ -adapted, see, e.g. [105, Theorem 1] and the forthcoming Theorem 2.2.8. Furthermore let

$$\mathbb{G} = \{\mathcal{G}_t, t \in [0, T]\} \text{ be generated by } \mu(\Delta), \Delta \subset [0, t] \times \mathcal{Z} \text{ and } \Lambda(\Delta), \Delta \subset [0, T] \times \mathbb{R}. \quad (1.1.10)$$

Note the difference in the time-intervals, $\mathcal{G}_t$ is the knowledge of the *entire history* of Λ and the history of μ up to time t . Remark that $\mathcal{G}_T = \mathcal{F}_T$, $\mathcal{G}_0 = \mathcal{F}^\Lambda$, while $\mathcal{F}_0$ is trivial.

We can easily show that μ (1.1.8) is a martingale random field per Definition 1.1.1 with respect to *both* $\mathbb{F}$ and $\mathbb{G}$ as defined above. In the rest of this introduction we assume that $\mathcal{I}$ is as described in Section 1.1.1 with μ given by (1.1.8) and $\mathbb{F}$ and $\mathbb{G}$ are as in (1.1.9) and (1.1.10) respectively (unless otherwise noted).

1.1.4 Time-changed Lévy processes

Here we detail the connection between the processes with CII and time-changed Lévy processes and discuss their application to mathematical finance. Assume that $\mathcal{Z}_o = \mathbb{R}_0$ and Λ (1.1.6) is given by

$$\begin{aligned} \Lambda^B(ds, \{0\}) &= \lambda_s^B ds \\ \Lambda^H(ds, dz) &= \nu(dz) \lambda_s^H ds, \quad z \in \mathbb{R}_0, \end{aligned}$$

where λ^W and λ^N are positive stochastic processes satisfying some additional conditions described in the forthcoming Section 3.2.1. Additionally ν is the Lévy measure of some Levy process, i.e. deterministic and satisfying $\int_{\mathbb{R}_0} z^2 \nu(dz) < \infty$.

Define $B_t := B([0, t] \times \{0\})$, $\eta_t := \int_0^t \int_{\mathbb{R}_0} z \tilde{H}(ds, dz)$ and $\hat{\Lambda}_t^H := \int_0^t \lambda_s^H ds$, for $t \in [0, T]$. The distribution of B is connected to the Brownian motion and the distribution of η is connected to a centered pure jump Lévy process with Lévy measure ν in the following sense [100, Theorem 3.1] (see also [51]):

Theorem 1.1.4. *Let $W_t, t \in [0, T]$ be a Brownian motion and $N_t, t \in [0, T]$ be a centered pure jump Lévy process with Levy measure ν . Assume that both W and N are independent of Λ . Then B satisfies A1) and A2) if and only if, for any $t \geq 0$,*

$$B_t \stackrel{d}{=} W_{\Lambda_t^B},$$

and η satisfies A3) and A4) if and only if, for any $t \geq 0$,

$$\eta_t \stackrel{d}{=} N_{\hat{\Lambda}_t^H}.$$

The time-changed Lévy processes are used in mathematical finance for modeling the asset

price S by

$$\begin{aligned} dS_t &= S_t \left(\int_{\mathbb{R}} \psi_t(z) \mu(dt, dz) \right) \\ &= S_t \left(\psi_t(0) dB_t + \int_{\mathbb{R}_0} \psi_t(z) \tilde{H}(dt, dz) \right) \quad S_0 > 0. \end{aligned} \quad (1.1.11)$$

for some suitable $\psi \in \mathcal{I}$. Examples with the DSPP occur in [25, 74], and with time-changed Brownian motions in [8, 52, 54, 102]. Thus the calculus and maximum principles for the PCII has relevance for the study of these asset models, in particular for portfolio optimization.

1.2 On chaos representation and orthogonal polynomials for the doubly stochastic Poisson process

Let N be a Poisson random field on $[0, T] \times \mathcal{Z}$ generating a filtration (augmented with the zero measure sets) $\mathbb{F}^N : \{\mathcal{F}_t^N, t \in [0, T]\}$ on some probability space $(\tilde{\Omega}, \mathcal{F}_T^N, \tilde{\mathbb{P}})$. The space $(\tilde{\Omega}, \mathcal{F}_T^N, \tilde{\mathbb{P}})$ can be decomposed into orthogonal subspaces of iterated integrals or similar products of the values of N , see [36, 43, 84]. From here the Skorohod and the Malliavin derivative (in the sense of the difference operator) can be defined. Of particular interest to us is the Clark-Ocone formula, for $\xi \in L_2(\tilde{\Omega})$ we have

$$\xi = E[\xi] + \int_0^T \int_{\mathcal{Z}} \mathbb{E}[D_{s,z}\xi | \tilde{\mathcal{F}}_s] \tilde{N}(ds, dz)$$

where D is the Malliavin derivative. Note however that this requires ξ to be Malliavin differentiable or one has to apply white noise techniques.

The orthogonal structures and subsequent calculus depend on the independent increments of the Lévy process. Thus we cannot apply the same type of calculus directly for the DSPP, but a similar framework can be built since the DSPP has conditionally independent increments. This has been studied in [107], where Skorohod type integration, Malliavin calculus and a Clark-Ocone type formula for random variables that are “Malliavin differentiable” are developed in the setting of processes with CII.

In chapter 2 we expand the study of orthogonal subspaces on the probability space generated by the DSPP. The key to this study is the α -multilinear forms, which are random variables of type

$$\beta \prod_{j=i}^p \tilde{H}(\Delta_j), \quad \Delta_j \subset [0, T] \times \mathcal{Z}.$$

Here the number p denotes the order of the α -multilinear form, $\tilde{H}$ is given by (1.1.7), $\Delta_1, \dots, \Delta_p$ are pairwise disjoint sets and β is an $\mathcal{F}^\Lambda$ -measurable random variable. The α -multilinear forms of order 0 are the $\mathcal{F}^\Lambda$ -measurable random variables with finite moments of all orders.

All α -multilinear forms of different orders are orthogonal. This can be used to create orthogonal subspaces $\mathbb{H}^p$, $0 \leq p < \infty$, generated by linear combinations of α -multilinear forms of

order p . Moreover, $\bigcup_{p=0}^{\infty} \mathbb{H}^p = L_2(\Omega)$. Thus we find the chaos expansion as a sum of elements from the orthogonal subspaces, for any $\xi \in L_2(\Omega)$,

$$\xi = \sum_{j=0}^{\infty} \xi_j, \quad \xi_j \in \mathbb{H}^j.$$

On these subspaces we investigate the non-anticipating stochastic derivative $\mathcal{D}$ and another anticipative derivative operator D^a . The operator D^a will turn out to coincide with the Malliavin operator D of [107], and the non-anticipating stochastic derivative $\mathcal{D}$ is the projection of D^a on $\mathbb{G}$. Besides allowing for an alternate representation of the Malliavin derivative via D^a we have also extended the Clark Ocone formula for DSPP to all square integrable random variables (without the use of white noise techniques).

1.3 BSDEs for time-changed Lévy processes and application to optimal control

Chapter 3 is based on the accepted version of the paper, however Section 3.7 on necessary maximum principles for optimal control is new.

In chapter 3 we show existence and uniqueness for BSDEs driven by processes with CII. The traditional arguments for the existence of the BSDE relies on a martingale representation (or equivalently an integral representation for all square integrable random variables), see, e.g. the seminal papers [88, 104]. Such a representation exists for filtrations generated by Lévy processes see, e.g. [79] and point processes [19, 21, 31, 55]. It does not, however, exist in general.

Let μ , $\mathbb{G}$ and $\mathbb{F}$ be as in Section 1.1.3. Then there is a martingale representation for the filtration $\mathbb{G}$ [58, Theorem 3.4.34]. In Chapter 3 we also show a constructive proof of the integral representation which shows how and why the filtration $\mathbb{G}$ occur naturally in the integrand. Every square integrable $\mathbb{G}$ -martingale M has representation

$$M_t = \mathbb{E}[M_T | \mathcal{F}^\Lambda] + \int_0^t \int_{\mathbb{R}_0} \phi_s(z) \mu(ds, dz), \quad t \in [0, T],$$

(for some $\phi \in \mathcal{I}$ that depend on M). Note however that M_0 is not deterministic, but $\mathcal{F}^\Lambda$ -measurable since $\mathcal{G}_0 = \mathcal{F}^\Lambda$. Using the martingale representation we can prove the existence of a BSDE of type

$$Y_t = \xi + \int_t^T g_s(\lambda_s, Y_s, \phi_s) ds - \int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz), \quad t \in [0, T]. \quad (1.3.1)$$

Here g is the driver, ξ is a square integrable $\mathcal{G}_T$ -measurable random variable, $\phi \in \mathcal{I}$ and Y is $\mathbb{G}$ -adapted. The uniqueness of the solution can be proved by the conditions set on g . The

element ϕ can be described via the non-anticipating derivative as the solution to

$$\phi_t(z) = \mathcal{D}_{t,z} \left(\xi + \int_0^T g_s(\lambda_s, Y_s, \phi_s) ds \right)$$

When μ is purely of DSPP type, this also means that the results of Chapter 2 applies to describing the structure of ϕ .

The filtration $\mathbb{F}$ is more natural for modeling purposes, it is the filtration generated by the noise. A highly useful fact about the BSDE framework is that we can use the BSDE for the calculus, and then project the results to $\mathbb{F}$ in applications. In particular, a BSDE of type (1.3.1) appears in the adjoint equation when considering problems of type (1.0.1) for $\mathbb{G}$ -adapted controls. The projection on $\mathbb{F}$ of the same BSDE appears when considering $\mathbb{F}$ -adapted controls. We use this fact to find a sufficient maximum principle using concavity conditions and a necessary maximum principle by considering perturbations of the optimal control. As an application we show an explicit solution to the mean-variance portfolio selection problem.

We also find a representation theorem for linear BSDEs and develop a comparison theorem.

1.4 Maximum principles for martingale random fields

Chapter 4 deals with maximum principles for martingale random fields. The goal is performance functionals of type (1.0.1). The essential idea is a family of optimization techniques that require an evaluation expectation of products of the form

$$\mathbb{E} \left[\xi \int_0^t \phi(s) \mu(ds) \right]$$

where μ is martingale random field as in Section 1.1.1 (but we do not assume that μ has CII) and ξ a random variable. Malliavin calculus (see [43] for a detailed description of the Malliavin calculus and the Skorohod integral) has been proposed [41, 82]

$$\mathbb{E} \left[\xi \int_0^t \phi(s) \mu(ds) \right] = \mathbb{E} \left[\int_0^t (D_s(\xi)) \phi(s) ds \right],$$

here D is the Malliavin operator. The Malliavin operator D is the dual of the Skorohod integral. The Skorohod integral is an extension of the Itô integral defined for Lévy processes, but the domain of Malliavin operator, $\mathbb{D}_{1,2}$, is strictly smaller than $L_2(\Omega)$. Indeed, whether a random variable is an element of $\mathbb{D}_{1,2}$ depends on the finite expectation of an infinite series expansion, which is difficult to check in practice. Using instead the non-anticipating stochastic derivative, as the dual operator of the Itô integral, enables us to work with the more natural L_2 integrals and L_2 random variables. Also, the results does not depend on Lévy type noise and goes beyond the DSPP, as any martingale random field can be considered.

Using a perturbation approach we find necessary conditions for a solution to (1.0.1). We let the state process X be $\mathbb{F}$ adapted and apply the non-anticipating derivative with respect to a

filtration $\mathbb{G}$. However we look for controls adapted to a filtration $\mathbb{F}$ which is a subfiltration of $\mathbb{G}$. This enables problems of partial information (in the sense of filtrations, not noisy observations) and is also useful for problems with martingale random fields with CII.

For the perturbation approach, we also discuss an issue with a set of assumptions that has been repeated in the literature. We show that these assumptions have implications for how we can choose the set $\mathcal{U}$ where the admissible controls are taking their values, namely that we must have $\mathcal{U} = \mathbb{R}^n$. To avoid this issue we state our maximum principle with less restrictive assumptions.

Indeed, we deduce sufficient conditions for a unique optimal control for martingale random fields with CII when $\mathbb{G}$ and $\mathbb{F}$ are as in section 1.1.3. Furthermore, in this setting the study in Chapter 2 on the DSPP and the non-anticipating stochastic derivative in applies. As an example we solve a utility maximization problem for bonds with credit risk modeled by the DSPP.

1.5 Optimal portfolio problems in the presence of default and the value of information

In chapter 5 we discuss utility maximization for a defaultable asset. We find sufficient and necessary conditions for optimal control without assuming the semi-martingale property on the noises in the asset model. The default is modelled by a general point process H .

Let S_t , $t \in [0, T]$, be the price of a tradeable asset at time t . We assume that the price model is driven by a Wiener process B and/or a Poisson random measure N . However, we also want to model “default events” with the point process H . We model S with the following SDE:

$$dS_t = S_t \left(\sigma dB_t + \int_{\mathbb{R}_0} \theta \tilde{N}(dt, dz) + \kappa dH_t \right).$$

Here $\tilde{N}$ is the centered Poisson random measure and σ , θ and κ are stochastic processes to be described further in Chapter 5 (for the purposes of this introduction we may assume that they are constants). Remark that B , H and $\mathbb{G}$ are *not* as discussed in the earlier sections.

Let $\mathbb{G}$ be the filtration that represents the information available to the investor and assume that H is $\mathbb{G}$ -adapted. The following important modeling issue now arise:

Are B and (the integral over) N semi-martingales with respect to $\mathbb{G}$?

If B and (the integral over) N are not semi-martingales, we cannot define integration in the usual way for $\mathbb{G}$ -adapted integrands. This is important because $\mathbb{G}$ -adapted integrands naturally occur when modeling the investor’s wealth and solving the utility maximization problem. We could set conditions on H and $\mathbb{G}$ which ensure that B and (the integral over) N remain semi-martingales under $\mathbb{G}$. This “enlargement of filtration”-techniques has been used in, e.g [14, 27, 33, 71, 62, 60]. It is, for instance, sufficient that H is independent of B and N and that $\mathbb{G}$ contains no “anticipating information”.

Here we instead study the utility maximization without assuming that B and (the integral over) N are semi-martingales. We do so by modeling the integrals as *forward integrals*, an extension of the Itô integral that do not require the integrands to be adapted to a specific filtration.

See, e.g. [96, 97, 95] on integration with respect to continuous stochastic processes and [39] for the case of integration with respect to the centered Poisson random measure.

Without the a priori semi-martingale assumption we can freely consider

- Models with insider information.
- Models with a complicated dependence between H and B and N (that include anticipating information).

Our result is sufficient and necessary conditions for an investment strategy to be either a local maximum (see Definition 5.3.2) or in fact optimal. We also provide several examples for the logarithmic utility. Moreover we prove the following: *If a local maximum for the utility maximization problem exist, then B and (the integral over) N are semi-martingales under $\mathbb{G}$.* Thus that B and (the integral over) N are semi-martingales (with respect to $\mathbb{G}$) is necessary for the existence of a solution to the utility maximization problem. Or equivalently, the existence of a solution to the utility maximization problem implies the semi-martingale property of the noises.

Chapter 2

On chaos representation and orthogonal polynomials for the doubly stochastic Poisson process

Abstract

In an L_2 -framework, we study various aspects of stochastic calculus with respect to the centered doubly stochastic Poisson process. We introduce an orthogonal basis via multilinear forms of the value of the random measure and we analyze the chaos representation property. We review the structure of non-anticipating integration for martingale random fields and in this framework we study non-anticipating differentiation. We present integral representation theorems where the integrand is explicitly given by the non-anticipating derivative.

Stochastic derivatives of anticipative nature are also considered: The Malliavin type derivative is put in relationship with another anticipative derivative operator here introduced. This gives a new structural representation of the Malliavin derivative based on simple functions. Finally we exploit these results to provide a Clark-Ocone type formula for the computation of the non-anticipating derivative.

2.1 Introduction

The doubly stochastic Poisson process (DSPP) also known as the Cox process, was introduced in [28] as a generalization of the Poisson process in the sense that the intensity is stochastic. Models based on DSPP's are used in risk theory, in the study of ruin probabilities in insurance and insurance-linked securities pricing, and for stochastic volatility see e.g. [14, 25, 47, 74].

For a given doubly stochastic Poisson process H with intensity α , we investigate some elements of stochastic calculus for $\tilde{H} := H - \alpha$, i.e. the *centered doubly stochastic Poisson process* (CDSPP) on a quite general Hausdorff topological space X . The stochastic intensity α is assumed non-atomic. The paper is dedicated to the study of the structure of L_2 -spaces generated by the noise and the non-anticipating integration and differentiation schemes with stochastic integral representations in view. One foreseen application of such integral representations is in the study of backward stochastic differential equations, which are presented in Chapter 3.

First we show that the observations of $\tilde{H} = H - \alpha$ give complete information on both H and α . Specifically, the σ -algebra generated by $\tilde{H}$ coincides with the one generated by H and α . With respect to the space $L_2(\Omega, \mathcal{F}, \mathbb{P})$ generated by the CDSPP, we suggest an orthogonal system of polynomials which lead to a chaos expansion type of result. This orthogonal system is based on what we call α -multilinear forms. These prove to be key constructive elements in our proofs.

After this analysis on a general X , we specify the study to the time-space $X = (0, T] \times \mathcal{Z}$ with the total ordering induced by time. Here we introduce an information structure associated to the CDSPP. We consider the filtration $\mathbb{G}$ generated by the CDSPP augmented by the knowledge of the whole intensity α . Note that, with respect to $\mathbb{G}$, the CDSPP is a stochastic measure with conditionally independent values. In this setup we study elements of stochastic integration and differentiation. We find a stochastic integral representation for all elements in $L_2(\Omega, \mathcal{F}, \mathbb{P})$ by interpreting the CDSPP as a martingale random field (see e.g. [38]) and applying the corresponding Itô stochastic integration scheme. The representation of $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ is explicit in the sense that the integrand is uniquely determined as the non-anticipating stochastic derivative $\mathcal{D}\xi$ with respect to the CDSPP. The non-anticipating derivative, introduced in [34] and later developed to include Lévy type random measures (see [35, 37]) and martingale random fields as integrators (see [38]) is defined by the linear operator adjoint to the Itô stochastic integral. A general formula for the calculus is here given in terms of limit of specific simple stochastic functions. In particular, non-anticipating derivatives are a natural tool to study mean-variance hedging, see [11, 34, 38].

When discussing non-anticipating differentiation, the connections with the well-known correspondent of the Clark-Ocone formulae have to be taken into account. A first study of chaos expansions in terms of iterated integrals for processes with conditionally independent increments can be found in [107]. Starting from this set up, a Malliavin derivative operator is defined. In the present paper we discuss explicitly the relationship between the orthogonal polynomials here suggested and the Itô type iterated integrals and we retrace the relationships between the non-anticipating derivative and the Clark-Ocone formula based on the Malliavin derivative operator given in the literature. Our study however takes a different approach to Malliavin calculus. In fact we introduce a new anticipative derivative operator D^a as a limit of specific simple stochastic functions. Because of its particular structure, it is immediate to see that the non-anticipative derivative $\mathcal{D}\xi$ at time t is the projection of D^a on the information $\mathcal{G}_t$. On the other side we prove that this operator coincides with the Malliavin derivative $D\xi$ as introduced in [107]. These arguments provide a new structural approach to the Malliavin derivative.

We have partially considered integration with respect to the smaller filtration generated by the CDSPP only. Based on the martingale structure the non-anticipative differentiation can be carried through. However, we remark that there is no structure of conditional independence in this case and the study of anticipative differentiation is rather different. The study of stochastic differentiation in this setting will be developed separately.

To conclude we remark that stochastic integral representations have been investigated in [19, 21, 31, 50, 55] for general point processes. Our contribution differs because we consider the filtration generated by the CDSPP, which is larger than the filtration generated by the DSPP alone.

The paper is organized as follows. Section 2 provides the basic information on DSPP and

CDSPP on a general space X . Multilinear forms and chaos expansions are studied in section 3. For $X = (0, T] \times \mathcal{Z}$, stochastic non-anticipating integration and martingale random fields are discussed in section 4. Section 5 presents the non-anticipating derivative $\mathcal{D}$. A review on iterated integrals and their connection with multilinear forms is detailed in section 6. Finally section 7 presents the anticipative derivatives D^a and D , their computation, and their relationship with the non-anticipating derivative $\mathcal{D}$ via a Clark-Ocone type formula.

2.2 The doubly stochastic Poisson process

Let X be a locally compact, second countable Hausdorff topological space. Under these conditions, there exists a complete, separable metric μ generating the topology on X . In particular this implies that X is σ -compact, i.e. that it admits representation as a countable union of compact sets, and that the topology on X has a countable basis consisting of precompact sets, i.e. sets with compact closure. We denote $\mathcal{B}_X$ the Borel σ -algebra of X and $\mathcal{B}_X^c$ the precompacts of $\mathcal{B}_X$. The stochastic elements considered in the paper are related to the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Let α be a (positive) random measure on X . We assume that α is non-atomic, meaning that $\mathbb{P}(\alpha(\{x\}) = 0 \text{ for all } x \in X) = 1$. For later use in the study of the polynomials we also assume that

$$\mathbb{E}[e^{c\alpha(\Delta)}] < \infty, \quad c \in \mathbb{R}, \Delta \in \mathcal{B}_X^c, \quad (2.2.1)$$

i.e. the moment generating function of $\alpha(\Delta)$ is well defined on the whole real line when Δ is precompact. This will in turn imply

$$\mathbb{E}[\alpha(\Delta)^k] < \infty \quad \text{for all } \Delta \in \mathcal{B}_X^c, k = 1, 2, \dots \quad (2.2.2)$$

We remark that condition (2.2.1) is satisfied if $X = [0, \infty) \times \mathbb{R}_0$ and $\alpha(dt, dz) = \nu(dz)dt$ with ν a Lévy measure on $\mathbb{R}_0$ and dt the Lebesgue measure on $[0, \infty)$, i.e. the case of pure jump Lévy processes.

Let us define

$$V(\Delta) := \mathbb{E}[\alpha(\Delta)], \quad \Delta \in \mathcal{B}_X.$$

We note that V is a non-atomic σ -finite measure (see e.g. [64, chapter 1.2]), which is finite at least on all precompact sets. The σ -algebra generated by α will be denoted $\mathcal{F}^\alpha$.

Let H be a random measure on X and let $\mathcal{F}_\Delta^H$ denote the σ -algebra generated by $H(\Delta')$, $\Delta' \in \mathcal{B}_X : \Delta' \subset \Delta$ (with $\Delta \in \mathcal{B}_X$). Set $\mathcal{F}^H$ to be the σ -algebra generated by all the values of H .

Definition 2.2.1. *The random measure H is a doubly stochastic Poisson process (DSPP) if*

$$A1) \quad \mathbb{P}(H(\Delta) = k \mid \mathcal{F}^\alpha) = \frac{\alpha(\Delta)^k}{k!} e^{-\alpha(\Delta)},$$

$$A2) \quad \mathcal{F}_{\Delta_1}^H \text{ and } \mathcal{F}_{\Delta_2}^H \text{ are conditionally independent given } \mathcal{F}^\alpha \text{ whenever } \Delta_1 \text{ and } \Delta_2 \text{ are disjoint sets.}$$

In particular, the conditional independence A2) implies that

$$\mathbb{E}\left[f(H(\Delta_1)) \mid \mathcal{F}_{\Delta_2}^H \vee \mathcal{F}^\alpha\right] = \mathbb{E}\left[f(H(\Delta_1)) \mid \mathcal{F}^\alpha\right],$$

whenever $\Delta_1, \Delta_2 \in \mathcal{B}_X$, $\Delta_1 \cap \Delta_2 = \emptyset$ and for $f : \mathbb{R} \rightarrow \mathbb{R}$ such that the conditional expectation is well defined. From A1) we have

$$\mathbb{E}\left[f(H(\Delta)) \mid \mathcal{F}^\alpha\right] = \sum_{k=0}^{\infty} f(k) \frac{\alpha(\Delta)^k}{k!} e^{-\alpha(\Delta)}, \quad \Delta \in \mathcal{B}_X, \quad (2.2.3)$$

and in particular

$$\mathbb{E}\left[H(\Delta) \mid \mathcal{F}^\alpha\right] = \alpha(\Delta), \quad \Delta \in \mathcal{B}_X. \quad (2.2.4)$$

From the above formulae the following ones are obtained [50, Lemma 3a p23]:

$$\begin{aligned} \mathbb{E}\left[H(\Delta)\right] &= \mathbb{E}[\alpha(\Delta)] = V(\Delta) \\ \text{Var}\left(H(\Delta)\right) &= \mathbb{E}[\alpha(\Delta)] + \text{Var}(\alpha(\Delta)) \\ \mathbb{E}\left[H(\Delta)^2\right] &= \mathbb{E}[\alpha(\Delta)] + \mathbb{E}[\alpha(\Delta)^2] \end{aligned}$$

for $\Delta \in \mathcal{B}_X$ with $V(\Delta) < \infty$. In the case $V(\Delta) = \infty$, the above relationships hold but clearly $\text{Var}(H(\Delta)) = \mathbb{E}[H(\Delta)^2] = \mathbb{E}[H(\Delta)] = \infty$.

Definition 2.2.2. *The centered doubly stochastic Poisson process (CDSPP) is the signed random measure $\tilde{H} := H - \alpha$, ie*

$$\tilde{H}(\Delta) := H(\Delta) - \alpha(\Delta), \quad \Delta \in \mathcal{B}_X. \quad (2.2.5)$$

We denote $\mathcal{F}^{\tilde{H}}$ the filtration generated by $\tilde{H}$. For any $\Delta \in \mathcal{B}_X$ with $V(\Delta) < \infty$, the conditional first moment is

$$\mathbb{E}\left[\tilde{H}(\Delta) \mid \mathcal{F}^\alpha\right] = 0.$$

and the conditional second moment is

$$\mathbb{E}\left[\tilde{H}(\Delta)^2 \mid \mathcal{F}^\alpha\right] = \alpha(\Delta).$$

and thus

$$\mathbb{E}\left[\tilde{H}(\Delta)^2\right] = \text{Var}(\tilde{H}(\Delta)) = \mathbb{E}[\alpha(\Delta)] = V(\Delta). \quad (2.2.6)$$

For the remaining conditional moments, the following recurrence formula holds:

Proposition 2.2.3.

$$\begin{aligned} \mathbb{E}\left[\tilde{H}(\Delta)^3 \mid \mathcal{F}^\alpha\right] &= \alpha(\Delta) \\ \mathbb{E}\left[\tilde{H}(\Delta)^n \mid \mathcal{F}^\alpha\right] &= \alpha(\Delta) + \alpha(\Delta) \sum_{k=2}^{n-2} \binom{n}{k} \mathbb{E}\left[\tilde{H}(\Delta)^k \mid \mathcal{F}^\alpha\right], \quad n \geq 4. \end{aligned} \quad (2.2.7)$$

Proof. The formulae are obtained by induction for the Poisson distribution in [91, Section 3].

Those computations can easily be adapted to our case using (2.2.3). $\square$

Corollary 2.2.4. *For $n \geq 4$, we have that*

$$\mathbb{E}[\tilde{H}(\Delta)^n] < \infty$$

if and only if

$$\begin{aligned} \mathbb{E}[\alpha(\Delta)^{n/2}] &< \infty \quad \text{for } n \text{ even,} \\ \mathbb{E}[\alpha(\Delta)^{(n-1)/2}] &< \infty \quad \text{for } n \text{ odd.} \end{aligned}$$

Proof. The result follows from an argument by induction using (2.2.7). $\square$

Remark 2.2.5. We remark that, in view of Corollary 2.2.4, the assumption (2.2.1) is sufficient to ensure that $\tilde{H}(\Delta)$ has finite moments of all orders for $\Delta \in \mathcal{B}_X^c$.

For the arguments presented in the sequel it is crucial to investigate the relationship between the σ -algebras $\mathcal{F}^{\tilde{H}}$ and $\mathcal{F}^H \vee \mathcal{F}^\alpha$. While it is immediate to see that $\mathcal{F}^{\tilde{H}} \subseteq \mathcal{F}^H \vee \mathcal{F}^\alpha$, the opposite relationship is more delicate. Here after we introduce a dissecting system on X which is instrumental in the study of the considered random measures and associated structures. Recall that $\mathcal{B}_X^c$ is a ring generating the topology on X and that X is a Hausdorff topological space such that $X = \bigcup_{n=1}^\infty X_n$, where X_n , $n = 1, 2, \dots$ is a growing sequence of compacts. Hence $V(X_n) < \infty$. Denote $|\Delta| := \sup_{x,y \in \Delta} \mu(x,y)$, $\Delta \subset X$, where μ is the metric in X . Then $|X_n| < \infty$ for all n .

Being V non-atomic, for every n and $\epsilon_n > 0$, there exists a partition of X_n , i.e. a finite family of pairwise disjoint sets:

$$\Delta_{n,1}, \dots, \Delta_{n,K_n} \in \mathcal{B}_X^c : \quad X_n = \bigcup_{k=1}^{K_n} \Delta_{n,k} \quad (2.2.8)$$

such that $\sup_{k=1,\dots,K_n} V(\Delta_{n,k}) \leq \epsilon_n$ and $\sup_{k=1,\dots,K_n} |\Delta_{n,k}| \leq \epsilon_n$.

Let us consider a decreasing sequence $\epsilon_n \searrow 0$, $n \rightarrow \infty$. Then, based on (2.2.8), we give the following definition.

Definition 2.2.6. *A dissecting system of X is the sequence of partitions of X :*

$$\Delta_{n,1}, \dots, \Delta_{n,K_n+1}, \quad n = 1, 2, \dots \quad (2.2.9)$$

with $\bigcup_{k=1}^{K_n} \Delta_{n,k} = X_n$ from (2.2.8) and $\Delta_{n,K_n+1} := X \setminus X_n$, satisfying the nesting property:

$$\Delta_{n,k} \cap \Delta_{n+1,j} = \Delta_{n+1,j} \text{ or } \emptyset \quad (2.2.10)$$

for all $k = 1, \dots, K_n + 1$ and $j = 1, \dots, K_{n+1} + 1$.

We remark that, from (2.2.8) and (2.2.10), we have

$$\sup_{k=1,\dots,K_n} V(\Delta_{n,k}) \leq \epsilon_n \rightarrow 0, \text{ and } \sup_{k=1,\dots,K_n} |\Delta_{n,k}| \leq \epsilon_n \rightarrow 0 \quad n \rightarrow \infty \quad (2.2.11)$$

We can refer to e.g. [64] and [29] for more on dissecting systems and partitions.

Lemma 2.2.7. *For any $\Delta \in \mathcal{B}_X$ such that $\alpha(\Delta) < \infty$ $\mathbb{P}$ -a.s. we have that*

$$\sup_{k=1, \dots, K_n+1} \alpha(\Delta \cap \Delta_{n,k}) \longrightarrow 0, \quad n \rightarrow \infty \quad \mathbb{P}\text{-a.s.} \quad (2.2.12)$$

Proof. The sets

$$\tilde{\Delta}_{n,k} := \Delta \cap \Delta_{n,k}, \quad k = 1, \dots, K_n, \quad n = 1, 2, \dots$$

constitute a dissecting system of Δ . Note that $\alpha(\Delta_{k,n}) < \infty$ $\mathbb{P}$ -a.s. for all k and n . Let $\tilde{\Omega}$ be the event where α is non-atomic and $\alpha(\Delta) < \infty$. Then $\mathbb{P}(\tilde{\Omega}) = 1$. From (2.2.9) we have

$$\alpha(\tilde{\Delta}_{n+1,j}, \omega) \leq \sup_{k=1, \dots, K_n+1} \alpha(\tilde{\Delta}_{n,k}, \omega) \leq \alpha(\Delta, \omega), \quad \omega \in \tilde{\Omega}$$

for all $j = 1, \dots, K_n + 1$. Hence, for every $n = 1, 2, \dots$, we have

$$\sup_{j=1, \dots, K_{n+1}+1} \alpha(\tilde{\Delta}_{n+1,j}, \omega) \leq \sup_{k=1, \dots, K_n+1} \alpha(\tilde{\Delta}_{n,k}, \omega), \quad \omega \in \tilde{\Omega}.$$

We denote $A(\omega) := \lim_{n \rightarrow \infty} \sup_{k=1, \dots, K_n+1} \alpha(\tilde{\Delta}_{n,k}, \omega)$, $\omega \in \tilde{\Omega}$. Naturally $A(\omega) \geq 0$, but we need to prove $A(\omega) = 0$. We proceed by contradiction. Set $\tilde{\Omega}_0 := \{\omega \in \tilde{\Omega} \mid A(\omega) > 0\}$ and suppose $\mathbb{P}(\tilde{\Omega}_0) > 0$. For each n there exists a set $\tilde{\Delta}_{n,\delta(n)}$ such that $\alpha(\tilde{\Delta}_{n,\delta(n)}, \omega) \geq A(\omega) > 0$, $\omega \in \tilde{\Omega}_0$. Comparing $\tilde{\Delta}_{n,\delta(n)}$ with the sets $\tilde{\Delta}_{n-1,j}$, $j = 1, \dots, K_{n-1} + 1$, we see that there is a set $\tilde{\Delta}_{n-1,\delta(n-1)}$ such that $\tilde{\Delta}_{n-1,\delta(n-1)} \supseteq \tilde{\Delta}_{n,\delta(n)}$.

Hence there exists a decreasing sequence of sets

$$\tilde{\Delta}_{n,\delta(n)}, \quad n = 1, 2, \dots$$

such that for every n , $\tilde{\Delta}_{n,\delta(n)}$ is an element of the dissecting system of Δ and $0 < A(\omega) \leq \alpha(\tilde{\Delta}_{n,\delta(n)}, \omega)$ ($\omega \in \tilde{\Omega}_0$). On the other side, from the property (2.2.11) of the dissecting system on X , and hence on Δ , the limit of a decreasing sequence of sets is either empty or a singleton. Thus we have

$$\lim_{n \rightarrow \infty} \alpha(\tilde{\Delta}_{n,\delta(n)}, \omega) = 0, \quad \omega \in \tilde{\Omega}_0,$$

since α is a non-atomic measure for $\omega \in \tilde{\Omega}_0$. This is a contradiction, and hence $A(\omega) = 0$ for all $\omega \in \tilde{\Omega}_0$. $\square$

Theorem 2.2.8. *The following equality holds:*

$$\mathcal{F}^{\tilde{H}} = \mathcal{F}^H \vee \mathcal{F}^\alpha.$$

Proof. It is sufficient to show that $H(\Delta)$ and $\alpha(\Delta)$ are $\mathcal{F}^{\tilde{H}}$ -measurable for any $\Delta \in \mathcal{B}_X^c$. Let

$\Delta \in \mathcal{B}_X^c$ and recall its representation

$$\Delta = \bigcup_k^{K_n+1} \tilde{\Delta}_{n,k} = \bigcup_k^{K_n+1} (\Delta \cap \Delta_{n,k}), \quad n = 1, 2, \dots,$$

as a pairwise disjoint union of sets obtained from the dissecting system (2.2.9) of X . Consider

$$g_n(\Delta) := \sum_{k=1}^{K_n+1} \text{ceil}(\tilde{H}(\tilde{\Delta}_{n,k})) = \sum_{k=1}^{K_n+1} \text{ceil}(H(\tilde{\Delta}_{n,k}) - \alpha(\tilde{\Delta}_{n,k})),$$

where $\text{ceil}(y)$ is the smallest integer greater than y . The random variables $g_n(\Delta)$, $n = 1, \dots$, are clearly $\mathcal{F}^{\tilde{H}}$ -measurable. From Lemma 2.2.7 there exists for $\mathbb{P}$ -a.a. ω , a $N(\omega) \in \mathbb{N}$ such that $\sup_{k=1, \dots, K_n+1} \alpha(\tilde{\Delta}_{n,k}, \omega) < 1$ for $n > N(\omega)$. Then we have

$$\lim_{n \rightarrow \infty} \text{ceil}(H(\tilde{\Delta}_{n,k}) - \alpha(\tilde{\Delta}_{n,k})) = H(\tilde{\Delta}_{n,k}) \quad \mathbb{P}\text{-a.s.}$$

Thus

$$\lim_{n \rightarrow \infty} g_n(\Delta) = \lim_{n \rightarrow \infty} \sum_{k=1}^{K_n+1} \text{ceil}(\tilde{H}(\tilde{\Delta}_{n,k})) = H(\Delta) \quad \mathbb{P}\text{-a.s.}$$

and $H(\Delta)$ is a pointwise limit of $\mathcal{F}^{\tilde{H}}$ -measurable functions. Since $\alpha(\Delta) = H(\Delta) - \tilde{H}(\Delta)$, we also have that $\alpha(\Delta)$ is $\mathcal{F}^{\tilde{H}}$ -measurable. $\square$

Note that the initial assumption that α is $\mathbb{P}$ -a.s. non-atomic is crucial for this result. On the other side we remark that the assumption (2.2.2) is here not required.

Theorem 2.2.8 can be regarded as an extension of a result proved for a time-changed Lévy processes with independent time-change in [105].

2.3 Multilinear forms, polynomials, and chaos expansions

In this section we construct a system of multilinear forms and show how they describe the intrinsic orthogonal structures in $L_2(\Omega, \mathcal{F}, \mathbb{P})$. Here and in the sequel we set $\mathcal{F} = \mathcal{F}^{\tilde{H}} = \mathcal{F}^H \vee \mathcal{F}^\alpha$, see Theorem 2.2.8.

Definition 2.3.1. *For any group of pairwise disjoint sets $\Delta_1, \dots, \Delta_p \in \mathcal{B}_X^c$, an α -multilinear form of order p is a random variable of type*

$$\beta \prod_{j=1}^p \tilde{H}(\Delta_j), \quad p \geq 1, \tag{2.3.1}$$

where β is an $\mathcal{F}^\alpha$ -measurable random variables with finite moments of all orders. The 0-order α -multilinear forms are the $\mathcal{F}^\alpha$ -measurable random variable with finite moments of all orders.

This definition is a generalization of the one given in [36, page 7]: A p -order multilinear

form of the values $\tilde{H}(\Delta_j)$, $j = 1, \dots, p$, is a random variable of type

$$\prod_{j=1}^p \tilde{H}(\Delta_j), \quad p \geq 1. \quad (2.3.2)$$

The 0-order multilinear forms are the constants.

Note that any α -multilinear form is an element of $L_2(\Omega, \mathcal{F}, \mathbb{P})$. In fact, by assumption (2.2.2), the following holds:

$$\mathbb{E}[\xi^2] = \mathbb{E}\left[\beta^2 \prod_{j=1}^p \mathbb{E}[\tilde{H}(\Delta_j)^2 | \mathcal{F}^\alpha]\right] = \mathbb{E}\left[\beta^2 \prod_{j=1}^p \alpha(\Delta_j)\right] < \infty. \quad (2.3.3)$$

In the sequel we will consider multilinear forms on the sets (2.2.8)-(2.2.9) of the dissecting system of X .

The present section completes and extends to the CDSPP the results presented in [36] in which measure based multilinear forms were introduced for the study of stochastic calculus for Lévy stochastic measures. In that case the structure of independence of the random measure values was heavily exploited. In particular we stress that the space $\mathbb{H}^p$ via (2.3.4) here below is a substantial element of novelty and it is crucial for the forthcoming analysis.

Definition 2.3.2. For $p \geq 1$ we write $\mathbb{H}^p$ for the subspace in $L_2(\Omega, \mathcal{F}, \mathbb{P})$ generated by the finite linear combinations of p -order α -multilinear form:

$$\sum_i \beta_i \prod_{j=1}^p \tilde{H}(\Delta_j^i). \quad (2.3.4)$$

Here above the sets Δ_j^i , $j = 1, \dots, p$, are pairwise disjoint and belong to the dissecting system (2.2.8)-(2.2.9) on X . The subspace $\mathbb{H}^0$ is the $\mathcal{F}^\alpha$ -measurable random variables with finite variance.

Remark 2.3.3. We may consider the multipliers β in Definition 2.3.2 to be finite products of the form $\prod_{i=1}^n \alpha(\Delta_i)$ with Δ_i , $i = 1, \dots, n$ pairwise disjoint sets from the dissecting system (2.2.8)-(2.2.9).

Remark 2.3.4. Let $p \geq 1$. By definition, for any $\xi \in \mathbb{H}^p$ there exists a sequence $\{\xi_m\}_m$ such that $\xi_m \rightarrow \xi$, $m \rightarrow \infty$ in $L_2(\Omega, \mathcal{F}, \mathbb{P})$, with $\xi_m = \sum_{l=1}^{L_m} \xi_{ml}$ and ξ_{ml} p -order α -multilinear forms. Note that we can always choose approximating sequences where the ξ_{ml} , $l = 1, \dots, L_m$ are orthogonal.

Lemma 2.3.5. For $p' \neq p''$, the subspaces $\mathbb{H}^{p'}$ and $\mathbb{H}^{p''}$ are orthogonal in $L_2(\Omega, \mathcal{F}, \mathbb{P})$.

Proof. We assume that $p'' > p'$. It is sufficient to prove the statement for $\xi' \in \mathbb{H}^{p'}$, $\xi'' \in \mathbb{H}^{p''}$ of type

$$\xi' = \beta' \prod_{i=1}^{p'} \tilde{H}(\Delta^i), \quad \xi'' = \beta'' \prod_{j=1}^{p''} \tilde{H}(\Delta^j)$$

where Δ^i , $i = 1, \dots, p'$ and Δ^j , $j = 1 \dots p'$ are two groups of pairwise disjoint sets (2.2.9) of the dissecting system of X . Note that, in view of the nesting property (2.2.10), there exists

$n \in \mathbb{N}$ such that all the sets above can be represented in terms of finite disjoint unions of elements from the same n 'th partition (2.2.9)-(2.2.10). Thus we can represent ξ' and ξ'' by finite sums of p' -order and p'' -order α -multilinear forms respectively over sets (2.2.9) in the same n 'th partition (2.2.10):

$$\begin{aligned}\xi' &= \beta' \sum_k \prod_{i=1}^{p'} \tilde{H}(\Delta_{n,k}^i) \\ \xi'' &= \beta'' \sum_l \prod_{j=1}^{p''} \tilde{H}(\Delta_{n,l}^j).\end{aligned}$$

To prove the statement it is then enough to verify that for all k, l ,

$$\mathbb{E} \left[\beta' \prod_{i=1}^{p'} \tilde{H}(\Delta_{n,k}^i) \beta'' \prod_{j=1}^{p''} \tilde{H}(\Delta_{n,l}^j) \right] = 0.$$

We remark that being $p'' > p'$, there is at least one set among $\Delta_{n,l}^j, j = 1, \dots, p''$ that is different from $\Delta_{n,k}^i, i = 1, \dots, p'$. Denote such a set by $\Delta_{n,l}^{\hat{j}}$. We have

$$\begin{aligned}\mathbb{E} \left[\beta' \prod_{i=1}^{p'} \tilde{H}(\Delta_{n,k}^i) \beta'' \tilde{H}(\Delta_{n,l}^{\hat{j}}) \prod_{\substack{j=1 \\ j \neq \hat{j}}}^{p''} \tilde{H}(\Delta_{n,l}^j) \right] \\ = \mathbb{E} \left[\beta' \beta'' \mathbb{E} \left[\prod_{i=1}^{p'} \tilde{H}(\Delta_{n,k}^i) \prod_{\substack{j=1 \\ j \neq \hat{j}}}^{p''} \tilde{H}(\Delta_{n,l}^j) \middle| \mathcal{F}^\alpha \right] \mathbb{E} \left[\tilde{H}(\Delta_{n,l}^{\hat{j}}) \middle| \mathcal{F}^\alpha \right] \right] = 0.\end{aligned}$$

By this we end the proof. □

Definition 2.3.6. We write $\mathbb{H}_p$ for the subspaces of $L_2(\Omega, \mathcal{F}, \mathbb{P})$ defined by:

$$\mathbb{H}_p := \sum_{q=0}^p \oplus \mathbb{H}^q.$$

Namely the subspaces generated by the linear combinations of α -multilinear forms:

$$\sum_i \beta_i \prod_{j=1}^{p_i} \tilde{H}(\Delta_j^i), \quad p_i \leq p. \quad (2.3.5)$$

We set

$$\mathbb{H} := \sum_{q=0}^{\infty} \oplus \mathbb{H}^q.$$

Lemma 2.3.7. Let $\Delta', \Delta'' \in \mathcal{B}_X : \Delta' \cap \Delta'' = \emptyset$. Consider $\mathcal{F}_{\Delta'}$ and $\mathcal{F}_{\Delta''}$ as the σ -algebras generated by $\tilde{H}(\Delta)$, $\Delta \in \mathcal{B}_X : \Delta \subset \Delta'$ and $\Delta \subset \Delta''$, respectively. Let $\xi' \in \mathbb{H}_{p'}$ be $\mathcal{F}_{\Delta'}$ -measurable and $\xi'' \in \mathbb{H}_{p''}$ be $\mathcal{F}_{\Delta''}$ -measurable. The product $\xi' \xi''$ is measurable with respect to $\mathcal{F}_{\Delta' \cup \Delta''}$ and belongs to $\mathbb{H}_{p'+p''}$.

Proof. If ξ' and ξ'' are of type (2.3.4), then clearly the product $\xi'\xi'' \in \mathbb{H}_{p'+p''}$ and it is $\mathcal{F}_{\Delta' \cup \Delta''}$ -measurable. In the general case, ξ' and ξ'' are approximated in $L_2(\Omega, \mathcal{F}, \mathbb{P})$ by sequences of elements ξ'_n and ξ''_n , $n = 1, 2, \dots$ of type (2.3.4):

$$\begin{aligned}\xi' &= \lim_{n \rightarrow \infty} \xi'_n = \lim_{n \rightarrow \infty} \sum_i \beta'_{ni} \prod_{j=1}^{p'} \tilde{H}(\Delta_{n,i}^j) \\ \xi'' &= \lim_{n \rightarrow \infty} \xi''_n = \lim_{n \rightarrow \infty} \sum_k \beta'_{nk} \prod_{l=1}^{p''} \tilde{H}(\Delta_{n,k}^l).\end{aligned}$$

Note that in view of the measurability assumptions we have $\Delta_{n,i}^j \subset \Delta'$, $j = 1, \dots, p'$ and $\Delta_{n,i}^j \subset \Delta''$, $j = 1, \dots, p''$ and also β'_{ni} are $\mathcal{F}_{\Delta'}^\alpha$ -measurable, while β'_{nk} are $\mathcal{F}_{\Delta''}^\alpha$ -measurable. Then it is easy to see that the statement holds. $\square$

We remark that the result still holds true if we consider the σ -algebras $\mathcal{F}_{\Delta'}^H \vee \mathcal{F}^\alpha$ and $\mathcal{F}_{\Delta''}^H \vee \mathcal{F}^\alpha$, for $\Delta' \cap \Delta'' = \emptyset$.

The *polynomials* of the values of $\tilde{H}$ of degree p are the random variables ξ admitting representation as

$$\xi = \sum_{m=1}^M c_m \prod_{j=1}^{J_m} \tilde{H}(\Delta_j^m)^{p_j}, \quad c_m \in \mathbb{R}, \quad M, J_m, p_j \in \mathbb{N} \quad (2.3.6)$$

such that $\sum_{j=1}^{J_m} p_j \leq p$, $m = 1, 2, \dots, M$ and $\Delta_j^m \in \mathcal{B}_X^c$, $j = 1, \dots, J_m$ are pairwise disjoint.

Theorem 2.3.8. *All the polynomials of values of $\tilde{H}$ of degree less or equal to p belong to the subspace $\mathbb{H}_p$.*

Proof. Let ξ be a polynomial of degree p as in (2.3.6). We proceed by induction. If $p = 0$ then $\xi \in \mathbb{H}_0$ and if $p = 1$ then $\xi \in \mathbb{H}_1$. Suppose the statement holds for $q < p$, we verify this for p . For each m , let us consider elements

$$\xi_m := \prod_{j=1}^{J_m} \tilde{H}(\Delta_j^m)^{p_j}, \quad \sum_{j=1}^{J_m} p_j \leq p.$$

If $p_j < p$ for $j = 1, \dots, J_m$

- i) and $\sum_{j=1}^{J_m} p_j < p$ then the induction hypothesis holds and $\xi_m \in \mathbb{H}_p$.
- ii) and $\sum_{j=1}^{J_m} p_j = p$ with $J_m > 1$, then for any j we have $\tilde{H}(\Delta_j^m)^{p_j} \in \mathbb{H}_{p_j}$ by the induction hypothesis. Furthermore, being the sets disjoint, from Lemma 2.3.7 we have that $\prod_{j=1}^{J_m} \tilde{H}(\Delta_j^m)^{p_j} \in \mathbb{H}_{\sum p_j}$.

Hence we only have to verify the case $J_m = 1$. Namely

$$\xi = \tilde{H}(\Delta_j^m)^p \in \mathbb{H}_p, \quad \text{for } p > 1.$$

Set $\Delta := \Delta_j^m$. For all n , we can represent Δ in terms of the sets (2.2.9) of the dissecting system

of X ,

$$\Delta = \bigcup_{k=1}^{K_n+1} (\Delta \cap \Delta_{n,k}) := \bigcup_{k=1}^{K_n+1} \tilde{\Delta}_{n,k},$$

hence we have that

$$\xi = \tilde{H}(\Delta)^p = \left(\sum_{k=1}^{K_n+1} \tilde{H}(\tilde{\Delta}_{n,k}) \right)^p.$$

Let

$$\xi_n^{(1)} := \xi - \sum_{k=1}^{K_n+1} \tilde{H}(\tilde{\Delta}_{n,k})^p = \left(\sum_{k=1}^{K_n+1} \tilde{H}(\tilde{\Delta}_{n,k}) \right)^p - \sum_{k=1}^{K_n+1} \tilde{H}(\tilde{\Delta}_{n,k})^p$$

and

$$\xi_n^{(2)} := \sum_{k=1}^{K_n+1} \tilde{H}(\tilde{\Delta}_{n,k})^p.$$

For all n we have $\xi = \xi_n^{(1)} + \xi_n^{(2)}$ and thus $\xi = \lim_{n \rightarrow \infty} \xi_n^{(1)} + \xi_n^{(2)}$ in $L_2(\Omega, \mathcal{F}, \mathbb{P})$. Let us investigate $\xi_n^{(1)}$ and $\xi_n^{(2)}$ separately. First of all we note that $\xi_n^{(1)}$ is a polynomial, as in (2.3.6), with $p_j < p$ for all $j = 1, \dots, K_n+1$. Thus $\xi_n^{(1)} \in \mathbb{H}_p$. Hence we have $\xi^{(1)} := \lim_{n \rightarrow \infty} \xi_n^{(1)} \in \mathbb{H}_p$ since $\mathbb{H}_p$ is closed in $L_2(\Omega, \mathcal{F}, \mathbb{P})$.

Consider the following limit in $L_2(\Omega, \mathcal{F}, \mathbb{P})$:

$$\begin{aligned} \xi^{(2)} &:= \lim_{n \rightarrow \infty} \xi_n^{(2)} \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^{K_n+1} \left(H(\tilde{\Delta}_{n,k}) - \alpha(\tilde{\Delta}_{n,k}) \right)^p \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^{K_n+1} \sum_{j=0}^p \binom{p}{j} H(\tilde{\Delta}_{n,k})^{p-j} (-1)^j \alpha(\tilde{\Delta}_{n,k})^j \\ &= \sum_{j=0}^p \lim_{n \rightarrow \infty} \sum_{k=1}^{K_n+1} \binom{p}{j} H(\tilde{\Delta}_{n,k})^{p-j} (-1)^j \alpha(\tilde{\Delta}_{n,k})^j. \end{aligned} \tag{2.3.7}$$

Since $\mathbb{P}(H(\{x\}) \neq 0, 1 \text{ for some } x \in X) = 0$ see [50, Theorem 1.3 page 19], we will ultimately have $H(\tilde{\Delta}_{n,k}) = 0$ or 1 $\mathbb{P}$ -a.s. as $n \rightarrow \infty$. Thus for the first term ($j = 0$) in (2.3.7), using dominated convergence, we have:

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{K_n+1} \left(H(\tilde{\Delta}_{n,k}) \right)^p = \sum_{k=1}^{H(\Delta)} 1^p = H(\Delta) = \tilde{H}(\Delta) + \alpha(\Delta).$$

For the remaining terms ($j > 0$) in (2.3.7) the following estimate applies

$$\begin{aligned}
\left| \sum_{k=1}^{K_n+1} (-1)^j H(\tilde{\Delta}_{n,k})^p \alpha(\tilde{\Delta}_{n,k})^{p-j} \right| &\leq \sum_{k=1}^{K_n+1} \mathbf{1}_{\{H(\tilde{\Delta}_{n,k}) > 0\}} H(\tilde{\Delta}_{n,k})^p \alpha(\tilde{\Delta}_{n,k})^{p-j} \\
&\leq \sup_k \alpha(\tilde{\Delta}_{n,k})^{p-j} \sum_{k=1}^{K_n} \mathbf{1}_{\{H(\tilde{\Delta}_{n,k}) > 0\}} H(\tilde{\Delta}_{n,k})^p \\
&\leq H(\tilde{\Delta})^{p+1} \sup_k \alpha(\tilde{\Delta}_{n,k})^{p-j} \longrightarrow 0, \quad n \rightarrow \infty,
\end{aligned}$$

by Lemma 2.2.7. Thus $\xi^{(2)} = \tilde{H}(\Delta) + \alpha(\Delta) \in \mathbb{H}_1 \subseteq \mathbb{H}_p$. Hence $\xi = \xi^{(1)} + \xi^{(2)} \in \mathbb{H}_p$. $\square$

The following statement is a direct consequence of the theorem above.

Corollary 2.3.9. All the polynomials of all degrees of the values of $\tilde{H}$ belong to $\mathbb{H}$.

Remark 2.3.10. We note that if the sets in (2.3.6) were not disjoint, then one could always represent the same polynomials via disjoint sets by applying the additivity of the measure $\tilde{H}$, but the degree would naturally change.

By assumption (2.2.1),

$$= \prod_{j=1}^J \mathbb{E} \left[e^{(e^{2^J c_j} - 1 - 2^J c_j) \alpha(\Delta_j)} \right] < \infty.$$

Following classical arguments via Fourier transforms (see e.g. [85, Lemma 4.3.1 and Lemma 4.3.2]) one can see that the random variables

$$\exp \left\{ \sum_{j=1}^J x_j \tilde{H}(\Delta_j) \right\}, \quad j = 1, 2, \dots, J; \quad x = (x_1, \dots, x_J) \in \mathbb{R}^J,$$

with $\Delta_j, j = 1, \dots, J$ pairwise disjoint sets in $\mathcal{B}_X^c$, constitute a complete system in $L_2(\Omega, \mathcal{F}, \mathbb{P})$. By Taylor approximations of the analytic extension on $\mathbb{C}^J$ we have that

$$\mathbb{E} \left[\left| \exp \left\{ \sum_{j=1}^J x_j \tilde{H}(\Delta_j) \right\} - \sum_{p=0}^q \frac{\sum_{j=1}^J i x_j \tilde{H}(\Delta_j)^p}{p!} \right|^2 \right] \longrightarrow 0, \quad q \rightarrow \infty$$

(see e.g. [16, Eq. (26.4)] for an estimate of the quantity here above justifying the convergence.) Hence we can conclude:

Lemma 2.3.11. The polynomials of the values of $\tilde{H}(\Delta)$, $\Delta \in \mathcal{B}_X^c$ are dense in $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Theorem 2.3.12 (Chaos expansion). The following equality holds:

$$\mathbb{H} = L_2(\Omega, \mathcal{F}, \mathbb{P}).$$

Namely, any $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ can be written as

$$\xi = \sum_{p=0}^{\infty} \xi_p, \quad \text{where } \xi_p \in \mathbb{H}^p \text{ for } p = 1, 2, \dots$$

Proof. The polynomials of the values of $\tilde{H}(\Delta)$ are dense in $L_2(\Omega, \mathcal{F}, \mathbb{P})$, see Lemma 2.3.11. By Theorem 2.3.8 and Corollary 2.3.9 all the polynomials are in $\mathbb{H}$. Since $\mathbb{H}$ is closed, we must have $L_2(\Omega, \mathcal{F}, \mathbb{P}) \subseteq \mathbb{H}$. On the other side we recall that by construction $\mathbb{H} \subseteq L_2(\Omega, \mathcal{F}, \mathbb{P})$, see (2.3.3) and Definitions 2.3.2 and 2.3.6. $\square$

Remark 2.3.13. Definitions 2.3.2 and 2.3.6 describe the spaces generated by α -multilinear forms. We can also consider analogous spaces generated only by the multilinear forms as in (2.3.2). However we have to stress that in this case the multilinear forms are not dense in $L_2(\Omega, \mathcal{F}, \mathbb{P})$ with the only exception made when H is a Poisson random measure, i.e. if α is deterministic. Indeed write $\tilde{\mathbb{H}}^p$ for the subspace in $L_2(\Omega, \mathcal{F}, \mathbb{P})$ generated by the finite linear combinations of p -order multilinear forms:

$$\sum_i c_i \prod_{j=1}^p \tilde{H}(\Delta_j^i). \quad (2.3.8)$$

The sets Δ_j^i , $j = 1, \dots, p$, are pairwise disjoint and the c_i are constants. Set $\tilde{\mathbb{H}}^0 = \mathbb{R}$ and $\tilde{\mathbb{H}} := \sum_{p=0}^{\infty} \oplus \tilde{\mathbb{H}}_p$. It is easily seen that $(\beta - \mathbb{E}[\beta]) \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ is orthogonal to $\tilde{\mathbb{H}}$ whenever β is $\mathcal{F}^\alpha$ -measurable. There are also α -multilinear forms of higher orders that are orthogonal to $\tilde{\mathbb{H}}$, one example is $(\mathbb{E}[\beta|\mathcal{F}_\Delta^\alpha] - \mathbb{E}[\beta])\tilde{H}(\Delta)$ ($\Delta \in \mathcal{B}_X^c$). Thus, in general, $\tilde{\mathbb{H}} \neq L_2(\Omega, \mathcal{F}, \mathbb{P})$. The case when H is a Poisson random measure was studied in [36] as a particular Lévy random field, and there we do have $\tilde{\mathbb{H}} = L_2(\Omega, \mathcal{F}, \mathbb{P})$.

2.4 Non-anticipating stochastic integration

In the sequel we consider $X = [0, T] \times \mathcal{Z}$ for $T < \infty$ and $\mathcal{Z}$ a locally compact, second countable Hausdorff topological space. Being interested in integration, without loss of generality we assume that

$$\alpha(\{0\} \times \mathcal{Z}) = 0 \quad \mathbb{P}\text{-a.s.}$$

Hence we can restrict the attention to $X = (0, T] \times \mathcal{Z}$.

We chose a dissecting system of X to be given by partitions (2.2.8)-(2.2.9) of the form

$$\Delta_{n,k} = (s_{n,k}, u_{n,k}] \times B_{n,k}, \quad s_{n,k} < u_{n,k}, \quad B_{n,k} \in \mathcal{B}_{\mathcal{Z}}^c, \quad (2.4.1)$$

for $n = 1, 2, \dots$ and $k = 1, 2, \dots, K_n$, see Definition 2.2.6. Here $\mathcal{B}_{\mathcal{Z}}$ denotes the Borel σ -algebra on $\mathcal{Z}$ and $\mathcal{B}_{\mathcal{Z}}^c$ the family of precompacts for the topology in $\mathcal{Z}$. The set X is ordered with the natural ordering given by time in $[0, T]$. Two filtrations naturally appear in the present setting:

- $\mathbb{F} := \{\mathcal{F}_t, t \in [0, T]\}$ where $\mathcal{F}_t$ is generated by $\{\tilde{H}(\Delta) : \Delta \in \mathcal{B}_{[0,t] \times \mathcal{Z}}\}$,
- $\mathbb{G} := \{\mathcal{G}_t, t \in [0, T]\}$ with $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{F}^\alpha$.

Clearly we have that $\mathcal{F}_t \subseteq \mathcal{G}_t$, $\mathcal{F}_0$ is trivial but $\mathcal{G}_0 = \mathcal{F}^\alpha$ and $\mathcal{F}_T = \mathcal{G}_T = \mathcal{F}$.

We define a martingale random field as in [38], see in particular Remark [38, Remark 2.3] for historical notes. We can also refer to the work of [106] and [23] as pioneering in the use of martingale random fields in stochastic calculus, though mostly related to Brownian sheet.

Hence we can see that the stochastic set function $\tilde{H}(\Delta)$, $\Delta \in \mathcal{B}_X$ is a martingale random field (with square integrable values) with respect to $\mathbb{F}$ and $\mathbb{G}$ as it satisfies the following properties:

i) $\tilde{H}$ has a σ -finite variance measure $V(\Delta) = E[\tilde{H}(\Delta)^2]$, $\Delta \in \mathcal{B}_X$, recall (2.2.6).

ii) $\tilde{H}$ is additive, i.e. for pairwise disjoint sets $\Delta_1, \dots, \Delta_K$: $V(\Delta_k) < \infty$

$$\tilde{H}\left(\bigcup_{k=1}^K \Delta_k\right) = \sum_{k=1}^K \tilde{H}(\Delta_k)$$

and σ -additive in L_2 , i.e. for pairwise disjoint sets $\Delta_1, \Delta_2, \dots$: $V(\Delta_k) < \infty$

$$\tilde{H}\left(\bigcup_{k=1}^{\infty} \Delta_k\right) = \sum_{k=1}^{\infty} \tilde{H}(\Delta_k)$$

with convergence in $L_2(\Omega, \mathcal{F}, \mathbb{P})$.

iii) $\tilde{H}$ is adapted to $\mathbb{F}$ and $\mathbb{G}$.

iv) $\tilde{H}$ has the *martingale property*. Consider $\Delta \subseteq (t, T] \times \mathcal{Z}$. Then, from (2.2.4) we have:

$$\mathbb{E}[\tilde{H}(\Delta) \mid \mathcal{F}_t] = \mathbb{E}[\mathbb{E}[\tilde{H}(\Delta) \mid \mathcal{G}_t] \mid \mathcal{F}_t] = \mathbb{E}[\mathbb{E}[\tilde{H}(\Delta) \mid \mathcal{F}^\alpha] \mid \mathcal{F}_t] = 0.$$

v) $\tilde{H}$ has conditionally orthogonal values. For any $\Delta_1, \Delta_2 \subseteq (t, T] \times \mathcal{Z}$ such that $\Delta_1 \cap \Delta_2 = \emptyset$ and. Then, from A2), we have:

$$\begin{aligned} \mathbb{E}[\tilde{H}(\Delta_1)\tilde{H}(\Delta_2) \mid \mathcal{F}_t] &= \mathbb{E}[\mathbb{E}[\tilde{H}(\Delta_1)\tilde{H}(\Delta_2) \mid \mathcal{G}_t] \mid \mathcal{F}_t] \\ &= \mathbb{E}[\mathbb{E}[\tilde{H}(\Delta_1) \mid \mathcal{F}^\alpha] \mathbb{E}[\tilde{H}(\Delta_2) \mid \mathcal{F}^\alpha] \mid \mathcal{F}_t] = 0. \end{aligned}$$

Given the martingale structure of the CDSPP $\tilde{H}$ with respect to the filtrations $\mathbb{G}$ and $\mathbb{F}$, we can construct a stochastic integration of Itô type according to the classical scheme, as retraced in [38]. Hereafter we consider $\mathbb{G}$ as the reference information flow. Recall that α is a positive random measure.

We define the $\mathbb{G}$ -predictable σ -algebra $\mathcal{P}_{\mathbb{G}}$ as the σ -algebra generated by $\{F \times (s, u] \times B : F \in \mathcal{G}_s, s < u, B \in \mathcal{B}_{\mathcal{Z}}\}$ and, as usual, we will say that a stochastic process ϕ is $\mathbb{G}$ -predictable if the mapping $\phi = \phi(\omega, t, z)$, $\omega \in \Omega$, $(t, z) \in X$, is $\mathcal{P}_{\mathbb{G}}$ -measurable. Hence we define

$$\|\phi\|_{\mathcal{I}} := \left(\mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} \phi^2(t, z) \alpha(dt, dz) \right] \right)^{1/2}$$

and set $\mathcal{I}$ to be the L_2 -subspace of stochastic processes ϕ admitting a $\mathbb{G}$ -predictable modification and such that $\|\phi\|_{\mathcal{I}} < \infty$.

Lemma 2.4.1. $\mathcal{F}^\alpha \times \mathcal{B}_X \subset \mathcal{P}_{\mathbb{G}}$, and α is the $\mathbb{G}$ -predictable compensator of H .

We take the $\mathbb{G}$ -predictable compensator to be as in [65], a predictable, locally integrable random measure such that $\mathbb{E}[H(\Delta)] = \mathbb{E}[\alpha(\Delta)]$.

Proof. For the first claim it is sufficient to show that $A \times (a, b] \times B$ with $A \in \mathcal{F}^\alpha$, $a < b$ and $B \in \mathcal{B}_{\mathcal{Z}}^c$ is an element of $\mathcal{P}_{\mathbb{G}}$. Recall that $A \in \mathcal{G}_s$ for all s and the claim follows. Since $\mathbb{E}[H(\Delta)] = \mathbb{E}[\alpha(\Delta)]$ for all $\Delta \in \mathcal{B}_X$, and α is $\mathbb{G}$ -predictable, it is the $\mathbb{G}$ -predictable compensator of H . $\square$

The non-anticipating stochastic integral with respect to $\tilde{H}$ under $\mathbb{G}$ is the isometric operator I mapping:

$$I : \text{dom} I \Longrightarrow L_2(\Omega, \mathcal{F}, \mathbb{P})$$

such that

$$I(\phi) := \int_0^T \int_{\mathcal{Z}} \phi(t, z) \tilde{H}(dt, dz) := \sum_{k=1}^K \phi_k \tilde{H}(\Delta_k) \quad (2.4.2)$$

for any

$$\phi(t, z) = \sum_{k=1}^K \phi_k \mathbf{1}_{\Delta_k}(t, z) \quad (2.4.3)$$

with $\Delta_k = (s_k, u_k] \times B_k \in \mathcal{B}_X^c$ and ϕ_k a $\mathcal{G}_{s_k}$ -measurable random variable such that $\|\phi\|_{\mathcal{I}} < \infty$. In fact,

$$\mathbb{E}[I(\phi)^2] = \mathbb{E}\left[\left(\sum_{k=1}^K \phi_k \tilde{H}(\Delta_k)\right)^2\right] = \mathbb{E}\left[\sum_{k=1}^K \phi_k^2 \alpha(\Delta_k)\right] = \|\phi\|_{\mathcal{I}}^2. \quad (2.4.4)$$

Naturally the integrands are given by $\text{dom} I \subseteq L_2(\Omega \times X)$, with $L_2(\Omega \times X) := L_2(\Omega \times X, \mathcal{F} \times \mathcal{B}_X, \mathbb{P} \times \alpha)$, which is the linear closure of the stochastic processes (2.4.3) and the integral is characterized in a standard manner exploiting the isometry (2.4.4).

Actually $\text{dom} I = \mathcal{I}$. In fact the following result holds true.

Lemma 2.4.2. For any $\phi \in \mathcal{I}$ there exists an approximating sequence of stochastic processes ϕ_n , $n = 1, 2, \dots$, of type (2.4.3) having the form:

$$\phi_n(t, z) = \sum_{k=1}^{K_n} \mathbb{E}\left[\frac{1}{\alpha(\Delta_{n,k})} \int_{\Delta_{n,k}} \phi(\tau, \zeta) \alpha(d\tau, d\zeta) \mid \mathcal{G}_{s_{n,k}}\right] \mathbf{1}_{\Delta_{n,k}}(t, z),$$

where $\Delta_k = (s_{n,k}, u_{n,k}] \times B_{n,k}$ are the sets (2.4.1) of the dissecting system of $X = (0, T] \times \mathcal{Z}$.

Proof. The arguments of [38, Lemma 3.1] can be easily adapted to the present framework. $\square$

Then, by isometry, it is clear that for any $\phi \in \mathcal{I}$, $I(\phi) = \lim_{n \rightarrow \infty} I(\phi_n)$ in $L_2(\Omega, \mathcal{F}, \mathbb{P})$ where $\phi_n \in \mathcal{I}$ are processes of type (2.4.3) approximating ϕ in $L_2(\Omega \times X)$.

From the construction of the stochastic integral, it follows that for any $\phi \in \mathcal{I}$, the stochastic set function

$$\mu(\phi, \Delta) := \int_{\Delta} \phi(t, z) \tilde{H}(dt, dz), \quad \Delta \in \mathcal{B}_X, \quad (2.4.5)$$

is again a martingale random field [38, Remark 3.2] under $\mathbb{G}$ with variance measure

$$m(\phi, \Delta) := \mathbb{E} \left[\int_{\Delta} \phi^2(t, z) \alpha(dt, dz) \right], \quad \Delta \in \mathcal{B}_X.$$

Proposition 2.4.3. *Consider the $\mathcal{F}^\alpha$ -measurable $\beta \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ and $\phi \in \mathcal{I}$. Then*

$$\beta I(\phi) = I(\beta \phi)$$

if either side of the equality exists as an element of $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Proof. Assume $\phi \in \mathcal{I}$ is of type (2.4.3) and β is bounded. Then, for every k , $\beta \phi_k$ is $\mathcal{G}_{s_k}$ -measurable and

$$\beta I(\phi) = \sum_{k=1}^K \beta \phi_k \tilde{H}((s_k, u_k] \times B_k) = I(\beta \phi).$$

The general case follows by taking limits. □

The classical calculus rules hold: for any $\phi \in \mathcal{I}$ we have

$$\mathbb{E} \left[\int_{\Delta} \phi(t, z) \tilde{H}(dt, dz) \mid \mathcal{G}_t \right] = 0, \quad \Delta \in \mathcal{B}_{(s, T] \times \mathcal{Z}},$$

and

$$\begin{aligned} & \mathbb{E} \left[\int_{\Delta} \phi_1(t, z) \tilde{H}(dt, dz) \int_{\Delta} \phi_2(t, z) \tilde{H}(dt, dz) \mid \mathcal{G}_s \right] \\ &= \mathbb{E} \left[\int_{\Delta} \phi_1(t, z) \phi_2(t, z) \alpha(dt, dz) \mid \mathcal{G}_s \right] \\ &= \int_{\Delta} \mathbb{E} [\phi_1(t, z) \phi_2(t, z) \mid \mathcal{G}_s] \alpha(dt, dz), \quad \Delta \in \mathcal{B}_{(s, T] \times \mathcal{Z}}, \end{aligned}$$

and in particular we have

$$\mathbb{E} [I(\phi)^2 \mid \mathcal{F}^\alpha] = \int_0^T \int_{\mathcal{Z}} \mathbb{E} [\phi^2(t, z) \mid \mathcal{F}^\alpha] \alpha(dt, dz). \quad (2.4.6)$$

Remark 2.4.4. In the same way as for the case of information flow $\mathbb{G}$, we can define the $\mathbb{F}$ -predictable σ -algebra $\mathcal{P}_{\mathbb{F}}$ and consider the associated $\mathbb{F}$ -predictable random fields. Being any $\mathbb{F}$ -predictable stochastic process also $\mathbb{G}$ -predictable, the integration can be done in the same setting as above with the result that the corresponding stochastic functions of type (2.4.5) will be martingale random fields under $\mathbb{F}$. Clearly results as in Proposition 2.4.3 fail in general in

this context. In fact β is $\mathcal{G}_0$ -measurable, but not $\mathcal{F}_0$ -measurable in general.

In the sequel we study integral representations with respect to the filtration $\mathbb{G}$ where the integrands are explicitly characterized via stochastic derivatives. Some of our results can be adapted to the case of the filtration $\mathbb{F}$. However the stochastic calculus is more delicate in the case of $\mathbb{F}$ and is a matter for future research.

2.5 Non-anticipating stochastic derivative and representation theorem

In this section we discuss stochastic differentiation in the context of non-anticipative calculus. We will use the terminology non-anticipating derivative to emphasize the fact that the operator introduced embeds the information flow associated with the framework as time evolves. This differs from other concepts of stochastic differentiation, as the Malliavin type derivative. We consider the relationships with anticipative derivatives in section 2.7.

The *non-anticipating stochastic derivative* is the adjoint linear operator $\mathcal{D} = I^*$ of the stochastic integral:

$$\mathcal{D} : L_2(\Omega, \mathcal{F}, \mathbb{P}) \Longrightarrow \mathcal{I}.$$

We can see that the non-anticipating derivative can be computed as the limit

$$\mathcal{D}\xi = \lim_{n \rightarrow \infty} \phi_n \quad (2.5.1)$$

with convergence in $\mathcal{I}$ of the stochastic functions of type (2.4.3) given by

$$\phi_n(t, z) := \sum_{k=1}^{K_n} \mathbb{E} \left[\xi \frac{\tilde{H}(\Delta_{n,k})}{\alpha(\Delta_{n,k})} \middle| \mathcal{G}_{s_{n,k}} \right] \mathbf{1}_{\Delta_{n,k}}(t, z) \quad (2.5.2)$$

where $\Delta_{n,k} = (s_{n,k}, u_{n,k}] \times B_{n,k}$ refers to the n 'th partition sets (2.4.1) in the dissecting system of $X = (0, T] \times \mathcal{Z}$ (as per Definition 2.2.6). In fact we have the following result:

Theorem 2.5.1. *All $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ have representation*

$$\xi = \xi_0 + \int_0^T \int_{\mathcal{Z}} \mathcal{D}_{t,z} \xi \tilde{H}(dt, dz). \quad (2.5.3)$$

Moreover $\mathcal{D}\xi_0 = 0$. In particular we have $\xi_0 = \mathbb{E}[\xi | \mathcal{F}^\alpha]$.

Proof. We proceed using arguments as those in [34, Theorem 2.1]. Set

$$\phi_{n,k} := \mathbb{E} \left[\xi \frac{\tilde{H}(\Delta_{n,k})}{\alpha(\Delta_{n,k})} \middle| \mathcal{G}_{s_{n,k}} \right].$$

First note that

$$\mathbb{E} \left[\left| \phi_{n,k} \tilde{H}(\Delta_{n,k}) \right|^2 \middle| \mathcal{G}_{s_{n,k}} \right] \leq \mathbb{E} [\xi^2 | \mathcal{G}_{s_{n,k}}]$$

by application of the conditional Hölder inequality. Thus $\mathbb{E}[|\phi_{n,k}\tilde{H}(\Delta_{n,k})|^2] < \infty$. Moreover, we have that

$$\mathbb{E}\left[(\xi - \phi_{n,k}\tilde{H}(\Delta_{n,k}))\psi\tilde{H}(\Delta_{n,k})\right] = 0 \quad (2.5.4)$$

for all $\mathcal{G}_{s_{n,k}}$ -measurable $\psi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$. In fact, we have

$$\begin{aligned} \mathbb{E}\left[(\xi - \phi_{n,k}\tilde{H}(\Delta_{n,k}))\psi\tilde{H}(\Delta_{n,k})\middle|\mathcal{G}_{s_{n,k}}\right] \\ = \psi\mathbb{E}[\xi\tilde{H}(\Delta_{n,k})\middle|\mathcal{G}_{s_{n,k}}] - \psi\phi_{n,k}\alpha(\Delta_{n,k}) = 0. \end{aligned}$$

Then, from (2.5.4), we conclude that

$$\hat{\xi}_n := \sum_{k=1}^{K_n} \phi_{n,k}\tilde{H}(\Delta_{n,k}) = \int_0^T \int_{\mathcal{Z}} \phi_n(s, z) \tilde{H}(ds, dz)$$

is the projection of ξ onto the subspace of the stochastic integrals of type (2.4.3). Moreover, $\hat{\xi} := \lim_{n \rightarrow \infty} \hat{\xi}_n$ in $L_2(\Omega, \mathcal{F}, \mathbb{P})$ is the projection of ξ onto the subspace of all the stochastic integrals with respect to $\tilde{H}$. Indeed, for any integral $I(\psi)$ with $\psi \in \mathcal{I}$, and $\psi := \lim_{n \rightarrow \infty} \sum_{k=1}^{K_n} \psi_{n,k} \mathbf{1}_{\Delta_{n,k}} \in \mathcal{I}$, we have

$$\mathbb{E}\left[(\xi - \hat{\xi})I(\psi)\right] = \lim_{n \rightarrow \infty} \sum_{k=1}^{K_n} \mathbb{E}\left[(\xi - \phi_{n,k}\tilde{H}(\Delta_{n,k}))\psi_{n,k}\tilde{H}(\Delta_{n,k})\right] = 0$$

(with convergence in $L_2(\Omega, \mathcal{F}, \mathbb{P})$). Denote by $\hat{\phi}$ the integrand representing $\hat{\xi}$, i.e.

$$\hat{\xi} = \int_0^T \int_{\mathcal{Z}} \hat{\phi}(s, z) \tilde{H}(ds, dz).$$

Then, by isometry, we have

$$\|\hat{\phi} - \phi_n\|_{\mathcal{I}}^2 = \|\hat{\xi} - \hat{\xi}_n\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})}^2 \rightarrow 0, \quad n \rightarrow \infty.$$

Hence $\hat{\phi} = \mathcal{D}\xi$. Moreover, being the difference $\xi^0 := \xi - \hat{\xi}$ orthogonal to all stochastic integrals, we have $\mathcal{D}\xi = 0$. In addition we also have that

$$\xi_0 = \mathbb{E}[\xi|\mathcal{G}_0] = \mathbb{E}[\xi|\mathcal{F}^\alpha].$$

By this we end the proof. □

Remark 2.5.2. Note that the non-anticipating derivative is continuous in L_2 . Namely, if $\xi = \lim_{n \rightarrow \infty} \xi_n$ in $L_2(\Omega, \mathcal{F}, \mathbb{P})$, then

$$\mathcal{D}\xi = \lim_{n \rightarrow \infty} \mathcal{D}\xi_n \quad \text{in } \mathcal{I}.$$

In fact $\|\mathcal{D}\xi - \mathcal{D}\xi_n\|_{\mathcal{I}}^2 \leq \mathbb{E}[(\xi - \xi_n)^2] \rightarrow 0, \quad n \rightarrow \infty$.

Example 2.5.3. Let $\xi \in \mathbb{H}^p$ be an α -multilinear form $\xi = \beta \prod_{j=1}^p \tilde{H}(\Delta_j)$ with $\Delta_1 = (s_1, u_1] \times B_1$, $\Delta_2 = (s_2, u_2] \times B_2, \dots, \Delta_p = (s_p, u_p] \times B_p$ and $0 \leq s_1 < u_1 \leq s_2 < u_2 < \dots < u_p \leq T$. Then

$$\mathcal{D}_{t,z}\xi = \beta \prod_{j=1}^{p-1} \tilde{H}(\Delta_j) \mathbf{1}_{\Delta_p}(t, z)$$

and

$$\xi = \int_{\Delta_p} \beta \prod_{j=1}^{p-1} \tilde{H}(\Delta_j) \tilde{H}(dt, dz).$$

Example 2.5.4. If $\beta \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ is $\mathcal{F}^\alpha$ -measurable, then $\mathcal{D}\beta = 0$. In fact $\mathbb{E}\left[\beta \frac{\tilde{H}(\Delta_{n,k})}{\alpha(\Delta_{n,k})} \middle| \mathcal{G}_{t_{n,k}}\right] = 0$ for all $\Delta_{n,k}$.

In general we have the following formula:

Proposition 2.5.5. Let ξ be an α -multilinear form, $\xi = \beta \prod_{j=1}^p \tilde{H}(\Delta_j)$. Then

$$\xi = \int_0^T \int_{\mathcal{Z}} \mathcal{D}_{s,z}\xi \tilde{H}(ds, dz)$$

with

$$\mathcal{D}_{s,z}\xi = \beta \sum_{\substack{1 \leq i \leq p \\ \Delta_i \subseteq \Delta'}} \mathbf{1}_{\Delta_i}(s, z) \prod_{j \neq i}^p \tilde{H}(\Delta_j \cap [0, s) \times \mathcal{Z}) \quad (2.5.5)$$

Here the set Δ' is given by

$$\Delta' = \bigcup_{j \notin \mathcal{K}} \Delta_j \quad (2.5.6)$$

where $\mathcal{K} = \{1 \leq i \leq p \mid \Delta_i \subset [0, t) \times \mathcal{Z} \text{ and } \Delta_j \subset [t, T] \times \mathcal{Z} \text{ for some } 1 \leq j \leq p \text{ and } t \in [0, T]\}$.

To explain the set $\mathcal{K}$ in Proposition 2.5.5, in Example 2.5.3 we would have $\mathcal{K} = \{1, \dots, p-1\}$, corresponding to the sets $\Delta_1, \dots, \Delta_{p-1}$, i.e. the elements of the multilinear form that “are entirely before the last set”.

Proof. Let ξ be a multilinear form of order $p \geq 1$, $\xi = \beta \prod_{j=1}^p \tilde{H}(\Delta_j)$. For simplicity assume $\Delta_j \cap \Delta_{n,k} = \emptyset$ or $\Delta_{n,k}$. Denote

$$\psi(n, k) = \mathbb{E}\left[\beta \prod_{j=1}^p \tilde{H}(\Delta_j) \frac{\tilde{H}(\Delta_{n,k})}{\alpha(\Delta_{n,k})} \middle| \mathcal{G}_{t_{n,k}}\right].$$

The computation of $\psi(n, k)$ is divided into three cases.

i) If $(\bigcup_{j=1}^p \Delta_j) \cap \Delta_{n,k} = \emptyset$ then $\psi(n, k) = 0$.

ii) If there exists i such that $\Delta_i \subset (t_{n,k}, T] \times \mathcal{Z}$ and $\Delta_i \cap \Delta_{n,k} = \emptyset$ then

$$\psi(n, k) = \mathbb{E} \left[\beta \prod_{j \neq i} \tilde{H}(\Delta_j) \frac{\tilde{H}(\Delta_{n,k})}{\alpha(\Delta_{n,k})} \middle| \mathcal{G}_{t_{n,k}} \right] \mathbb{E} \left[\tilde{H}(\Delta_i) \middle| \mathcal{F}^\alpha \right] = 0.$$

iii) Neither case i) or ii) is true. This implies that $\Delta_{n,k} \subset \Delta'$. By assumption there exists $1 \leq i \leq p$ such that $\Delta_i \cap \Delta_{n,k} = \Delta_{n,k}$. We have

$$\begin{aligned} \psi(n, k) &= \mathbb{E} \left[\beta \prod_{j \neq i} \left(\tilde{H}(\Delta_j \cap [0, t_{n,k}] \times \mathcal{Z}) + \tilde{H}(\Delta_j \cap (t_{n,k}, T] \times \mathcal{Z}) \right) \right. \\ &\quad \left. \left(\tilde{H}(\Delta_i \cap \Delta_{n,k}) + \tilde{H}(\Delta_i \cap \Delta_{n,k}^c) \right) \frac{\tilde{H}(\Delta_{n,k})}{\alpha(\Delta_{n,k})} \middle| \mathcal{G}_{t_{n,k}} \right] \\ &= \mathbb{E} \left[\beta \prod_{j \neq i} \tilde{H}(\Delta_j \cap [0, t_{n,k}] \times \mathcal{Z}) \middle| \mathcal{G}_{t_{n,k}} \right] \\ &= \beta \prod_{j \neq i} \tilde{H}(\Delta_j \cap [0, t_{n,k}] \times \mathcal{Z}). \end{aligned}$$

Thus

$$\psi(n, k) = \mathbf{1}_{\{\Delta_{n,k} \cap \Delta' \neq \emptyset\}}(n, k) \beta \prod_{\substack{j \\ \Delta_j \cap \Delta_{n,k} = \emptyset}} \tilde{H}(\Delta_j \cap [0, t_{n,k}] \times \mathcal{Z}), \quad (2.5.7)$$

and with Δ' as above, $\mathcal{D}\xi$ is given by (2.5.5). Since $\mathbb{E}[\xi | \mathcal{F}^\alpha] = 0$ the representation is

$$\xi = \mathbb{E}[\xi | \mathcal{F}^\alpha] + I(\mathcal{D}\xi) = I(\mathcal{D}\xi).$$

□

The doubly stochastic Poisson process H is an example of a point process. For point processes in general, some integral representations have been developed in [19, 21, 31, 55], see also the survey [32]. Note that the filtration of reference in these studies is $\mathbb{F}^H = \{\mathcal{F}_t^H, t \in [0, T]\}$. As an illustration consider [21, Theorem 8.8]:

Theorem 2.5.6. *Let $\xi \in L_2(\Omega, \mathcal{F}_T^H, \mathbb{P})$. Let Λ be the $\mathbb{F}^H$ -predictable compensator of H . Then there exists an $\mathbb{F}^H$ -predictable process ϕ such that*

$$\xi = \mathbb{E}[\xi] + \int_0^T \int_{\mathcal{Z}} \phi(s, z) (H(ds, dz) - \Lambda(ds, dz)) \quad (2.5.8)$$

and $\mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} \phi(s, z)^2 \Lambda(ds, dz) \right] < \infty$.

We remark that Theorem 2.5.1 allows the representation of random variables that are $\mathcal{F}_T = \mathcal{F}_T^\alpha \vee \mathcal{F}_T^H$ -measurable, which is a larger σ -algebra than $\mathcal{F}_T^H$. The function ϕ in (2.5.8) can be described in explicit terms depending on conditional expectations. This approach exploits the fact that the filtration $\mathbb{F}^H$ can be fully characterized by the jump times. This is not the case for

the filtration $\mathbb{G}$ in which case we consider additional random noise such as the one generated by α . Theorem 2.5.1 provides an explicit characterization of the integrand in this setting.

2.6 Iterated integrals and chaos expansions

In this section, we revise the notion of Itô-type iterated integrals, with the intent to relate them with the α -multilinear forms of section 2.3. With this in mind, the iterated integrals are developed without any symmetrization schemes. These iterated integrals will later help us connect the α -multilinear forms with the Malliavin-type derivatives developed in [107] using symmetrization schemes and multiple integrals. In particular, Theorem 2.6.4 resembles [107, Corollary 14], but our construction is better suited for an analysis starting from α -multilinear forms. Let

$$S_p := \left\{ (s_1, z_1 \dots s_p, z_p) \in ([0, T] \times \mathcal{Z})^p \mid 0 \leq s_1 \leq s_2 \leq \dots \leq s_p \leq T \right\}$$

Denote $\mathcal{I}_\alpha^p$ the set of $\mathcal{F}^\alpha$ -measurable functions, $\phi : \Omega \times S_p \rightarrow \mathbb{R}$, such that

$$\|\phi\|_{\mathcal{I}_\alpha^p} := \left(\mathbb{E} \left[\int_{S_p} \phi^2(s_1, z_1, s_2 \dots s_p, z_p) \alpha(ds_1 dz_1) \dots \alpha(ds_p dz_p) \right] \right)^{\frac{1}{2}} < \infty. \quad (2.6.1)$$

For any $\phi \in \mathcal{I}_\alpha^p$, the p 'th iterated integral is defined as

$$J_p(T, \phi) := \int_0^T \int_{\mathcal{Z}} \int_0^{s_p} \int_{\mathcal{Z}} \dots \int_0^{s_2} \int_{\mathcal{Z}} \phi(s_1, z_1 \dots s_p, z_p) \tilde{H}(ds_1 dz_1) \dots \tilde{H}(ds_p dz_p), \quad (2.6.2)$$

and we set $J_p := \{J_p(T, \phi), \phi \in \mathcal{I}_\alpha^p\}$. From (2.4.4) and (2.4.6) we have

$$\begin{aligned} \mathbb{E} \left[J_p^2(T, \phi) \right] &= \mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} \mathbb{E} \left[\left(\int_0^{s_p} \int_{\mathcal{Z}} \dots \int_0^{s_2} \int_{\mathcal{Z}} \phi(s_1, z_1 \dots s_p, z_p) \right. \right. \right. \\ &\quad \left. \left. \left. \tilde{H}(ds_1 dz_1) \dots \tilde{H}(ds_{p-1} dz_{p-1}) \right)^2 \middle| \mathcal{F}^\alpha \right] \alpha(ds_p dz_p) \right] \\ &= \mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} \left(\int_0^{s_p} \int_{\mathcal{Z}} \dots \int_0^{s_2} \int_{\mathcal{Z}} \phi^2(s_1, z_1 \dots s_p, z_p) \alpha(ds_1, dz_1) \dots \alpha(ds_p dz_p) \right) \right] \\ &= \|\phi\|_{\mathcal{I}_\alpha^p}^2. \end{aligned} \quad (2.6.3)$$

The iterated integrals J_p are in correspondence with the space of α -multilinear forms $\mathbb{H}^p$ (see Definition 2.3.6). An example is instructive before considering the general case.

Example 2.6.1. Let ξ be a p -order α -multilinear form with pairwise disjoint time-intervals, i.e.

$$\xi = \beta \prod_{j=1}^p \tilde{H}(\Delta_j),$$

with $\Delta_1 = (s_1, u_1] \times Z_1, \Delta_2 = (s_2, u_2] \times Z_2, \dots$ and $0 \leq s_1 < u_1 \cdots < s_p \leq u_1$. Then

$$\begin{aligned}
\xi &= \beta \prod_{j=1}^p \tilde{H}(\Delta_j) = \beta \prod_{j=1}^p I(\mathbf{1}_{\Delta_j}) \\
&= \int_0^T \int_{\mathcal{Z}} \mathbf{1}_{\Delta_p}(s_p, z_p) \beta \prod_{j=1}^{p-1} I(\mathbf{1}_{\Delta_j}) \tilde{H}(ds_p, dz_p) \\
&= \int_0^T \int_{\mathcal{Z}} \mathbf{1}_{\Delta_p}(s_p, z_p) \int_0^{s_p} \int_{\mathcal{Z}} \mathbf{1}_{\Delta_{p-1}}(s_{p-1}, z_{p-1}) \beta \\
&\quad \prod_{j=1}^{p-2} I(\mathbf{1}_{\Delta_j}) \tilde{H}(ds_{p-1}, dz_{p-1}) \tilde{H}(ds_p, dz_p) \\
&= \int_0^T \int_{\mathcal{Z}} \int_0^{s_p} \int_{\mathcal{Z}} \dots \int_0^{s_2} \int_{\mathcal{Z}} \beta \mathbf{1}_{\Delta_p}(s_p, z_p) \dots \mathbf{1}_{\Delta_1}(s_1, z_1) \tilde{H}(ds_1 dz_1) \dots \tilde{H}(ds_p dz_p). \quad (2.6.4)
\end{aligned}$$

Next we expand this representation to the case when the sets are “overlapping in time”. It is possible to investigate this using Itô’s formula or symmetric functions, but instead we exploit the explicit result from Proposition 2.5.5.

Theorem 2.6.2. *If $\xi \in \mathbb{H}^p$, $p \geq 1$, then ξ can be represented as a p ’th iterated integral, ie*

$$\xi = \int_0^T \int_{\mathcal{Z}} \int_0^{s_p^-} \int_{\mathcal{Z}} \dots \int_0^{s_2^-} \int_{\mathcal{Z}} \phi(s_p, z_p, \dots, s_1, z_1) \tilde{H}(ds_1 dz_1) \dots \tilde{H}(ds_p dz_p), \quad (2.6.5)$$

where $\phi \in \mathcal{I}_{\alpha}^p$. Furthermore we have

$$\|\xi\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})} = \|\phi\|_{\mathcal{I}_{\alpha}^p}. \quad (2.6.6)$$

Proof. First we prove this for the α -multilinear forms by induction. The result is true for α -multilinear forms of order $p = 1$. Consider $p \geq 2$. Assume, as induction hypothesis, that a representation of type (2.6.5) holds for all multilinear forms of order $p - 1$. Let

$$\xi' = \prod_{j=1}^{p-1} \tilde{H}(\Delta_j \cap [0, t) \times \mathcal{Z}) \quad (2.6.7)$$

Being a $(p - 1)$ -order α -multilinear, it has representation

$$\xi' = \int_0^t \int_{\mathcal{Z}} \dots \int_0^{s_{2-}} \int_{\mathcal{Z}} \phi'_{p-1} \tilde{H}(ds_1, dz_1) \dots \tilde{H}(dz_{p-1} ds_{p-1}),$$

with means of $\phi'_{p-1} \in \mathcal{I}_{\alpha}^{p-1}$. Denote this integral as $J_{p-1}(t, \phi'_{p-1})$.

Let ξ be an α -multilinear form of order p , $\xi = \beta \prod_{j=1}^p \tilde{H}(\Delta_j)$. From Proposition 2.5.5, we know that $\xi = I(\mathcal{D}\xi)$, with

$$\mathcal{D}_{s,z}\xi = \beta \sum_{\substack{1 \leq i \leq p \\ \Delta_i \subseteq \Delta'}} \mathbf{1}_{\Delta_i}(s, z) \prod_{j \neq i}^p \tilde{H}(\Delta_j \cap [0, s) \times \mathcal{Z}).$$

Hence, by (2.6.7) we have

$$\begin{aligned} \xi &= I(\mathcal{D}\xi) \\ &= \int_0^T \int_{\mathcal{Z}} \left[\beta \sum_{\substack{1 \leq i \leq p \\ \Delta_i \subseteq \Delta'}} \mathbf{1}_{\Delta_i}(s, z) \prod_{j \neq i}^p \tilde{H}(\Delta_j \cap [0, s) \times \mathcal{Z}) \right] \tilde{H}(ds, dz) \\ &= \int_0^T \int_{\mathcal{Z}} \left[\beta \sum_{\substack{1 \leq i \leq p \\ \Delta_i \subseteq \Delta'}} \mathbf{1}_{\Delta_i}(s, z) J_{p-1}(s, \phi_i) \right] \tilde{H}(ds, dz) \\ &= \int_0^T \int_{\mathcal{Z}} \int_0^{s_p^-} \dots \int_0^{s_2^-} \left[\beta \sum_{\substack{1 \leq i \leq p \\ \Delta_i \subseteq \Delta'}} \mathbf{1}_{\Delta_i}(s_p, z_p) \phi_i(s_1, \dots, s_{p-1}, z_{p-1}) \right] \\ &\quad \tilde{H}(ds_1, dz_1) \dots \tilde{H}(ds_p, dz_p) \end{aligned}$$

which is an iterated integral of order p .

Any ξ in $\mathbb{H}^p$ is the limit in $L_2(\Omega, \mathcal{F}, \mathbb{P})$ of a finite sums of multilinear forms of order p . Let ξ_n be such a sequence. Any finite sum of multilinear forms can be written as a p 'th iterated integral, let

$$\xi_n = J_p(\phi_n), \quad \phi_n \in \mathcal{I}_\alpha^p.$$

Since ξ_n is a Cauchy sequence in $L_2(\Omega, \mathcal{F}, \mathbb{P})$, ϕ_n is a Cauchy sequence in $\mathcal{I}_\alpha^p$ by the isometry in (2.6.3). Hence there exists a unique $\phi \in \mathcal{I}_\alpha^p$ such that $\phi_n \rightarrow \phi$ as $n \rightarrow \infty$ and we must have $\xi = J_p(\phi)$. Finally, equation (2.6.6) follows directly from (2.6.3). $\square$

Remark 2.6.3. From (2.6.4), we can see that if $\xi = \beta \prod_{j=1}^p \tilde{H}(\Delta_j)$ is an α -multilinear form, with $\beta \in \mathbb{R}$ then $\xi = J_p(\phi)$ with ϕ deterministic. For general $\xi \in \tilde{\mathbb{H}}^p$ (Remark 2.3.13), we can use the same arguments as in Theorem 2.6.2 to conclude that $\xi = J_p(\phi)$, where $\phi \in \mathcal{I}_\alpha^p$ is deterministic. Thus $\tilde{\mathbb{H}}^p$ is the space spanned by iterated integrals of order p with deterministic integrands.

Theorem 2.6.4 (Chaos expansion). *For any $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$, there is unique sequence of integrands $\phi_p \in \mathcal{I}_\alpha^p$, $p = 1, 2, \dots$ such that the following representation holds:*

$$\xi = \mathbb{E}[\xi | \mathcal{F}^\alpha] + \sum_{p=1}^{\infty} J_p(\phi_p).$$

Proof. Theorem 2.3.12 shows that any $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ has orthogonal decomposition

$$\xi = \sum_{p=0}^{\infty} \xi_p$$

with $\xi_p \in \mathbb{H}^p$, $p = 0, 1, \dots$. Any ξ_p , $p \geq 1$ can be written as a p 'th iterated integral by Theorem 2.6.2 and $\xi_0 = \mathbb{E}[\xi | \mathcal{G}_0] = \mathbb{E}[\xi | \mathcal{F}^\alpha]$ is the projection of ξ on $\mathbb{H}^0$. $\square$

Directly from Theorem 2.6.4 we can formulate an integral representation theorem.

Corollary 2.6.5. *For any $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ there exists a unique $\phi \in \mathcal{I}$ such that*

$$\xi = \mathbb{E}[\xi | \mathcal{F}^\alpha] + \int_0^T \int_{\mathcal{Z}} \phi(s, z) \tilde{H}(ds, dz)$$

Note that this corollary is in line with classical stochastic integral representation theorems with respect to square integrable martingales as integrators. Corollary 2.6.5 offers no immediate way of computing the integrand ϕ since only the existence of the representation via the kernel functions of the iterated integrals is given. Corollary 2.6.5 deeply differs from Theorem 2.5.1 and the following Theorem 2.7.9. The last ones characterize the integrand ϕ in terms of derivative operators.

2.7 Anticipative stochastic derivatives and integral representations

Motivated by Clark-Ocone type formulae we study ways to compute the non-anticipating derivative and to have stochastic integral representations. We introduce a new anticipative derivative operator D^a , partially inspired by [37]. We study this operator in relation with a Malliavin-type derivative for processes with conditionally independent increments developed in [107].

Let $\mathcal{G}_{\Delta^c}$ be the minimal complete σ -algebra containing $\mathcal{F}^\alpha$ and the sets $\{\tilde{H}(\Delta') | \Delta' \subset \Delta^c\}$, where Δ^c is the complement of Δ .

Definition 2.7.1. *For $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ we set $D_{s,z}^a \xi(n)$ as the element of $L_2(\Omega \times X)$ given by*

$$D_{s,z}^a \xi(n) := \sum_{k=1}^{K_n} \mathbb{E} \left[\xi \frac{\tilde{H}(\Delta_{n,k})}{\alpha(\Delta_{n,k})} \middle| \mathcal{G}_{\Delta_{n,k}^c} \right] \mathbf{1}_{\Delta_{n,k}}(s, z), \quad (2.7.1)$$

with the n referring to the n 'th partition of the dissecting system. Denote by $D^a \xi$ the limit in $L_2(\Omega \times X)$ (if it exists) given by

$$D^a \xi = \lim_{n \rightarrow \infty} D_{s,z}^a \xi(n). \quad (2.7.2)$$

We define $\mathbb{D}^a$ as the subset of $L_2(\Omega, \mathcal{F}, \mathbb{P})$ where the limit (2.7.2) exists and we define

$$\|\xi\|_{\mathbb{D}^a} := \left(\mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} (D_{s,z}^a \xi)^2 \alpha(ds, dz) \right] \right)^{\frac{1}{2}} < \infty.$$

Then D^a is a linear operator, $D^a : \mathbb{D}^a \rightarrow L_2(\Omega \times X)$.

We remark that for any $\beta \in \mathbb{H}^0$ (recall Definition 2.3.2), $\beta \in \mathbb{D}^a$ and $D^a \beta = 0$.

Lemma 2.7.2. For $p \geq 1$, let ξ be a p -order α -multilinear form, i.e. $\xi = \beta \prod_{j=1}^p \tilde{H}(\Delta_j)$. Then

$$D_{s,z}^a \xi = \beta \sum_{i=1}^p \mathbf{1}_{\Delta_i}(s, z) \prod_{j \neq i} \tilde{H}(\Delta_j), \quad (2.7.3)$$

and

$$D^a \xi(n) = \sum_{i=1}^p \sum_{k=1}^{K_n} \frac{\alpha(\Delta_i \cap \Delta_{n,k})}{\alpha(\Delta_{n,k})} \prod_{j \neq i} \tilde{H}(\Delta_j) \mathbf{1}_{\{\Delta_{n,k} \cap \Delta_j = \emptyset\}}(k, j) \mathbf{1}_{\Delta_{n,k}}(s, z). \quad (2.7.4)$$

Furthermore

$$\|\xi\|_{\mathbb{D}^a} = \|D^a \xi\|_{L_2(\Omega \times X)} = \sqrt{p} \|\xi\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})}.$$

Proof. For any n and $k = 1, \dots, K_n$, denote

$$\psi(n, k) = \mathbb{E} \left[\beta \prod_{j=1}^p \tilde{H}(\Delta_j) \frac{\tilde{H}(\Delta_{n,k})}{\alpha(\Delta_{n,k})} \middle| \mathcal{G}_{\Delta_{n,k}^c} \right]$$

If $\Delta_{n,k} \cap \left(\bigcup_{j=1}^p \Delta_j \right) = \emptyset$ or if $\Delta_{n,k}$ intersects with *more than one* of the sets Δ_j 's, then $\psi(n, k)$ is equal to zero by direct computation. If $\Delta_{n,k} \subset \Delta_i$ for some i , then

$$\begin{aligned} \psi(n, k) &= \mathbb{E} \left[\beta \prod_{j=1}^p \tilde{H}(\Delta_j) \frac{\tilde{H}(\Delta_{n,k})}{\alpha(\Delta_{n,k})} \middle| \mathcal{G}_{\Delta_{n,k}^c} \right] \\ &= \beta \prod_{j \neq i} \tilde{H}(\Delta_j) \mathbb{E} \left[\left(\tilde{H}(\Delta_i \cap \Delta_{n,k}) + \tilde{H}(\Delta_i \setminus \Delta_{n,k}) \right) \frac{\tilde{H}(\Delta_{n,k})}{\alpha(\Delta_{n,k})} \middle| \mathcal{G}_{\Delta_{n,k}^c} \right] \\ &= \beta \prod_{j \neq i} \tilde{H}(\Delta_j). \end{aligned}$$

If $\Delta_i \subsetneq \Delta_{n,k}$ for some i and $\Delta_{n,k} \cap \Delta_j = \emptyset$ for all $j \neq i$, then

$$\psi(n, k) = \beta \frac{\alpha(\Delta_i)}{\alpha(\Delta_{n,k})} \prod_{j \neq i} \tilde{H}(\Delta_j).$$

Combining the above cases we conclude that

$$\psi(n, k) = \sum_{i=1}^p \frac{\alpha(\Delta_i \cap \Delta_{n,k})}{\alpha(\Delta_{n,k})} \prod_{j \neq i} \tilde{H}(\Delta_j) \mathbf{1}_{\{\Delta_{n,k} \cap \Delta_j = \emptyset\}}(k, j).$$

and (2.7.4) follows. Passing to the limit in $L_2(\Omega \times X)$ we have

$$D^a \xi = \lim_{n \rightarrow \infty} D^a \xi(n) = \beta \sum_{i=1}^p \mathbf{1}_{\Delta_i}(s, z) \prod_{j \neq i} \tilde{H}(\Delta_j).$$

Moreover

$$\begin{aligned} \|\xi\|_{\mathbb{D}^a}^2 &= \mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} \left(\beta \sum_{i=1}^p \mathbf{1}_{\Delta_i}(s, z) \prod_{j \neq i} \tilde{H}(\Delta_j) \right)^2 \alpha(ds, dz) \right] \\ &= \mathbb{E} \left[\beta^2 \sum_{i=1}^p \alpha(\Delta_i) \prod_{j \neq i} \tilde{H}(\Delta_j)^2 \right] \\ &= \mathbb{E} \left[\beta^2 p \prod_{j=1}^p \alpha(\Delta_j) \right] \\ &= p \|\xi\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})}^2. \end{aligned}$$

□

Comparing (2.7.3) and (2.7.4) we can see that, for the p -order α -multilinear form ξ , the following estimate holds for all n :

$$\|D^a \xi(n)\|_{L_2(\Omega \times X)} \leq \|D^a \xi\|_{L_2(\Omega \times X)} = \sqrt{p} \|\xi\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})}. \quad (2.7.5)$$

The following statements are an immediate consequence in Lemma 2.7.2.

Corollary 2.7.3. *Let $p \geq 1$. Let ξ_1 and ξ_2 be orthogonal p -order α -multilinear forms. Then, for all n , $D^a \xi_1(n)$ and $D^a \xi_2(n)$ are orthogonal in $L_2(\Omega \times X)$. The same holds for $D^a \xi_1$ and $D^a \xi_2$.*

Corollary 2.7.4. *For $p_1 > p_2 \geq 1$, let $\xi_1 \in \mathbb{H}^{p_1}$ and $\xi_2 \in \mathbb{H}^{p_2}$ be α -multilinear forms. Then $D^a \xi_1$ and $D^a \xi_2$ are orthogonal in $L_2(\Omega \times X)$.*

Finally we have the following result:

Proposition 2.7.5. *For $p \geq 1$, if $\xi \in \mathbb{H}^p$ then $\xi \in \mathbb{D}^a$ with*

$$\|\xi\|_{\mathbb{D}^a} = \sqrt{p} \|\xi\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})} < \infty. \quad (2.7.6)$$

Proof. Any $\xi \in \mathbb{H}^p \subset L_2(\Omega, \mathcal{F}, \mathbb{P})$ can be approximated by a sequence ξ_m , $m = 1, 2, \dots$, of finite sums of α -multilinear forms of order p : $\lim_{m \rightarrow \infty} \|\xi_m - \xi\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})} = 0$. First of all we

observe that, from Remark 2.3.4, ξ_m can be represented as finite sums of orthogonal p -order α -multilinear forms. From Lemma 2.7.2 and Corollary 2.7.3 we can see that $D^a \xi_m$ is a Cauchy sequence in $L_2(\Omega \times X)$ with limit ϕ such that $\|\phi\|_{L_2(\Omega \times X)} = \sqrt{p} \|\xi\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})}$.

We show that indeed $\phi = D^a \xi := \lim_{n \rightarrow \infty} D^a \xi(n)$ in $L_2(\Omega \times X)$. By application of Corollary 2.7.3 and (2.7.5) we have

$$\|D^a \xi_m(n)\|_{L_2(\Omega \times X)} \leq \sqrt{p} \|\xi_m\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})}. \quad (2.7.7)$$

Moreover we note that

$$\begin{aligned} & \|D^a \xi_m(n) - D^a \xi(n)\|_{L_2(\Omega \times X)}^2 \\ &= \mathbb{E} \left[\int_0^T \int_Z \sum_{k=1}^{K_n} \left(\mathbb{E} \left[(\xi_m - \xi) \frac{\tilde{H}(\Delta_{n,k})}{\alpha(\Delta_{n,k})} \middle| \mathcal{G}_{\Delta_{n,k}^c} \right] \right)^2 \mathbf{1}_{\Delta_{n,k}}(s, z) \alpha(ds, dz) \right] \\ &\leq \mathbb{E} \left[\sum_{k=1}^{K_n} \mathbb{E} \left[(\xi_m - \xi)^2 \middle| \mathcal{G}_{\Delta_{n,k}^c} \right] \mathbb{E} \left[\tilde{H}(\Delta_{n,k})^2 \middle| \mathcal{G}_{\Delta_{n,k}^c} \right] \frac{1}{\alpha(\Delta_{n,k})} \right] \\ &= K_n \|\xi_m - \xi\|^2. \end{aligned}$$

Hence we have

$$\begin{aligned} & \lim_{n \rightarrow \infty} \|\phi - D^a \xi(n)\|_{L_2(\Omega \times X)} \\ &\leq \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \left\{ \|\phi - D^a \xi_m\|_{L_2(\Omega \times X)} + \|D^a \xi_m - D^a \xi_m(n)\|_{L_2(\Omega \times X)} \right. \\ &\quad \left. + \|D^a \xi_m(n) - D^a \xi(n)\|_{L_2(\Omega \times X)} \right\} = 0. \end{aligned}$$

In fact

$$\begin{aligned} & \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \|D^a \xi_m(n) - D^a \xi(n)\|_{L_2(\Omega \times X)} \\ &\leq \lim_{n \rightarrow \infty} \left\{ \sqrt{K_n} \lim_{m \rightarrow \infty} \|\xi_m - \xi\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})} \right\} = 0 \end{aligned}$$

and by (2.7.7)

$$\begin{aligned} & \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \|D^a \xi_m - D^a \xi_m(n)\|_{L_2(\Omega \times X)} \\ &\leq \lim_{q \rightarrow \infty} \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \left\{ \|D^a \xi_m - D^a \xi_q(n)\|_{L_2(\Omega \times X)} \right. \\ &\quad \left. + \|D^a \xi_q(n) - D^a \xi_m(n)\|_{L_2(\Omega \times X)} \right\} \\ &\leq \lim_{q \rightarrow \infty} \|\phi - D^a \xi_q\|_{L_2(\Omega \times X)} + \sqrt{p} \lim_{q \rightarrow \infty} \lim_{m \rightarrow \infty} \|\xi_q - \xi_m\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})} = 0. \end{aligned}$$

□

The Malliavin calculus for processes with conditionally independent increments was developed in [107], this include the CDSPP. The results and developments therein are close to those of [83, Chapter 1]. We summarize some of those results with the aim of showing how these

operators relate to the operator D^a and the non-anticipating derivative $\mathcal{D}$.

Let $f_p : L_2(\Omega \times ((0, T] \times \mathcal{Z})^p) \rightarrow \mathbb{R}$ where f_p is $\mathcal{F}^\alpha \times \mathcal{B}_X$ -measurable. Remark that f_p is *not* defined on $\mathcal{I}_\alpha^p$, which is a smaller space. We say that f_p is *simple* if

$$f_p = \sum_{i=1}^n \beta_i(\omega) \mathbf{1}_{\Delta_1}(s_1, z_1) \dots \mathbf{1}_{\Delta_p}(s_p, z_p)$$

where $\beta_i, i = 1, \dots, n$ is a bounded $\mathcal{F}^\alpha$ -measurable random variable and the sets $\Delta_1, \dots, \Delta_p$ are pairwise disjoint. The multiple integrals of order p of a simple function are then

$$I_p(T, f_p) := \sum_i \beta_i \prod_{j=1}^p \tilde{H}(\Delta_j),$$

i.e. the multiple integrals of simple functions of order p are sums of α -multilinear forms of order p . These multiple integrals are extended to integrals of general $\mathcal{F}^\alpha \times \mathcal{B}_X$ -measurable functions $f_p : L_2(\Omega \times (0, T] \times \mathcal{Z})^p \rightarrow \mathbb{R}$ by taking limits of simple functions. We conclude that the space spanned by multiple integrals of order p on the functions above coincide with $\mathbb{H}^p$.

Any $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ has representation (as per Theorem 2.6.4 and [107, Corollary 14])

$$\xi = \mathbb{E}[\xi | \mathcal{F}^\alpha] + \sum_{p=1}^{\infty} I_p(\tilde{f}_p), \quad (2.7.8)$$

by means of a sequence $\tilde{f}_p, p \geq 1$, of symmetric functions in $L_2(\Omega \times ((0, T] \times \mathcal{Z})^p)$.

Denote the symmetrization of f_p by

$$\tilde{f}_p := \frac{1}{p!} \sum_{\sigma} f(s_{\sigma(1)}, z_{\sigma(1)}, \dots, s_{\sigma(p)}, z_{\sigma(p)})$$

where σ is running over all permutations of $1, \dots, p$. Let $\phi_p \in \mathcal{I}_\alpha^p$ (see (2.6.1)) and $f_p = \mathbf{1}_{S_p} \phi_p$. Then the following equalities hold [107, Section 3]:

$$J_p(T, \phi_p) = I_p(T, f_p) = I_p(T, \tilde{f}_p) = p! J_p(T, \tilde{f}_p).$$

The *Malliavin derivative* $D : \mathbb{D}_{1,2} \rightarrow L_2(\Omega \times X)$ is given by

$$D_{s,z} \xi := \sum_{p=1}^{\infty} p I_{p-1}(\tilde{f}_p(\cdot, s, z)) \quad (2.7.9)$$

for all $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ with $\xi = \mathbb{E}[\xi | \mathcal{F}^\alpha] + \sum_{p=1}^{\infty} I_p(\tilde{f}_p)$, such that

$$\|\xi\|_{\mathbb{D}_{1,2}} := \left(\sum_{p=1}^{\infty} p p! \|\tilde{f}_p\|_{\mathcal{I}_\alpha^p}^2 \right)^{\frac{1}{2}} < \infty.$$

Indeed, $\|D\xi\|_{L_2(\Omega \times X)}^2 = \sum_{p=1}^{\infty} p p! \|\tilde{f}_p\|_{\mathcal{I}_\alpha^p}^2$.

Lemma 2.7.6. For $p \geq 1$, let $\xi = \beta \prod_{j=1}^p \tilde{H}(\Delta_j)$ be an α -multilinear form. Then

$$D_{s,z}\xi = \beta \sum_{i=1}^p \mathbf{1}_{\Delta_i}(s, z) \prod_{j \neq i} \tilde{H}(\Delta_j) \quad (2.7.10)$$

and

$$\|\xi\|_{\mathbb{D}_{1,2}} = \sqrt{p} \|\xi\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})}.$$

Proof. Since $\xi = J_p(T, \beta \mathbf{1}_{\Delta_1} \dots \mathbf{1}_{\Delta_p})$, $\xi = I_p(\tilde{f}_p)$ with

$$\tilde{f}_p = \beta \frac{1}{p!} \sum_{\sigma} \mathbf{1}_{\Delta_1}(s_{\sigma(1)}, z_{\sigma(1)}) \dots \mathbf{1}_{\Delta_p}(s_{\sigma(p)}, z_{\sigma(p)}).$$

Thus, from (2.7.9), we have

$$\begin{aligned} D_{s,z}\xi &= \beta \frac{1}{p!} p \sum_{\sigma} \mathbf{1}_{\Delta_p}(s_{\sigma(p)}, z_{\sigma(p)}) \\ &\quad I_{p-1} \left(\mathbf{1}_{\Delta_1}(s_{\sigma(1)}, z_{\sigma(1)}) \dots \mathbf{1}_{\Delta_p}(s_{\sigma(p-1)}, z_{\sigma(p-1)}) \right) \Big|_{\substack{s_{\sigma(p)}=s \\ z_{\sigma(p)}=z}} \\ &= \frac{p}{p!} \beta \sum_{i=1}^p \mathbf{1}_{\Delta_i}(s, z) (p-1)! \prod_{j \neq i} \tilde{H}(\Delta_j) \\ &= \beta \sum_{i=1}^p \mathbf{1}_{\Delta_i}(s, z) \prod_{j \neq i} \tilde{H}(\Delta_j). \end{aligned}$$

Let us compute the norm of ξ in $\mathbb{D}_{1,2}$. Note that

$$\|\tilde{f}_p\|_{\mathcal{I}_\alpha^p}^2 = \mathbb{E} \left[\beta^2 \frac{1}{p!} \prod_{j=1}^p \alpha(\Delta_j) \right].$$

Hence

$$\|\xi\|_{\mathbb{D}_{1,2}}^2 = pp! \|\tilde{f}_p\|_{\mathcal{I}_\alpha^p}^2 = p \mathbb{E} \left[\beta^2 \prod_{j=1}^p \alpha(\Delta_j) \right] = p \|\xi\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})}^2.$$

□

We observe that if $\xi \in \mathbb{H}^p$, $p \geq 1$, then by the closability of D , [107, Lemma 21], and Lemma 2.7.6 it follows that $\xi \in \mathbb{D}_{1,2}$ with

$$\|\xi\|_{\mathbb{D}_{1,2}} = \sqrt{p} \|\xi\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})} < \infty. \quad (2.7.11)$$

Moreover, if β is $\mathcal{F}^\alpha$ measurable we have $D\beta = 0$ by [107, Proposition 25].

Recall that $L_2(\Omega, \mathcal{F}, \mathbb{P}) = \mathbb{H} = \sum_{p=0}^{\infty} \oplus \mathbb{H}^p$ (see Theorem 2.3.12).

Proposition 2.7.7. *If $\xi \in \mathbb{H}$ then $\xi \in \mathbb{D}^a$ if and only if*

$$\sum_{p=1}^{\infty} \sqrt{p} \|\xi_p\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})} < \infty. \quad (2.7.12)$$

Here $\xi_0, \xi_1, \dots$ is the orthogonal decomposition of ξ in the chaos expansion of Theorem 2.3.12. Indeed we have $\|D^a \xi\|_{\mathbb{D}^a} = \|D\xi\|_{\mathbb{D}_{1,2}} = \sum_{p=1}^{\infty} \sqrt{p} \|\xi_p\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})}$.

Proof. This is a direct application of Lemma 2.7.2, Corollary 2.7.4 for D^a and of Theorem 2.6.2, Lemma 2.7.6, and (2.7.10), for D . $\square$

We conclude that the spaces $\mathbb{D}^a$ and $\mathbb{D}_{1,2}$ coincide but are not equal to the whole of $L_2(\Omega, \mathcal{F}, \mathbb{P})$, i.e.

$$\mathbb{D}^a = \mathbb{D}_{1,2} \subsetneq L_2(\Omega, \mathcal{F}, \mathbb{P}).$$

Moreover, we stress that for any $\xi \in \mathbb{D}^a$, there exists a sequence $\xi_m, m = 1, 2, \dots$, of finite sums of α -multilinear forms approximating ξ . Then $D^a \xi_m$ and $D\xi_m$ are two identical converging sequences by Lemma 2.7.2 and Lemma 2.7.6. These two sequences must have the same limit in $L_2(\Omega \times X)$.

We summarize the above arguments into the following statement:

Theorem 2.7.8. *The operators D^a and D coincide, i.e. $\mathbb{D}^a = \mathbb{D}_{1,2}$ and*

$$D^a \xi = D\xi \quad \text{in } L_2(\Omega \times X).$$

After the above result we can also interpret the operator D^a as an alternative approach to describe the Malliavin derivative which shows the anticipative dependence of the operator on the information in a much more structural and explicit way than the classical approach via chaos expansions of iterated integrals.

The following theorem is a Clark-Ocone type result which provides an alternative way to compute the non-anticipating derivative in the integral representation of Theorem 2.5.1. Denote $\mathbb{E}[D^a \xi | \mathcal{G}]$ the stochastic process given by $\phi(s, z) = \mathbb{E}[D_{s,z}^a \xi | \mathcal{G}_{s-}]$, $s \in [0, T]$, $z \in \mathcal{Z}$.

Theorem 2.7.9. *For any $\xi \in \mathbb{D}^a$ we have*

$$\mathbb{E}[D^a \xi | \mathcal{G}] = \mathbb{E}[D\xi | \mathcal{G}] = \mathcal{D}\xi \quad \mathbb{P} \times \alpha \text{ a.e.}$$

Proof. The first equality follows from Theorem 2.7.8. Assume $\xi \in \mathbb{H}^p$ is a p -order α -multi-

linear form, $\xi = \beta \prod_{j=1}^p \tilde{H}(\Delta_j)$. From (2.7.3),

$$\begin{aligned} \mathbb{E}[D_{s,z}^a \xi | \mathcal{G}_{s-}] &= \beta \sum_{i=1}^p \mathbb{E}\left[\prod_{j \neq i} \tilde{H}(\Delta_j) \middle| \mathcal{G}_{s-}\right] \mathbf{1}_{\Delta_i}(s, z) \\ &= \beta \sum_{\substack{1 \leq i \leq p \\ \Delta_i \subseteq \Delta'}} \mathbf{1}_{\Delta_i}(s, z) \prod_{j \neq i}^p \tilde{H}(\Delta_j \cap [0, s) \times \mathcal{Z}) \\ &= \mathcal{D}_{s,z} \xi \end{aligned}$$

by comparing to (2.5.5). The set Δ' is as described in (2.5.6). By approximation we obtain the statement first for the general $\xi \in \mathbb{H}^p$ and then for $\xi \in \mathbb{H}$: $\xi = \lim_{q \rightarrow \infty} \sum_{p=0}^q \xi_p$ with $\xi_p \in \mathbb{H}^p$. $\square$

Corollary 2.7.10. *For any $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ there exists a sequence $\xi_q \in \mathbb{D}^a$, $q = 1, \dots$ such that $\xi_q \rightarrow \xi$ in $L_2(\Omega, \mathcal{F}, \mathbb{P})$ and*

$$\mathcal{D}\xi_q = \mathbb{E}[D^a \xi_q | \mathcal{G}] \longrightarrow \mathcal{D}\xi \quad \text{as } q \rightarrow \infty \quad \text{in } \mathcal{I}.$$

Thus

$$\xi = \mathbb{E}[\xi | \mathcal{F}^\alpha] + \lim_{q \rightarrow \infty} \int_0^T \int_{\mathcal{Z}} \mathbb{E}[D_{s,z}^a \xi_q | \mathcal{G}_{s-}] \tilde{H}(ds, dz)$$

with convergence in $L_2(\Omega, \mathcal{F}, \mathbb{P})$.

Proof. Take ξ_q to be the projection of ξ on $\mathbb{H}_q = \sum_{p=0}^q \oplus \mathbb{H}^p$, this is $\xi_q = \mathbb{E}[\xi | \mathcal{F}^\alpha] + \sum_{p=1}^q \xi_p$, and apply Remark 2.5.2 and Theorem 2.7.9. $\square$

Chapter 3

BSDEs for time-changed Lévy processes and application to optimal control

Abstract

We study backward stochastic differential equations (BSDEs) for time-changed Lévy noises when the time-change is independent of the Lévy process. We prove existence and uniqueness of the solution and we obtain an explicit formula for linear BSDEs and a comparison principle. BSDEs naturally appear in control problems. Here we prove sufficient and necessary maximum principles for a general optimal control problem of a system driven by a time-changed Lévy noise. As an application we solve the mean-variance portfolio selection problem.

3.1 Introduction

We establish a framework for the study of backward stochastic differential equations (BSDEs) driven by a conditional Brownian motion and a doubly stochastic Poisson random field. Indeed the structure of these noises can be strongly related to the corresponding time-changed Brownian motion and the time-changed Poisson random measure when the time-change is independent of the Brownian motion and Poisson field.

In the framework of the non-anticipating integration for martingale random fields, we prove the existence and uniqueness of the solution of a general BSDE of the form

$$\begin{aligned} Y_t &= \xi + \int_t^T g_s(\lambda_s, Y_s, \phi_s) ds - \int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz) \\ &= \xi + \int_t^T g_s(\lambda_s, Y_s, \phi_s) ds - \int_t^T \phi_s(0) dB_s - \int_t^T \int_{\mathbb{R}_0} \phi_s(z) \tilde{H}(ds, dz) \end{aligned} \quad (3.1.1)$$

where μ is the mixture of a conditional Brownian measure B on $[0, T] \times \{0\}$ and a centered doubly stochastic Poisson measure $\tilde{H}$ on $[0, T] \times \mathbb{R}_0$ ($\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$). Namely

$$\mu(\Delta) := B(\Delta \cap [0, T] \times \{0\}) + \tilde{H}(\Delta \cap [0, T] \times \mathbb{R}_0), \quad (3.1.2)$$

for any Borel measurable set Δ in $[0, T] \times \mathbb{R}$. Moreover we specifically study linear BSDEs achieving a closed form solution for the process Y and use this solution to obtain a comparison theorem.

These results rely strongly on the stochastic integral representation of square integrable random variables and martingales. In the language of time-change, we can formulate the result as follows: Given the time-change processes Λ^B and Λ^H , the complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{G})$ where $\mathbb{G}$ is the filtration generated by μ and the whole of Λ^B and Λ^H , any L_2 -martingale M can be represented as

$$M_t = M_0 + \int_0^t \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz) \quad (3.1.3)$$

where ϕ is proved to exist and M_0 is a random element measurable with respect to Λ^B and Λ^H .

In Chapter 2 a detailed study on the structure of the spaces generated by the measure $\tilde{H}$ is carried though achieving chaos decompositions via orthogonal polynomials and also integral representation results of type (3.1.3) in which the integrand is given in closed form via the non-anticipating stochastic derivative in first place and then via Clark-Ocone type formulae and anticipating stochastic derivatives. These results hold for very general choices of Λ^H also beyond the present paper. Here we give an alternative slimmer proof for representation (3.1.3) which will provide only existence of the integrand ϕ . This is enough for the study of (3.1.1).

We remark that (3.1.3) shows that martingales M of the type considered do not have a (full) predictable representation property as described in [22, 83, 93] since the initial value M_0 is not a constant in general. Indeed the predictable representation property depends on the combination of integrator and the information flow. In [44, Theorem 2.2] it is proved that the predictable property with respect to the class of random measures μ with independent values *if and only if* μ is given as a mixture of Gaussian and centered Poisson random measures.

The integration and the representation results are developed with respect to the filtration $\mathbb{G}$, the filtration generated by μ and *the entire history* of Λ^B and Λ^H . It is with respect to Λ^B and Λ^H that H and B have conditionally independent increments. From the point of view of modeling and applications $\mathbb{G}$ is not a natural choice of filtration since it includes “anticipating information”, the future values of Λ^B and Λ^H . However we can still apply our results in the problems related to models where the reference filtration is $\mathbb{F}$, the smallest right-continuous filtration to which μ is adapted. Indeed we show sufficient conditions for solving an optimal control problem with a classical performance functional for both $\mathbb{G}$ - and $\mathbb{F}$ -predictable controls. This is achieved by projecting the results obtained for the $\mathbb{G}$ -predictable case onto the $\mathbb{F}$ -predictable one.

The framework proposed based on specific integral representation under $\mathbb{G}$ is a novel framework for problems related to time-changed processes. The work [77] considers BSDEs with doubly stochastic Poisson processes, where the intensity of the doubly stochastic Poisson process depends on a Brownian motion in a specific way. Our setting does not overlap with that of [77] due to a different relationship between the noises considered. Our BSDE also differs from another approach to BSDEs beyond Lévy processes, [24, 18, 86, 61, 73], where an extra martingale N is inserted to the backward stochastic differential equation so that Y attains the terminal

condition $Y_t = \xi$ and Y_0 is a real number. BSDEs with random measures is discussed in [63] assuming a martingale representation exist. We however prove the martingale representation and explicitly link the random measures, the martingale representation, and the conditions on the driver.

Taking a view towards applications we sketch some of the uses of the time-changed Lévy processes in mathematical finance and the relevance of our BSDE-framework. This is not meant as a comprehensive review. The time-changed Lévy processes occur in mathematical finance in the modeling of asset prices as follows:

$$\begin{aligned} dS_t &= S_t \left(\int_{\mathbb{R}} \psi_t(z) \mu(ds, dz) \right. \\ &= S_t \left(\psi_t(0) dB_t + \int_{\mathbb{R}_0} \psi_t(z) \tilde{H}(dst, dz) \right) \quad S_0 > 0. \end{aligned} \quad (3.1.4)$$

The well-known stochastic exponentiation model of [25, Section 4.3], where stocks are modeled as time-changed pure jump Lévy processes, can be described in our terminology as

$$S_t = \exp \left\{ \int_0^t \int_{\mathbb{R}_0} z \tilde{H}(ds, dz) - \int_0^t \int_{\mathbb{R}_0} [e^z - 1 - z] \lambda_s^H \nu(dz) ds \right\} \quad (3.1.5)$$

which in differential form is

$$dS_t = S_t \left(\int_{\mathbb{R}_0} (e^z - 1) \tilde{H}(ds, dz) \right). \quad (3.1.6)$$

Here the jump measure ν and time-change intensity λ^H determine the properties of S .

A popular class of stochastic volatility models with Brownian filtrations including [8, 52, 54, 102] is

$$dS_t = \rho S_t dt + \sigma S_t \lambda_t^B dW_t^{(1)} \quad (3.1.7)$$

$$d\lambda_t^B = M(\lambda_t^B) dt + K(\lambda_t^B) dW_t^{(2)} \quad (3.1.8)$$

where M and K are real functions, $\rho, \sigma \in \mathbb{R}$ and $W^{(1)}$ and $W^{(2)}$ are Brownian motions. Here S is the asset price and λ^B the stochastic volatility. Whenever $W^{(1)}$ and $W^{(2)}$ are *independent*, $B_t := \int_0^t \lambda_s^B dW_s^{(1)}$ is a conditional independent Brownian motion as in Definition 3.2.1 and our framework applies.

In credit risk, the jump times of the doubly stochastic Poisson process are used to signify the occurrence of downwards abrupt price movements and default. A classical example [74] is the case of an integer valued stochastic process H_t , $t \in [0, T]$, with $\nu(dz) = \mathbf{1}_{\{z=1\}}(z)$ and λ^H given. Then $\tilde{H}_t = H_t - \Lambda_t^H$. The default time τ is the first jump of H , i.e. $\tau = \inf_t \{H(t) > 0\}$. This is then used to model bonds or derivatives of the form $P \mathbf{1}_{\tau > T}$, where P is a random variable, so that $P \mathbf{1}_{\tau > T}$ is a payoff which is received only if there is no default. An example of

type (3.1.4) is the zero coupon bond which can be modeled as

$$dS_t = S_t \left(\lambda_t^H dt - d\tilde{H}_t \right), \quad S_0 = 1, \quad \text{for } t \leq \tau.$$

To the best of our knowledge, the present work is the first to detail BSDEs for time-changed Lévy processes in general form, which opens up for studies on risk measures and filtration-consistent expectations as in [49, 94] via our comparison theorem. Moreover we explicitly treat general optimal control problems with time-changed Lévy processes, see e.g. (3.6.1), via the present BSDE. Indeed the BSDE can be used to investigate mean-variance hedging, utility maximization and optimal consumption problems for assets modeled as in (3.1.4) via Theorems 3.6.3 and 3.6.4. Utility maximization for time-changed Lévy processes is studied in [66, 67] for the power utility. Mean-variance hedging (for stochastic volatility and credit risk) has been discussed in terms of affine models [69, 68] and with BSDEs for general semi-martingales [18, 61, 73]. However [18] only consider continuous semi-martingales, [61] requires a system of several BSDEs while [73] requires a martingale representation result which is not true in our setting.

The present paper is organized as follows. In the next section the details about the noises considered and the integration framework are set into place. Section 3.3 is dedicated to the martingale representation type of result while Section 3.4 deals with existence and uniqueness of the solution of the BSDEs (3.1.1). The study of explicit solutions of linear BSDEs and their applications to prove a comparison theorem is given in Section 3.5. We show a sufficient maximum principle in Section 3.6 and a necessary maximum principle in Section 3.7. Finally use the maximum principles in some optimal control problems in Section 3.8. There we study expected utility of the final wealth, for which we find a characterization of the optimal portfolio, and a mean-variance portfolio selection problem for which we give an explicit formula of the optimal portfolio.

3.2 The framework

3.2.1 The random measures and their properties

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $X := [0, T] \times \mathbb{R}$, we will consider $X = ([0, T] \cup \{0\}) \cup ([0, T] \times \mathbb{R}_0)$, where $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$ and $T > 0$. Denote $\mathcal{B}_X$ the Borel σ -algebra on X . Throughout this presentation $\Delta \subset X$ denotes an element Δ in $\mathcal{B}_X$.

Let $\lambda := (\lambda^B, \lambda^H)$ be a two dimensional stochastic process such that each component λ^l , $l = B, H$, satisfies

- i) $\lambda_t^l \geq 0$ $\mathbb{P}$ -a.s. for all $t \in [0, T]$,
- ii) $\lim_{h \rightarrow 0} \mathbb{P}(|\lambda_{t+h}^l - \lambda_t^l| \geq \epsilon) = 0$ for all $\epsilon > 0$ and almost all $t \in [0, T]$,
- iii) $\mathbb{E} \left[\int_0^T \lambda_t^l dt \right] < \infty$,

We denote $\mathcal{L}$ as the space of all processes $\lambda := (\lambda^B, \lambda^H)$ satisfying i)-ii)-iii) above.

Define the random measure Λ on X by

$$\Lambda(\Delta) := \int_0^T \mathbf{1}_{\{(t,0) \in \Delta\}}(t) \lambda_t^B dt + \int_0^T \int_{\mathbb{R}_0} \mathbf{1}_\Delta(t, z) \nu(dz) \lambda_t^H dt, \quad (3.2.1)$$

as the mixture of measures on disjoint sets. Here ν is a deterministic, σ -finite measure on the Borel sets of $\mathbb{R}_0$ satisfying

$$\int_{\mathbb{R}_0} z^2 \nu(dz) < \infty.$$

We denote the σ -algebra generated by the values of Λ by $\mathcal{F}^\Lambda$. Furthermore, Λ^H denotes the restriction of Λ to $[0, T] \times \mathbb{R}_0$ and Λ^B the restriction of Λ to $[0, T] \times \{0\}$. Hence $\Lambda(\Delta) = \Lambda^B(\Delta \cap [0, T] \times \{0\}) + \Lambda^H(\Delta \cap [0, T] \times \mathbb{R}_0)$, $\Delta \subseteq X$. Here below we introduce the noises driving (3.1.1).

Definition 3.2.1. *B is a signed random measure on the Borel sets of $[0, T] \times \{0\}$ satisfying,*

$$A1) \mathbb{P}\left(B(\Delta) \leq x \mid \mathcal{F}^\Lambda\right) = \mathbb{P}\left(B(\Delta) \leq x \mid \Lambda^B(\Delta)\right) = \Phi\left(\frac{x}{\sqrt{\Lambda^B(\Delta)}}\right), \quad x \in \mathbb{R}, \Delta \subseteq [0, T] \times \{0\},$$

A2) $B(\Delta_1)$ and $B(\Delta_2)$ are conditionally independent given $\mathcal{F}^\Lambda$ whenever Δ_1 and Δ_2 are disjoint sets.

Here Φ stands for the cumulative probability distribution function of a standard normal random variable.

H is a random measure on the Borel sets of $[0, T] \times \mathbb{R}_0$ satisfying

$$A3) \mathbb{P}\left(H(\Delta) = k \mid \mathcal{F}^\Lambda\right) = \mathbb{P}\left(H(\Delta) = k \mid \Lambda^H(\Delta)\right) = \frac{\Lambda^H(\Delta)^k}{k!} e^{-\Lambda^H(\Delta)}, \quad k \in \mathbb{N}, \Delta \subseteq [0, T] \times \mathbb{R}_0,$$

A4) $H(\Delta_1)$ and $H(\Delta_2)$ are conditionally independent given $\mathcal{F}^\Lambda$ whenever Δ_1 and Δ_2 are disjoint sets.

Furthermore we assume that

A5) B and H are conditionally independent given $\mathcal{F}^\Lambda$.

Conditions A1) and A3) mean that conditional on Λ , B is a Gaussian random measure and H is Poisson a random measure. In particular, if λ^B and λ^H are deterministic then B is a Wiener process and H is a Poisson random random measure.

We refer to [51] or [65] for the existence of conditional distributions as in Definition 3.2.1.

Let $\tilde{H} := H - \Lambda^H$ be the signed random measure given by

$$\tilde{H}(\Delta) = H(\Delta) - \Lambda^H(\Delta), \quad \Delta \subset [0, T] \times \mathbb{R}_0.$$

Definition 3.2.2. *We define the signed random measure μ on the Borel subsets of X by*

$$\mu(\Delta) := B\left(\Delta \cap [0, T] \times \{0\}\right) + \tilde{H}\left(\Delta \cap [0, T] \times \mathbb{R}_0\right), \quad \Delta \subseteq X. \quad (3.2.2)$$

Clearly, from A1) we have that the conditional first moment of B is $\mathbb{E}[B(\Delta)|\mathcal{F}^\Lambda] = 0$ and from A3) the conditional first moment of H is $\mathbb{E}[H(\Delta)|\mathcal{F}^\Lambda] = \Lambda^H(\Delta)$ so that $\mathbb{E}[\tilde{H}(\Delta)|\mathcal{F}^\Lambda] = 0$. Thus

$$\mathbb{E}[\mu(\Delta)|\mathcal{F}^\Lambda] = 0. \quad (3.2.3)$$

The second conditional moments of B and $\tilde{H}$ are given by

$$\begin{aligned} \mathbb{E}[B(\Delta)^2|\mathcal{F}^\Lambda] &= \Lambda^B(\Delta), \\ \mathbb{E}[\tilde{H}(\Delta)^2|\mathcal{F}^\Lambda] &= \Lambda^H(\Delta). \end{aligned}$$

By the conditional independence A2), A4) and A5) we have

$$\mathbb{E}[\mu(\Delta)^2|\mathcal{F}^\Lambda] = \Lambda(\Delta)$$

and

$$\mathbb{E}[\mu(\Delta_1)\mu(\Delta_2)|\mathcal{F}^\Lambda] = \mathbb{E}[\mu(\Delta_1)|\mathcal{F}^\Lambda]\mathbb{E}[\mu(\Delta_2)|\mathcal{F}^\Lambda] = 0 \quad (3.2.4)$$

for Δ_1 and Δ_2 disjoint. Hence $\mu(\Delta_1)$ and $\mu(\Delta_2)$ are conditionally orthogonal.

The random measures B and H are related to a specific form of time-change for Brownian motion and pure jump Lévy process. More specifically define $B_t := B([0, t] \times \{0\})$, $\Lambda_t^B := \int_0^t \lambda_s^B ds$, $\eta_t := \int_0^t \int_{\mathbb{R}_0} z \tilde{H}(ds, dz)$ and $\hat{\Lambda}_t^H := \int_0^t \lambda_s^H ds$, for $t \in [0, T]$.

We can immediately see the role that the time-change processes Λ^B and $\hat{\Lambda}^H$ play, studying the characteristic function of B and η . In fact, from A1) and A3) we see that the conditional characteristic functions of B_t and η_t are given by

$$\mathbb{E}[e^{icB_t}|\mathcal{F}^\Lambda] = \exp \left\{ \int_0^t \frac{1}{2} c^2 \lambda_s^B ds \right\} = \exp \left\{ \frac{1}{2} c^2 \Lambda_t^B \right\}, \quad c \in \mathbb{R} \quad (3.2.5)$$

$$\begin{aligned} \mathbb{E}[e^{i\eta_t}|\mathcal{F}^\Lambda] &= \exp \left\{ \int_0^t \int_{\mathbb{R}_0} [e^{icz} - 1 - icz] \nu(dz) \lambda_t^H dt \right\} \\ &= \exp \left\{ \left(\int_{\mathbb{R}_0} [e^{icz} - 1 - icz] \nu(dz) \right) \hat{\Lambda}_t^H \right\}, \quad c \in \mathbb{R}. \end{aligned} \quad (3.2.6)$$

Indeed there is a strong connection between the distributions of B and the Brownian motion and between η and a centered pure jump Lévy process with the same jump behavior. The relationship is based on a random distortion of the time scale. The following characterization is due to [100, Theorem 3.1] (see also [51]).

Theorem 3.2.3. *Let W_t , $t \in [0, T]$ be a Brownian motion and N_t , $t \in [0, T]$ be a centered pure jump Lévy process with Levy measure ν . Assume that both W and N are independent of Λ . Then B satisfies A1)-(3.2.5) and A2) if and only if, for any $t \geq 0$,*

$$B_t \stackrel{d}{=} W_{\Lambda_t^B},$$

and η satisfies A3)-(3.2.6) and A4) if and only if, for any $t \geq 0$,

$$\eta_t \stackrel{d}{=} N_{\hat{\Lambda}_t^H}.$$

In addition, B is infinitely divisible if Λ^B is infinitely divisible and η is infinitely divisible if $\hat{\Lambda}^H$ is infinitely divisible, see [7, Theorem 7.1].

3.2.2 Stochastic non-anticipating integration

Let us define $\mathbb{F}^\mu = \{\mathcal{F}_t^\mu, t \in [0, T]\}$ as the filtration generated by $\mu(\Delta)$, $\Delta \subset [0, t] \times \mathbb{R}$. In view of (3.2.2), A1), and A3) we can see, that for any $t \in [0, T]$,

$$\mathcal{F}_t^\mu = \mathcal{F}_t^B \vee \mathcal{F}_t^H \vee \mathcal{F}_t^\Lambda,$$

where $\mathcal{F}_t^B$ is generated by $B(\Delta \cap [0, T] \times \{0\})$, $\mathcal{F}_t^H$ by $H(\Delta \cap [0, T] \times \mathbb{R}_0)$, and $\mathcal{F}_t^\Lambda$ by $\Lambda(\Delta)$, $\Delta \in [0, t] \times \mathbb{R}$. This is an application of [105, Theorem 1] and Theorem 2.2.8. Set $\mathbb{F} = \{\mathcal{F}_t, t \in [0, T]\}$ where

$$\mathcal{F}_t = \bigcap_{r>t} \mathcal{F}_r^\mu$$

Furthermore, we set $\mathbb{G} = \{\mathcal{G}_t, t \in [0, T]\}$ where $\mathcal{G}_t = \mathcal{F}_t^\mu \vee \mathcal{F}^\Lambda$. Remark that $\mathcal{G}_T = \mathcal{F}_T$, $\mathcal{G}_0 = \mathcal{F}^\Lambda$, while $\mathcal{F}_0^\mu$ is trivial. From now on we set $\mathcal{F} = \mathcal{F}_T$.

Lemma 3.2.4. *The filtration $\mathbb{G}$ is right-continuous.*

Proof. This can be shown using classical arguments from the Lévy case (as in e.g. [5, Theorem 2.1.9]). $\square$

For $\Delta \subset (t, T] \times \mathbb{R}$, the conditional independence A2), A4) means that

$$\mathbb{E}[\mu(\Delta) | \mathcal{G}_t] = \mathbb{E}[\mu(\Delta) | \mathcal{F}_t \vee \mathcal{F}^\Lambda] = \mathbb{E}[\mu(\Delta) | \mathcal{F}^\Lambda] = 0. \quad (3.2.7)$$

Thus μ has the martingale property with respect to $\mathbb{G}$ from (3.2.3). Hence μ is a martingale random field with respect to $\mathbb{G}$ in the sense of [38] since

- μ has a σ -finite variance measure

$$m(\Delta) := \mathbb{E}[\mu(\Delta)^2] = \mathbb{E}[\Lambda(\Delta)],$$

- μ is $\mathbb{G}$ -adapted,
- μ has conditionally orthogonal values, if $\Delta_1, \Delta_2 \subset (t, T] \times \mathbb{R}$ such that $\Delta_1 \cap \Delta_2 = \emptyset$ then, combining the arguments in (3.2.4) and (3.2.7),

$$\mathbb{E}[\mu(\Delta_1)\mu(\Delta_2) | \mathcal{G}_t] = \mathbb{E}[\mu(\Delta_1) | \mathcal{F}^\Lambda] \mathbb{E}[\mu(\Delta_2) | \mathcal{F}^\Lambda] = 0. \quad (3.2.8)$$

Denote $\mathcal{I}$ as the subspace of $L_2([0, T] \times \mathbb{R} \times \Omega, \mathcal{B}_X \times \mathbb{P}, \Lambda \times \mathbb{P})$ of the random fields admitting a $\mathbb{G}$ -predictable modification, in particular

$$\|\phi\|_{\mathcal{I}} := \left(\mathbb{E} \left[\int_0^T \phi_s(0)^2 \lambda_s^B ds + \int_0^T \int_{\mathbb{R}_0} \phi_s(z)^2 \nu(dz) \lambda_s^H ds \right] \right)^{\frac{1}{2}} < \infty. \quad (3.2.9)$$

For any $\phi \in \mathcal{I}$ we define the (Itô type) non-anticipative stochastic integral $I : \mathcal{I} \Rightarrow L_2(\Omega, \mathcal{F}, \mathbb{P})$ by

$$I(\phi) := \int_0^T \phi_s(0) dB_s + \int_0^T \int_{\mathbb{R}_0} \phi_s(z) \tilde{H}(ds, dz).$$

We refer to [38] for details on the integration with respect to martingale random fields of the type discussed here. In particular, I is a linear isometric operator:

$$\left(E[I(\phi)^2] \right)^{\frac{1}{2}} = \|I(\phi)\|_{L_2(\Omega, \mathcal{F}, \mathbb{P})} = \|\phi\|_{\mathcal{I}}. \quad (3.2.10)$$

Because of the structure of the filtration considered we have:

Lemma 3.2.5. *Consider $\xi \in L_2(\Omega, \mathcal{F}^\Lambda, \mathbb{P})$ and $\phi \in \mathcal{I}$. Then*

$$\xi I(\phi) = I(\xi \phi),$$

whenever either side of the equality exists as an element in $L_2(\Omega, \mathcal{F}, \mathbb{P})$.

Proof. Assume that ξ is bounded and $\phi \in \mathcal{I}$ is simple, i.e.

$$\phi_s(z, \omega) = \sum_{j=1}^J \phi_j(\omega) \mathbf{1}_{\Delta_j}(s, z),$$

where, for $j = 1, \dots, J$, we have $\Delta_j = (d_j, u_j] \times Z_j$, $0 \leq d_j \leq u_j$, $Z_j \subseteq \mathbb{R}$. Then

$$\xi I(\phi) = \xi \sum_{j=1}^J \phi_j \mu(\Delta_j) = \sum_{j=1}^J \xi \phi_j \mu(\Delta_j) = I(\xi \phi),$$

where $\xi \phi_j$ is $\mathcal{G}_{d_j}$ -measurable since ξ is $\mathcal{F}^\Lambda$ -measurable. The general case follows by taking limits. $\square$

Remark 3.2.6. The random field μ is also a martingale random field with respect to $\mathbb{F}$ and integration can be done as for $\mathbb{G}$. However, results such as Lemma 3.2.5 and the forthcoming representation would not hold. See also Remark 2.4.4.

3.3 Integral and martingale representation theorems

In this section we prove an integral representation theorem for a random variable $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ in the setting described above. We freshly prove this result here for the sake of

completeness. There are other similar results in the literature available. We refer for example to [58, Theorem III.4.34]. See Remark 3.3.4 for further details.

Recall that $\mathcal{G}_T = \mathcal{F}_T$. Here we remark that $\mathcal{F}_T = \sigma\{\mu(\Delta), \Delta \subseteq X\} = \sigma\{I(\phi), \phi \in \mathcal{I}\}$ (indeed $\mu(\Delta) = I(\mathbf{1}_\Delta)$). Denote $\mathcal{K} := \{\phi \in \mathcal{I} \mid \phi \text{ is } \mathcal{F}^\Lambda\text{-measurable, } \phi \mathbf{1}_{\mathbb{R}_0} \text{ is bounded a.e., and } \int_0^T \int_{\mathbb{R}} \phi_s(z)^2 \Lambda(ds, dz) \text{ is a bounded random variable}\}$.

Lemma 3.3.1. *For any $\phi \in \mathcal{K}$ we have*

$$\begin{aligned} \exp\{I(\phi)\} &\in L_2(\Omega, \mathcal{F}, \mathbb{P}), \quad \text{and} \\ \frac{\exp\{I(\phi)\}}{\mathbb{E}[\exp\{I(\phi)\} | \mathcal{F}^\Lambda]} &\in L_2(\Omega, \mathcal{F}, \mathbb{P}). \end{aligned}$$

Furthermore, the random variables $\{e^{I(\phi)}, \phi \in \mathcal{K}\}$ form a total subset of $L_2(\Omega, \mathcal{F}, \mathbb{P})$.

Proof. The first claim is proved in [107, Lemma 6], the second can be shown using arguments as in the proofs of [107, Lemmas 4 and 6]. The last claim is proved in [107, Lemma 9]. $\square$

Lemma 3.3.2. *Assume $\phi \in \mathcal{K}$. Define, for $t \in [0, T]$,*

$$\zeta_t = \exp \left\{ \int_0^t \phi_s(0) dB_s + \int_0^t \int_{\mathbb{R}_0} \phi_s(z) \tilde{H}(ds, dz) \right\}.$$

Then the following representation holds:

$$\begin{aligned} \zeta_T &= \mathbb{E}[\zeta_T | \mathcal{F}^\Lambda] + \int_0^T \left[\mathbb{E}\left[\frac{\zeta_T}{\zeta_s} \mid \mathcal{F}^\Lambda\right] \zeta_s \phi_s(0) \right] dB_s \\ &\quad + \int_0^T \int_{\mathbb{R}_0} \left[\mathbb{E}\left[\frac{\zeta_T}{\zeta_s} \mid \mathcal{F}^\Lambda\right] \zeta_s (e^{\phi_s(z)} - 1) \right] \tilde{H}(ds, dz). \end{aligned} \quad (3.3.1)$$

Note that the integrands in (3.3.1) are $\mathbb{G}$ -predictable.

Proof. Let

$$\begin{aligned} Y_t &= \frac{\zeta_t}{\mathbb{E}[\zeta_t | \mathcal{F}^\Lambda]} \\ &= \exp \left\{ \int_0^t \phi_s(0) dB_s + \int_0^t \int_{\mathbb{R}_0} \phi_s(z) \tilde{H}(ds, dz) - \int_0^t \frac{1}{2} \phi_s(0)^2 \lambda_s^B ds \right. \\ &\quad \left. - \int_0^t \int_{\mathbb{R}_0} [e^{\phi_s(z)} - 1 - \phi_s(z)] \nu(dz) \lambda_s^H ds \right\}. \end{aligned} \quad (3.3.2)$$

Note that both Y_t and ζ_t are elements of $L_2(\Omega, \mathcal{F}, \mathbb{P})$ by Lemma 3.3.1. By Itô's formula

$$dY_t = Y_t \left(\phi_t(0) dB_t + \int_{\mathbb{R}_0} (e^{\phi_t(z)} - 1) \tilde{H}(dt, dz) \right), \quad (3.3.3)$$

$$Y_0 = 1.$$

Combining (3.3.2) and (3.3.3) the above equalities yield

$$\begin{aligned} \zeta_T &= \mathbb{E}[\zeta_T | \mathcal{F}^\Lambda] Y_T \\ &= \mathbb{E}[\zeta_T | \mathcal{F}^\Lambda] \left(1 + \int_0^T Y_s \phi_s(0) dB_s + \int_0^T \int_{\mathbb{R}_0} Y_s (e^{\phi_s(z)} - 1) \tilde{H}(ds, dz) \right) \\ &= \mathbb{E}[\zeta_T | \mathcal{F}^\Lambda] + \int_0^T \mathbb{E}[\zeta_T | \mathcal{F}^\Lambda] Y_s \phi_s(0) dB_s \\ &\quad + \int_0^T \int_{\mathbb{R}_0} \left[\mathbb{E}[\zeta_T | \mathcal{F}^\Lambda] Y_s (e^{\phi_s(z)} - 1) \right] \tilde{H}(ds, dz) \\ &= \mathbb{E}[\zeta_T | \mathcal{F}^\Lambda] + \int_0^T \mathbb{E}\left[\frac{\zeta_T}{\zeta_s} | \mathcal{F}^\Lambda\right] \zeta_s \phi_s(0) dB_s \\ &\quad + \int_0^T \int_{\mathbb{R}_0} \left[\mathbb{E}\left[\frac{\zeta_T}{\zeta_s} | \mathcal{F}^\Lambda\right] \zeta_s (e^{\phi_s(z)} - 1) \right] \tilde{H}(ds, dz) \end{aligned}$$

where we used Lemma 3.2.5 and the equations

$$Y_s \mathbb{E}[\zeta_T | \mathcal{F}^\Lambda] = Y_s \mathbb{E}[\zeta_s | \mathcal{F}^\Lambda] \mathbb{E}\left[\frac{\zeta_T}{\zeta_s} | \mathcal{F}^\Lambda\right] = \zeta_s \mathbb{E}\left[\frac{\zeta_T}{\zeta_s} | \mathcal{F}^\Lambda\right].$$

□

Theorem 3.3.3. Assume $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$. Then there exists a unique $\phi \in \mathcal{I}$ such that

$$\xi = \mathbb{E}[\xi | \mathcal{F}^\Lambda] + \int_0^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz). \quad (3.3.4)$$

Note that the two summands in (3.3.4) are orthogonal. Here $\mathbb{E}[\xi | \mathcal{F}^\Lambda]$ represents the stochastic component of ξ which cannot be recovered by integration on μ .

Proof. At first let $\xi = \zeta(T)$, where

$$\zeta(T) = \exp \left\{ \int_0^T \int_{\mathbb{R}} \kappa_s(z) \mu(ds, dz) \right\}.$$

From Lemma 3.3.2 the representation (3.3.4) holds in this case.

Consider a general $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$. Then ξ can be approximated by a sequence of linear combinations of the form (3.3.4) by Lemma 3.3.1. Let $\{\xi_n\}_{n \geq 1}$ be such a sequence. Then, by (3.2.10), we have

$$\mathbb{E}[(\xi_n - \xi_m)^2] = \mathbb{E}\left[\left(\mathbb{E}[\xi_n - \xi_m | \mathcal{F}^\Lambda]\right)^2\right] + \int_0^T \int_{\mathbb{R}} (\phi_s^{(n)}(z) - \phi_s^{(m)}(z))^2 \Lambda(ds, dz).$$

Thus $\{\phi^{(n)}\}_{n \geq 1}$ is a Cauchy-sequence in $\mathcal{I}$, which proves existence. To prove uniqueness, suppose

$$\begin{aligned} \xi &= \mathbb{E}[\xi | \mathcal{F}^\Lambda] + \int_0^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz) \\ &= \mathbb{E}[\xi | \mathcal{F}^\Lambda] + \int_0^T \int_{\mathbb{R}} \psi_s(z) \mu(ds, dz). \end{aligned}$$

Then, using (3.2.10), $\mathbb{E}\left[\int_0^T \int_{\mathbb{R}} (\phi_s(z) - \psi_s(z))^2 \Lambda(ds, dz)\right] = 0$.

□

Remark 3.3.4. We have here chosen to prove the above result using classical arguments well established for integrators as the Brownian motion, see e.g. [85, Section 4] and the Poisson random measure, see e.g. [79]. The existence of such a representation is a topic of [58, Chapter 3]. There the result is obtained after a discussion on the solution of *the martingale problem* (see [58, Chapter 3]).

In Chapter 2 we have instead proven this result for $\tilde{H}$ using orthogonal polynomials and we have derived an *explicit formula* for the integrand ϕ using the non-anticipating derivative, see Theorem 2.5.1. This result holds for more general choices of Λ^H , but requires an assumption on the moments of $\tilde{H}$.

There are other related results in the literature. In [107, Proposition 41] the same representation is proved for a class of Malliavin differentiable random variables (à la Clark-Ocone type results).

If $\mathcal{F}_T^H$ -measurable ξ are considered then representation is given in the general context of (marked) point processes, see for instance [21, Theorem 4.12 and 8.8] or [31, 19, 55]. Our result differs in the choice of filtration, which also leads to slightly different integrals. In [21, 31, 19, 55] the integrator in the representation theorem are given by $H - \vartheta$ where ϑ is $\mathbb{F}^H$ -predictable compensator of H . Our Λ^H is not $\mathbb{F}^H$ -predictable.

Theorem 3.3.5. *Assume M_t , $t \in [0, T]$, is a $\mathbb{G}$ -martingale. Then there exists a unique $\phi \in \mathcal{I}$ such that*

$$M_t = \mathbb{E}[M_T | \mathcal{F}^\Lambda] + \int_0^t \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz), \quad t \in [0, T].$$

Proof. The proof follows classical arguments as in [85, Theorem 4.3.4] using Theorem 3.3.3. $\square$

3.4 BSDE: Existence and uniqueness of the solution

Hereafter we tackle directly the question of existence and uniqueness of the solution of (3.1.1):

$$Y_t = \xi + \int_t^T g_s(\lambda_s, Y_s, \phi_s) ds - \int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz), \quad t \in [0, T].$$

Indeed for the given terminal condition ξ and driver (or generator) g , a solution is given by the couple of $\mathbb{G}$ -adapted processes (Y, ϕ) on $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying the equation above. In the sequel we characterize explicitly the functional spaces in use and the elements of the BSDE to obtain a solution. In the following section we study explicitly the case when the driver g is linear.

Let S be the space of $\mathbb{G}$ -adapted stochastic processes $Y(t, \omega)$, $t \in [0, T]$, $\omega \in \Omega$, such that

$$\|Y\|_S := \sqrt{\mathbb{E} \left[\sup_{0 \leq t \leq T} |Y_t|^2 \right]} < \infty,$$

and $\mathcal{H}_2^{\mathcal{G}}$ be the space of $\mathbb{G}$ -predictable stochastic processes $f(t, \omega)$, $t \in [0, T]$, $\omega \in \Omega$, such that

$$\mathbb{E} \left[\int_0^T f_s^2 ds \right] < \infty.$$

Recall the definition of $\mathcal{I}$ in (3.2.9) and denote Φ the space of functions $\phi : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$|\phi(0)|^2 + \int_{\mathbb{R}_0} \phi(z)^2 \nu(dz) < \infty. \quad (3.4.1)$$

Definition 3.4.1. We say that (ξ, g) are standard parameters when $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ and $g : [0, T] \times [0, \infty)^2 \times \mathbb{R} \times \Phi \times \Omega \rightarrow \mathbb{R}$ such that g satisfies (for some $K_g > 0$)

$$g(\lambda, Y, \phi, \cdot) \text{ is } \mathbb{G}\text{-adapted for all } \lambda \in \mathcal{L}, Y \in S, \phi \in \mathcal{I}, \quad (3.4.2)$$

$$g(\lambda, 0, 0, \cdot) \in \mathcal{H}_2^{\mathcal{G}}, \text{ for all } \lambda \in \mathcal{L} \quad (3.4.3)$$

$$|g_t((\lambda^B, \lambda^H), y_1, \phi^{(1)}) - g_t((\lambda^B, \lambda^H), y_2, \phi^{(2)})| \leq K_g \left(|y_1 - y_2| \right. \quad (3.4.4)$$

$$\left. + |\phi^{(1)}(0) - \phi^{(2)}(0)| \sqrt{\lambda^B} + \sqrt{\int_{\mathbb{R}_0} |\phi^{(1)}(z) - \phi^{(2)}(z)|^2 \nu(dz) \sqrt{\lambda^H}} \right),$$

for all $(\lambda^B, \lambda^H) \in [0, \infty)^2$, $y_1, y_2 \in \mathbb{R}$, and $\phi^{(1)}, \phi^{(2)} \in \Phi$ $dt \times d\mathbb{P}$ a.e.

We recall the fundamental inequality $(a_1 + a_2 + \dots + a_n)^2 \leq n(a_1^2 + a_2^2 + \dots + a_n^2)$, for any $n \in \mathbb{N}$ and $a_1, a_2, \dots, a_n \in \mathbb{R}$, playing an important role in the technical lemmas below.

Lemma 3.4.2. Consider $(Y, \phi), (U, \psi) \in S \times \mathcal{I}$. Let $g : [0, T] \times [0, \infty)^2 \times \mathbb{R} \times \Phi \times \Omega \rightarrow \mathbb{R}$ satisfy (3.4.3) and (3.4.4). Then, for any $t \in [0, T]$, we have

$$\begin{aligned} \mathbb{E} \left[\left(\int_t^T g_s(\lambda_s, Y_s, \phi_s) - g_s(\lambda_s, U_s, \psi_s) ds \right)^2 \right] &\leq 3K_g^2(T-t) \\ \mathbb{E} \left[(T-t) \sup_{t \leq r \leq T} |Y_r - U_r|^2 + \int_t^T \int_{\mathbb{R}} |\phi_s(z) - \psi_s(z)|^2 \Lambda(ds, dz) \right] &\end{aligned} \quad (3.4.5)$$

and

$$\begin{aligned} \mathbb{E} \left[\left(\int_t^T |g_s(\lambda_s, U_s, \psi_s)| ds \right)^2 \right] &\leq (T-t) \mathbb{E} \left[2 \int_t^T |g_s(\lambda_s, 0, 0)|^2 ds \right. \\ &\quad \left. + 6K_g^2 \left((T-t) \sup_{t \leq r \leq T} |U_r|^2 + \int_t^T \int_{\mathbb{R}} |\psi_s(z)|^2 \Lambda(ds, dz) \right) \right]. \end{aligned} \quad (3.4.6)$$

Proof. Let $t \in [0, T]$. Inequality (3.4.5) follows from the Lipschitz conditions (3.4.4):

$$\begin{aligned} &\mathbb{E} \left[\left(\int_t^T g_s(\lambda_s, Y_s, \phi_s) - g_s(\lambda_s, U_s, \psi_s) ds \right)^2 \right] \\ &\leq K_g^2 \mathbb{E} \left[\left(\int_t^T |Y_s - U_s| + |\phi_s(0) - \psi_s(0)| \sqrt{\lambda_s^B} \right. \right. \\ &\quad \left. \left. + \sqrt{\int_{\mathbb{R}_0} |\phi_s(z) - \psi_s(z)|^2 \nu(dz) \sqrt{\lambda_s^H} ds} \right)^2 \right] \\ &\leq 3K_g^2(T-t) \mathbb{E} \left[\int_t^T |Y_s - U_s|^2 + |\phi_s(0) - \psi_s(0)|^2 \lambda_s^B \right. \\ &\quad \left. + \int_{\mathbb{R}_0} |\phi_s(z) - \psi_s(z)|^2 \nu(dz) \lambda_s^H ds \right] \\ &\leq 3K_g^2(T-t) \mathbb{E} \left[(T-t) \sup_{t \leq r \leq T} |Y_r - U_r|^2 \right. \\ &\quad \left. + \int_t^T \int_{\mathbb{R}} |\phi_s(z) - \psi_s(z)|^2 \Lambda(ds, dz) \right]. \end{aligned}$$

For inequality (3.4.6) we have

$$\begin{aligned}
& \mathbb{E} \left[\left(\int_t^T \left| g_s(\lambda_s, U_s, \psi_s) \right| ds \right)^2 \right] \\
& \leq (T-t) \mathbb{E} \left[\int_t^T \left| g_s(\lambda_s, U_s, \psi_s) \right|^2 ds \right] \\
& \leq (T-t) \mathbb{E} \left[\int_t^T \left(\left| g_s(\lambda_s, 0, 0) \right| + \left| g_s(\lambda_s, U_s, \psi_s) - g_s(\lambda_s, 0, 0) \right| \right)^2 ds \right] \\
& \leq 2(T-t) \mathbb{E} \left[\int_t^T \left| g_s(\lambda_s, 0, 0) \right|^2 + \left| g_s(\lambda_s, U_s, \psi_s) - g_s(\lambda_s, 0, 0) \right|^2 ds \right]
\end{aligned}$$

The result now follows from (3.4.4) by proceeding as in the proof of (3.4.5) above. $\square$

Lemma 3.4.3. Consider $U \in S$, $\psi, \phi \in \mathcal{I}$ and let (ξ, g) be standard parameters. Define a stochastic process Y_t , $t \in [0, T]$, by

$$Y_t = \xi + \int_t^T g_s(\lambda_s, U_s, \psi_s) ds - \int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz). \quad (3.4.7)$$

Then $Y \in S$. In particular we have

$$\begin{aligned}
\mathbb{E} \left[\sup_{t \leq r \leq T} |Y_r|^2 \right] & \leq \mathbb{E} \left[3\xi^2 + 3 \left(\int_t^T \left| g_s(\lambda_s, U_s, \psi_s) \right| ds \right)^2 \right. \\
& \quad \left. + 30 \int_t^T \int_{\mathbb{R}} |\phi_s(z)|^2 \Lambda(ds, dz) \right]. \quad (3.4.8)
\end{aligned}$$

Proof. Directly from (3.4.7), taking the square, we have

$$|Y_t|^2 \leq 3\xi^2 + 3 \left(\int_t^T \left| g_s(\lambda_s, U_s, \psi_s) \right| ds \right)^2 + 3 \left(\int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz) \right)^2.$$

In the next step we take the supremum and obtain

$$\begin{aligned}
\mathbb{E} \left[\sup_{t \leq r \leq T} |Y_r|^2 \right] & \leq \mathbb{E} \left[3\xi^2 + 3 \left(\int_t^T \left| g_s(\lambda_s, U_s, \psi_s) \right| ds \right)^2 \right] \\
& \quad + \mathbb{E} \left[\sup_{t \leq r \leq T} 3 \left(\int_r^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz) \right)^2 \right].
\end{aligned}$$

We have

$$\begin{aligned}
& \mathbb{E} \left[\sup_{t \leq r \leq T} \left(\int_r^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz) \right)^2 \right] \\
&= \mathbb{E} \left[\sup_{t \leq r \leq T} \left(\int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz) - \int_t^r \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz) \right)^2 \right] \\
&\leq \mathbb{E} \left[2 \left(\int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz) \right)^2 + 2 \sup_{t \leq r \leq T} \left(\int_t^r \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz) \right)^2 \right] \\
&\leq 10 \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} \phi_s(z)^2 \Lambda(ds, dz) \right]
\end{aligned}$$

by application of Doob's martingale inequality, see e.g. [5, Theorem 2.1.5]. Equation (3.4.8) follows, and we conclude that $Y \in S$ by (3.4.6). $\square$

Now let (g, ξ) be standard parameters. Define the mapping

$$\Theta : S \times \mathcal{I} \rightarrow S \times \mathcal{I}, \quad \Theta(U, \psi) := (Y, \phi) \quad (3.4.9)$$

as follows. The component ϕ is given by Theorem 3.3.5 as the unique element in $\mathcal{I}$ that provides the stochastic integral representation

$$M_t = M_0 + \int_0^t \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz), \quad t \in [0, T],$$

of the martingale

$$M_t = \mathbb{E} \left[\xi + \int_0^T g_s(\lambda_s, U_s, \psi_s) ds \mid \mathcal{G}_t \right], \quad t \in [0, T].$$

Note that $M_0 = \mathbb{E} \left[\xi + \int_0^T g_s(\lambda_s, U_s, \psi_s) ds \mid \mathcal{F}^\Lambda \right]$. The component Y in (3.4.9) is defined by

$$Y_t = \mathbb{E} \left[\xi + \int_t^T g_s(\lambda_s, U_s, \psi_s) ds \mid \mathcal{G}_t \right], \quad t \in [0, T]. \quad (3.4.10)$$

Note that

$$\begin{aligned} Y_t &= M_t - \int_0^t g_s(\lambda_s, U_s, \psi_s) ds \\ &= M_0 + \int_0^t \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz) - \int_0^t g_s(\lambda_s, U_s, \psi_s) ds. \end{aligned}$$

Since $Y_T = \xi$, we also have $Y_t = \xi - Y_T + Y_t$ so that

$$Y_t = \xi + \int_t^T g_s(\lambda_s, U_s, \psi_s) ds - \int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz). \quad (3.4.11)$$

Hence $Y \in S$ by Lemma 3.4.3 and the mapping (3.4.9) is well-defined.

We use the mapping Θ to prove that the BSDE of type (3.1.1) admits a unique solution for the given standard parameters (ξ, g) .

Lemma 3.4.4. *Let $(U^{(1)}, \psi^{(1)}), (U^{(2)}, \psi^{(2)}) \in S \times \mathcal{I}$ and define $(Y^{(1)}, \phi^{(1)}) = \Theta(U^{(1)}, \psi^{(1)})$ and $(Y^{(2)}, \phi^{(2)}) = \Theta(U^{(2)}, \psi^{(2)})$. Set $\bar{U} = U^{(1)} - U^{(2)}$, $\bar{\psi} = \psi^{(1)} - \psi^{(2)}$, $\bar{Y} = Y^{(1)} - Y^{(2)}$ and $\bar{\phi} = \phi^{(1)} - \phi^{(2)}$. There exists a $K > 0$ such that*

$$\begin{aligned} &\mathbb{E} \left[\sup_{t \leq r \leq T} |\bar{Y}_r|^2 + \int_t^T \int_{\mathbb{R}} |\bar{\phi}_s(z)|^2 \Lambda(ds, dz) \right] \\ &\leq K(T-t) \mathbb{E} \left[(T-t) \sup_{t \leq r \leq T} |\bar{U}_r|^2 + \int_t^T \int_{\mathbb{R}} |\bar{\psi}_s(z)|^2 \Lambda(ds, dz) \right], \quad t \in [0, T]. \end{aligned} \quad (3.4.12)$$

Proof. From (3.4.11), for any $t \in [0, T]$, we have

$$\begin{aligned} \bar{Y}_t &= \int_t^T g_s(\lambda_s, U_s^{(1)}, \psi_s^{(1)}) ds - \int_t^T g_s(\lambda_s, U_s^{(2)}, \psi_s^{(2)}) ds \\ &\quad - \int_t^T \int_{\mathbb{R}} \bar{\phi}_s(z) \mu(ds, dz). \end{aligned}$$

Since

$$E \left[\bar{Y}_t \int_t^T \int_{\mathbb{R}} \bar{\phi}_s(z) \mu(ds, dz) \right] = E \left[\bar{Y}_t \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} \bar{\phi}_s(z) \mu(ds, dz) | \mathcal{G}_t \right] \right] = 0,$$

we have

$$\begin{aligned} \mathbb{E} \left[\left(\bar{Y}_t + \int_t^T \int_{\mathbb{R}} \bar{\phi}_s(z) \mu(ds, dz) \right)^2 \right] &= \mathbb{E} \left[|\bar{Y}_t|^2 + \int_t^T \int_{\mathbb{R}} |\bar{\phi}_s(z)|^2 \Lambda(ds, dz) \right] \\ &= \mathbb{E} \left[\left(\int_t^T g_s(\lambda_s, U_s^{(1)}, \psi_s^{(1)}) ds - \int_t^T g_s(\lambda_s, U_s^{(2)}, \psi_s^{(2)}) ds \right)^2 \right]. \end{aligned} \quad (3.4.13)$$

We apply (3.4.5) and obtain

$$\begin{aligned} \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} |\bar{\phi}_s(z)|^2 \Lambda(ds, dz) \right] &\leq \mathbb{E} \left[|\bar{Y}_t|^2 + \int_t^T \int_{\mathbb{R}} |\bar{\phi}_s(z)|^2 \Lambda(ds, dz) \right] \\ &\leq 3K_g^2(T-t) \mathbb{E} \left[(T-t) \sup_{t \leq r \leq T} |\bar{U}_r|^2 + \int_t^T \int_{\mathbb{R}} |\bar{\psi}_s(z)|^2 \Lambda(ds, dz) \right]. \end{aligned} \quad (3.4.14)$$

By (3.4.5), (3.4.8) and (3.4.14) we have

$$\begin{aligned} \mathbb{E} \left[\sup_{t \leq r \leq T} |\bar{Y}_r|^2 \right] &\leq \mathbb{E} \left[0 + 3 \left(\int_t^T \left| g_s(\lambda_s, U_s^{(1)}, \psi_s^{(1)}) - g_s(\lambda_s, U_s^{(2)}, \psi_s^{(2)}) \right| ds \right)^2 \right. \\ &\quad \left. + 30 \left(\int_t^T \int_{\mathbb{R}} |\bar{\phi}_s(z)|^2 \Lambda(ds, dz) \right) \right] \\ &\leq (9+90) K_g^2(T-t)^2 \mathbb{E} \left[\sup_{t \leq r \leq T} |\bar{U}_r(z)|^2 \right] \\ &\quad + (9+90) K_g^2(T-t) \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} |\bar{\psi}_s(z)|^2 \Lambda(ds, dz) \right]. \end{aligned} \quad (3.4.15)$$

Combining (3.4.15) and (3.4.14) gives (3.4.12). $\square$

The existence and uniqueness for the BSDE now follow from the above estimates:

Theorem 3.4.5. *Let (g, ξ) be standard parameters. Then there exists a unique couple $(Y, \phi) \in S \times \mathcal{I}$ such that*

$$\begin{aligned} Y_t &= \xi + \int_t^T g_s(\lambda_s, Y_s, \phi_s) ds - \int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz) \\ &= \xi + \int_t^T g_s(\lambda_s, Y_s, \phi_s) ds - \int_t^T \phi_s(0) dB_s - \int_t^T \int_{\mathbb{R}_0} \phi_s(z) \tilde{H}(ds, dz). \end{aligned} \quad (3.4.16)$$

Proof. Let K be as in (3.4.12). Choose $t_1 \in [0, T)$ such that $\max \{K(T-t_1)^2, K(T-t_1)\} < 1$. Denote $S(u, v)$ as the space consisting of the elements of S equipped with the

norm $\|Y\|_{S(u,v)}^2 = \mathbb{E}[\sup_{u \leq r \leq v} |Y_r|^2]$ and $\mathcal{I}(u, v)$ as the space of the elements of $\mathcal{I}$ equipped with the norm $\|\phi\|_{\mathcal{I}(u,v)}^2 = \mathbb{E}[\int_u^v \int_{\mathbb{R}} |\phi_s(z)|^2 \Lambda(ds, dz)]$. From (3.4.12), Θ is a contraction on $S(t_1, T) \times \mathcal{I}(t_1, T)$, and thus there exists a unique $(Y^{(1)}, \phi^{(1)}) \in S(t_1, T) \times \mathcal{I}(t_1, T)$ such that $\Theta(Y^{(1)}, \phi^{(1)}) = (Y^{(1)}, \phi^{(1)})$ on $[t_1, T]$, i.e.

$$Y_t^{(1)} = \xi + \int_t^T g_s(\lambda_s, Y_s^{(1)}, \phi_s^{(1)}) ds - \int_t^T \int_{\mathbb{R}} \phi_s^{(1)}(z) \mu(ds, dz), \quad t \in [t_1, T]$$

Take $t_2 \in [0, t_1)$ so that $\max\{K(t_1 - t_2)^2, K(t_1 - t_2)\} < 1$. Next, $\tilde{\phi} \in \mathcal{I}(t_2, t_1)$ is given by Theorem 3.3.5, which is depending on $\tilde{U}$ and $\tilde{\psi}$, i.e.

$$\begin{aligned} E\left[Y_{t_1}^{(1)} + \int_0^{t_1} g_s(\lambda_s, \tilde{U}_s, \tilde{\psi}_s) ds \mid \mathcal{G}_t\right] &= E\left[Y_t^{(1)} + \int_0^{t_1} g_s(\lambda_s, \tilde{U}_s, \tilde{\psi}_s) ds \mid \mathcal{G}_{t_2}\right] \\ &\quad + \int_{t_2}^t \int_{\mathbb{R}} \tilde{\phi}_s(z) \mu(ds, dz), \quad t \in [t_2, t_1], \end{aligned}$$

In addition, $\tilde{Y}_t$ is defined as

$$\tilde{Y}_t = E\left[Y_{t_1}^{(1)} + \int_t^{t_1} g_s(\lambda_s, \tilde{U}_s, \tilde{\psi}_s) ds \mid \mathcal{G}_t\right], \quad t \in [t_2, t_1].$$

Then, $\tilde{\Theta}$ can be defined by $\tilde{\Theta}(\tilde{U}, \tilde{\psi}) = (\tilde{Y}, \tilde{\phi})$ for $(\tilde{U}, \tilde{\psi}) \in S(t_2, t_1) \times \mathcal{I}(t_2, t_1)$.

Following the same arguments as above we conclude that $\tilde{\Theta}$ is a contraction on $S(t_2, t_1) \times \mathcal{I}(t_2, t_1)$ so that there exists a unique element $(Y^{(2)}, \phi^{(2)}) \in S(t_2, t_1) \times \mathcal{I}(t_2, t_1)$ such that $(Y^{(2)}, \phi^{(2)}) = \tilde{\Theta}(Y^{(2)}, \phi^{(2)})$. Then we have

$$Y_t^{(2)} = Y_{t_1}^{(1)} + \int_t^{t_1} g_s(\lambda_s, Y_s^{(2)}, \phi_s^{(2)}) ds - \int_t^{t_1} \int_{\mathbb{R}} \phi_s^{(2)}(z) \mu(ds, dz), \quad t \in [t_2, t_1]. \quad (3.4.17)$$

Now consider

$$\begin{aligned} Y_t &= Y_t^{(1)} \mathbf{1}_{t_1 < t \leq T}(t) + Y_t^{(2)} \mathbf{1}_{t_2 < t \leq t_1}(t), \quad t \in [t_2, T], \\ \phi_t &= \phi_t^{(1)} \mathbf{1}_{t_1 < t \leq T}(t) + \phi_t^{(2)} \mathbf{1}_{t_2 < t \leq t_1}(t), \quad t \in [t_2, T]. \end{aligned} \quad (3.4.18)$$

We can see that

$$Y_t = \xi + \int_t^T g_s(\lambda_s, Y_s, \phi_s) ds - \int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz), \quad \text{for } t \in [t_2, T]. \quad (3.4.19)$$

In fact, clearly (3.4.19) holds for $t \in [t_1, T]$. Assume $t \in (t_2, t_1]$, then

$$\begin{aligned}
Y_t &= Y_{t_1}^{(1)} + \int_t^{t_1} g_s(\lambda_s, Y_s^{(2)}, \phi_s) ds - \int_t^{t_1} \int_{\mathbb{R}} \phi_s^{(2)}(z) \mu(ds, dz) \\
&= \xi + \int_{t_1}^T g_s(\lambda_s, Y_s^{(1)}, \phi_s) ds - \int_{t_1}^T \int_{\mathbb{R}} \phi_s^{(1)}(z) \mu(ds, dz) \\
&\quad + \int_t^{t_1} g_s(\lambda_s, Y_s^{(2)}, \phi_s) ds - \int_t^{t_1} \int_{\mathbb{R}} \phi_s^{(2)}(z) \mu(ds, dz) \\
&= \xi + \int_t^T g_s(\lambda_s, Y_s, \phi_s) ds - \int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz).
\end{aligned}$$

Proceed iteratively. Eventually, there is a step n such that $\max \{K(t_n - t_{n+1})^2, K(t_n - t_{n+1})\} < 1$ for $t_{n+1} = 0$ (here $t_0 = T$). Then we conclude and have found a (unique) couple $(Y, \phi) \in S(0, T) \times \mathcal{I}(0, T) = S \times \mathcal{I}$ such that (3.4.16) holds. $\square$

Remark 3.4.6. The initial point Y_0 of the solution Y is *not* necessarily a (deterministic) constant. From the definition of $\mathbb{G}$ and (3.4.10), we see that Y_0 is a square integrable $\mathcal{F}^\Lambda$ -measurable random variable. To be specific we have:

$$Y_0 = \mathbb{E} \left[\xi + \int_0^T g_s(\lambda_s, Y_s, \phi_s) ds \middle| \mathcal{F}^\Lambda \right].$$

3.5 Linear BSDEs and a comparison theorem

In the case of a linear driver the BSDE with Brownian motion or Lévy processes have an explicit representation. A similar representation holds in our case.

Theorem 3.5.1. *Assume we have the following BSDE:*

$$\begin{aligned}
-dY_t &= \left[A_t Y_t + C_t + E_t(0) \phi_t(0) \sqrt{\lambda_t^B} + \int_{\mathbb{R}_0} E_t(z) \phi_t(z) \nu(dz) \sqrt{\lambda_t^H} \right] dt \\
&\quad - \phi_t(0) dB_t - \int_{\mathbb{R}_0} \phi_t(z) \tilde{H}(dt, dz), \quad Y_T = \xi,
\end{aligned} \tag{3.5.1}$$

where the coefficients satisfy

- i) A is a bounded stochastic process, there exists $K_A > 0$ such that $|A_t| \leq K_A$ for all $t \in [0, T]$ $\mathbb{P}$ -a.s.,
- ii) $C \in \mathcal{H}_2^G$,
- iii) $E \in \mathcal{I}$,

iv) There exists $K_E > 0$ such that $0 \leq E_t(z) < K_E z$ for $z \in \mathbb{R}_0$, and $|E_t(0)| < K_E dt \times d\mathbb{P}$ -a.e.

Then (3.5.1) has a unique solution (Y, ϕ) in $S \times \mathcal{I}$ and Y has representation

$$Y_t = \mathbb{E} \left[\xi \Gamma_T(t) + \int_t^T \Gamma_s(t) C_s ds \middle| \mathcal{G}_t \right], \quad t \in [0, T],$$

where

$$\begin{aligned} \Gamma_s(t) = & \exp \left\{ \int_t^s A_u - \frac{1}{2} E_u(0)^2 \mathbf{1}_{\{\lambda_u^B \neq 0\}} du + \int_t^s E_u(0) \frac{\mathbf{1}_{\{\lambda_u^B \neq 0\}}}{\sqrt{\lambda_u^B}} dB_u \right. \\ & + \int_t^s \int_{\mathbb{R}_0} \left[\ln \left(1 + E_u(z) \frac{\mathbf{1}_{\{\lambda_u^H \neq 0\}}}{\sqrt{\lambda_u^H}} \right) - E_u(z) \frac{\mathbf{1}_{\{\lambda_u^H \neq 0\}}}{\sqrt{\lambda_u^H}} \right] \nu(dz) \lambda_u^H du \\ & \left. + \int_t^s \int_{\mathbb{R}_0} \ln \left(1 + E_u(z) \frac{\mathbf{1}_{\{\lambda_u^H \neq 0\}}}{\sqrt{\lambda_u^H}} \right) \tilde{H}(du, dz) \right\}. \end{aligned}$$

Note that $\Gamma_s(t) = \frac{\Gamma_s(0)}{\Gamma_t(0)}$.

Proof. The proof follows classical arguments, see e.g. [89, Theorem 6.2.2]. Condition (3.4.3) is guaranteed by ii). From Hölder's inequality

$$\begin{aligned} \int_{\mathbb{R}_0} |E_t(z) \phi_t(z)| \nu(dz) \sqrt{\lambda_t^H} & \leq \sqrt{\int_{\mathbb{R}_0} E_t^2(z) \nu(dz)} \sqrt{\int_{\mathbb{R}_0} \phi_t^2(z) \nu(dz)} \sqrt{\lambda_t^H} \\ & \leq K_E \sqrt{\int_{\mathbb{R}_0} z^2 \nu(dz)} \sqrt{\int_{\mathbb{R}_0} \phi_t^2(z) \nu(dz)} \sqrt{\lambda_t^H}, \end{aligned} \quad (3.5.2)$$

so from i) and iv) we obtain (3.4.4). It follows from Theorem 3.4.5 that (3.5.1) has a unique solution $(Y, \phi) \in S \times \mathcal{I}$.

Denote $\Gamma_t = \Gamma_t(0)$. We have $\Gamma_0 = 1$ and Itô's formula gives us

$$d\Gamma_t = \Gamma_t \left(A_t dt + E_t(0) \frac{\mathbf{1}_{\{\lambda_t^B \neq 0\}}}{\sqrt{\lambda_t^B}} dB_t + \int_{\mathbb{R}_0} E_t(z) \frac{\mathbf{1}_{\{\lambda_t^H \neq 0\}}}{\sqrt{\lambda_t^H}} \tilde{H}(dt, dz) \right). \quad (3.5.3)$$

Starting from (3.5.3),

$$\begin{aligned} \mathbb{E} \left[|\Gamma_t|^2 \right] & \leq 4\mathbb{E} \left[1 + \left(\int_0^t \Gamma_s A_s ds \right)^2 + \left(\int_0^t \Gamma_s E_s(0) \frac{\mathbf{1}_{\{\lambda_s^B \neq 0\}}}{\sqrt{\lambda_s^B}} dB_s \right)^2 \right. \\ & \quad \left. + \left(\int_0^t \int_{\mathbb{R}_0} \Gamma_s E_s(z) \frac{\mathbf{1}_{\{\lambda_s^H \neq 0\}}}{\sqrt{\lambda_s^H}} \tilde{H}(ds, dz) \right)^2 \right] \end{aligned}$$

$$\begin{aligned}
&\leq 4\mathbb{E}\left[1 + T \int_0^t |\Gamma_s|^2 A_s^2 ds + \int_0^t |\Gamma_s E_s(0)|^2 ds \right. \\
&\quad \left. + \int_0^t \int_{\mathbb{R}_0} |\Gamma_s|^2 K_E^2 z^2 \nu(dz) ds \right] \\
&\leq K_\Gamma \mathbb{E}\left[1 + \int_0^t |\Gamma_s|^2 ds\right]
\end{aligned}$$

for some $K_\Gamma > 0$, since A and $E(0)$ are bounded and z^2 is integrable with respect to ν . We conclude that $\Gamma_t \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ for all $t \in [0, T]$ by Gronwall's inequality. By Itô's formula we

have

$$\begin{aligned}
d(Y_t \Gamma_t) &= \Gamma_t \left(\left[-A_t Y_t - C_t - E_t(0) \phi_t(0) \sqrt{\lambda_t^B} \right] dt + \phi_t(0) dB_t \right. \\
&\quad \left. - \int_{\mathbb{R}_0} E_t(z) \phi_t(z) \nu(dz) \sqrt{\lambda_t^H} dt + \int_{\mathbb{R}_0} \phi_t(z) \tilde{H}(dt, dz) \right) \\
&\quad + Y_t \Gamma_t \left(A_t dt + E_t(0) \frac{\mathbf{1}_{\{\lambda_t^B \neq 0\}}}{\sqrt{\lambda_t^B}} dB_t + \int_{\mathbb{R}_0} E_t(z) \frac{\mathbf{1}_{\{\lambda_t^H \neq 0\}}}{\sqrt{\lambda_t^H}} \tilde{H}(dt, dz) \right) \\
&\quad + \Gamma_t \left(E_t(0) \phi_t(0) \frac{\mathbf{1}_{\{\lambda_t^B \neq 0\}}}{\sqrt{\lambda_t^B}} \lambda_t^B dt + \int_{\mathbb{R}_0} E_t(z) \phi_t(z) \frac{\mathbf{1}_{\{\lambda_t^H \neq 0\}}}{\sqrt{\lambda_t^H}} H(dt, dz) \right) \\
&= -\Gamma_t C_t dt + \left[\Gamma_t \phi_t(0) + Y_t E_t(0) \frac{\mathbf{1}_{\{\lambda_t^B \neq 0\}}}{\sqrt{\lambda_t^B}} \right] dB_t + \int_{\mathbb{R}_0} \left[\phi_t(z) \Gamma_t \right. \\
&\quad \left. + Y_t \Gamma_t E_t(z) \frac{\mathbf{1}_{\{\lambda_t^H \neq 0\}}}{\sqrt{\lambda_t^H}} + \Gamma_t \phi_t(z) E_t(z) \frac{\mathbf{1}_{\{\lambda_t^H \neq 0\}}}{\sqrt{\lambda_t^H}} \right] \tilde{H}(dt, dz). \tag{3.5.4}
\end{aligned}$$

Hence $Y_t \Gamma_t + \int_0^t \Gamma_s C_s ds$, $t \in [0, T]$, is a $\mathbb{G}$ -martingale so that

$$\begin{aligned}
Y_t \Gamma_t + \int_0^t \Gamma_s C_s ds &= \mathbb{E} \left[Y_T \Gamma_T + \int_0^T \Gamma_s C_s ds \middle| \mathcal{G}_t \right] \\
Y_t \Gamma_t &= \mathbb{E} \left[\xi \Gamma_T + \int_t^T \Gamma_s C_s ds \middle| \mathcal{G}_t \right] \\
Y_t &= \mathbb{E} \left[\xi \Gamma_T^t + \int_t^T \Gamma_s(t) C_s ds \middle| \mathcal{G}_t \right].
\end{aligned}$$

(Recall that $\Gamma_s(t) = \frac{\Gamma_s}{\Gamma_t}$).

□

Theorem 3.5.2. *Let $(g^{(1)}, \xi^{(1)})$ and $(g^{(2)}, \xi^{(2)})$ be two sets of standard parameters for the BSDEs*

with solutions $(Y^{(1)}, \phi^{(1)}), (Y^{(2)}, \phi^{(2)}) \in S \times \mathcal{I}$. Assume that

$$g_t^{(2)}(\lambda, y, \phi, \omega) = f_t\left(y, \phi(0)\kappa_t(0)\sqrt{\lambda^B}, \int_{\mathbb{R}_0} \phi(z)\kappa_t(z)\nu(dz)\sqrt{\lambda^H}, \omega\right)$$

where $\kappa \in \mathcal{I}$ satisfies condition iv) from Theorem 3.5.1 and f is a function $f : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \Omega \rightarrow \mathbb{R}$ which satisfies, for some $K_f > 0$,

$$|f_t(y, b, h) - f_t(y', b', h')| \leq K_f(|y - y'| + |b - b'| + |h - h'|), \quad (3.5.5)$$

$dt \times d\mathbb{P}$ a.e. and

$$\mathbb{E}\left[\int_0^T |f_t(0, 0, 0)|^2 dt\right] < \infty.$$

If $\xi^{(1)} \leq \xi^{(2)}$ $\mathbb{P}$ -a.s. and $g_s^{(1)}(\lambda_s, Y_s^{(1)}, \phi_s^{(1)}) \leq g_s^{(2)}(\lambda_s, Y_s^{(1)}, \phi_s^{(1)})$ $dt \times d\mathbb{P}$ -a.e., then

$$Y_t^{(1)} \leq Y_t^{(2)} \quad dt \times d\mathbb{P}\text{-a.e.}$$

It can be shown that $g^{(2)}$ does indeed satisfy conditions (3.4.2)-(3.4.3)-(3.4.4) in Definition 3.4.1, recall in particular (3.5.2).

Proof. Define $\bar{Y}_t := Y_t^{(2)} - Y_t^{(1)}$, $\bar{\phi}_t := \phi_t^{(2)} - \phi_t^{(1)}$,

$$\phi_t^{(2,H)}(z) := \phi_t^{(2)}\mathbf{1}_{\{z \neq 0\}} + \phi_t^{(1)}\mathbf{1}_{\{z=0\}},$$

$$\phi_t^{(2,B)}(z) := \phi_t^{(2)}\mathbf{1}_{\{z=0\}} + \phi_t^{(1)}\mathbf{1}_{\{z \neq 0\}},$$

and

$$C_t := g_t^{(2)}(\lambda_t, Y_t^{(1)}, \phi_t^{(1)}) - g_t^{(1)}(\lambda_t, Y_t^{(1)}, \phi_t^{(1)}),$$

$$A_t := \frac{g_t^{(2)}(\lambda_t, Y_t^{(2)}, \phi_t^{(1)}) - g_t^{(2)}(\lambda_t, Y_t^{(1)}, \phi_t^{(1)})}{\bar{Y}_t} \mathbf{1}_{\{\bar{Y}_t \neq 0\}},$$

$$D_t := \frac{g_t^{(2)}(\lambda_t, Y_t^{(2)}, \phi_t^{(2,H)}) - g_t^{(2)}(\lambda_t, Y_t^{(2)}, \phi_t^{(1)})}{\int_{\mathbb{R}_0} \kappa_t(z)\bar{\phi}_t(z)\nu(dz)\sqrt{\lambda_t^H}} \mathbf{1}_{\{\int_{\mathbb{R}_0} \kappa_t(z)\bar{\phi}_t(z)\nu(dz)\sqrt{\lambda_t^H} \neq 0\}},$$

$$F_t := \frac{g_t^{(2)}(\lambda_t, Y_t^{(2)}, \phi_t^{(2,B)}) - g_t^{(2)}(\lambda_t, Y_t^{(2)}, \phi_t^{(1)})}{\kappa_t(0)\bar{\phi}_t(0)\sqrt{\lambda_t^B}} \mathbf{1}_{\{\kappa_t(0)\bar{\phi}_t(0)\sqrt{\lambda_t^B} \neq 0\}}.$$

Then

$$\begin{aligned} -d\bar{Y}_t &= \left[A_t \bar{Y}_t + C_t + F_t \kappa_t(0) \bar{\phi}_t(0) \sqrt{\lambda_t^B} + D_t \int_{\mathbb{R}_0} \kappa_t(z) \bar{\phi}_t(z) \nu(dz) \sqrt{\lambda_t^H} \right] dt \\ &\quad - \bar{\phi}_t(0) dB_t - \int_{\mathbb{R}_0} \bar{\phi}_t(z) \tilde{H}(dt, dz), \\ \bar{Y}_T &= \xi_2 - \xi_1. \end{aligned} \quad (3.5.6)$$

The processes A , D and F are bounded due to the Lipschitz condition (3.5.5), and $C \in \mathcal{H}_2^G$ since it is a difference of functions in $\mathcal{H}_2^G$. It follows that $F\kappa(0) + D\kappa(z)\mathbf{1}_{\mathbb{R}_0}(z)$ satisfies iv) in Theorem 3.5.1.

Thus the assumptions of Theorem 3.5.1 are satisfied. The BSDE in (3.5.6) has solution

$$\bar{Y}_t = \mathbb{E} \left[\bar{\xi} \Gamma_T^t + \int_t^T \Gamma_s(t) C_s ds \middle| \mathcal{G}_t \right]$$

which is positive a.s. since $\bar{\xi}$, Γ and C are all positive a.s. □

3.6 Sufficient stochastic maximum principle

Here we show an application of the BSDE, proving sufficient conditions for an optimal control problem with both $\mathbb{G}$ and $\mathbb{F}$ -predictable controls. This problem cannot be solved with dynamic programming methods since the state process is, in general, not Markovian. We consider the optimization problem associated to the performance functional

$$J(u) = \mathbb{E} \left[\int_0^T f_t(\lambda_t, u_t, X_t) dt + l(X_T) \right], \quad (3.6.1)$$

where $l(x, \omega)$, $x \in \mathbb{R}$, $\omega \in \Omega$ is a stochastic function concave and differentiable in x a.s. and $f_t(\lambda, u, x, \omega)$, $t \in [0, T]$, $\lambda \in [0, \infty)^2$, $u \in \mathcal{U}$, $x \in \mathbb{R}$, $\omega \in \Omega$ is a stochastic function differentiable in x a.s. Here $\mathcal{U} \subseteq \mathbb{R}$ is a closed, convex set. The state process X_t , $t \in [0, T]$, has the form

$$dX_t = b_t(\lambda_t, u_t, X_t) dt + \int_{\mathbb{R}} \kappa_t(z, \lambda_t, u_t, X_t) \mu(dt, dz), \quad (3.6.2)$$

$$X_0 \in \mathbb{R},$$

where $b_t(\lambda, u, x, \omega)$ and $\kappa_t(z, \lambda, u, x, \omega)$, $t \in [0, T]$, $\lambda \in [0, \infty)^2$, $z \in \mathbb{R}$, $u \in \mathcal{U}$, $x \in \mathbb{R}$, $\omega \in \Omega$ are $\mathbb{F}$ -adapted stochastic processes differentiable in x a.s. We denote these derivatives $\frac{\partial b}{\partial x}$ and $\frac{\partial \kappa}{\partial x}$ respectively. The stochastic process u_t , $t \in [0, T]$, is the control. We have the following definition

Definition 3.6.1. *The admissible controls are càglàd stochastic processes $u : [0, T] \times \Omega \rightarrow \mathcal{U}$, such that X (3.6.2) has a unique strong solution,*

$$\mathbb{E} \left[\int_0^T |f_t(\lambda_t, u_t, X_t)|^2 dt + |l(X_T)| + |l'(X_T)|^2 \right] < \infty, \quad (3.6.3)$$

where $l'(x) = \frac{d}{dx} l(x)$. The admissible controls are either $\mathbb{G}$ -predictable or $\mathbb{F}$ -predictable and we denote these sets as $\mathcal{A}^G$ and $\mathcal{A}^F$ respectively. The couple (u, X) is called an admissible pair.

Naturally $\mathcal{A}^{\mathcal{F}} \subset \mathcal{A}^{\mathcal{G}}$. Remark also that X is $\mathbb{G}$ -adapted if $u \in \mathcal{A}^{\mathcal{G}}$ and X is $\mathbb{F}$ -adapted if $u \in \mathcal{A}^{\mathcal{F}}$. Given the performance functional J (3.6.1) we aim to find an optimal control depending on the information available:

$$\sup_{u \in \mathcal{A}^{\mathcal{G}}} J(u) \quad (3.6.4)$$

$$\sup_{u \in \mathcal{A}^{\mathcal{F}}} J(u). \quad (3.6.5)$$

For a detailed discussion on the existence of a strong solution to (3.6.2) we refer to [57, 56]. However the following conditions are sufficient [57]: for u a càglàd stochastic process there exists a $K_2 > 0$ such that

$$|\kappa_t(0, \lambda_t, u_t, x) - \kappa_t(0, \lambda_t, u_t, x')| \leq K_2 |x - x'| \quad \mathbb{P}\text{-a.s.}, \quad (3.6.6)$$

$$|\kappa_t(z, \lambda_t, u_t, x) - \kappa_t(z, \lambda_t, u_t, x')| \leq K_2 |z| |x - x'| \quad \text{for } z \neq 0 \quad \mathbb{P}\text{-a.s.}, \quad (3.6.7)$$

$$|b_t(\lambda_t, u_t, x) - b_t(\lambda_t, u_t, x')| \leq K_2 |x - x'| \quad \mathbb{P}\text{-a.s.}, \quad (3.6.8)$$

$$\int_0^T \int_{\mathbb{R}} |\kappa_s(z, \lambda_s, u_s, a)|^2 \Lambda(ds, dz) < \infty \quad \mathbb{P}\text{-a.s.}, \quad (3.6.9)$$

$$\int_0^T |b_s(\lambda_s, u_s, a)| ds \leq \infty \quad \mathbb{P}\text{-a.s.}, \quad (3.6.10)$$

for some $a \in \mathbb{R}$, all $t \in [0, T]$ and all $x, x' \in \mathbb{R}$.

In addition we assume, for computation purposes, the following:

Assumption 3.6.2. There exist a $K_1 > 0$ such that all $u \in \mathcal{A}^{\mathcal{G}}$

$$\left| \frac{\partial \kappa_t}{\partial x}(0, \lambda_t, u_t, X_t) \right| \sqrt{\lambda_t^B} \leq K_1 \quad dt \times d\mathbb{P}\text{-a.e.}, \quad (3.6.11)$$

$$\int_{\mathbb{R}_0} \left(\frac{\partial \kappa_t}{\partial x}(z, \lambda_t, u_t, X_t) \right)^2 \nu(dz) \sqrt{\lambda_t^H} \leq K_1 \quad dt \times d\mathbb{P}\text{-a.e.}, \quad (3.6.12)$$

$$\left| \frac{\partial b_t}{\partial x}(\lambda_t, u_t, X_t) \right| \leq K_1 \quad dt \times d\mathbb{P}\text{-a.e.} \quad (3.6.13)$$

We define the Hamiltonian, $\mathcal{H} : [0, T] \times [0, \infty)^2 \times \mathcal{U} \times \mathbb{R} \times \mathbb{R} \times \Phi \times \Omega \rightarrow \mathbb{R}$ (where Φ is defined in (3.4.1)), by

$$\begin{aligned} \mathcal{H}_t(\lambda, u, x, y, \phi) &= f_t(\lambda, u, x) + b_t(\lambda, u, x)y + \kappa_t(0, \lambda, u, x)\phi(0)\lambda^B \\ &\quad + \int_{\mathbb{R}_0} \kappa_t(z, \lambda, u, x)\phi(z) \lambda^H \nu(dz). \end{aligned} \quad (3.6.14)$$

Corresponding to the admissible pair (u, X) is the couple (Y, ϕ) , which is the solution to the

BSDE of type (3.1.1)

$$\begin{aligned} Y_t &= l'(X_T) + \int_t^T \frac{\partial \mathcal{H}_s}{\partial x}(\lambda, u_s, X_s, Y_s, \phi_s) ds - \int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz), \\ Y_T &= l'(X_T). \end{aligned} \quad (3.6.15)$$

Here $\frac{\partial \mathcal{H}_t}{\partial x} = \frac{\partial}{\partial x} \mathcal{H}_t(\lambda, u, x, y, \phi)$ and we note that $\mathcal{H}$ is differentiable in x by the assumptions on f , g and κ . The above conditions, (3.6.3)-(3.6.11)-(3.6.12)-(3.6.13), ensure that the pair $(\frac{\partial \mathcal{H}}{\partial x}, l'(X_T))$ are standard parameters (Definition 3.4.1). By Theorem 3.4.5 the BSDE (3.6.15) has a unique solution (Y, ϕ) .

In the sequel we set $\hat{b}_s = b_s(\lambda_s, \hat{u}_s, \hat{X}_s)$, etc. for the coefficients associated with the admissible pair $(\hat{u}, \hat{X})$ with solution $(\hat{Y}, \hat{\phi})$ of the adjoint equation (3.6.15). Set $b_s = b_s(\lambda_s, u_s, X_s)$ and so forth for the coefficients associated to another arbitrary, admissible pair (u, X) . In addition $\hat{\mathcal{H}}_s(u, x) = \mathcal{H}_s(\lambda_s, u, x, \hat{Y}_s, \hat{\phi}_s)$.

Theorem 3.6.3. *Let $\hat{u} \in \mathcal{A}^{\mathcal{G}}$. Assume that*

$$\mathbb{E} \left[\int_0^T \int_{\mathbb{R}} |\hat{Y}_s(\hat{\kappa}_s(z) - \kappa_s(z))|^2 + |(\hat{X}_s - X_s)\hat{\phi}_s(z)|^2 \Lambda(ds, dz) \right] < \infty \quad (3.6.16)$$

for all $u \in \mathcal{A}^{\mathcal{G}}$ and that Assumption 3.6.2 holds. If

$$h_t(x) = \max_{u \in \mathcal{U}} \mathcal{H}_t(\lambda_t, u, x, \hat{Y}_t, \hat{\phi}_t) \quad (3.6.17)$$

exists and is a concave function in x for all $t \in [0, T]$ $\mathbb{P}$ -a.s., and

$$\mathcal{H}_t(\lambda_t, \hat{u}_t, \hat{X}_t, \hat{Y}_t, \hat{\phi}_t) = h_t(\hat{X}_t) \quad (3.6.18)$$

for all $t \in [0, T]$, then $\hat{u}$ is optimal for (3.6.4) and $(\hat{u}, \hat{X})$ is an optimal pair.

Proof. We proceed as in [48]. Recall that for l concave and differentiable we have $l(x_2) - l(x_1) \geq l'(x_2)(x_2 - x_1)$, $x_1, x_2 \in \mathbb{R}$. Thus, by Itô's formula, (3.6.16) and the fact that $\hat{X}_0 - X_0 = 0$, we have

$$\begin{aligned} & \mathbb{E} \left[l(\hat{X}_T) - l(X_T) \right] \\ & \geq \mathbb{E} \left[l'(\hat{X}_T)(\hat{X}_T - X_T) \right] \\ & = \mathbb{E} \left[\hat{Y}_T(\hat{X}_T - X_T) \right] \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E} \left[\int_0^T -(\hat{X}_s - X_s) \frac{\partial \hat{\mathcal{H}}_s}{\partial x}(\hat{u}_s, \hat{X}_s) + \hat{Y}_s(\hat{b}_s - b_s) ds \right. \\
&\quad + \int_0^T \int_{\mathbb{R}} \left\{ \hat{Y}_s(\hat{\kappa}_s(z) - \kappa_s(z)) + (\hat{X}_s - X_s) \hat{\phi}_s(z) \right\} \mu(ds, dz) \\
&\quad + \int_0^T (\hat{\kappa}_s(0) - \kappa_s(0)) \hat{\phi}_s(0) \lambda_s^B ds + \int_0^T \int_{\mathbb{R}_0} (\hat{\kappa}_s(z) - \kappa_s(z)) \hat{\phi}_s(z) H(ds, dz) \Big] \\
&= \mathbb{E} \left[\int_0^T -(\hat{X}_s - X_s) \frac{\partial \hat{\mathcal{H}}_s}{\partial x}(\hat{u}_s, \hat{X}_s) + \hat{Y}_s(\hat{b}_s - b_s) ds \right. \\
&\quad + \left. \int_0^T \int_{\mathbb{R}} (\hat{\kappa}_s(z) - \kappa_s(z)) \hat{\phi}_s(z) \Lambda(ds, dz) \right].
\end{aligned}$$

We remark that ϕ is integrable with respect to $H \times \mathbb{P}$ by (3.6.7), Cauchy's inequality and (3.6.16). Furthermore, from the Hamiltonian, we have

$$\begin{aligned}
&\mathbb{E} \left[\int_0^T \{ \hat{f}_s - f_s \} ds \right] \\
&= \mathbb{E} \left[\int_0^T \left\{ \hat{\mathcal{H}}_s(\hat{u}_s, \hat{X}_s) - \hat{\mathcal{H}}_s(u_s, X_s) - (\hat{b}_s - b_s) \hat{Y}_s \right. \right. \\
&\quad \left. \left. - (\hat{\kappa}_s(0) - \kappa_s(0)) \hat{\phi}_s(0) \lambda_s^B - \int_{\mathbb{R}_0} (\hat{\kappa}_s(z) - \kappa_s(z)) \hat{\phi}_s(z) \nu(dz) \lambda_s^H \right\} ds \right] \\
&= \mathbb{E} \left[\int_0^T \left\{ \hat{\mathcal{H}}_s(\hat{u}_s, \hat{X}_s) - \hat{\mathcal{H}}_s(u_s, X_s) - (\hat{b}_s - b_s) \hat{Y}_s \right\} ds \right. \\
&\quad \left. - \int_0^T \int_{\mathbb{R}} (\hat{\kappa}_s(z) - \kappa_s(z)) \hat{\phi}_s(z) \Lambda(ds, dz) \right].
\end{aligned}$$

Hence

$$\begin{aligned}
J(\hat{u}) - J(u) &\geq \mathbb{E} \left[\int_0^T \left\{ \hat{\mathcal{H}}_s(\hat{u}_s, \hat{X}_s) \right. \right. \\
&\quad \left. \left. - \hat{\mathcal{H}}_s(u_s, X_s) - (\hat{X}_s - X_s) \frac{\partial \hat{\mathcal{H}}_s}{\partial x}(\hat{u}_s, \hat{X}_s) \right\} ds \right]. \tag{3.6.19}
\end{aligned}$$

The integrand in (3.6.19) is non-negative $dt \times d\mathbb{P}$ -a.e. by the maximality of $\hat{u}$ (3.6.18) and the concavity of h_t , see [99, page 108]. Hence $\hat{u}$ is also an optimal control by inequality (3.6.19). We sketch the last part of the argument for completeness. From (3.6.17) and (3.6.18) we have

$h_t(\hat{X}_t) = \hat{\mathcal{H}}_t(\hat{u}_t, \hat{X}_t)$. Thus

$$\hat{\mathcal{H}}_t(u, x) - \hat{\mathcal{H}}_t(\hat{u}_t, \hat{X}_t) \leq h_t(x) - h_t(\hat{X}_t), \quad \text{for all } (t, u, x). \quad (3.6.20)$$

To prove that the integrand in (3.6.19) is non-negative it is sufficient to show that almost surely

$$h_t(X_t) - h_t(\hat{X}_t) - \frac{\partial \hat{\mathcal{H}}_t}{\partial x}(\hat{u}_t, \hat{X}_t)(X_t - \hat{X}_t) \leq 0. \quad (3.6.21)$$

Fix $t \in [0, T]$. Since $x \rightarrow h_t(x)$ is concave, it follows by a separating hyperplane argument that there exists $a \in \mathbb{R}$ such that

$$h_t(x) - h_t(\hat{X}_t) - a(x - \hat{X}_t) \leq 0, \quad \text{for all } x. \quad (3.6.22)$$

Define

$$\rho(x) := \hat{\mathcal{H}}_t(u_t, x) - \hat{\mathcal{H}}_t(\hat{u}_t, \hat{X}_t) - a(x - \hat{X}_t).$$

By (3.6.20) and (3.6.22) $\rho(x) \leq 0$ for all x . Clearly $\rho(\hat{X}_t) = 0$. Hence $\rho'(\hat{X}_t) = 0$ so that $\frac{\partial \hat{\mathcal{H}}_t}{\partial x}(\hat{X}_t, \hat{u}_t) = a$. Substituting into (3.6.22) we obtain (3.6.21). $\square$

Recall that the solution of the BSDE (3.6.15) is $\mathbb{G}$ -adapted. However, the other coefficients in (3.6.18) are $\mathbb{F}$ -adapted whenever $u \in \mathcal{A}^{\mathcal{F}}$. We use this fact to find an optimal $\mathbb{F}$ -predictable control via projections. We keep the notation used in the proof of Theorem 3.6.3.

Theorem 3.6.4. *Let $\hat{u} \in \mathcal{A}^{\mathcal{F}}$. Denote the corresponding state process as $\hat{X}$ with solution $(\hat{Y}, \hat{\phi})$ of the adjoint equation (3.6.15). Assume (3.6.16) and Assumption 3.6.2 holds. Denote*

$$\begin{aligned} \mathcal{H}_t^{\mathbb{F}}(\lambda_t, u, x, \hat{Y}_t, \hat{\phi}_t) &:= \mathbb{E} \left[\mathcal{H}_t(\lambda_t, u, x, \hat{Y}_t, \hat{\phi}_t) \mid \mathcal{F}_t \right] \\ &= f_t(\lambda_t, u, x) + b_t(\lambda_t, u, x) \mathbb{E}[\hat{Y}_t \mid \mathcal{F}_t] + \kappa_t(0, \lambda_t, u, x) \mathbb{E}[\hat{\phi}_t(0) \mid \mathcal{F}_t] \\ &\quad + \int_{\mathbb{R}_0} \kappa_t(z, \lambda_t, u, x) \mathbb{E}[\phi_t(z) \mid \mathcal{F}_t] \lambda_t^H \nu(dz) \end{aligned}$$

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

$$h_t^{\mathbb{F}}(x) = \max_{u \in \mathcal{U}} \mathcal{H}_t^{\mathbb{F}}(\lambda_t, u, x, \hat{Y}_t, \hat{\phi}_t) \quad (3.6.23)$$

exists and is a concave function in x for all $t \in [0, T]$, and

$$\mathcal{H}_t^{\mathbb{F}}(\lambda_t, \hat{u}_t, \hat{X}_t, \hat{Y}_t, \hat{\phi}_t) = h_t^{\mathbb{F}}(\hat{X}_t), \quad (3.6.24)$$

then $(\hat{u}, \hat{X})$ is an optimal pair for (3.6.5).

Proof. The arguments in the proof of Theorem 3.6.3 leading to

$$J(\hat{u}) - J(u) \geq \mathbb{E} \left[\int_0^T \hat{\mathcal{H}}_s(\hat{u}_s, \hat{X}_s) - \hat{\mathcal{H}}_s(u_s, X_s) - (\hat{X}_s - X_s) \frac{\partial \hat{\mathcal{H}}_s}{\partial x}(\hat{u}_s, \hat{X}_s) ds \right] \quad (3.6.25)$$

still hold. Since $\hat{u}$ and u are $\mathbb{F}$ -predictable controls the only coefficients in the integrand in

(3.6.25) that are not $\mathbb{F}$ -adapted are the solution of the adjoint equation $(\hat{Y}, \hat{\phi})$ so that

$$\begin{aligned}
& \mathbb{E} \left[\int_0^T \hat{\mathcal{H}}_s(\hat{u}_s, \hat{X}_s) - \hat{\mathcal{H}}_s(u_s, X_s) - \frac{\partial \hat{\mathcal{H}}_s}{\partial x}(\hat{u}_s, \hat{X}_s)(\hat{X}_s - X_s) ds \right] \\
&= \mathbb{E} \left[\int_0^T \hat{f}_s - f_s + (\hat{b}_s - b_s) \mathbb{E}[\hat{Y}_s | \mathcal{F}_s] + (\hat{X}_s - X_s) \frac{\partial f_s}{\partial x} + \frac{\partial b_s}{\partial x} \mathbb{E}[\hat{Y}_s | \mathcal{F}_s] ds \right] \\
&\quad + \mathbb{E} \left[\int_0^T \int_{\mathbb{R}} \left\{ (\hat{\kappa}_s(z) - \kappa_s(z)) \mathbb{E}[\hat{\phi}_s(z) | \mathcal{F}_s] \right. \right. \\
&\quad \left. \left. + (\hat{X}_s - X_s) \frac{\partial \hat{\kappa}_s}{\partial x}(z) \mathbb{E}[\hat{\phi}_s(z) | \mathcal{F}_s] \right\} \Lambda(ds, dz) \right] \\
&= \mathbb{E} \left[\int_0^T \hat{\mathcal{H}}_s^{\mathbb{F}}(\hat{u}_s, \hat{X}_s) - \hat{\mathcal{H}}_s^{\mathbb{F}}(u_s, X_s) - \frac{\partial \hat{\mathcal{H}}_s^{\mathbb{F}}}{\partial x}(\hat{u}_s, \hat{X}_s)(\hat{X}_s - X_s) ds \right]. \tag{3.6.26}
\end{aligned}$$

The integrand in (3.6.26) is non-negative $dt \times d\mathbb{P}$ -a.e. by the maximality of $\hat{u}$ (3.6.24) and the concavity of $h_t^{\mathbb{F}}$ (3.6.23). The argument is the same as in the proof of Theorem 3.6.3. $\square$

3.7 Necessary maximum principle

Here we show necessary conditions for optimal control for the performance functional (3.6.1), i.e.

$$J(u) = \mathbb{E} \left[\int_0^T f_t(\lambda_t, u_t, X_t) dt + l(X_T) \right].$$

Conditions on the coefficients are as in Section 3.6 unless noted otherwise here. We assume that $f_t(\lambda, u, x, \omega)$ is also differentiable in u a.s. and no longer assume that l is concave. We set $\mathcal{U} \subseteq \mathbb{R}$ to be an open, convex set.

The state process X_t , $t \in [0, T]$ is given by (3.6.2) but with additional assumptions on differentiability, namely that $b_t(\lambda, u, x, \omega)$ and $\kappa_t(z, \lambda, u, x, \omega)$, are also differentiable in u a.s. Furthermore we assume $\frac{\partial \kappa}{\partial u} \in \mathcal{I}$. The set of admissible controls is as in Definition 3.6.3.

To emphasize that the state process X depends on the control u we will sometimes write X^u .

Assumption 3.7.1. The following conditions are assumed. For all $u \in \mathcal{A}^{\mathcal{F}}$ if β is a bounded $\mathbb{F}$ -predictable process such that

$$u_t \pm \beta_t \in \mathcal{U} \quad dt \times d\mathbb{P}\text{-a.e.} \tag{3.7.1}$$

then

- i) $u + y\beta \in \mathcal{A}^{\mathcal{F}}$, for all $|y| \leq 1$.

ii) The family

$$\left\{ \frac{\partial f_t}{\partial x}(\lambda_t, u_t + y\beta_t, X_t^{u+y\beta}) \frac{\partial}{\partial y} X_t^{u+y\beta} + \frac{\partial f_t}{\partial u}(\lambda_t, u_t + y\beta_t, X_t^{u+y\beta}) \beta_t \right\}_{y \in (-1,1)} \quad (3.7.2)$$

is uniformly $dt \times d\mathbb{P}$ -integrable, and the family

$$\left\{ l'(X_T^{u+y\beta}) \frac{\partial}{\partial y} X_T^{u+y\beta} \right\}_{y \in (-1,1)} \quad (3.7.3)$$

is uniformly $\mathbb{P}$ -integrable.

iii) The process $\zeta_t^{(u,\beta)} = \frac{\partial}{\partial y} X_t^{u+y\beta} \big|_{y=0}$ exists as an element of $L_2(\Omega, \mathcal{F}, \mathbb{P})$ for all $t \in [0, T]$ and satisfies

$$\begin{aligned} \zeta_t &:= \zeta_t^{(u,\beta)} = \frac{\partial}{\partial y} X_t^{u+y\beta} \big|_{y=0} \\ &= \int_0^t \left[\frac{\partial b_s}{\partial x}(\lambda_s, u_s, X_s) \zeta_s + \frac{\partial b_s}{\partial u}(\lambda_s, u_s, X_s) \beta_s \right] ds \\ &\quad + \int_0^t \int_{\mathbb{R}_0} \left[\frac{\partial \kappa_s}{\partial x}(z, \lambda_s, u_s, X_s) \zeta_s + \frac{\partial \kappa_s}{\partial u}(z, \lambda_s, u_s, X_s) \beta_s \right] \mu(ds, dz). \end{aligned} \quad (3.7.4)$$

The Hamiltonian $\mathcal{H}$ is as in (3.6.14) and the corresponding BSDE (Y, ϕ) is given by (3.6.15), both rewritten here below for convenience:

$$\begin{aligned} \mathcal{H}_t(\lambda, u, x, y, \phi) &= f_t(\lambda, u, x) + b_t(\lambda, u, x)y + \kappa_t(0, \lambda, u, x)\phi(0)\lambda^B \\ &\quad + \int_{\mathbb{R}_0} \kappa_t(z, \lambda, u, x)\phi(z)\lambda^H \nu(dz), \\ Y_t &= l'(X_T) + \int_t^T \frac{\partial \mathcal{H}_s}{\partial x}(\lambda, u_s, X_s, Y_s, \phi_s) ds - \int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz). \end{aligned}$$

A control $\hat{u} \in \mathcal{A}^{\mathcal{F}}$ is a “local maximum” if

$$J(\hat{u}) \geq J(\hat{u} + y\beta), \quad |y| < 1,$$

for all bounded $\mathbb{F}$ -predictable β such that $u_t \pm \beta_t \in \mathcal{U}$ $dt \times d\mathbb{P}$ -a.e. Meaning that we cannot improve $J(\hat{u})$ by making “bounded” perturbations of $\hat{u}$. Any solution to (3.6.5), i.e. $J(\hat{u}) = \sup_{u \in \mathcal{A}^{\mathcal{F}}} J(u)$, must also be a local maximum. If $\hat{u}$ is a local maximum, we must naturally have

$$\frac{\partial}{\partial y} J(\hat{u} + y\beta) \big|_{y=0} = 0. \quad (3.7.5)$$

The converse conclusions are not true however. Not every u satisfying (3.7.5) is a local maximum and a local maximum is not necessarily the optimal solution to (3.6.5).

We keep the notation of shortening $f_t(\lambda_t, u_t, X_t)$ to f_t and similar for the other stochastic processes when convenient as in Section 3.6. Furthermore $\hat{Y}$, $\hat{\phi}$, etc refers to the solutions corresponding to the control $\hat{u}$.

Theorem 3.7.2. *Let $\hat{u} \in \mathcal{A}^{\mathcal{F}}$. Suppose Assumption 3.7.1 holds and that for all β with $\hat{u}_t \pm \beta_t \in \mathcal{U}$ $dt \times d\mathbb{P}$ -a.e we have*

$$\mathbb{E} \left[\int_0^T \int_{\mathbb{R}} \left| \hat{Y}_s \left(\frac{\partial \hat{\kappa}_s}{\partial x}(z) \zeta_s^{\hat{u}, \beta} + \frac{\partial \hat{\kappa}_s}{\partial u}(z) \right)^2 + |\zeta_s^{\hat{u}, \beta} \hat{\phi}_s(z)|^2 \Lambda(ds, dz) \right] < \infty. \quad (3.7.6)$$

If $\hat{u}$ is a critical point for $J(u)$, in the sense that

$$\frac{\partial}{\partial y} J(\hat{u} + y\beta) \Big|_{y=0} = 0$$

for all bounded $\mathbb{F}$ -predictable processes β such that $u \pm \beta$ takes values in $\mathcal{U}$, then

$$\begin{aligned} \frac{\partial \mathcal{H}_t^{\mathbb{F}}}{\partial u}(\lambda_t, \hat{u}_t, \hat{X}_t, \hat{Y}_t, \hat{\phi}_t) &= \frac{\partial f_t}{\partial u}(\lambda_t, \hat{u}, \hat{X}_t) + \frac{\partial b_t}{\partial u}(\lambda_t, \hat{u}, \hat{X}_t) \mathbb{E}[\hat{Y}_t | \mathcal{F}_t] \\ &\quad + \frac{\partial \kappa_t}{\partial u}(0, \lambda_t, \hat{u}, \hat{X}_t) \mathbb{E}[\hat{\phi}_t(0) | \mathcal{F}_t] + \int_{\mathbb{R}_0} \frac{\partial \kappa_t}{\partial u}(z, \lambda_t, \hat{u}, \hat{X}_t) \mathbb{E}[\phi_t(z) | \mathcal{F}_t] \lambda_t^H \nu(dz) \\ &= 0, \quad dt \times d\mathbb{P}\text{-a.e.} \end{aligned} \quad (3.7.7)$$

Proof. Assume $\hat{u}$ is a critical point. Let β be a bounded $\mathbb{F}$ -predictable process such that

$$\hat{u}_t \pm \beta_t \in \mathcal{U} \quad dt \times d\mathbb{P}\text{-a.e.} \quad (3.7.8)$$

Consider

$$\begin{aligned} \frac{d}{dy} J(\hat{u} + y\beta) &= \mathbb{E} \left[\int_0^T \frac{d}{dy} f_s(\hat{u}_s + y\beta_s, X_s^{\hat{u} + y\beta}) ds + \frac{d}{dy} l(X_T^{\hat{u} + y\beta}) \right] \\ &= \mathbb{E} \left[\int_0^T \left[\frac{\partial f_s}{\partial u}(\hat{u}_s + y\beta_s, X_s^{\hat{u} + y\beta}) \beta_s + \frac{\partial f_s}{\partial x}(\hat{u}_s + y\beta_s, X_s^{\hat{u} + y\beta}) \frac{d}{dy} X_s^{\hat{u} + y\beta} \right] ds \right. \\ &\quad \left. + l'(X_T^{\hat{u} + y\beta}) \frac{d}{dy} X_T^{\hat{u} + y\beta} \right]. \end{aligned}$$

Evaluating $\frac{d}{dy} J(u + y\beta)$ at $y = 0$ gives (where $\zeta_t = \zeta_t^{\hat{u}, \beta}$)

$$0 = \mathbb{E} \left[\int_0^T \left(\frac{\partial \hat{f}_s}{\partial u} \beta_s + \frac{\partial \hat{f}_s}{\partial x} \zeta_s \right) ds + l'(\hat{X}_T) \zeta_T \right]. \quad (3.7.9)$$

By Itô's formula

$$\begin{aligned}
d(\hat{Y}_t \zeta_t) &= \hat{Y}_t \left[\frac{\partial \hat{b}_t}{\partial x} \zeta_t + \frac{\partial \hat{b}_t}{\partial u} \beta_t \right] dt + \hat{Y}_t \int_{\mathbb{R}} \left[\frac{\partial \hat{\kappa}_t}{\partial x}(z) \zeta_t + \frac{\partial \hat{\kappa}_t}{\partial u}(z) \beta_t \right] \mu(dt, dz) \\
&\quad - \zeta_t \frac{\partial \hat{\mathcal{H}}_t}{\partial x} dt + \zeta_t \int_{\mathbb{R}} \left[\hat{\phi}_t(z) \mu(dt, dz) + \left(\zeta_t \frac{\partial \hat{\kappa}_t}{\partial x}(0) + \frac{\partial \hat{\kappa}_t}{\partial u}(0) \right) \hat{\phi}_t(0) \right] \lambda_t^B dt \\
&\quad + \int_{\mathbb{R}_0} \left[\left(\zeta_t \frac{\partial \hat{\kappa}_t}{\partial x}(z) + \frac{\partial \hat{\kappa}_t}{\partial u}(z) \right) \right] \nu(dz) \lambda_t^H dt.
\end{aligned}$$

Since $\zeta_0 = 0$ and by the definition of $\frac{\partial \hat{\mathcal{H}}}{\partial x}$, we have

$$\begin{aligned}
\mathbb{E}[\hat{Y}_T \zeta_T] &= \mathbb{E} \left[\int_0^T \left[\hat{Y}_s \frac{\partial \hat{b}_s}{\partial u} \beta_s - \zeta_s \frac{\partial \hat{f}_s}{\partial x} \right] ds \right. \\
&\quad \left. + \int_0^T \int_{\mathbb{R}} \frac{\partial \hat{\kappa}_s}{\partial u}(0) \hat{\phi}_s(0) \beta_s \Lambda(ds, dz) \right]. \tag{3.7.10}
\end{aligned}$$

Let $\beta_s(\omega) = \alpha(\omega) \mathbf{1}_{t, t+h}(s)$, where α is a $\mathcal{F}_t$ -measurable random variable such that (3.7.8) is satisfied and $0 \leq t < t+h \leq T$. By (3.7.9)-(3.7.10)

$$0 = \mathbb{E} \left[\int_t^{t+h} \left(\frac{\partial \hat{f}_s}{\partial u} + \hat{Y}_s \frac{\partial \hat{b}_s}{\partial u} \right) \alpha ds + \int_t^{t+h} \int_{\mathbb{R}} \frac{\partial \hat{\kappa}_s}{\partial u}(z) \hat{\phi}_s(z) \alpha \Lambda(ds, dz) \right]. \tag{3.7.11}$$

Taking the derivative of h at $h = 0$ in equation (3.7.11) gives

$$\mathbb{E} \left[\alpha \frac{\partial \mathcal{H}_t}{\partial u} (\lambda_t, \hat{u}_t, \hat{X}_t, \hat{Y}_t, \hat{\phi}_t) \right] = 0. \tag{3.7.12}$$

Hence (3.7.12) holds for any bounded $\mathcal{F}_t$ -measurable random variable α such that $u_t + \alpha \in \mathcal{U}$ a.s. Define

$$\delta(\omega) = \sup_{c \in \mathbb{R}} \{u_t(\omega) + c \in \mathcal{U} \text{ and } u_t(\omega) - c \in \mathcal{U}\} \wedge 1.$$

Let A be an $\mathcal{F}_t$ -measurable random variable bounded by some constant $K_A > 0$. Then

$$u_t + \frac{1}{2K_A} \xi \delta \in \mathcal{U} \text{ a.s.}$$

Thus

$$\mathbb{E} \left[\frac{1}{2K_A} A \delta \frac{\partial \mathcal{H}_t}{\partial u} (\lambda_t, \hat{u}_t, \hat{X}_t, \hat{Y}_t, \hat{\phi}_t) \right] = 0.$$

We multiply by $2K_A$ to get

$$\mathbb{E} \left[A \delta \frac{\partial \mathcal{H}_t}{\partial u} (\lambda_t, \hat{u}_t, \hat{X}_t, \hat{Y}_t, \hat{\phi}_t) \right] = 0.$$

Note that δ does not depend on A . Let $A^{(n)} = (\frac{1}{\delta}A) \wedge n$. Then $A^{(n)} \rightarrow A$ a.s. and we must have

$$\mathbb{E}\left[A \frac{\partial \mathcal{H}_t}{\partial u^{(j)}}(\hat{X}_t, \hat{u}_t)\right] = 0.$$

Since this holds for all bounded $\mathcal{F}_t$ measurable A we conclude

$$\mathbb{E}\left[\frac{\partial \mathcal{H}_t}{\partial u}(\lambda_t, \hat{u}_t, \hat{X}_t, \hat{Y}_t, \hat{\phi}_t) \middle| \mathcal{F}_t\right] = 0.$$

By this we end the proof. □

A similar technique to the proof of Theorem 3.7.2 was used in [6] to prove necessary conditions in partial information framework for Lévy processes. Note however that our Assumption 3.7.1 is significantly different from the assumptions made in [6] for reasons that will be discussed in Section 4.6.

Hereafter is our analysis on the uniqueness and optimality of the critical points.

Theorem 3.7.3. *Let $J(u) = \mathbb{E}[l(X_T)]$ and assume*

i) All bounded $\mathbb{F}$ -predictable processes taking values in $\mathcal{U}$ are admissible controls.

ii) For all $u \in \mathcal{A}^{\mathcal{F}}$ and bounded, $\mathbb{F}$ -predictable processes β such that $u_t \pm \beta_t \in \mathcal{U}$ $dt \times d\mathbb{P}$ -a.e. we have

$$\mathbb{E}\left[l''(X_T^u) \frac{\partial}{\partial y} X_t^{u+y\beta} \Big|_{y=0} + l'(X_T^u) \frac{\partial^2}{\partial y^2} X_t^{u+y\beta} \Big|_{y=0}\right] < 0. \quad (3.7.13)$$

iii) For all $u \in \mathcal{A}^{\mathcal{F}}$ and bounded, $\mathbb{F}$ -predictable processes β such that $u_t \pm \beta_t \in \mathcal{U}$ $dt \times d\mathbb{P}$ -a.e. the family

$$\left\{ l''(X_T^{u+y\beta}) \frac{\partial}{\partial y} X_T^{u+y\beta} + l'(X_T^{u+y\beta}) \frac{\partial^2}{\partial y^2} X_T^{u+y\beta} \right\}_{\{y \in (-1,1)\}} \quad (3.7.14)$$

is uniformly $\mathbb{P}$ -integrable.

iv) Assumption 3.7.1 holds.

Then

v) for all $u \in \mathcal{A}^{\mathcal{F}}$ and bounded, $\mathbb{F}$ -predictable processes β such that $u_t \pm \beta_t \in \mathcal{U}$ $dt \times d\mathbb{P}$ -a.e. the mapping $y \rightarrow J(u + y\beta)$, $y \in [-1, 1]$, is concave,

vi) there is at most one bounded $u \in \mathcal{A}^{\mathcal{F}}$ such that u is a critical point (in the sense of Theorem 3.7.2),

vii) if $\mathcal{U}$ is bounded and a critical point $\hat{u} \in \mathcal{A}^{\mathcal{F}}$ exist, then $\hat{u}$ is optimal, i.e. $J(\hat{u}) = \sup_{u \in \mathcal{A}^{\mathcal{F}}} J(u)$.

Proof. To prove the concavity of $y \rightarrow J(u + y\beta)$ it is sufficient that $\frac{\partial^2}{\partial y^2} J(u + y\beta)|_{y=a} < 0$ for all $a \in [-1, 1]$. Let $a \in [-1, 1]$ be given and set $v = u + a\beta$. Note that $v \in \mathcal{A}^F$ by condition i) in Assumption 3.7.1. We have

$$\begin{aligned} \frac{\partial^2}{\partial y^2} J(u + y\beta)|_{y=a} &= \frac{\partial^2}{\partial y^2} J(v + y\beta)|_{y=0} \\ &= \mathbb{E} \left[l''(X_T^v) \frac{\partial}{\partial y} X_t^{v+y\beta} \Big|_{y=0} + l'(X_T^v) \frac{\partial^2}{\partial y^2} X_T^{v+y\beta} \Big|_{y=0} \right] < \infty, \end{aligned}$$

where the last inequality follows from (3.7.13).

Next we want to show that there is at most one bounded $u \in \mathcal{A}^F$ such that u is a critical point. Suppose $\bar{u}, \hat{u} \in \mathcal{A}^F$ are both bounded and critical points. Set $\beta_t = \bar{u}_t - \hat{u}_t$. Consider the control $\hat{u} + \frac{1}{2}\beta \in \mathcal{A}^F$, and the mapping

$$h(y) \rightarrow J\left(\hat{u} + \frac{1}{2}\beta + y\frac{1}{2}\beta\right), \quad y \in [-1, 1] \quad (3.7.15)$$

Note that $h(1) = J(\bar{u})$ and $h(-1) = J(\hat{u})$. Since h is concave at most one of $h(-1)$ and $h(1)$ can be a maximum.

The optimality of the bounded critical point $\hat{u}$ when $\mathcal{U}$ is bounded follows from the uniqueness of the bounded critical points and the concavity of the mappings described above. $\square$

3.8 Optimal portfolio problems

Here we show how investments in financial assets can be modeled within our framework with a state process suitable for various optimization problems. First we setup the general framework, then consider the specific problems of mean-variance hedging and utility maximization.

We consider two assets, a risk free asset R and a risky asset S defined by

$$\begin{aligned} dR_t &= \rho_t R_t dt, & R_0 &= 1, \\ dS_t &= \alpha_t S_t dt + S_t \int_{\mathbb{R}} \psi_t(z) \mu(dt, dz), & S_0 &> 0. \end{aligned}$$

Models of this type include (3.1.6) (the model from [25]) and (3.1.7). Here α and ρ are $\mathbb{F}$ -adapted stochastic processes with $\alpha, \rho : [0, T] \times \Omega \rightarrow \mathbb{R}$ and $\psi \in \mathcal{I}$ is an $\mathbb{F}$ -adapted random field. We assume that ρ is bounded. Let $z_t^{(R)}$ denote the units of R held at time t and $z_t^{(S)}$ the number of units of S held at time t . The wealth process $X_t, t \in [0, T]$, is the value of the assets held,

$$X_t = z_t^{(R)} R_t + z_t^{(S)} S_t. \quad (3.8.1)$$

We assume that the portfolio is self-financing, i.e. that

$$dX_t = z_t^{(R)} dR_t + z_t^{(S)} dS_t. \quad (3.8.2)$$

Let $u_t = z_t^{(S)} S_t, t \in [0, T]$, denote the amount of wealth invested in the risky asset S . By (3.8.1)

and (3.8.2) the wealth equation is given by

$$dX_t = [\rho_t X_t + (\alpha_t - \rho_t)u_t] dt + u_t \int_{\mathbb{R}} \psi_t(z) \mu(dt, dz). \quad (3.8.3)$$

Clearly assumptions (3.6.11), (3.6.12) and (3.6.13) are satisfied. We assume that (3.8.3) admits a strong solution. For this we see that the sufficient conditions (3.6.6), (3.6.7), (3.6.8) are satisfied and together with

$$\begin{aligned} \int_0^T |\alpha_s - \rho_s| |u_s| ds &< \infty \\ \int_0^T \int_{\mathbb{R}_0} |u_s \psi_s(z)|^2 \Lambda(ds, dz) &< \infty \end{aligned}$$

$\mathbb{P}$ -a.s. also (3.6.9) and (3.6.10). The SDE (3.8.3) is of Ornstein-Uhlenbeck type and has solution

$$\begin{aligned} X_t = e^{\int_0^t \rho_r dr} &\left(X_0 + \int_0^t e^{-\int_0^s \rho_r dr} (\alpha_s - \rho_s) u_s ds \right. \\ &\left. + \int_0^t \int_{\mathbb{R}} e^{-\int_0^s \rho_r dr} u_s \psi_s(z) \mu(ds, dz) \right). \end{aligned} \quad (3.8.4)$$

We consider portfolio problems of type (3.6.4)-(3.6.5) associated

$$J(u) = \mathbb{E}[l(X_T)] \quad (3.8.5)$$

with l as in (3.6.1). Problems of this type include utility maximization and mean-variance portfolio selection. Hedging problems are also included, since l is a function of both ω and X we can consider e.g. the mean-variance hedge by $l(X_T) = -(X_T - Z)^2$ for a square integrable random variable Z . The Hamiltonian for this class of problems is

$$\begin{aligned} \mathcal{H}_t(\lambda, u, x, y, \phi) = & [\rho_t x + (\alpha_t - \rho_t)u] y + u \psi_t(0) \phi(0) \lambda_t^B \\ & + \int_{\mathbb{R}_0} [u \psi_t(z) \phi(z)] \lambda_t^H \nu(dz) \end{aligned} \quad (3.8.6)$$

and the associated BSDE is given by

$$Y_t = \partial_x l(X_T) + \int_t^T Y_s \rho_s ds - \int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz). \quad (3.8.7)$$

By Theorem 3.5.1 we also have the representation

$$Y_t = \mathbb{E} \left[\partial_x l(X_T) \exp \left\{ \int_t^T \rho_s ds \right\} \middle| \mathcal{G}_t \right]. \quad (3.8.8)$$

3.8.1 The mean-variance portfolio selection

Here we discuss mean-variance portfolio selection starting from an initial wealth $x \in \mathbb{R}$, i.e. solve $\inf_u \mathbb{E}[(X_T - E[X_T])^2]$ with $E[X_T] = k$ for some $k \in \mathbb{R}$ and controls taking values in $\mathcal{U} = \mathbb{R}$. For notational convenience we consider the equivalent formulation $J(u) = \mathbb{E}[-\frac{1}{2}(X_T - k)^2]$ and want to find

$$\sup_{u \in \mathcal{A}^{\mathcal{F}}} J(u) = \sup_{u \in \mathcal{A}^{\mathcal{F}}} \mathbb{E} \left[-\frac{1}{2}(X_T - k)^2 \right]. \quad (3.8.9)$$

To solve this problem we first consider the optimization on $u \in \mathcal{A}^{\mathcal{G}}$ with deterministic coefficients and apply Theorem 3.6.3. To avoid trivial solutions we assume $\alpha_t > \rho_t dt \times d\mathbb{P}$ a.e.

Theorem 3.8.1. *Assume that ρ and α are deterministic. Consider the feedback control $\hat{u}^{\mathbb{G}} \in \mathcal{A}^{\mathcal{G}}$ given by*

$$\hat{u}_t^{\mathbb{G}} = \frac{-(\alpha_t - \rho_t)(A_t \hat{X}_t + C_t)}{A_t(|\psi_t(0)|^2 \lambda_t^B + \int_{\mathbb{R}_0} |\psi_t(z)|^2 \lambda_t^H \nu(dz))},$$

where $\hat{X}$ refers to (3.8.4) with $\hat{u}^{\mathbb{G}}$ and

$$A_t = -\exp \left\{ -\int_t^T \frac{(\alpha_s - \rho_s)^2}{|\psi_s(0)|^2 \lambda_s^B + \int_{\mathbb{R}_0} |\psi_s(z)|^2 \lambda_s^H \nu(dz)} - 2\rho_s ds \right\} \quad (3.8.10)$$

$$C_t = k \exp \left\{ -\int_t^T \frac{(\alpha_s - \rho_s)^2}{|\psi_s(0)|^2 \lambda_s^B + \int_{\mathbb{R}_0} |\psi_s(z)|^2 \lambda_s^H \nu(dz)} - \rho_s ds \right\}. \quad (3.8.11)$$

If (3.6.16) holds then $\hat{u}^{\mathbb{G}}$ is optimal for $\sup_{u \in \mathcal{A}^{\mathcal{G}}} J(u)$.

Remark that, from (3.8.10)-(3.8.11), the processes A and C depend on future values of λ^B and λ^H , hence they are $\mathbb{G}$ -adapted, but in general not $\mathbb{F}$ -adapted. Observe that equation (3.8.8) gives the useful characterization of the adjoint equation

$$\begin{aligned} Y_t &= \mathbb{E} \left[-(X_T - k) \exp \left\{ \int_t^T \rho_s ds \right\} \middle| \mathcal{G}_t \right] \\ &= (k - X_t) \mathbb{E} \left[\exp \left\{ \int_t^T \rho_s ds \right\} \middle| \mathcal{G}_t \right] + \mathbb{E} \left[(X_t - X_T) \exp \left\{ \int_t^T \rho_s ds \right\} \middle| \mathcal{G}_t \right]. \end{aligned}$$

The study of the above representation hints that the process Y takes the form $Y_t = A_t X_t + C_t$ where A and C are some $\mathbb{G}$ -adapted processes of finite variation and $A_T = -1$, $C_T = k$.

Together with some results in the Lévy case, see e.g. [48], we can take a guess of the structure of Y for an optimal control candidate and use our verification theorem to actually determine its optimality.

Proof of Theorem 3.8.1. Denote a and c as of the derivatives of A and C with respect to t , so that $dA_t = a_t dt$ and $dC_t = c_t dt$. Set

$$\hat{u}_t := -\frac{\rho_t(A_t\hat{X}_t + C_t) + A_t\rho_t\hat{X}_t + \hat{X}_ta_t + c_t}{A_t(\alpha_t - \rho_t)}, \quad (3.8.12)$$

$$\hat{Y}_t := A_t\hat{X}_t + C_t, \quad (3.8.13)$$

where A and C are given by (3.8.10)-(3.8.11). We need to show

i) that (3.8.13) indeed is the process Y in the solution of (3.8.7) corresponding to the control $\hat{u}$.

ii) that $\hat{u}$ (3.8.12) satisfies (3.6.18) in Theorem 3.6.3.

To prove i), we combine (3.8.3) and (3.8.13) to get

$$\begin{aligned} d\hat{Y}_t &= A_t(\rho_t\hat{X}_t + (\alpha_t - \rho_t)\hat{u}_t) dt + A_t\hat{u}_t \int_{\mathbb{R}} \psi_t(z) \mu(dt, dz) \\ &\quad + \hat{X}_ta_t dt + c_t dt; \quad \hat{Y}_T = k - \hat{X}_T. \end{aligned} \quad (3.8.14)$$

Inserting (3.8.12) in (3.8.14) and by the fact that $A_T = -1$ and $C_T = k$, we see that

$$\hat{Y}_t = k - \hat{X}_T + \int_t^T \hat{Y}_s \rho_s ds - \int_t^T \int_{\mathbb{R}} A_s \hat{u}_s \psi_s(z) \mu(ds, dz). \quad (3.8.15)$$

By the uniqueness of the BSDE-solution (Theorem 3.4.5) we see that $(\hat{Y}, \hat{\phi})$ with

$$\hat{\phi}_t(z) = A_t\hat{u}_t\psi_t(z), \quad (3.8.16)$$

solves (3.8.7) with $\hat{u}$. Hence i) is proved.

Next we study ii). The Hamiltonian (3.8.6) is a linear function in $u \in \mathcal{U} = \mathbb{R}$ and, composed with $(\hat{Y}, \hat{\phi})$, it is:

$$\begin{aligned} \mathcal{H}_t(\lambda, u, x, \hat{Y}_t, \hat{\phi}_t) &= \rho_t x \hat{Y}_t \\ &\quad + u \left[(\alpha_t - \rho_t) \hat{Y}_t + \psi_t(0) \hat{\phi}_t(0) \lambda_t^B + \int_{\mathbb{R}_0} [\psi_t(z) \hat{\phi}_t(z)] \lambda_t^H \nu(dz) \right]. \end{aligned}$$

To have (3.6.17) well defined and (3.6.18) satisfied for the control $\hat{u}$ we see that

$$(\alpha_t - \rho_t) \hat{Y}_t + \psi_t(0) \hat{\phi}_t(0) \lambda_t^B + \int_{\mathbb{R}_0} [\psi_t(z) \hat{\phi}_t(z)] \lambda_t^H \nu(dz) = 0 \quad (3.8.17)$$

is necessary and sufficient. Indeed we can see that equation (3.8.17) is satisfied. Substituting (3.8.13) and (3.8.16) into (3.8.17), we obtain the equation:

$$\hat{u}_t = \frac{-(\alpha_t - \rho_t)(A_t \hat{X}_t + C_t)}{A_t(|\psi_t(0)|^2 \lambda_t^B + \int_{\mathbb{R}_0} |\psi_t(z)|^2 \lambda_t^H \nu(dz))}. \quad (3.8.18)$$

Substituting the definition of $\hat{u}$ (3.8.12), we have

$$\begin{aligned} & (\alpha_t - \rho_t)^2 (A_t \hat{X}_t + C_t) \\ &= \left(2\rho_t A_t \hat{X}_t + \rho_t C_t + \hat{X}_t a_t + c_t \right) \left(|\psi_t(0)|^2 \lambda_t^B + \int_{\mathbb{R}_0} |\psi_t(z)|^2 \lambda_t^H \nu(dz) \right), \end{aligned}$$

which is verified once the definitions of A and C (3.8.10)-(3.8.11) are inserted. Hence, in the setting of the theorem, we conclude that $\hat{u}^G = \hat{u}$ (3.8.12)-(3.8.18) is optimal. $\square$

We can now solve problem (3.8.9) using similar arguments.

Theorem 3.8.2. *Consider the feedback control $\hat{u}^\mathbb{F} \in \mathcal{A}^\mathcal{F}$ given by*

$$\hat{u}_t^\mathbb{F} = - \frac{(\alpha_t - \rho_t) \left(\mathbb{E}[A_t | \mathcal{F}_t] \hat{X}_t + \mathbb{E}[C_t | \mathcal{F}_t] \right)}{\mathbb{E}[A_t | \mathcal{F}_t] \left(|\psi_t(0)|^2 \lambda_t^B + \int_{\mathbb{R}_0} |\psi_t(z)|^2 \lambda_t^H \nu(dz) \right)}, \quad (3.8.19)$$

where $\hat{X}$ refers to (3.8.4) with $\hat{u}^\mathbb{F}$ and the processes A and C are given by (3.8.10)-(3.8.11). If (3.6.16) holds then $\hat{u}^\mathbb{F}$ solves (3.8.9).

Apart from a technical point involving conditional expectations, the proof is similar to Theorem 3.8.1.

Proof. Let $\hat{u}$ be given by (3.8.19), i.e.

$$\hat{u}_t = - \frac{(\alpha_t - \rho_t) \left(\mathbb{E}[A_t | \mathcal{F}_t] \hat{X}_t + \mathbb{E}[C_t | \mathcal{F}_t] \right)}{\mathbb{E}[A_t | \mathcal{F}_t] \left(|\psi_t(0)|^2 \lambda_t^B + \int_{\mathbb{R}_0} |\psi_t(z)|^2 \lambda_t^H \nu(dz) \right)}.$$

The adjoint equation is given by

$$\begin{aligned} \hat{Y}_t &= \mathbb{E}[A_t | \mathcal{F}_t] \hat{X}_t + \mathbb{E}[C_t | \mathcal{F}_t], \\ \hat{\phi}_t(z) &= \mathbb{E}[A_t | \mathcal{F}_t] \hat{u}_t \psi_t(z), \end{aligned}$$

with A and C given by (3.8.10) and (3.8.11). We can verify that $\hat{Y}$, $\hat{\phi}$ is indeed the solution of (3.8.7) using the same arguments as in the proof of Theorem 3.8.1. First we show that $d\mathbb{E}[A_t | \mathcal{F}_t] = \mathbb{E}[a_t | \mathcal{F}_t] dt$ and $d\mathbb{E}[C_t | \mathcal{F}_t] = \mathbb{E}[c_t | \mathcal{F}_t] dt$, where

$$\begin{aligned} \mathbb{E}[a_t | \mathcal{F}_t] &= \left(\frac{(\alpha_t - \rho_t)^2}{|\psi_t(0)|^2 \lambda_t^B + \int_{\mathbb{R}_0} |\psi_t(z)|^2 \lambda_t^H \nu(dz)} - 2\rho_t \right) \mathbb{E}[A_t | \mathcal{F}_t], \\ \mathbb{E}[c_t | \mathcal{F}_t] &= \left(\frac{(\alpha_t - \rho_t)^2}{|\psi_t(0)|^2 \lambda_t^B + \int_{\mathbb{R}_0} |\psi_t(z)|^2 \lambda_t^H \nu(dz)} - \rho_t \right) \mathbb{E}[C_t | \mathcal{F}_t]. \end{aligned}$$

The argument is described for A , the case of C is identical. Define

$$J_t := \frac{(\alpha_t - \rho_t)^2}{|\psi_t(0)|^2 \lambda_t^B + \int_{\mathbb{R}_0} |\psi_t(z)|^2 \lambda_t^H \nu(dz)} - 2\rho_t.$$

Remark that

$$A_t = -\exp \left\{ - \int_t^T J_s ds \right\}.$$

Thus, we have

$$\begin{aligned} \mathbb{E}[A_{t+\Delta t} | \mathcal{F}_{t+\Delta t}] &= \mathbb{E} \left[-\exp \left\{ - \int_{t+\Delta t}^T J_s ds \right\} \middle| \mathcal{F}_{t+\Delta t} \right] \\ &= \exp \left\{ \int_t^{t+\Delta t} J_s ds \right\} \mathbb{E} \left[-\exp \left\{ - \int_t^T J_s ds \right\} \middle| \mathcal{F}_{t+\Delta t} \right]. \end{aligned}$$

Hence, since $\mathbb{E}[\exp \{ - \int_t^T J_s ds \} | \mathcal{F}_{t+\Delta t}] \rightarrow \mathbb{E}[\exp \{ - \int_t^T J_s ds \} | \mathcal{F}_t] dt \times d\mathbb{P}$ a.e as $\Delta t \rightarrow 0$, we obtain

$$\begin{aligned} \lim_{\Delta t \rightarrow 0^+} \frac{\mathbb{E}[A_{t+\Delta t} | \mathcal{F}_{t+\Delta t}] - \mathbb{E}[A_t | \mathcal{F}_t]}{\Delta t} &= \lim_{\Delta t \rightarrow 0^+} \frac{\exp \left\{ \int_t^{t+\Delta t} J_s ds \right\} - 1}{\Delta t} \mathbb{E}[A_t | \mathcal{F}_t] \\ &= J_t \mathbb{E}[A_t | \mathcal{F}_t] = \mathbb{E}[a_t | \mathcal{F}_t]. \end{aligned}$$

Following Theorem 3.6.4 we define

$$\begin{aligned} \mathcal{H}_t^{\mathbb{F}}(\lambda, u, x, \hat{Y}_t, \hat{\phi}_t) &= \rho_t x \mathbb{E}[\hat{Y}_t | \mathcal{F}_t] + u \left\{ (\alpha_t - \rho_t) \mathbb{E}[\hat{Y}_t | \mathcal{F}_t] \right. \\ &\quad \left. + \psi_t(0) \mathbb{E}[\hat{\phi}_t(0) | \mathcal{F}_t] \lambda^B + \int_{\mathbb{R}_0} \psi_t(z) \mathbb{E}[\hat{\phi}_t(z) | \mathcal{F}_t] \lambda^H \nu(dz) \right\}. \end{aligned}$$

For $\hat{u}$ to be optimal, from (3.6.23)-(3.6.24), it is sufficient to show that

$$(\alpha_t - \rho_t) \mathbb{E}[\hat{Y}_t | \mathcal{F}_t] + \psi_t(0) \mathbb{E}[\hat{\phi}_t(0) | \mathcal{F}_t] \lambda_t^B + \int_{\mathbb{R}_0} \psi_t(z) \mathbb{E}[\hat{\phi}_t(z) | \mathcal{F}_t] \lambda_t^H \nu(dz) = 0. \quad (3.8.20)$$

Replacing $\hat{\phi}_t(z) = A_t \hat{u}_t \psi_t(z)$ and inserting (3.8.19) in (3.8.20), we obtain the desired result. $\square$

3.8.2 Utility maximization of final wealth

For the utility maximization problem of the final wealth we set

$$J(u) = E[U(X_T)]$$

where $U : \mathbb{R} \rightarrow \mathbb{R}$ is a differentiable utility function, increasing and concave. The BSDE is given by

$$Y_t = U'(X_T) + \int_t^T Y_s \rho_s ds - \int_t^T \int_{\mathbb{R}} \phi_s(z) \mu(ds, dz)$$

(where $U'(x) = \frac{d}{dx}U(x)$). By arguments as in the mean-variance portfolio problem, the sufficient condition from the maximum principle for optimal $\hat{u} \in \mathcal{A}^G$ can be reduced to

$$(\alpha_t - \rho_t)\hat{Y}_t = \psi_t(0)\hat{\phi}_t(0)\lambda_t^B + \int_{\mathbb{R}_0} [\psi_t(z)\hat{\phi}_t(z)] \lambda_t^H \nu(dz),$$

or

$$\begin{aligned} (\alpha_t - \rho_t) & \left(U'(\hat{X}_T) + \int_t^T \hat{Y}_s \rho_s ds - \int_t^T \int_{\mathbb{R}_0} \hat{\phi}_s(z) \mu(ds, dz) \right) \\ & = \psi_t(0)\hat{\phi}_t(0)\lambda_t^B + \int_{\mathbb{R}_0} [\psi_t(z)\hat{\phi}_t(z)] \lambda_t^H \nu(dz), \end{aligned}$$

where $\hat{Y}, \hat{\phi}$ depend on $\hat{u}$.

Chapter 4

Maximum principles for martingale random fields via non-anticipating stochastic derivatives

Abstract

We find a maximum principle for processes driven by martingale random fields. We do so by describing the adjoint processes with non-anticipating stochastic derivatives. In the case of the Lévy processes this mimics maximum principles with Malliavin derivatives, but we replace Malliavin differentiability conditions with L_2 -conditions.

As an application we use the maximum principle to solve a portfolio optimization problem for assets with credit risk modeled by doubly stochastic Poisson processes.

4.1 Introduction

We find a maximum principle for an optimization problem when the state process depends on a martingale random field [23, 38], a generalization of the martingale. We do so in a perturbation based approach, using the *non-anticipating stochastic derivative* [34, 35, 38] to describe the adjoint processes.

We consider the performance functional

$$J(u) = \mathbb{E} \left[\int_0^T f_t(u_t, X_t) dt + g(X_T) \right] \quad (4.1.1)$$

and the associated optimal stochastic control problem, where u is the control and the state process is given by the semi-martingale X ,

$$X_t^{(u)} = X_0 + \int_0^t b_s(u_s, X_s) ds + \int_0^t \int_{\mathcal{Z}} \phi_s(z, u_s, X_s) \mu(ds, dz), \quad t \in [0, T],$$

where the last integral is over the martingale random field μ on $[0, T] \times \mathcal{Z}$.

The goal is to find $\sup_u J(u)$ for controls adapted to the filtration $\mathbb{F}$, where X is adapted to the filtration $\mathbb{G}$ and $\mathbb{F} \subseteq \mathbb{G}$, i.e. for all $t \in [0, T]$ we have $\mathcal{F}_t \subseteq \mathcal{G}_t$. This is a problem with *partial information* if X is not $\mathbb{F}$ -adapted. We find (candidates for) optimal solutions by investigating

$$\frac{\partial}{\partial y} J(u + y\beta) \Big|_{y=0}, \quad u, u + y\beta \text{ are admissible controls and } |y| < \delta, \quad (4.1.2)$$

for some $\delta > 0$. The controls are taking values in an open, convex set $\mathcal{U} \subseteq \mathbb{R}^n$. In the literature (4.1.2) has sometimes been evaluated using a set of assumptions that requires $\mathcal{U} = \mathbb{R}^n$. We explain this issue in Section 4.6, and state our maximum principle with weaker assumptions so that we can overcome this problem and indeed allow for $\mathcal{U} \subset \mathbb{R}^n$.

Key to our approach is the non-anticipating derivative $\mathcal{D}$, an operator from $L_2(\Omega, \mathcal{G}, \mathbb{P})$ to the space of integrable random fields, see, e.g. [34, 38]. The operator $\mathcal{D}$ coincides with the dual of the Itô non-anticipating stochastic integral with respect to a general martingale random field. Indeed we have that, for $\xi \in L_2(\Omega, \mathcal{G}, \mathbb{P})$, that

$$\mathbb{E} \left[\xi \int_0^T \int_{\mathcal{Z}} \kappa(s, z) \mu(ds, dz) \right] = \mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} (\mathcal{D}_{s,z} \xi) \kappa(s, z) \Lambda(ds, dz) \right].$$

Here Λ represents the conditional variance measure associated to μ . These concepts are further detailed in the forthcoming sections 4.2 and 4.3.

Maximum principles using the duality relation of the Malliavin derivative with the Skorohod integral have been studied in [41, 82]. This limits the study to Lévy processes and, additionally, some restrictions are imposed to match the domains of the Malliavin derivative. The key novelty in our paper is the use of the non-anticipating stochastic derivative, which enables us to treat very general martingale noises. Furthermore, in the case of Lévy noise, we reduce assumptions of Malliavin differentiable random variables to square integrability. Since the non-anticipating derivative coincide with the Malliavin derivative when both are well defined, this extends previous results. An additional benefit of the martingale random field structure is an easy extension to multi-dimensional controls.

For the portfolio problem with default risk, the main result is extended to a simpler sufficient condition for optimal control.

The benefit of the duality type approach used herein compared to HJB-type equations is that we can treat problems of partial information, which are of non-Markovian nature. Maximum principles for partial information have also been studied in a BSDE approach in e.g. [6, 3] but again limited to the case of Lévy noise. Note that partial information in the sense of the filtrations $\mathbb{F}$ and $\mathbb{G}$ differs from *partial observation* problems with noisy observations of the state process as treated in e.g. [10, 70, 103].

In this paper, the maximum principle is studied in Section 4.5. The specific discussion of a Lévy type martingale μ is considered in Section 4.7. Section 4.8 presents an application to portfolio optimization in a market with assets subject to default risk.

4.2 The martingale random field

We now retrieve the stochastic integration and the non-anticipating stochastic derivative over a martingale random field μ . We refer to [38] for a detailed discussion on these concepts.

Let $(\Omega, \mathcal{G}, \mathbb{P})$ be a complete probability space equipped with a right-continuous filtration $\mathbb{G} := \{\mathcal{G}_t, t \in [0, T]\}$. Let $\mathcal{Z}$ be a separable topological space. We denote $\mathcal{B}_{\mathcal{Z}}$ as the Borel σ -algebra on $\mathcal{Z}$ and $\mathcal{B}_{[0, T] \times \mathcal{Z}}$ as the Borel σ -algebra on the product space $[0, T] \times \mathcal{Z}$. Note that $\mathcal{B}_{[0, T] \times \mathcal{Z}}$ is generated by a semi-ring of sets of type

$$\Delta = (t, s] \times Z, \quad 0 \leq t < s \leq T, Z \in \mathcal{B}_{\mathcal{Z}}.$$

We say that the stochastic set function $\mu(\Delta)$, $\Delta \in \mathcal{B}_{[0, T] \times \mathcal{Z}}$ is a martingale random field in $L_2(\Omega, \mathcal{G}, \mathbb{P})$ on $[0, T] \times \mathcal{Z}$ (with conditionally orthogonal values) with respect to $\mathbb{G}$ if it satisfies the following properties [38, Definition 2.1]:

i) μ has a tight, σ -finite variance measure $V(\Delta) = E[\mu(\Delta)^2]$, $\Delta \in \mathcal{B}_{[0, T] \times \mathcal{Z}}$, which satisfies $V(\{0\} \times \mathcal{Z}) = 0$.

ii) μ is additive, i.e. for pairwise disjoint sets $\Delta_1, \dots, \Delta_K$: $V(\Delta_K) < \infty$

$$\mu\left(\bigcup_{k=1}^K \Delta_k\right) = \sum_{k=1}^K \mu(\Delta_k)$$

and σ -additive in L_2 .

iii) μ is $\mathbb{G}$ -adapted.

iv) μ has the *martingale property*. Consider $\Delta \subseteq (t, T] \times \mathcal{Z}$. We have:

$$\mathbb{E}[\mu(\Delta) \mid \mathcal{G}_t] = 0.$$

v) μ has conditionally orthogonal values. For any $\Delta_1, \Delta_2 \subseteq (t, T] \times \mathcal{Z}$ such that $\Delta_1 \cap \Delta_2 = \emptyset$ we have:

$$\mathbb{E}[\mu(\Delta_1)\mu(\Delta_2) \mid \mathcal{G}_t] = 0.$$

In particular, any finite sums of orthogonal, square integrable martingales would be a martingale random field in the sense of i)-ii)-iii)-iv)-v) above. In general, the filtration $\mathbb{G}$ does not need to be the one generated by μ .

The $\mathbb{G}$ -predictable σ -algebra on $\Omega \times [0, T] \times \mathcal{Z}$ is denoted by $\mathcal{P}_{[0, T] \times \mathcal{Z}}$ and is generated by sets of type

$$\Delta = A \times (t, s] \times Z, \quad 0 \leq t < s \leq T, Z \in \mathcal{B}_{\mathcal{Z}}, A \in \mathcal{G}_t.$$

The $\mathbb{G}$ -predictable σ -algebra $\Omega \times [0, T]$ is denoted by $\mathcal{P}_{[0, T]}$ and is generated by sets of type $\Delta = A \times (t, s]$, $0 \leq t < s \leq T, A \in \mathcal{G}_t$. On $(\Omega \times [0, T] \times \mathcal{Z}, \mathcal{P}_{[0, T] \times \mathcal{Z}})$ the random field μ has

a σ -finite conditional random variance measure [38, Theorem 2.1]. For martingale processes the conditional variance measure is the $\mathbb{G}$ -predictable compensator. We denote this conditional variance measure by Λ , and it has the following properties

$$\begin{aligned}\mathbb{E}[\mu(\Delta)^2 | \mathcal{G}_t] &= \Lambda(\Delta), \quad \text{in } L_1(\Omega, \mathcal{G}, \mathbb{P}) \text{ for } \Delta \subseteq (t, T] \times \mathcal{Z}, \\ \mathbb{E}[\mu(\Delta)^2] &= \mathbb{E}[\Lambda(\Delta)].\end{aligned}$$

For later purposes we assume that Λ is absolutely continuous with respect to the Lebesgue measure on $[0, T]$. Namely we assume that there exists a transition kernel λ from $(\Omega \times [0, T], \mathcal{P}_{[0, T]})$ to $(\mathcal{Z}, \mathcal{B}_{\mathcal{Z}})$ such that $\Lambda(\omega, dt, dz) = \lambda_t(\omega, dz) dt$. Meaning that the mapping $(\omega, t) \rightarrow \lambda_t(\omega, Z)$ is $\mathcal{P}_{[0, T]}$ measurable for every $Z \in \mathcal{B}_{\mathcal{Z}}$ and $\lambda_t(\omega, \cdot)$ is measure on $(\mathcal{Z}, \mathcal{B}_{\mathcal{Z}})$ for every $(\omega, t) \in \Omega \times [0, T]$. We refer to [26] for further details on transition kernels.

We denote $\mathcal{I}$ as the set of $\mathbb{G}$ -predictable random fields $\phi : \Omega \times [0, T] \times \mathcal{Z} \rightarrow \mathbb{R}$ satisfying

$$\|\phi\|_{\mathcal{I}} := \mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} \phi(s, z)^2 \lambda_s(dz) ds \right]^{\frac{1}{2}} < \infty.$$

We say that $\phi \in \mathcal{I}$ is a simple random field if it can be expressed as a finite sum of type

$$\phi(s, z, \omega) = \sum_{i=1}^N \phi_i(\omega) \mathbf{1}_{\Delta_i}(s, z), \quad (4.2.1)$$

where $\Delta_i = (t_i, s_i] \times Z_i$ and ϕ_i are bounded, $\mathcal{G}_{t_i}$ -measurable random variables for $i = 1, \dots, N < \infty$. Simple, $\mathbb{G}$ -predictable random fields are dense in $\mathcal{I}$ by the usual Itô integration type arguments and we have that, for every $\phi \in \mathcal{I}$:

$$\begin{aligned}\mathbb{E} \left[\left(\int_0^T \int_{\mathcal{Z}} \phi(s, z) \mu(ds, dz) \right)^2 \right] &= \mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} \phi(s, z)^2 \Lambda(ds, dz) \right] \\ &= \mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} \phi(s, z)^2 \lambda_s(dz) ds \right].\end{aligned} \quad (4.2.2)$$

Remark also that $\phi \in \mathcal{I}$ implies that

$$\int_{\mathcal{Z}} \phi_t(z)^2 \lambda_t(dz) < \infty, \quad dt \times d\mathbb{P} \text{ a.e.}$$

Note that $\int_0^t \int_{\mathcal{Z}} \phi(s, z) \mu(ds, dz)$, $t \in [0, T]$ is a $\mathbb{G}$ -martingale with values in L_2 .

4.3 Duality relations with non-anticipating stochastic derivatives

The *non-anticipating stochastic derivative* is a characterization of the integrand in the Kunita-Watanabe decomposition, developed in [34, 35, 44, 37, 38]. It is the adjoint (linear) operator $\mathcal{D} = I^*$ of the stochastic integral:

$$\mathcal{D} : L_2(\Omega, \mathcal{G}, \mathbb{P}) \Longrightarrow \mathcal{I}.$$

A full characterization is given in constructive form using the elements of the following dissecting system, a sequence of partitions of $[0, T] \times \mathcal{Z}$. Let $A_n \subseteq [0, T] \times \mathcal{Z}$ be an increasing sequence of Borel-sets such that $V(A_n) < \infty$ for all $n \in \mathbb{N}$ and $\cup_{n=1}^{\infty} A_n = [0, T] \times \mathcal{Z}$. For every n we chose a partition $\{\Delta_{n,k}\}$, $k = 1, \dots, K_n < \infty$, of A_n such that¹

$$\bigcup_{1 \leq k \leq K_n} \Delta_{n,k} = A_n, \quad (4.3.1)$$

$$\Delta_{n,k} = (t_{n,k}, s_{n,k}] \times Z_{n,k}, \quad 0 \leq t_{n,k} < s_{n,k} \leq T, \quad Z_{n,k} \in \mathcal{B}_{\mathcal{Z}} \quad (4.3.2)$$

$$\max_{1 \leq k \leq K_n} (s_{n,k} - t_{n,k}) < 1/n, \quad (4.3.3)$$

$$\max_{1 \leq k \leq K_n} V(\Delta_{n,k}) < 1/n, \quad (4.3.4)$$

$$\Delta_{n,k} \cap \Delta_{n,j} = \emptyset \text{ for } k \neq j. \quad (4.3.5)$$

Moreover, the partitions are nested in the sense that

$$\Delta_{n,k} \cap \Delta_{n+1,j} = \emptyset \text{ or } \Delta_{n+1,j}. \quad (4.3.6)$$

The non-anticipating stochastic derivative can be represented as the limit [38, Theorem 3.1]

$$\mathcal{D}\xi = \lim_{n \rightarrow \infty} \phi_n \quad (4.3.7)$$

with convergence in $\mathcal{I}$ of the stochastic functions of type (4.2.1) given by

$$\phi_n(t, z) := \sum_{k=1}^{K_n} \mathbb{E} \left[\xi \frac{\mu(\Delta_{n,k})}{\Lambda(\Delta_{n,k})} \middle| \mathcal{G}_{t_{n,k}} \right] \mathbf{1}_{\Delta_{n,k}}(t, z) \quad (4.3.8)$$

where $\Delta_{n,k} = (t_{n,k}, s_{n,k}] \times Z_{n,k}$ refers to the partion of A_n described in (4.3.1)-(4.3.6). We have the following result [38, Theorem 3.1]:

Theorem 4.3.1. *All $\xi \in L_2(\Omega, \mathcal{G}, \mathbb{P})$ have representation*

$$\xi = \xi_0 + \int_0^T \int_{\mathcal{Z}} \mathcal{D}_{t,z} \xi \mu(dt, dz). \quad (4.3.9)$$

Moreover $\mathcal{D}\xi_0 = 0$ and $\xi_0 \in L_2(\Omega, \mathcal{G}, \mathbb{P})$ is orthogonal to space generated by

¹Here it is possible to substitute $1/n$ with any sequence ϵ_n such that $\epsilon_n \rightarrow 0^+$ as $n \rightarrow \infty$.

$$\left\{ \int_0^T \int_{\mathcal{Z}} \phi(s, z) \mu(ds, dz) \mid \phi \in \mathcal{I} \right\}.$$

Indeed, by the orthogonality of the sum in (4.3.9), one can see that the following duality rule is verified: Let $\xi \in L_2(\Omega, \mathcal{G}, \mathbb{P})$ and $\kappa \in \mathcal{I}$, then

$$\begin{aligned} & \mathbb{E} \left[\xi \int_0^T \int_{\mathcal{Z}} \kappa(s, z) \mu(ds, dz) \right] \\ &= \mathbb{E} \left[\left(\xi_0 + \int_0^T \int_{\mathcal{Z}} \mathcal{D}_{s,z} \xi \mu(ds, dz) \right) \int_0^T \int_{\mathcal{Z}} \kappa(s, z) \mu(ds, dz) \right] \\ &= \mathbb{E} \left[\int_0^T \int_{\mathcal{Z}} (\mathcal{D}_{s,z} \xi) \kappa(s, z) \Lambda(ds, dz) \right]. \end{aligned} \quad (4.3.10)$$

4.4 Optimization problem

Define the state process $X_t, t \in [0, T]$ by $X_0 = a \in \mathbb{R}$ and

$$X_t^{(u)} = X_0 + \int_0^t b_s(u_s, X_s) ds + \int_0^t \int_{\mathcal{Z}} \phi_s(z, u_s, X_s) \mu(ds, dz).$$

Here $b : \Omega \times [0, T] \times \mathcal{U} \times \mathbb{R} \rightarrow \mathbb{R}$ and $\phi : \Omega \times [0, T] \times \mathcal{Z} \times \mathcal{U} \times \mathbb{R} \rightarrow \mathbb{R}$ are $\mathbb{G}$ -predictable. Moreover $\phi \in \mathcal{I}$. We assume that X has a unique strong solution and note that X is $\mathbb{G}$ -adapted. The stochastic process u is the control taking values in an open and convex set $\mathcal{U} \subseteq \mathbb{R}^n$.

In the performance functional (4.1.1),

$$J(u) = \mathbb{E} \left[\int_0^T f_t(u_t, X_t) dt + g(X_T) \right], \quad (4.4.1)$$

we have $f : \Omega \times [0, T] \times \mathcal{U} \times \mathbb{R} \rightarrow \mathbb{R}$ and $g : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$. Remark that we have allowed for g and f to depend on additional randomness besides u and X , and assume that they are both measurable.

We assume f and b are continuously differentiable in $x \in \mathbb{R}$ and $u \in \mathcal{U}$ for all $t \in [0, T]$ and almost all $\omega \in \Omega$. We denote these derivatives $\frac{\partial f_s}{\partial x}, \frac{\partial f_s}{\partial u}$, similarly for b and ϕ . Remark that $\frac{\partial f_s}{\partial u} \in \mathbb{R}^n$ since u is n -dimensional. We will denote $\cdot$ as the inner product in $\mathbb{R}^n$ when appropriate. Furthermore g is continuously differentiable with respect to $x \in \mathbb{R}$ a.s., and we denote this derivative by g' .

The random field ϕ is continuously differentiable in $x \in \mathbb{R}$ and $u \in \mathcal{U}$ for almost all $(\omega, t, z) \in \Omega \times [0, T] \times \mathcal{Z}$. We assume that $\frac{\partial \phi}{\partial x} \in \mathcal{I}$ and with $u = (u^1, \dots, u^n) \in \mathbb{R}^n$, $\frac{\partial \phi}{\partial u^j} \in \mathcal{I}$ for $j = 1, \dots, n$. Finally we define the $\mathbb{G}$ -semi-martingale

$$M_s := \int_0^s \frac{\partial b_r}{\partial x}(u_r, X_r) dr + \int_0^s \int_{\mathcal{Z}} \frac{\partial \phi_r}{\partial x}(u_r, X_r) \mu(dr, dz), \quad s \in [0, T].$$

The first variation process $G_s(t)$, $s \in [0, T]$, is the solution to the equation

$$\begin{aligned} G_s(t) &:= 1 + \int_t^s G_r(t) dM_r, \quad s \in [t, T], \\ &= 1 + \int_t^s G_r(t) \frac{\partial b_r}{\partial x}(u_r, X_r) dr + \int_t^s \int_{\mathcal{Z}} G_r(t) \frac{\partial \phi_r}{\partial x}(z, u_r, X_r) \mu(dr, dz). \end{aligned} \quad (4.4.2)$$

The solution of (4.4.2) is given as follows ([92, Theorem II.37])

$$G_s(t) = \exp \left\{ M_s(t) - \frac{1}{2} [M(t), M(t)]_s \right\} \prod_{t < s \leq T} (1 + \Delta M_s(t)) \exp \{-\Delta M_s(t)\}$$

where for any t , $M(t)$ is the $\mathbb{G}$ -semi-martingale defined by $M_s(t) = \int_t^s dM_r$ for $t < s \leq T$ and $M_s(t) = 0$ for $s \leq t$. Furthermore we define, where $t \in [0, T]$,

$$K_t := K_t^{(u, X)} = g'(X_T) + \int_t^T \frac{\partial f_s}{\partial x}(u_s, X_s) ds, \quad (4.4.3)$$

$$\mathcal{D}_{t,z} K_t := \mathcal{D}_{t,z} g'(X_T) + \mathcal{D}_{t,z} \left(\int_t^T \frac{\partial f_s}{\partial x}(u_s, X_s) ds \right), \quad (4.4.4)$$

$$F_t(u, X_t) = K_t \frac{\partial b_t}{\partial x}(u_t, X_t) + \int_{\mathcal{Z}} (\mathcal{D}_{t,z} K_t) \frac{\partial \phi_t}{\partial x}(z, u_t, X_t) \lambda_t(dz), \quad (4.4.5)$$

$$p_t := p_t^{(u, X)} = K_t + \int_t^T F_s(u_s, X_s) G_s(t) ds, \quad (4.4.6)$$

$$\kappa_t := \kappa_t^{(u, X)} = \mathcal{D}_{t,z} p_t. \quad (4.4.7)$$

In order to have the above quantities well-defined the following requirements are needed:

Assumption 4.4.1. The control u with state process $X^{(u)}$ satisfies

$$\mathbb{E}[g'(X_T)^2] < \infty, \quad (4.4.8)$$

$$\mathbb{E} \left[\int_0^T \frac{\partial f_t}{\partial x}(u_t, X_t)^2 dt \right] < \infty, \quad (4.4.9)$$

$$\mathbb{E} \left[\int_t^T (F_s G_s(t))^2 ds \right] < \infty, \quad \text{for all } t \in [0, T]. \quad (4.4.10)$$

Remark 4.4.2. Suppose μ is a Brownian or Lévy noise and $\mathbb{G}$ is generated by μ with $\mathcal{G} = \mathcal{G}_T$. If using the duality relation of Malliavin calculus (4.4.8)-(4.4.9)-(4.4.10) would be stated in terms of Malliavin differentiability, see [82, Equation 3.5]. Meaning that both $g'(X_T)$ and

$\int_t^T (F_s G_s(t))^2 ds$ need to be in the domain of the Malliavin derivative, a space strictly smaller than $L_2(\Omega, \mathcal{G}, \mathbb{P})$. In addition, (4.4.9) would be replaced by the Malliavin differentiability of $\frac{\partial f_t}{\partial x}(u_t, X_t)$ and the integrability of $D_t \frac{\partial f_t}{\partial x}(u_t, X_t)$ so that $\int_0^T D_t \frac{\partial f_t}{\partial x}(u_t, X_t) dt$ is well defined (where D is the Malliavin derivative) since the arguments in the forthcoming (4.5.12) does not apply.

For a given control u with state process $X = X^{(u)}$, we define the Hamiltonian by

$$\begin{aligned} \mathcal{H}_t(v, x) &= \mathcal{H}_t^{(u, X)}(v, x) \\ &:= f_t(v, x) + b_t(v, x) p_t^{(u, X)} + \int_{\mathcal{Z}} \kappa_t^{(u, X)}(z) \phi_t(z, v, x) \lambda_t(dz), \end{aligned} \quad (4.4.11)$$

where $t \in [0, T]$, $v \in \mathcal{U}$ and $x \in \mathbb{R}$.

4.5 Maximum principle

Let $\mathbb{F} := \{\mathcal{F}_t, t \in [0, T]\}$ be a right continuous filtration such that $\mathcal{F}_t \subseteq \mathcal{G}_t$ for all $t \in [0, T]$. We state the optimization result for $\mathbb{F}$, naturally we can have $\mathbb{F} = \mathbb{G}$.

Definition 4.5.1. We say that u is an admissible control if $u : \Omega \times [0, T] \rightarrow \mathcal{U}$ is $\mathbb{F}$ -predictable, Assumption 4.4.1 holds and

$$\mathbb{E} \left[\int_0^T f_t(u_t, X_t)^2 dt + g(X_T)^2 \right] < \infty. \quad (4.5.1)$$

We denote the set of admissible controls by $\mathcal{A}^{\mathbb{F}}$.

The following assumption is needed on controls on which we apply the maximum principle.

Assumption 4.5.2. Let $u \in \mathcal{A}^{\mathbb{F}}$ be fixed. For this u we assume that for any $\mathbb{F}$ -predictable and bounded process β satisfying

$$u_t - \beta_t \in \mathcal{U} \quad \text{and} \quad u_t + \beta_t \in \mathcal{U} \quad dt \times d\mathbb{P} \text{ a.e.} \quad (4.5.2)$$

there exists a $\delta > 0$ such that

A1) $u + y\beta \in \mathcal{A}^{\mathbb{F}}$ for all $|y| \leq \delta$.

A2) The family

$$\left\{ \frac{\partial f_t}{\partial x}(u_t + y\beta_t, X_t^{u+y\beta}) \frac{\partial}{\partial y} X_t^{u+y\beta} + \frac{\partial f_t}{\partial u}(u_t + y\beta_t, X_t^{u+y\beta}) \beta_t \right\}_{y \in (-\delta, \delta)} \quad (4.5.3)$$

is uniformly $dt \times d\mathbb{P}$ -integrable, and the family

$$\left\{ g'(X_T^{u+y\beta}) \frac{\partial}{\partial y} X_T^{u+y\beta} \right\}_{y \in (-\delta, \delta)} \quad (4.5.4)$$

is uniformly $\mathbb{P}$ -integrable.

A3) The process $Y_t^{(u,\beta)} = \frac{\partial}{\partial y} X_t^{u+y\beta}|_{y=0}$ exists as an element of $L_2(\Omega, \mathcal{G}, \mathbb{P})$ for all $t \in [0, T]$ and satisfies

$$\begin{aligned} Y_t &= Y_t^{(u,\beta)} = \frac{\partial}{\partial y} X_t^{u+y\beta}|_{y=0} \\ &= \int_0^t \left[\frac{\partial b_s}{\partial x}(u_s, X_s) Y_s + \frac{\partial b_s}{\partial u}(u_s, X_s) \cdot \beta_s \right] ds \\ &\quad + \int_0^t \int_{\mathcal{Z}} \left[\frac{\partial \phi_s}{\partial x}(z, u_s, X_s) Y_s + \frac{\partial \phi_s}{\partial u}(z, u_s, X_s) \cdot \beta_s \right] \mu(ds, dz). \end{aligned} \quad (4.5.5)$$

In a converse conclusion in the forthcoming maximum principle, we will also require the following assumption:

Assumption 4.5.3.

A4) If α is a random variable taking values in $\mathcal{U}$ a.s. then (with $0 \leq t < r \leq T$)

$$u_s(\omega) = \alpha(\omega) \mathbf{1}_{(t,r]}(s),$$

is an admissible control (i.e. $u \in \mathcal{A}^{\mathcal{F}}$).

A control $\hat{u} \in \mathcal{A}^{\mathcal{F}}$ is a “local maximum” if

$$J(\hat{u}) \geq J(\hat{u} + y\beta), \quad |y| \leq \delta, \quad (4.5.6)$$

for all bounded $\mathbb{F}$ -predictable β and some $\delta > 0$ that may depend on β . Meaning that we cannot improve $J(\hat{u})$ by making “bounded” perturbations of $\hat{u}$. Thus any solution to (4.4.1), $J(\hat{u}) = \sup_{u \in \mathcal{A}^{\mathcal{F}}} J(u)$, must also be a local maximum. If $\hat{u}$ is a local maximum, we must naturally have

$$\frac{\partial}{\partial y} J(\hat{u} + y\beta)|_{y=0} = 0. \quad (4.5.7)$$

The converse conclusions are not true however. Not every u satisfying (4.5.7) is a local maximum and a local maximum is not necessarily the optimal solution to (4.4.1).

Theorem 4.5.4. *Let $\hat{u}$ be an admissible control and suppose $\hat{u}$ satisfies Assumption 4.5.2. Denote*

$$\begin{aligned} \hat{X}_t &= X_t^{(\hat{u})} \\ \hat{\mathcal{H}}_t(v, \hat{X}_t) &= f_t(v, \hat{X}_t) + b_t(\lambda_t, v, \hat{X}_t) \hat{p}_t + \\ &\quad + \int_{\mathcal{Z}} \hat{\kappa}_t(z) \phi_t(z, v, \hat{X}_t) \lambda_t(dz), \quad v \in \mathcal{U} \subseteq \mathbb{R}, \end{aligned}$$

with

$$\begin{aligned}\hat{p}_t &= p_t^{(\hat{u}, \hat{X})}, \\ \hat{\kappa}_t &= \kappa_t^{(\hat{u}, \hat{X})}.\end{aligned}$$

If $\hat{u}$ is a critical point for $J(u)$, in the sense that

$$\frac{\partial}{\partial y} J(\hat{u} + y\beta) \Big|_{y=0} = 0$$

for all bounded, $\mathbb{F}$ -predictable processes β such that $\hat{u}_t \pm \beta_t \in \mathcal{U}$ $dt \times d\mathbb{P}$ -a.e., then

$$\mathbb{E} \left[\frac{\partial \mathcal{H}_t}{\partial u}(\hat{u}_t, \hat{X}_t) \Big| \mathcal{F}_t \right] = 0, \quad dt \times d\mathbb{P}\text{-a.e.} \quad (4.5.8)$$

If also Assumption 4.5.3 hold then the converse is also true: If $\hat{u}$ satisfies (4.5.8) then $\hat{u}$ is a critical point.

For ease of notation we use the short hand notation $b_s = b_s(\hat{u}_s, \hat{X}_s)$, $f_s = f_s(\hat{u}_s, \hat{X}_s)$, and similarly for the other coefficients. We proceed using similar arguments as done in [82] with Malliavin derivatives.

Proof. Suppose $\hat{u}$ is a critical point. Then

$$\begin{aligned}0 &= \frac{\partial}{\partial y} J(\hat{u} + y\beta) \Big|_{y=0} \\ &= \mathbb{E} \left[\int_0^T \frac{\partial f_s}{\partial x} Y_s + \frac{\partial f_s}{\partial u} \cdot \beta_s ds + g'(X_T) Y_T \right].\end{aligned} \quad (4.5.9)$$

By the duality formula (4.3.10) (and (4.4.8))

$$\begin{aligned}&\mathbb{E} [g'(X_T) Y_T] \\ &= \mathbb{E} \left[\int_0^T g'(X_T) \left[\frac{\partial b_s}{\partial x} Y_s + \frac{\partial b_s}{\partial u} \cdot \beta_s \right] ds \right. \\ &\quad \left. + \int_0^T \int_{\mathcal{Z}} [(\mathcal{D}_{s,z} g'(X_T)) \left(\frac{\partial \phi_s}{\partial x}(z) Y_s + \frac{\partial \phi_s}{\partial u}(z) \cdot \beta_s \right)] \Lambda(ds, dz) \right].\end{aligned} \quad (4.5.10)$$

By the Fubini theorem and the duality formula (4.3.10) (with integrability ensured by (4.5.3))

and the non-anticipating stochastic derivative is well defined by (4.4.9))

$$\begin{aligned}
& \mathbb{E} \left[\int_0^T \frac{\partial f_t}{\partial x} Y_t dt \right] \\
&= \int_0^T \mathbb{E} \left[\frac{\partial f_t}{\partial x} \left[\int_0^t \frac{\partial b_s}{\partial x} Y_s + \frac{\partial b_s}{\partial u} \cdot \beta_s ds \right] \right. \\
&\quad \left. + \frac{\partial f_t}{\partial x} \left[\int_0^t \int_{\mathcal{Z}} \left(\frac{\partial \phi_s}{\partial x} Y_s + \frac{\partial \phi_s}{\partial u} \cdot \beta_s \right) \mu(ds, dz) \right] \right] dt \\
&= \mathbb{E} \left[\int_0^T \left\{ \frac{\partial f_t}{\partial x} \left[\int_0^t \frac{\partial b_s}{\partial x} Y_s + \frac{\partial b_s}{\partial u} \cdot \beta_s ds \right] \right. \right. \\
&\quad \left. \left. + \left[\int_0^t \int_{\mathcal{Z}} \left(\mathcal{D}_{s,z} \frac{\partial f_t}{\partial x} \right) \left(\frac{\partial \phi_s}{\partial x} Y_s + \frac{\partial \phi_s}{\partial u} \cdot \beta_s \right) \Lambda(ds, dz) \right] \right\} dt \right] \\
&= \mathbb{E} \left[\int_0^T \left[\int_t^T \frac{\partial f_s}{\partial x} ds \right] \left(\frac{\partial b_t}{\partial x} Y_t + \frac{\partial b_t}{\partial u} \cdot \beta_t \right) dt \right. \\
&\quad \left. + \int_0^T \int_{\mathcal{Z}} \left[\int_t^T \mathcal{D}_{t,z} \frac{\partial f_s}{\partial x} ds \right] \left(\frac{\partial \phi_t}{\partial x} Y_t + \frac{\partial \phi_t}{\partial u} \cdot \beta_t \right) \Lambda(dt, dz) \right]. \tag{4.5.11}
\end{aligned}$$

By the continuity of $\mathcal{D}$ [38, Remark 3.4] and with sufficient integrability from (4.4.9) we have

$$\int_t^T \mathcal{D}_{t,z} \frac{\partial f_s}{\partial x} ds = \mathcal{D}_{t,z} \int_t^T \frac{\partial f_s}{\partial x} ds, \quad d\Lambda \times d\mathbb{P} \text{ a.e.} \tag{4.5.12}$$

We recall (4.4.3), (4.4.4), and by (4.5.9)-(4.5.10)-(4.5.11) conclude that

$$\begin{aligned}
& \mathbb{E} \left[\int_0^T K_s \left(\frac{\partial b_s}{\partial x} Y_s + \frac{\partial b_s}{\partial u} \cdot \beta_s \right) + \frac{\partial f_s}{\partial u} \cdot \beta_s ds \right. \\
&\quad \left. + \int_0^T \int_{\mathcal{Z}} (\mathcal{D}_{s,z} K_s) \left(\frac{\partial \phi_s}{\partial x} Y_s + \frac{\partial \phi_s}{\partial u} \cdot \beta_s \right) \Lambda(ds, dz) \right] = 0. \tag{4.5.13}
\end{aligned}$$

Let $\alpha = (0, \dots, \alpha^{(j)}, \dots, 0)$, be a random variable in $\mathbb{R}^n$ which is zero except at the index j , where $1 \leq j \leq n$. Set

$$\beta_s = \alpha \mathbf{1}_{(t, t+h]}(s) = (0, \dots, \alpha^{(j)}, \dots, 0) \mathbf{1}_{(t, t+h]}(s)$$

We assume $\alpha^{(j)}$ is bounded, $\mathcal{F}_t$ -measurable and such that, $u_t \pm \beta_t$ takes values in $\mathcal{U}$ $dt \times d\mathbb{P}$ a.e.

Then $Y_s = Y_s^{(u,\beta)} = 0$ for $s < t$ so that (4.5.13) can be rewritten as

$$A_1 + A_2 = 0 \quad (4.5.14)$$

where

$$\begin{aligned} A_1 &= \mathbb{E} \left[\int_t^T K_s \frac{\partial b_s}{\partial x} Y_s ds + \int_t^T \int_{\mathcal{Z}} (\mathcal{D}_{s,z} K_s) \frac{\partial \phi_s}{\partial x} Y_s \Lambda(ds, dz) \right], \\ A_2 &= \mathbb{E} \left[\alpha \cdot \left(\int_t^{t+h} \left[K_s \frac{\partial b_s}{\partial u} + \frac{\partial f_s}{\partial u} \right] ds + \int_t^{t+h} \int_{\mathcal{Z}} (\mathcal{D}_{s,z} K_s) \frac{\partial \phi_s}{\partial u} \Lambda(ds, dz) \right) \right]. \end{aligned}$$

From (4.4.5)

$$\begin{aligned} A_1 &= \mathbb{E} \left[\int_t^T F_s Y_s ds \right] \\ &= \int_t^{t+h} \mathbb{E} [F_s Y_s] ds + \int_{t+h}^T \mathbb{E} [F_s Y_s] ds. \end{aligned}$$

Since Y admits a càdlàg representative and $Y_t = 0$ we have

$$\frac{\partial}{\partial h} \int_t^{t+h} \mathbb{E} [F_s Y_s] ds \Big|_{h=0} = 0.$$

Recall (4.5.5) and (4.4.2). We have

$$Y_s = Y_{t+h} G_s(t+h) \quad \text{for } s \geq t+h.$$

Since $Y_t = 0$ (interchange of integration and expectation justified by (4.5.3), (4.5.4))

$$\begin{aligned} \frac{\partial}{\partial h} A_1 \Big|_{h=0} &= \frac{\partial}{\partial h} \int_{t+h}^T \mathbb{E} [F_s Y_s] ds \Big|_{h=0} \\ &= \int_t^T \frac{\partial}{\partial h} \left\{ \mathbb{E} [F_s Y_{t+h} G_s(t+h)] \right\} ds \Big|_{h=0} - F_t Y_t \\ &= \int_t^T \mathbb{E} \left[F_s \left\{ Y_{t+h} \frac{\partial}{\partial h} G_s(t+h) + G_s(t+h) \frac{\partial}{\partial h} Y_{t+h} \right\} \right] ds \Big|_{h=0} \\ &= \int_t^T \frac{\partial}{\partial h} \mathbb{E} [F_s G_s(t) Y_{t+h}] \Big|_{h=0} ds. \end{aligned}$$

By (4.5.5) we have

$$Y_{t+h} = \alpha \cdot \left(\int_t^{t+h} \frac{\partial b_s}{\partial u} ds + \int_t^{t+h} \int_{\mathcal{Z}} \frac{\partial \phi_s}{\partial u} \mu(ds, dz) \right) \\ + \int_t^{t+h} Y_s \frac{\partial b_s}{\partial x} ds + \int_t^{t+h} \int_{\mathcal{Z}} Y_s \frac{\partial \phi_s}{\partial x} \mu(ds, dz).$$

Denote $\frac{\partial}{\partial h} A_1|_{h=0} = B_1 + B_2$ with

$$B_1 = \int_t^T \frac{\partial}{\partial h} \mathbb{E} \left[F_s G_s(t) \left\{ \alpha \cdot \left(\int_t^{t+h} \frac{\partial b_r}{\partial u} dr + \int_t^{t+h} \int_{\mathcal{Z}} \frac{\partial \phi_r}{\partial u} \mu(dr, dz) \right) \right\} \right] \Big|_{h=0} ds, \\ B_2 = \mathbb{E} \left[\int_t^T \frac{\partial}{\partial h} \mathbb{E} \left[F_s G_s(t) \left\{ \int_t^{t+h} Y_r \frac{\partial b_r}{\partial x} dr + \int_t^{t+h} \int_{\mathcal{Z}} Y_r \frac{\partial \phi_r}{\partial x} \mu(dr, dz) \right\} \right] \Big|_{h=0} ds. \right.$$

By the duality formula (4.3.10) (well defined by (4.4.10))

$$B_1 = \int_t^T \frac{\partial}{\partial h} \mathbb{E} \left[F_s G_s(t) \left\{ \alpha \cdot \left(\int_t^{t+h} \frac{\partial b_r}{\partial u} dr + \int_t^{t+h} \int_{\mathcal{Z}} \frac{\partial \phi_r}{\partial u} \mu(dr, dz) \right) \right\} \right] \Big|_{h=0} ds \\ = \int_t^T \frac{\partial}{\partial h} \mathbb{E} \left[\left\{ \alpha \cdot \left(\int_t^{t+h} F_s G_s(t) \frac{\partial b_r}{\partial u} dr \right. \right. \right. \\ \left. \left. \left. + \int_t^{t+h} \int_{\mathcal{Z}} \mathcal{D}_{r,z}(F_s G_s(t)) \frac{\partial \phi_r}{\partial u} \lambda_r(dz) dr \right) \right\} \right] \Big|_{h=0} ds \\ = \int_t^T \mathbb{E} \left[\left\{ \alpha \cdot \left(F_s G_s(t) \frac{\partial b_t}{\partial u} + \int_{\mathcal{Z}} \mathcal{D}_{t,z}(F_s G_s(t)) \frac{\partial \phi_t}{\partial u} \lambda_t(dz) \right) \right\} \right] ds. \quad (4.5.15)$$

By the duality formula (4.3.10) (well defined by (4.4.10)) and since $Y_t = 0$ we have

$$B_2 = \int_t^T \frac{\partial}{\partial h} \mathbb{E} \left[F_s G_s(t) \left\{ \int_t^{t+h} Y_r \frac{\partial b_r}{\partial x} dr + \int_t^{t+h} \int_{\mathcal{Z}} Y_r \frac{\partial \phi_r}{\partial x} \mu(dr, dz) \right\} \right] \Big|_{h=0} ds \\ = \int_t^T \mathbb{E} \left[\frac{\partial}{\partial h} \left\{ \int_t^{t+h} F_s G_s(t) Y_r \frac{\partial b_r}{\partial x} dr \right. \right. \\ \left. \left. + \int_t^{t+h} \int_{\mathcal{Z}} \mathcal{D}_{r,z}(F_s G_s(t)) Y_r \frac{\partial \phi_r}{\partial x} \lambda_r(dz) dr \right\} \right] \Big|_{h=0} ds$$

$$\begin{aligned}
&= \int_t^T \mathbb{E} \left[\left\{ F_s G_s(t) Y_t \frac{\partial b_t}{\partial x} + \int_z \mathcal{D}_{r,z}(F_s G_s(t)) Y_t \frac{\partial \phi_t}{\partial x} \lambda_t(dz) \right\} \right] ds \\
&= 0.
\end{aligned} \tag{4.5.16}$$

We see immediately that (interchange of derivation and expectation justified by (4.5.3)-(4.5.4))

$$\frac{\partial}{\partial h} A_2 \Big|_{h=0} = \mathbb{E} \left[\alpha \cdot \left(K_t \frac{\partial b_t}{\partial u} + \frac{\partial f_t}{\partial u} + \int_z (\mathcal{D}_{t,z} K_t) \frac{\partial \phi_t}{\partial u} \lambda_t(dz) \right) \right]. \tag{4.5.17}$$

Recall that $\frac{\partial}{\partial h} A_1 = B_1 + B_2$ and the definition of p in (4.4.6). By (4.5.15)-(4.5.16)-(4.5.17) we have

$$\begin{aligned}
\frac{\partial}{\partial h} \{A_1 + A_2\}_{h=0} &= \mathbb{E} \left[\alpha \cdot \left\{ \frac{\partial f_t}{\partial u} + p_t \frac{\partial b_t}{\partial u} + \int_z (\mathcal{D}_{t,z} p_t) \frac{\partial \phi_t}{\partial u}(z) \lambda_t(dz) \right\} \right] \\
&= \mathbb{E} \left[\alpha \cdot \frac{\partial \mathcal{H}_t}{\partial u}(\hat{X}_t, \hat{u}_t) \right].
\end{aligned} \tag{4.5.18}$$

As a function of h , $A_1(h) + A_2(h) = 0$ for all $0 \leq h \leq T - t$ by (4.5.14). Hence $\frac{\partial}{\partial h} \{A_1(h) + A_2(h)\} = 0$ and thus

$$0 = \mathbb{E} \left[\alpha \cdot \frac{\partial \mathcal{H}_t}{\partial u}(\hat{X}_t, \hat{u}_t) \right] = \mathbb{E} \left[\alpha^{(j)} \frac{\partial \mathcal{H}_t}{\partial u^{(j)}}(\hat{X}_t, \hat{u}_t) \right].$$

Recall that here $u_t + \alpha$ is a $\mathcal{F}_t$ -measurable random variable taking values in $\mathcal{U}$ a.s. Define

$$D(\omega) = \sup_{c \in \mathbb{R}} \{u_t(\omega) + c \in \mathcal{U} \text{ and } u_t(\omega) - c \in \mathcal{U}\} \wedge 1.$$

Here D “measures the minimum distance” between $u^{(j)}$ and $\mathcal{U}$ ω -wise. Note that $0 < D \leq 1$ a.s. Let ζ be a $\mathcal{F}_t$ -measurable random variable bounded by $C > 0$. Then

$$u_t + \frac{1}{2C} \zeta D \in \mathcal{U}, \quad \text{a.s.}$$

We take $\alpha^{(j)} = \frac{1}{2C} \zeta D$ and from (4.5.18) get

$$\mathbb{E} \left[\frac{1}{2C} \zeta D \frac{\partial \mathcal{H}_t}{\partial u^{(j)}}(\hat{X}_t, \hat{u}_t) \right] = 0.$$

We multiply by $2C$ to find

$$\mathbb{E} \left[\zeta D \frac{\partial \mathcal{H}_t}{\partial u^{(j)}}(\hat{X}_t, \hat{u}_t) \right] = 0.$$

Let $\zeta^{(m)} = \left(\frac{1}{\delta} \zeta\right) \wedge m$. Then $\zeta^{(m)} \rightarrow \zeta$ when $m \rightarrow \infty$ a.s. and we must have

$$\mathbb{E} \left[\zeta \frac{\partial \mathcal{H}_t}{\partial u^{(j)}}(\hat{X}_t, \hat{u}_t) \right] = 0.$$

Since this holds for all $\mathcal{F}_t$ measurable ζ we conclude

$$\mathbb{E} \left[\frac{\partial \mathcal{H}_t}{\partial u^{(j)}}(\hat{X}_t, \hat{u}_t) \middle| \mathcal{F}_t \right] = 0. \quad (4.5.19)$$

The proof for the sufficient condition is complete as (4.5.19) holds for all $1 \leq j \leq n$.

Conversely, suppose (4.5.8). By reversing the above argument we get that (4.5.14) holds for all $\beta \in \mathcal{A}^{\mathcal{F}}$ of the form

$$\beta(s, \omega) = \alpha(\omega) \mathbf{1}_{(t, t+h]}(s),$$

where the random variable α is $\mathcal{F}_t$ -measurable, bounded and such that $u \pm \beta$, takes values in $\mathcal{U}$ $dt \times d\mathbb{P}$ a.e. Here $0 \leq t < t+h \leq T$. Hence (4.5.14) holds for all linear combinations of such β . Since any $\beta \in \mathcal{A}^{\mathcal{F}}$ can be approximated by such linear combinations it follows that (4.5.14) holds for all bounded $\beta \in \mathcal{A}^{\mathcal{F}}$. $\square$

4.6 A remark on the technique used

In this paper, the maximum principle relies on evaluating

$$\frac{d}{dy} J(u + y\beta) \quad (4.6.1)$$

where J is the performance functional (4.1.1). Here u is the control which is a “candidate” to be an optimal solution, $y \in \mathbb{R}$ and β is a perturbation of u . In this Section we discuss a technical point in how this technique has been presented in the literature, because some frequently used conditions have implications on how we can choose $\mathcal{U}$ (the space where the controls are taking their values). In several papers, e.g. [2, 4, 6, 82, 80, 81, 87], that evaluate (4.6.1) (for performance functionals of type (4.4.1), but with different assumptions on the noises) the following four assumptions are standard:

- i) The admissible controls u take values in an open, convex set $U \subseteq \mathbb{R}^n$.
- ii) The admissible controls satisfy some integrability conditions related to the problem and the corresponding state-process (given by a SDE) has a unique strong solution.
- iii) For all bounded and $\tilde{\mathcal{F}}_t$ -measurable random variables α , the control

$$u_s(\omega) = \alpha(\omega) \mathbf{1}_{(t, t+r]}(s), \quad 0 \leq t < t+r \leq T,$$

is admissible². Here $\tilde{\mathbb{F}} = \{\tilde{\mathcal{F}}_t, t \in [0, T]\}$ is an filtration relevant to the optimization problem.

- iv) If u and β are admissible controls, with β bounded, there exist $\delta > 0$ such that $u + y\beta$ is also an admissible control for all $|y| < \delta$.

For convenience we only discuss the case when $n = 1$ in Condition i). However the issue presented here can easily be generalized to any $n > 1$.

²In [6, 81] it is only assumed that α takes values in U .

Condition iii) implies that all the constants are elements of U , since $C\mathbf{1}_{(t,t+r]}(s)$, $C \in \mathbb{R}$ must be an admissible control. This can only be satisfied if $U = \mathbb{R}$. Meaning that U cannot be taken to be any open, convex set as described in i), but it is necessary that $U = \mathbb{R}$ for the maximum principle to apply.

We could attempt to change Condition iii) to

iii') For all bounded and $\tilde{\mathcal{F}}_t$ -measurable random variables α such that $\alpha \in U$ a.s., the control

$$u_s(\omega) = \alpha(\omega)\mathbf{1}_{(t,t+r]}(s) \quad 0 \leq t < t+r \leq T,$$

is admissible.

However, Condition iv) is still a problem. To explain, suppose $U = (c_1, c_2)$ for some $c_1 < c_2$ and let α be a bounded, $\tilde{\mathcal{F}}_t$ -measurable random variable taking values in U . If condition iii') holds, both $u_s(\omega) := \alpha(\omega)\mathbf{1}_{(t,t+r]}(s)$ and $\beta_s(\omega) := C\mathbf{1}_{(t,t+r]}(s)$, $C \in (c_1, c_2)$, are admissible controls. Even if the random variable α satisfies $\alpha < c_2$ a.s. we can have

$$\text{ess sup } \alpha = c_2$$

and thus $u_t + y\beta_t \in U$ a.s. is *not* possible for any $y > 0$. Hence $u_t + y\beta_t$ is not an admissible control for any $y > 0$, as it is not taking values in U , and Condition iv) fails.

The use of the “standard” assumptions i)-ii)-iii)-iv) is not a major issue, the resulting maximum principle will hold for $U = \mathbb{R}$. Indeed the technical conditions are correct even if opaque. Moreover, if one is only interested in bounded controls one can apply the maximum principle and then check whether the resulting control is in fact bounded. There will however be a problem, at least formally, if integrability conditions or other conditions (i.e. ii)) on the admissible controls require them to take values in a bounded set. Also, the study of the control problem with U bounded has independent interest. As an example, in the forthcoming Proposition 4.8.1 we show additional results on the uniqueness of the solution when U is bounded. Hence we used Assumption 4.5.2 in the maximum principle, Theorem 4.5.4.

4.7 Lévy Processes

Here we compare the non-anticipating stochastic derivative with the Malliavin operator in the case of Lévy processes. We refer to e.g. [44, 43, 83, 101] for a comprehensive treatment on Malliavin calculus on Lévy processes, or more precisely the mixture of Gaussian and Poisson random measures. In short, the stochastic non-anticipating derivative coincide with the projection of the Malliavin derivative D

$$\mathcal{D}_{s,z}\xi = \mathbb{E}[D_{s,z}\xi|\mathcal{G}_s] \quad (4.7.1)$$

whenever the right hand side is well defined. But the domain of the Malliavin derivative D is $\mathbb{D}_{1,2}$, a space strictly smaller than $L_2(\Omega, \mathcal{G}, \mathbb{P})$.

Let N be a Poisson random field on $[0, T] \times \mathbb{R}_0$ with $\mathbb{E}[N(dt, dz)] = \nu(dz)dt$ and denote $\tilde{N}(dt, dz) = N(dz, dz) - \nu(dz)dt$ as the centered Poisson random field. Furthermore

$B_t, t \in [0, T]$ is a Brownian motion. With $\mathbb{G}$ as the (completed) filtration generated by B and N , let the martingale random field μ be given by :

$$\mu(dt, dz) = \mathbf{1}_{\{0\}}(z)dB_t + \mathbf{1}_{\mathbb{R}_0}(z)\tilde{N}(dt, dz),$$

with $\lambda_t(dz) = \mathbf{1}_{\{0\}}(z) + \mathbf{1}_{\mathbb{R}_0}(z)\nu(dz)$. Then the process η defined by

$$\eta_t = B_t + \int_0^t \int_{\mathbb{R}_0} z \tilde{N}(ds, dz), \quad t \in [0, T],$$

is a Lévy process. If we assume that $(\Omega, \mathcal{G}, \mathbb{P}) = (\Omega, \mathcal{G}_T, \mathbb{P})$, i.e. that our probability space can be generated by B and N , then there exists a subspace $\mathbb{D}_{1,2} := \mathbb{D}_{1,2}^B \cap \mathbb{D}_{1,2}^N \subsetneq L_2(\Omega, \mathcal{G}, \mathbb{P})$ such that for any $\xi \in \mathbb{D}_{1,2}$ [11, Theorem 3.11]

$$\xi = \mathbb{E}[\xi] + \int_0^T \mathbb{E}[D_s \xi | \mathcal{G}_s] dB_s + \int_0^T \int_{\mathbb{R}_0} \mathbb{E}[D_{s,z} \xi | \mathcal{G}_s] \tilde{N}(ds, dz) \quad (4.7.2)$$

where D_s and $D_{s,z}$ are the Mallavin derivatives for the Brownian motion and Poisson random field with domains $\mathbb{D}_{1,2}^B$ and $\mathbb{D}_{1,2}^N$ respectively. This is also known as the Clark-Ocone formula. Furthermore $D_s \xi \in L_2(\Omega \times [0, T])$ and $D_{s,z} \in L_2(\Omega \times [0, T] \times \mathbb{R}_0)$.

Remark that the Clark-Ocone formula (4.7.2) can be extended to $L_2(\Omega, \mathcal{G}, \mathbb{P})$ in the setting of white noise and Hida-Malliavin derivatives, see [1, 42]. However this requires further assumptions on the probability space $(\Omega, \mathcal{G}, \mathbb{P})$.

4.8 Application to default risk

Here we show an application of the maximum principle to portfolio optimization. We choose a setting outside Lévy processes that has independent interest: Assets with credit risk modeled by doubly stochastic Poisson processes. Credit risk with doubly stochastic Poisson processes has been widely studied in the literature, see e.g. [59, 74, 46].

Let $\lambda_s = (\lambda_s^{(1)}, \dots, \lambda_s^{(n)})$, $s \in [0, T]$, be a positive, stochastic process in $\mathbb{R}^n$. Let $\Lambda_t^{(j)} = \int_0^t \lambda_s^{(j)} ds$, and denote the filtration generated by λ as $\mathbb{F}^\Lambda = \{\mathcal{F}_t^\Lambda, t \in [0, T]\}$. No assumptions of independence are required between $\Lambda^{(j)}$ and $\Lambda^{(k)}$ for any $j \neq k$.

The n -dimensional pure jump process $H_s = (H_s^{(1)}, \dots, H_s^{(n)})$ is a doubly stochastic Poisson process if, when conditioned on the λ 's, it is Poisson distributed. We assume that

$$\mathbb{P}(H_t^{(j)} = k | \mathcal{F}_T^\Lambda) = \mathbb{P}(H_t^{(j)} = k | \Lambda_t^{(k)}) = \frac{(\Lambda_t^{(j)})^k}{k!} e^{-\Lambda_t^{(j)}}$$

for all $1 \leq j \leq n$ and $k \in \mathbb{N}$. Let $\tilde{H}_t := H_t - \Lambda_t$, $t \in [0, T]$ and $\mathbb{F} := \{\mathcal{F}_t, t \in [0, T]\}$ the filtration generated by $\tilde{H}$. Let $\mathbb{G}$ be such that $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{F}_t^\Lambda$ and $\mathcal{Z} = \{1, \dots, n\}$, where $\mathcal{Z}$ is equipped with the discrete topology. Then μ defined by $\mu(dt, z) = d\tilde{H}_t^{(z)}$ is a martingale random field with respect to both $\mathbb{F}$ and $\mathbb{G}$ on $[0, T] \times \mathcal{Z}$.

The filtration $\mathbb{G}$ contains anticipating information in knowing the future values of λ . It is

however a natural candidate for computing the non-anticipating stochastic derivative for doubly stochastic Poisson processes, which has been studied in Chapter 2. Computational rules of Malliavin type can also be found in [107].

Let $\tau^{(z)}$ be the first jump of $H^{(z)}$, $z = 1, \dots, n$. We model each asset $S^{(z)}$ with a return $\rho^{(z)} + \lambda^{(z)}$ up to the time of default τ . In the case of default the asset $S^{(z)}$ become worthless, i.e. $S_{\tau^{(z)}}^{(z)} = 0$ (whenever $\tau^{(z)} < T$). The goal of the investor is to invest in the n assets maximizing expected utility of the wealth at terminal time T . In mathematical terms: Let

$$\begin{aligned} S_t^{(1)} &= S_{t-}^{(1)} \mathbf{1}_{\{\tau^{(1)} > t\}}(t) (\rho_t^{(1)} dt - d\tilde{H}_t^{(1)}), \\ &\vdots \\ S_t^{(n)} &= S_{t-}^{(n)} \mathbf{1}_{\{\tau^{(n)} > t\}}(t) (\rho_t^{(n)} dt - d\tilde{H}_t^{(n)}). \end{aligned}$$

Let X denote the total wealth of the investor and the control u denote the amount invested in the n assets:

$$X_t = \int_0^t \sum_{z=1}^n \mathbf{1}_{\{\tau^{(z)} > r\}}(r) u_r^{(z)} \rho_r^{(z)} dr - \int_0^t \sum_{z=1}^n \mathbf{1}_{\{\tau^{(z)} > r\}}(r) u_r^{(z)} d\tilde{H}_t^{(j)}$$

Remark that every asset $S^{(z)}$ and the wealth process X are $\mathbb{F}$ -adapted. Hence $\mathbb{F}$ is a natural model for the investor's information. With

$$J(u) = \mathbb{E} \left[U(X_T) \right] \quad (4.8.1)$$

where $U : \mathbb{R} \rightarrow \mathbb{R}$ is an utility function (differentiable, increasing and strictly concave), we look for

$$\sup_{u \in \mathcal{A}^{\mathcal{F}}} J(u).$$

We have

$$\begin{aligned} K_t &= U'(X_T), \\ F_t &= U'(X_T) \sum_{z=1}^n \mathbf{1}_{\{\tau^{(z)} > t\}}(t) u_t^{(z)} \rho_t^{(z)} + \sum_{z=1}^n (\mathcal{D}_{t,z} U'(X_t)) \mathbf{1}_{\{\tau^{(z)} > t\}}(t) u_t^{(z)} \lambda_t^{(z)}, \\ p_t &= U'(X_T), \\ \kappa_t &= \sum_{z=1}^n \mathcal{D}_{t,z} U'(X_t), \\ G_s(t) &= 0. \end{aligned}$$

Remark that under these assumptions, any $\mathbb{F}$ -predictable process u is an admissible control if

$$\mathbb{E} [U(X_T^u)^2 + U'(X_T^u)^2] < \infty. \quad (4.8.2)$$

Furthermore Assumption 4.5.2 only depends on verifying (4.8.2) for $u + y\beta$. The Hamiltonian

(4.4.11) is given by

$$\begin{aligned}\mathcal{H}_t(u, x) &= U'(X_T) \sum_{z=1}^n \mathbf{1}_{\{\tau^{(z)} > t\}}(t) u_t^{(z)} \rho_t^{(z)} \\ &\quad + \sum_{z=1}^n (\mathcal{D}_{t,z} U'(X_T)) \mathbf{1}_{\{\tau^{(z)} > t\}}(t) u_t^{(z)} \lambda_t^{(z)}.\end{aligned}$$

Hence

$$\frac{\partial \mathcal{H}_t}{\partial u}(v, x) = U'(X_T) \sum_{z=1}^n \mathbf{1}_{\{\tau^{(z)} > t\}}(t) \rho_t^{(z)} + \sum_{z=1}^n (\mathcal{D}_{t,z} U'(X_T)) \mathbf{1}_{\{\tau^{(z)} > t\}}(t) \lambda_t^{(z)}.$$

Theorem 4.5.4 finds critical points for (4.8.1). To ensure that a critical point $\hat{u}$ is a solution to (4.8) we need to know that 1) the critical point is a local maximum and 2) there are no other critical points $\bar{u}$ where $J(\bar{u}) > J(\hat{u})$. We investigate the exact properties of the critical points in Proposition 4.8.1 and sufficient conditions for a solution to (4.8) are given in Corollary 4.8.2.

Proposition 4.8.1. *Assume that*

- i) *U is twice continuously differentiable and concave,*
- ii) *All bounded $\mathbb{F}$ -predictable processes taking values in $\mathcal{U}$ are admissible controls,*
- iii) *For any $u \in \mathcal{A}^{\mathcal{F}}$ and $\mathbb{F}$ -predictable bounded process β such that*

$$u_t \pm \beta_t \in \mathcal{U}, \quad dt \times d\mathbb{P} \text{ a.e.} \quad (4.8.3)$$

then there exist $\varepsilon > 0$ such that

$$\begin{aligned}\left\{ U''(X_T^{u+y\beta}) \left(\int_0^t \sum_{z=1}^n \mathbf{1}_{\{\tau^{(z)} > r\}}(r) \beta_r^{(z)} \rho_r^{(z)} dr \right. \right. \\ \left. \left. - \sum_{z=1}^n \int_0^t \mathbf{1}_{\{\tau^{(z)} > r\}}(r) \beta_r^{(z)} d\tilde{H}_t^{(z)} \right)^2 \right\}_{y \in (-\varepsilon, \varepsilon)}\end{aligned} \quad (4.8.4)$$

is uniformly $\mathbb{P}$ -integrable,

- iv) *Assumption 4.5.2 holds for all bounded $u \in \mathcal{A}^{\mathcal{F}}$.*

Let $\epsilon = \min(\delta, \varepsilon)$, where δ is as in (4.5.4). Then the mapping $y \rightarrow J(u + y\beta)$, $y \in (-\epsilon, \epsilon)$, is strictly concave for all $u \in \mathcal{A}^{\mathcal{F}}$ and bounded $\mathbb{F}$ -predictable β satisfying (4.8.3). Furthermore, there is at most one bounded $u \in \mathcal{A}^{\mathcal{F}}$ such that u is a critical point (in the sense of Theorem 4.5.4).

Proof. First we prove the concavity of the mapping $y \rightarrow J(u + y\beta)$, $y \in (-\epsilon, \epsilon)$. We inter-

change the derivation and expectation and get

$$\begin{aligned}
\frac{\partial^2}{\partial y^2} J(u + y\beta) &= \mathbb{E} \left[\frac{\partial^2}{\partial y^2} U(X_T^{u+y\beta}) \right] \\
&= \mathbb{E} \left[U''(X_T^{u+y\beta_T}) \left(\int_0^t \sum_{z=1}^n \mathbf{1}_{\{\tau^{(z)} > r\}}(r) \beta^{(z)} \rho_r^{(z)} dr \right. \right. \\
&\quad \left. \left. - \sum_{z=1}^n \int_0^t \mathbf{1}_{\{\tau^{(z)} > r\}}(r) \beta^{(z)} d\tilde{H}_t^{(z)} \right)^2 \right] < 0,
\end{aligned}$$

where the last inequality follows by the concavity of U .

Next we want to show that there is at most one bounded $u \in \mathcal{A}^{\mathcal{F}}$ such that u is a critical point. First we show that when $u \in \mathcal{A}^{\mathcal{F}}$ is bounded and β is as in (4.8.3), we have $\epsilon > 1$, i.e. that $y \rightarrow J(u + y\beta)$ is a strictly concave mapping for $y \in (-\epsilon, \epsilon)$ with $\epsilon > 1$. The claim $\epsilon > 1$ follows from ii) and the uniform integrability conditions (4.5.3)-(4.5.4)-(4.8.4) since $\frac{\partial}{\partial y} J(u + y\beta)|_{y=a} = \frac{\partial}{\partial y} J(u + a\beta + y\beta)|_{y=0}$.

Suppose $\bar{u}, \hat{u} \in \mathcal{A}^{\mathcal{F}}$ are both bounded and critical points. Set $\beta_t = \bar{u}_t - \hat{u}_t$. Consider the control $\hat{u} + \frac{1}{2}\beta \in \mathcal{A}^{\mathcal{F}}$, and the mapping

$$h(y) \rightarrow J\left(\hat{u} + \frac{1}{2}\beta + y\frac{1}{2}\beta\right), \quad y \in (-\epsilon, \epsilon) \quad (4.8.5)$$

Note that $\epsilon > 1$, $h(1) = J(\bar{u})$ and $h(-1) = J(\hat{u})$. Since h is strictly concave at most one of $h(-1)$ and $h(1)$ can be a maximum. □

Corollary 4.8.2. *Suppose the Assumptions in Proposition 4.8.1 hold. If $\mathcal{U}$ is bounded and a critical point $\hat{u}$ exists, then $\hat{u}$ is optimal, i.e.*

$$J(\hat{u}) = \sup_{u \in \mathcal{A}^{\mathcal{F}}} J(u),$$

and optimal portfolio $\hat{u}$ is characterized by

$$\begin{aligned}
\mathbb{E} \left[\frac{\partial \mathcal{H}_t}{\partial u}(\hat{u}, X_t^{(\hat{u})}) | \mathcal{F}_t \right] &= \sum_{z=1}^n \mathbf{1}_{\{\tau^{(z)} > t\}}(t) \rho_t^{(z)} \mathbb{E} [U'(X_T) | \mathcal{F}_t] \\
&\quad + \sum_{z=1}^n (\mathcal{D}_{t,z} U'(X_T)) \mathbf{1}_{\{\tau^{(z)} > t\}}(t) \lambda_t^{(z)} = 0,
\end{aligned}$$

for all $t \in [0, T]$ a.s.

Proof. This is a restatement of Proposition 4.8.1. □

Chapter 5

Information and optimal investment in defaultable assets

Abstract

We study optimal investment in an asset subject to risk of default for investors that rely on different levels of information. The price dynamics can include noises both from a Wiener process and a Poisson random measure with infinite activity. The default events are modeled via a counting process in line with large part of the literature in credit risk. In order to deal with both cases of inside and partial information we consider the framework of the anticipating calculus of forward integration. This does not require a priori assumptions typical of the framework of enlargement of filtrations. We find necessary and sufficient conditions for the existence of a locally maximizing portfolio of the expected utility at terminal time. We consider a large class of utility functions. In addition we show that the existence of the solution implies the semi-martingale property of the noises driving the stock. Some discussion on unicity of the maxima is included.

5.1 Introduction: The model, the optimization problem, the streams of information

Occasionally, we observe that unexpected events wipe out shareholder values. We will generically call all these events *default events*. Inspired by default risk literature, we consider a model for stocks where there is a varying risk of instantaneous loss in the stock value.

Of particular interest here is when the default events are dependent on the noises driving the stock or when the investor has insider information. In these cases mathematical questions arise as to whether the driving noises are still (semi)-martingales and the relevant stochastic integrals can be interpreted in the Itô sense. Since this is not a priori certain, we choose to investigate this issue using forward integration in the modeling of stock prices. With this we do not need a priori assumptions or restrictions on the information available to the investor and we will be able to use a unique framework for all the situations of interest.

Our main result is a sufficient and necessary criteria for an optimal investment strategy maximizing the expected utility of the final portfolio value, for a portfolio involving the defaultable asset. We remark that this result also holds for optimization problems with partial or delayed information.

Furthermore we show that the existence of an optimal strategy yields the semi-martingale property of the noises. This would usually be assumed a priori if working in the framework of enlargement of filtrations see for instance [14, 27, 33, 71, 62, 60].

The defaultable stock is modeled with three random noises, a Wiener process W , a Poisson random measure N and a pure jump process H . The occurrence of defaults or catastrophic events is modelled by H . The intensity of H , as viewed by the investor, is stochastic and can either depend on current and future knowledge of W and N or be independent of the two.

Our model market on the time horizon $[0, T]$ ($T > 0$) consists of a (non-defaultable) bond S_0 serving as numéraire with dynamics:

$$\begin{aligned} dS_0(t) &= S_0(t)\rho(t)dt, \\ S_0(0) &= 1 \end{aligned} \tag{5.1.1}$$

and a defaultable asset S_1 with price dynamics:

$$\begin{aligned} d^- S_1(t) &= S_1(t-) \left(\mu(t)dt + \sigma(t)d^- W(t) \right. \\ &\quad \left. + \int_{\mathbb{R}_0} \theta(t, z) \tilde{N}(d^- t, dz) + \kappa(t)dH(t) \right), \quad S_1(0) > 0. \end{aligned} \tag{5.1.2}$$

Here $W(t)$, $t \geq 0$, is a standard Wiener process and $N(dt, dz)$, $t \geq 0$, $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$ is a Poisson random measure, independent of W and with $\mathbb{E}[N(dt, dz)] = \nu(dz)dt$. We denote $\tilde{N}(dt, dz) := N(dt, dz) - \nu(dz)dt$. Moreover $H(t)$, $t \geq 0$ is a càdlàg counting process, with

$$\mathbb{E}[H(T)] < \infty \quad \text{and} \quad \mathbb{P}(\Delta H(t) > 1 \text{ for any } t \in [0, T]) = 0.$$

We remark that H is not necessarily independent of N and W . Being H a process of finite variation the corresponding integral is intended path-wise. On the other side, the d^- indicates forward integration. The forward integral extends the Itô integral but does not require the integrand to be adapted to a specific filtration, see Section 5.2 for details.

The random processes considered live in a complete probability space $(\Omega, \mathcal{A}, \mathbb{P})$. In the sequel the following $\mathbb{P}$ -augmented filtrations appear

- $\mathbb{F}^H := \{\mathcal{F}_t^H \subset \mathcal{A}, t \geq 0\}$ where $\mathcal{F}_t^H = \sigma\{H(s), s \leq t\}$,
- $\mathbb{F} := \{\mathcal{F}_t \subset \mathcal{A}, t \geq 0\}$ where $\mathcal{F}_t = \sigma\{W(s), N((s, t], B), s \leq t, B \in \mathcal{B}(\mathbb{R}_0)\}$,
- $\mathbb{G} := \{\mathcal{G}_t \subset \mathcal{A}, t \geq 0\}$ where $\mathcal{G}_t$ is a right continuous filtration that represents the information available to the investor at time t .

We assume that the coefficients ρ , μ , σ , and κ are càglàd stochastic processes and θ is a càglàd random field, in the sense that $\theta(\cdot, z)$ is càglàd ν -a.e. ($\mathbb{P}$ -a.e.). Here ρ , μ , σ and κ are measurable with respect to $\mathcal{A} \times \mathcal{B}([0, T])$ while κ is $\mathcal{A} \times \mathcal{B}([0, T] \times \mathbb{R}_0)$ -measurable. The choice of the forward integral in (5.1.2) allows us to drop the usual requirements of adaptedness of the coefficients to the given information. Naturally, in the case of adaptedness (5.1.2) could be expressed in terms of Itô integration (see Section 5.2).

The Borel measure $\nu(dz)$ on $\mathbb{R}_0$ is σ -finite and satisfies $\int_{\mathbb{R}_0} z^2 \nu(dz) < \infty$. For modeling purposes κ would be taken to be negative though it is not a necessary condition for the optimization problem.

We denote Λ as the $\mathbb{G}$ -predictable intensity of H , i.e. the $\mathbb{G}$ -predictable random measure such that

$$\mathbb{E} \left[\int_0^t \kappa(s) dH(s) \right] = \mathbb{E} \left[\int_0^t \kappa(s) \Lambda(ds) \right],$$

for all $\mathbb{G}$ -predictable processes κ . In addition, we assume that H and N do not jump at the same time, i.e.

$$\mathbb{P}(\text{There exist } t \in [0, T] \text{ and } U \subset \mathbb{R}_0 \text{ compact such that } \Delta H(t) > 0 \text{ and } N(\Delta t, U) > 0) = 0. \quad (5.1.3)$$

We set σ forward integrable with respect to W , θ and $\ln(1 + \theta)$ forward integrable with respect to N and

$$\mathbb{E} \left[\int_0^T |\mu(s)| + |\sigma(s)|^2 + \int_{\mathbb{R}_0} |\theta(s, z)|^2 \nu(dz) ds + \int_0^T |\kappa(s)| \Lambda(ds) \right] < \infty. \quad (5.1.4)$$

To have S_1 well defined and non-negative at all times, we assume

$$-1 < \theta(t, z, \omega) \quad dt \times \nu(dz) \times d\mathbb{P} \text{ a.e.} \quad (5.1.5)$$

$$-1 \leq \kappa(t, \omega) \quad dt \times d\mathbb{P} \text{ a.e.} \quad (5.1.6)$$

Using an adequate version of the Itô formula (Theorem 5.2.4), we see that the solution of (5.1.2) is

$$\begin{aligned} S_1(t) = S_1(0) & \prod_{\substack{\Delta H(s) > 0 \\ s \leq t}} \left(1 + \kappa(s) \Delta H(s) \right) \exp \left\{ \int_0^t \left[\mu(s) - \frac{1}{2} \sigma^2(s) \right] ds \right. \\ & + \int_0^t \sigma(s) d^- W(s) - \int_0^t \int_{\mathbb{R}_0} \left[\ln(1 + \theta(s, z)) - \theta(s, z) \right] \nu(dz) ds \\ & \left. + \int_0^t \int_{\mathbb{R}_0} \ln(1 + \theta(s, z)) \tilde{N}(d^- s, dz) \right\} \end{aligned} \quad (5.1.7)$$

and it is easy to argue that this solution is unique. This can be achieved using similar arguments as in [92, Theorem 37] though adapted to forward integration.

The investor's optimization problem is to divide his money between the asset S_1 and the bond S_0 in order to achieve the maximum expected utility of the portfolio value at the end of the period allowed. The investor bases his decisions on the information available to him

represented by the filtration $\mathbb{G}$. The investor's wealth $\tilde{X}_\pi(t)$, $t \in [0, T]$, is given by:

$$d\tilde{X}(t) = (1 - \pi(t))dS_0(t) + \pi(t)d^-S_1(t) \quad (5.1.8)$$

with initial value $\tilde{X}(0) = x_0 > 0$. The process $\pi(t)$, $t \in [0, T]$, represents the fraction of wealth invested in S_1 . Note that π is a $\mathbb{G}$ -adapted stochastic process.

We aim for generality in how the optimization scheme ends. In particular we are interested in the two different scenarios:

- i) It is no longer possible to invest in the asset S_1 after the first jump of H . In this case, the jump of H signifies default or another catastrophic event. See, e.g. [17, 20, 33].
- ii) It is possible to invest in the asset S_1 even after several “default” events. The jumps of H signify the occurrence of these “default” events. and the dynamics of S_1 can possibly change. See, e.g. [71, 90].

To describe both the above scenarios, we assume that the it is longer possible to invest in S_1 after a $\mathbb{G}$ -stopping time $\tau \leq T$. For the period $(\tau, T]$ all the investor's wealth is invested in the bond. The stopping time τ must satisfy $\tau \leq T$, meaning that the optimization problem terminates in any case when the time horizon is reached, and $\tau \leq \inf\{t \in [0, T] : S_1(t) = 0\}$, meaning that the optimization problem ends if there is no value in the asset S_1 .

By application of the Itô formula, we can see that the (unique) solution of (5.1.8), for a given admissible π (see Definition 5.3.1), is:

$$\begin{aligned} \tilde{X}_\pi(t) = x_0 \exp \Big\{ & \int_0^t \left[\rho(s) + (\mu(s) - \rho(s))\pi(s) - \frac{1}{2}\sigma^2(s)\pi^2(s) \right] ds \\ & + \int_0^t \int_{\mathbb{R}_0} \left[\ln(1 + \pi(s)\theta(s, z)) - \pi(s)\theta(s, z)\nu(dz) \right] ds + \int_0^t \sigma(s)\pi(s)d^-W(s) \\ & + \int_0^t \int_{\mathbb{R}_0} \ln(1 + \pi(s)\theta(s, z))\tilde{N}(d^-s, dz) + \int_0^t \ln(1 + \kappa(s)\pi(s))dH(s) \Big\}. \end{aligned} \quad (5.1.9)$$

and set $X_\pi(T) := \tilde{X}_\pi(\tau)e^{\int_\tau^T \rho(s)ds}$.

In summary we study the optimal portfolio problem

$$\sup_{\pi \in \mathcal{A}_{\mathbb{G}}} \mathbb{E} \left[U \left(\tilde{X}_\pi(\tau) e^{\int_\tau^T \rho(s)ds} \right) \right] = \sup_{\pi \in \mathcal{A}_{\mathbb{G}}} \mathbb{E} \left[U \left(X_\pi(T) \right) \right], \quad (5.1.10)$$

of an investor having $\mathbb{G}$ as information flow at disposal and U as utility function. Here $\mathcal{A}_{\mathbb{G}}$ represents the set of admissible portfolios (see Definition 5.3.1).

The optimization scheme itself Theorem 5.3.3 and the related Theorem 5.4.2 are an extension of the results in [12, 39] to include a form of default risk. We refer to [96, 97] for the treatment of the forward integral with respect to the Wiener process, and to [39] for the case of integration with respect to the compensated Poisson random measure. The forward integral

is an extension of the Itô integral, but does not require the adaptedness of the integrands to the integral filtration. Applications of this type of integration to optimization problems and the justification of the use of these integrals from the modeling point of view have been studied. See, e.g. [12, 40, 30, 72]. We also refer to [43] for a unified presentation of the topics.

Related to our optimization problem is the optimization of investments under uncertain time-horizons, as done in [17, 33]. In [17], optimization ends at a stopping time τ related to the noise in stock price. In [33] both optimal consumption and investment are treated. Typically the problems are solved using some variants of Hamilton-Jacobi-Bellman (HJB) equations. Our approach differs from these works for several reasons. First we focus on different streams of information for the investor, second we consider that the loss in the case of default depends on the position in the risky asset. Moreover, our approach is different in framework and we do not use HJB type solutions. In [78] we find a study of a problem similar to ours. The approach is however entirely different as in this case backward stochastic differential equations are involved. Moreover we allow for a more general information structure and we consider a Lévy type of noise in the price dynamics.

Our work has some similarities to [9], where an optimization problem is considered when the stock dynamics include a jump component with an unknown intensity modeled by a continuous time Markov chain. But the filtering techniques therein may be less suited to default modeling since default is a jump happening only once. The methodology presented there relies on HJB equations and differs from ours.

Bielecki and coauthors consider various forms of optimal investments in, e.g. [14], [15] and [13], looking at optimality and hedging when there is a number of instruments, some of which are subject to default. However, their main focus is on the use of defaultable instruments for hedging purposes and the evaluation on whether to invest in defaultable bonds. In the same line is the study in [53].

As announced, in this paper we adopt the framework of anticipating stochastic calculus, specifically forward integration to tackle the optimization problem (5.1.10). Moreover, we consider the problem for various choices of investor's information flow $\mathbb{G}$. To the best of our knowledge it is the first time that the framework of forward integration is applied in optimization problems in presence of default.

In this paper we provide a characterization for the existence of locally optimal controls in a great generality both in the choice of utility function and in the amount of information available. Considerations on the meaning of locality and some examples are also provided. These topics are presented in Section 3. The key results of forward integration is summarized in Section 2. In Section 4 we reinterpret the results of section 3 in the context of semimartingale-integration.

5.2 Mathematical framework: Forward Integrals

Forward integrals were introduced by Russo and Valois in the articles [96] and [97] for continuous processes and in [39] for pure jump Lévy process, see also [43] for a systematic presentation.

The forward integral is a type of stochastic anticipating integration that does not require assumptions of adaptedness or predictability to some filtration related to the integrator. Moreover,

it is also an extension of the Itô integral in the sense that when the appropriate predictability is in place the two integrals coincide. This makes the forward integral especially suited for studying portfolio optimization problems under insider or partial information, where different filtrations are considered. See for, e.g. [12, 39] and [43].

We follow the idea of [72] and consider the forward integral with respect to the Wiener processes as a limit in $L^1(\mathbb{P})$. This would also imply forward integrability in the sense of Russo and Vallois, [96, 97, 98], who consider the same limit in probability.

Definition 5.2.1. *We say that the stochastic process $\sigma = \sigma(t, \omega)$, $t \in [0, T]$, $\omega \in \Omega$, is forward integrable over the interval $[0, T]$ with respect to W if there exists a process $I = I(\sigma, t)$, $t \in [0, T]$, such that*

$$\mathbb{E} \left[\sup_{t \in [0, T]} \left| \int_0^t \sigma(s) \frac{W(s + \epsilon) - W(s)}{\epsilon} ds - I(\sigma, t) \right| \right] \longrightarrow 0, \quad \text{as } \epsilon \rightarrow 0^+,$$

In this case we write

$$I(\sigma, t) = \int_0^t \sigma(s) d^- W(s), \quad t \in [0, T],$$

and call $I(\sigma, t)$ the forward integral of σ with respect to W on $[0, t]$.

Lemma 5.2.1 shows that the forward integral is an extension of the Itô integral.

Lemma 5.2.1. *Let $\mathbb{G} = \{\mathcal{G}_t, t \in [0, T]\}$ be a given filtration. Suppose that*

- i) *W is a semimartingale with respect to the filtration $\mathbb{G}$,*
- ii) *σ is $\mathbb{G}$ -predictable and the Itô integral $\int_0^T \sigma(t) dW(t)$ exists (in $L^1(\mathbb{P})$),*

then σ is forward integrable and

$$\int_0^T \sigma(t) d^- W(t) = \int_0^T \sigma(t) dW(t).$$

For proof we refer to, e.g. [43, Lemma 8.9].

Elementary processes are forward integrable, and have a natural interpretation as Riemann-like sums. Suppose the stochastic process $\sigma = \sigma(t, \omega)$, $t \in [0, T]$, $\omega \in \Omega$, is elementary, meaning that it has the form

$$\sigma(t, \omega) = \sum_{i=0}^{N-1} \sigma_i(\omega) \mathbf{1}_{(t_i, t_{i+1}]}(t), \quad (5.2.1)$$

where the σ_i are bounded random variables and $0 = t_0 < t_1 < \dots < t_N = T$. Then σ is forward

integrable, see [95, Remark 1], and

$$\int_0^t \sigma(s) d^-W(s) = \sum_{i=0}^{N-1} \sigma_i \left(W(t_{i+1} \wedge t) - W(t_i \wedge t) \right), \quad t \in [0, T]. \quad (5.2.2)$$

However, it is not obvious that one can approximate a general forward integrable function by elementary functions and in this way obtain also an approximation to the integral.

Example 5.2.2. Let $A = \{f \in L^\infty([0, 1] \times \Omega) : f \text{ is càglàd, } |f(t, \omega)| \leq 1 \text{ for all } (t, \omega) \in [0, 1] \times \Omega\}$. Then any $f \in A$ is forward integrable. But

$$\sup_{f \in A} \mathbb{E} \left[\int_0^1 f(s) d^-W(s) \right] = \infty. \quad (5.2.3)$$

So even though f is bounded, the forward integral with respect to d^-W can have arbitrarily large expectations. This would not happen with Itô integrals as it is a result of W having infinite total variation and using anticipating information.

To prove (5.2.3), let f_n be elementary functions of the form

$$f_n = \sum_{j=0}^{n-1} \text{sign} \left(W\left(\frac{j+1}{n}\right) - W\left(\frac{j}{n}\right) \right) \mathbf{1}_{\{t \in (\frac{j+1}{n}, \frac{j+1}{n}]\}},$$

where

$$\text{sign}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

Then $f_n \in A$ and

$$\begin{aligned} \int_0^1 f_n(s) d^-W(s) &= \sum_{j=0}^{n-1} \text{sign} \left(W\left(\frac{j+1}{n}\right) - W\left(\frac{j}{n}\right) \right) \left(W\left(\frac{j+1}{n}\right) - W\left(\frac{j}{n}\right) \right) \\ &= \sum_{j=0}^{n-1} \left| W\left(\frac{j+1}{n}\right) - W\left(\frac{j}{n}\right) \right|. \end{aligned}$$

We have $\mathbb{E}[|W(t) - W(s)|] = \sqrt{(t-s) \cdot 2/\pi}$, (see for instance [76]), so

$$\begin{aligned} \mathbb{E} \left[\int_0^1 f_n(s) d^-W(s) \right] &= \mathbb{E} \left[\sum_{j=0}^{n-1} \left| W\left(\frac{j+1}{n}\right) - W\left(\frac{j}{n}\right) \right| \right] \\ &= \sqrt{n \cdot 2/\pi} \longrightarrow \infty \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Additionally we can remark that letting $g_n = n^{-1/4} f_n$, we would get that $g_n \rightarrow 0$ pointwise and is bounded by the forward integrable function 1, thus proving that the dominated convergence theorem does not hold for forward integrals with respect to Brownian motions.

Characterizing when integrals are finite or limits do not explode is non-trivial, and from the remark above we see that the boundedness of the integrand is not enough. Thus we have to

be careful, even though a sequence of forward integrable functions converge in some suitable space, the corresponding forward integrals over these functions may not converge at all. See also the discussion in [92, Section 1.8].

Definition 5.2.3. *The forward integral*

$$J(\theta, t) := \int_0^t \int_{\mathbb{R}_0} \theta(s, z) \tilde{N}(d^-s, dz), \quad t \in [0, T]$$

with respect to the Poisson random measure $\tilde{N}$ of a càglàd random field $\theta(t, z, \omega)$, $t \in [0, T]$, $z \in \mathbb{R}_0$, $\omega \in \Omega$, is defined as

$$J(\theta, t) = \lim_{m \rightarrow \infty} \int_0^t \int_{\mathbb{R}_0} \theta(s, z) \mathbf{1}_{U_m} \tilde{N}(ds, dz)$$

if the limit exists in $L^2(\mathbb{P})$. Here, U_m , $m = 1, 2, \dots$, is an increasing sequence of compact sets $U_m \subset \mathbb{R}_0$ with $\nu(U_m) < \infty$ such that $\lim_{m \rightarrow \infty} U_m = \mathbb{R}_0$.

Also in this case the forward integral is an extension of the Itô integral [43, Remark 15.2]:

Remark 5.2.2. Let $\mathbb{G} = \{\mathcal{G}_t, t \in [0, T]\}$ be a given filtration such that

- i) The process $\eta(t) = \int_0^t \int_{\mathbb{R}_0} z \tilde{N}(ds, dz)$, $t \in [0, T]$, is a semimartingale with respect to $\mathbb{G}$.
- ii) The random field $\theta = \theta(t, z)$, $t \in [0, T]$, $z \in \mathbb{R}_0$, is $\mathbb{G}$ -predictable.
- iii) The integral $\int_0^t \int_{\mathbb{R}_0} \theta(t, z) \tilde{N}(ds, dz)$ exists as a classical Itô integral.

Then θ is forward integrable and we have

$$\int_0^T \int_{\mathbb{R}_0} \theta(s, z) \tilde{N}(d^-t, dz) = \int_0^T \int_{\mathbb{R}_0} \theta(s, z) \tilde{N}(dt, dz).$$

5.2.1 The Itô formula for forward integrals

An Itô formula for forward type integrals when the integrator is continuous was developed in [97, 98]. An Itô formula for forward integrals with Poisson random measures is found in [39], both the results are also summarized in [43]. In this paper we need a more general version that include processes of finite variation to guarantee the existence of solutions of (5.1.8) and (5.1.2). The proof can be seen as a continuation of the one presented in [43, Theorem 8.12], thus we only sketch the additional part.

Theorem 5.2.4.

Let

$$d^-X(t) = \mu(t)dt + \sigma(t)d^-W(t) + \int_{\mathbb{R}_0} \theta(t, z) \tilde{N}(d^-t, dz) + d\zeta(t),$$

where

- μ is a stochastic process satisfying $\int_0^T |\mu(s)| ds < \infty$ $\mathbb{P}$ -a.s.
- σ is forward integrable with respect to W .
- θ and $|\theta|$ are forward integrable with respect to $\tilde{N}$ and θ satisfies $\int_0^T \int_{\mathbb{R}_0} |\theta(s, z)|^2 \nu(dz) ds < \infty$ $\mathbb{P}$ -a.s.
- ζ is a càdlàg pure jump process of finite variation, with $\zeta(0) = 0$ and

$$\mathbb{P}(\text{There exist } t \in [0, T] \text{ and } U \subset \mathbb{R}_0 \text{ compact such that } \Delta\zeta(t) > 0 \text{ and } N(\Delta t, U) > 0) = 0. \quad (5.2.4)$$

for all $U \subset \mathbb{R}_0$ compact. Here $N(\Delta t, U) := N((0, t], U) - N((0, t), U)$ and $\Delta\zeta(t) := \zeta(t) - \zeta(t-)$.

Assume $f \in C^2(\mathbb{R})$ and let $Y(t) = f(X(t))$. Then

$$\begin{aligned} Y(t) = Y(0) &+ \int_0^t \left[f'(X(s-))\mu(s) + \frac{1}{2}f''(X(s-))\sigma^2(s) \right] ds \\ &+ \int_0^t \int_{\mathbb{R}_0} \left[f(X(s-) + \theta(s, z)) - f(X(s-)) - f'(X(s-))\theta(s, z) \right] \nu(dz) ds \\ &+ \int_0^t f'(X(s-))\sigma(s)d^-W(s) + \int_0^t \int_{\mathbb{R}_0} \left[f(X(s-) + \theta(s, z)) - f(X(s-)) \right] \tilde{N}(d^-s, dz) \\ &+ \sum_{\substack{0 < s < t \\ \Delta\zeta(s) \neq 0}} \left[f(X(s-) + \Delta\zeta(s)) - f(X(s-)) \right]. \end{aligned}$$

Remark 5.2.3. Condition (5.2.4) is for instance fulfilled if N and ζ are independent.

Proof. Let

$$X_m(t) = x + \int_0^t \mu(s)ds + \int_0^t \sigma(s)d^-W(s) + \int_0^t \int_{\mathbb{R}_0} \mathbf{1}_{U_m}(z)\theta(s, z)\tilde{N}(d^-s, dz) + \zeta(t),$$

where $\mathbf{1}_{U_m}$ is as in Definition 5.2.3. We denote $\alpha_i, i = 1, 2, \dots$ the times of the jumps of X_m . By condition (5.2.4) we can uniquely ($\mathbb{P}$ -a.s.) divide the sequence α_i by the jumps of either ζ or $\mathbf{1}_{U_m}(z)N(dt, dz)$ as α_i^ζ and α_i^N . We formally set $\alpha_0 = \alpha_0^\zeta = \alpha_0^N = 0$.

Then

$$\begin{aligned}
f(X_m(t)) - f(X_m(0)) &= \sum_i \left[f(X_m(\alpha_i \wedge t)) - f(X_m(\alpha_i \wedge t-)) \right] \\
&\quad + \sum_i \left[f(X_m(\alpha_i \wedge t-)) - f(X_m(\alpha_{i-1} \wedge t)) \right] \\
&= \sum_{\alpha_i^\zeta \leq t} \left[f(X_m(\alpha_i^\zeta)) - f(X_m(\alpha_i^\zeta-)) \right] \\
&\quad + \sum_{\alpha_i^N \leq t} \left[f(X_m(\alpha_i^N)) - f(X_m(\alpha_i^N-)) \right] \\
&\quad + \sum_{\alpha_i \leq t} \left[f(X_m(\alpha_i-)) - f(X_m(\alpha_{i-1})) \right] \\
&= J_1(t) + J_2(t) + J_3(t),
\end{aligned}$$

with

$$J_1(t) = \sum_{\substack{0 \leq s \leq t \\ \Delta\zeta(s) \neq 0}} \left[f(X_m(s-) + \Delta\zeta(s)) - f(X_m(s-)) \right]$$

and

$$\begin{aligned}
J_2(t) &= \sum_i \left[f(X_m(\alpha_i^N)) - f(X_m(\alpha_i^N-)) \right] \mathbf{1}_{\{\alpha_i^N \leq t\}} \tag{5.2.5} \\
&= \int_0^t \int_{\mathbb{R}_0} \left[f(X_m(s-) + \theta(s, z)) - f(X_m(s-)) \right] N(ds, dz) \\
&= \int_0^t \int_{\mathbb{R}_0} \left[f(X_m(s-) + \theta(s, z)) - f(X_m(s-)) \right] \tilde{N}(ds, dz) \\
&\quad + \int_0^t \int_{\mathbb{R}_0} \left[f(X_m(s-) + \theta(s, z)) - f(X_m(s-)) \right] \nu(dz) ds.
\end{aligned}$$

For the elements of the sum in $J_3(t)$ we use [43, Theorem 8.12]:

$$\begin{aligned}
J_3(t) &= \sum_i \left[\int_{\alpha_{i-1} \wedge t}^{\alpha_i \wedge t} \left[f'(X_m(s-)) \mu(s) ds - \int_{\mathbb{R}_0} f'(X_m(s-)) \mathbf{1}_{U_m} \theta(s, z) \nu(dz) \right] ds \right. \\
&\quad \left. + \int_{\alpha_{i-1} \wedge t}^{\alpha_i \wedge t} f'(X_m(s-)) \sigma(s) d^-W(s) + \int_{\alpha_{i-1} \wedge t}^{\alpha_i \wedge t} f''(X_m(s-)) \sigma^2(s) ds \right] \\
&= \int_0^t \left[f'(X_m(s-)) \mu(s) + \frac{1}{2} f''(X_m(s-)) \sigma^2(s) - \int_{\mathbb{R}_0} f'(X(s-)) \mathbf{1}_{U_m} \theta(s, z) \nu(dz) \right] ds \\
&\quad + \int_0^t f'(X_m(s-)) \sigma(s) d^-W(s).
\end{aligned}$$

Adding J_1 , J_2 and J_3 together and letting $m \rightarrow \infty$ the result follows. □

5.3 Optimization problem: local maxima

Now we are ready to tackle directly our stated optimization problem (5.1.10). First we give a description of the set of the investor's admissible portfolios.

Definition 5.3.1. *The set $\mathcal{A}_{\mathbb{G}}$ of admissible portfolios consists of stochastic processes $\pi = \pi(t, \omega)$, $t \in [0, T]$, $\omega \in \Omega$, such that*

i) π is càglàd and $\mathbb{G}$ -adapted,

ii) for every $\pi \in \mathcal{A}_{\mathbb{G}}$, there exists $\epsilon_\pi > 0$ such that for all t ,

$$\pi(t) \kappa(t) > -1 + \epsilon_\pi \tag{5.3.1}$$

and

$$\pi(t) \theta(t, z) > -1 + \epsilon_\pi, \tag{5.3.2}$$

iii)

$$\mathbb{E} \left[\int_0^T |(\mu(s) - \rho(s))| |\pi(s)| + \sigma^2(s) \pi^2(s) ds \right] < \infty$$

and

$$\mathbb{E} \left[\int_0^T \int_{\mathbb{R}_0} |\theta(s, z) \pi(s)|^2 \nu(dz) ds \right] < \infty,$$

iv) $\pi \sigma$ is càglàd and forward integrable with respect to W ,

v) $\pi\theta$, $\ln(1 + \pi\theta)$ and $\frac{\pi\theta}{1+\pi\theta}$ are càglàd and forward integrable with respect to $\tilde{N}$.

The subset $\mathcal{A}_{\mathbb{G}}^e$ of $\mathcal{A}_{\mathbb{G}}$ consists of all admissible portfolios that are representable as elementary integrands- see (5.2.1).

In particular we note that condition i) ensures that the portfolio choices correspond to the investors knowledge and that condition ii) ensures that the investor never reaches zero wealth from the jumps of H or $\tilde{N}$. In addition ii) means that fractions of the form $\frac{1}{1+\kappa\pi}$ are bounded, which is implicitly used in some forthcoming equations.

Note that if

$$\pi(s, \omega) = \alpha(\omega) \mathbf{1}_{(t, t+h]}(s),$$

where α is a bounded $\mathcal{G}_t$ -measurable random variable, then $\pi \in \mathcal{A}_{\mathbb{G}}^e \subset \mathcal{A}_{\mathbb{G}}$ as long as (5.3.1) and (5.3.2) are satisfied.

As announced we are interested in the problem

$$\sup_{\pi \in \mathcal{A}_{\mathbb{G}}} \mathbb{E}[U(X_{\pi}(T))]. \quad (5.3.3)$$

(We recall that $X_{\pi}(T) = \tilde{X}_{\pi}(\tau) e^{\int_{\tau}^T \rho(s) ds}$ is the value of the investor's wealth at T and the definition of $\tilde{X}_{\pi}$ is in (5.1.8)). We will search for solutions to (5.3.3) that are optimal in the sense that they cannot be improved by small perturbations.

Definition 5.3.2. We say that the stochastic process π is a local maximum for the problem (5.3.3) if

$$\mathbb{E}[U(X_{\pi+y\beta}(T))] \leq \mathbb{E}[U(X_{\pi}(T))] \quad (5.3.4)$$

for all bounded $\beta \in \mathcal{A}_{\mathbb{G}}$ and $|y| < \delta_{\pi, \beta}$ for some $\delta_{\pi, \beta} > 0$ that may depend on β . We say that π is a weak local maximum for (5.3.3) if (5.3.4) is true for all $\beta \in \mathcal{A}_{\mathbb{G}}^e$.

From the terminology point of view, when we say that a property holds under $(\mathbb{Q}, \mathbb{G})$, we mean that the property holds under the measure $\mathbb{Q}$ with respect to the filtration $\mathbb{G}$. Moreover, we say that a stochastic process Y has the martingale property under $(\mathbb{Q}, \mathbb{G})$ if

$$\mathbb{E}_{\mathbb{Q}}[Y(t+h) - Y(t) | \mathcal{G}_t] = 0$$

for all $0 < t < t+h < \infty$. We stress that Y does not need to be a $(\mathbb{Q}, \mathbb{G})$ -martingale despite having the martingale property under $(\mathbb{Q}, \mathbb{G})$. In fact no statement is given about Y being adapted to $\mathbb{G}$.

Following the techniques in [12, 40], we consider perturbations of stochastic controls to find necessary and sometimes sufficient criteria to characterize local maximums. We will need the following assumption for a differentiable utility function U .

Assumption $A_{u.i.}$ We say that assumption $A_{u.i.}$ holds for $\pi \in \mathcal{A}_{\mathbb{G}}$ if

$$\text{i) } \mathbb{E}[U(X_{\pi}(T))] < \infty,$$

$$\text{ii) } 0 < \mathbb{E}[U'(X_{\pi}(T))X_{\pi}(T)] < \infty, \text{ with } U'(x) = \frac{dU}{dx}(x),$$

iii) For all $\beta \in \mathcal{A}_G$ with β bounded, there exists $\delta_{\pi,\beta} > 0$ that may depend on β such that the family

$$\left\{ U'(X_{\pi+y\beta}(T)) X_{\pi+y\beta}(T) |\Psi(y, \beta, \pi)| \right\}_{y \in (-\delta_{\pi,\beta}, \delta_{\pi,\beta})} \quad (5.3.5)$$

is uniformly integrable, where

$$\begin{aligned} \Psi(y, \beta, \pi) := & \int_0^\tau \beta(s) [\mu(s) - \rho(s) - (\pi(s) + y\beta(s))\sigma^2(s)] ds \\ & + \int_0^\tau \int_{\mathbb{R}_0} \left[\frac{\beta(s)\theta(s, z)}{1 + (\pi(s) + y\beta(s))\theta(s, z)} - \beta(s)\theta(s, z) \right] \nu(dz) ds \\ & + \int_0^\tau \beta(s)\sigma(s)d^-W(s) + \int_0^\tau \int_{\mathbb{R}_0} \frac{\beta(s)\theta(s, z)}{1 + (\pi(s) + y\beta(s))\theta(s, z)} \tilde{N}(d^-z, ds) \\ & + \int_0^\tau \frac{\beta(s)\kappa(s)}{1 + \kappa(s)(\pi(s) + y\beta(s))} dH(s). \end{aligned} \quad (5.3.6)$$

Assumption $A_{u.i.}$ depends strongly on the utility function U . Condition i) is related to the optimization problem (5.3.3) and ii) is used in the definition of (5.3.8). Condition iii), uniform integrability, is the minimal condition for taking limits under the integral sign. It is unfortunate in that it stems from mathematical rather than modeling necessities, but we cannot do without it. There is a good discussion in when uniform integrability conditions like Assumption $A_{u.i.}$ is fulfilled in [43, section 16.5]. The conclusions from [43, section 16.5] can be transferred to our model.

Theorem 5.3.3. Suppose the utility function U is increasing and differentiable, $\pi \in \mathcal{A}_G$ and $A_{u.i.}$ holds.

i) If π is a local maximum for (5.3.3), then the process $M_\pi(t)$, $t \in [0, T]$, has the martingale property under $(\mathbb{Q}_\pi, \mathbb{G})$. Where M_π is defined as

$$\begin{aligned} M_\pi(t) := & \int_0^{t \wedge \tau} \left[\mu(s) - \rho(s) - \pi(s)\sigma^2(s) - \int_{\mathbb{R}_0} \frac{\pi(s)\theta^2(s, z)}{1 + \pi(s)\theta(s, z)} \nu(dz) \right] ds \\ & + \int_0^{t \wedge \tau} \sigma(s)d^-W(s) + \int_0^{t \wedge \tau} \int_{\mathbb{R}_0} \frac{\theta(s, z)}{1 + \pi(s)\theta(s, z)} \tilde{N}(d^-s, dz) \\ & + \int_0^{t \wedge \tau} \frac{\kappa(s)}{1 + \kappa(s)\pi(s)} dH(s), \end{aligned} \quad (5.3.7)$$

and the measure $\mathbb{Q}_\pi$ is defined by $d\mathbb{Q}_\pi = F_\pi(T)d\mathbb{P}$, with

$$F_\pi(T) = \frac{U'(X_\pi(T)) X_\pi(T)}{\mathbb{E}[U'(X_\pi) X_\pi(T)]}. \quad (5.3.8)$$

ii) Suppose the mappings

$$y \rightarrow \mathbb{E}[U(X_{\pi+y\beta}(T))], \quad y \in (-\delta_{\pi,\beta}, \delta_{\pi,\beta}), \quad (\delta_{\pi,\beta} > 0) \quad (5.3.9)$$

are concave for all controls $\beta \in \mathcal{A}_{\mathbb{G}}^c$ and $|y| < \delta$. If M_π has the martingale property under $(\mathbb{Q}_\pi, \mathbb{G})$ then π is a weak local maximum for (5.3.3)

iii) Suppose M_π is $\mathbb{G}$ -adapted and the conditions in ii) are satisfied. If M_π is a $(\mathbb{Q}_\pi, \mathbb{G})$ -martingale then π is a local maximum for (5.3.3).

Proof. Part i) If π is a local maximum, then for all bounded β we have

$$0 = \frac{d}{dy} \mathbb{E}[U(X_{\pi+y\beta}(T))]_{|y=0} = \mathbb{E}[U'(X_{\pi+y\beta}(T)) \frac{d}{dy} X_{\pi+y\beta}(T)]_{|y=0}. \quad (5.3.10)$$

Here assumption $A_{u.i.}$ is used, see for instance [45, Appendix A]. With some calculations we obtain

$$\begin{aligned} 0 &= \mathbb{E} \left[U'(X_\pi(T)) X_\pi(T) \left\{ \int_0^\tau \beta(s) [\mu(s) - \rho(s) - \pi(s)\sigma^2(s)] ds \right. \right. \\ &\quad + \int_0^\tau \int_{\mathbb{R}_0} \beta(s) \frac{-\pi(s)\theta^2(s, z)}{1 + \pi(s)\theta(s, z)} \nu(dz) ds + \int_0^\tau \beta(s) \sigma(s) d^-W(s) \\ &\quad \left. \left. + \int_0^\tau \int_{\mathbb{R}_0} \frac{\beta(s)\theta(s, z)}{1 + \pi(s)\theta(s, z)} \tilde{N}(d^-s, dz) + \int_0^\tau \frac{\beta(s)\kappa(s)}{1 + \pi(s)\kappa(s)} dH(s) \right\} \right] \\ &= \mathbb{E} \left[U'(X_\pi(T)) X_\pi(T) \Psi(0, \beta, \pi) \right]. \end{aligned} \quad (5.3.11)$$

We now let $\beta(s) = \alpha \mathbf{1}_{(t, t+h]}(s)$, where α is a $\mathcal{G}_t$ -measurable bounded random variable. We can put α outside the forward integrals, see for instance [43, Lemma 8.7] and [43, Remark 15.3] to get

$$\begin{aligned} \mathbb{E} \left[U'(X_\pi(T)) X_\pi(T) \left\{ \int_{t \wedge \tau}^{(t+h) \wedge \tau} [\mu(s) - \rho(s) - \pi(s)\sigma^2(s)] \right. \right. \\ \left. \left. - \int_{\mathbb{R}_0} \frac{\pi(s)\theta^2(s, z)}{1 + \pi(s)\theta(s, z)} \nu(dz) \right] ds + \int_{t \wedge \tau}^{(t+h) \wedge \tau} \sigma(s) d^-W(s) \right. \\ \left. + \int_{t \wedge \tau}^{(t+h) \wedge \tau} \int_{\mathbb{R}_0} \frac{\theta(s, z)}{1 + \pi(s)\theta(s, z)} \tilde{N}(d^-z, ds) + \int_{t \wedge \tau}^{(t+h) \wedge \tau} \frac{\kappa(s)}{1 + \kappa(s)\pi(s)} dH(s) \right\} \alpha \right] = 0. \end{aligned} \quad (5.3.12)$$

Hence we conclude that

$$\mathbb{E}[F_\pi(T)(M_\pi(t+h) - M_\pi(t))|\mathcal{G}_t] = 0$$

with $F_\pi(T)$ and M_π defined as in (5.3.8) and (5.3.7) respectively. Since $\mathbb{E}[F_\pi(T)] = 1$, we can define a new probability measure on $(\Omega, \mathcal{A})$ by

$$d\mathbb{Q}_\pi = F_\pi(T)d\mathbb{P}. \quad (5.3.13)$$

We thus have that if π is a local maximum, M_π has the martingale property under $(\mathbb{Q}_\pi, \mathbb{G})$.

Part ii). Suppose M_π has the martingale property under $(\mathbb{Q}_\pi, \mathbb{G})$. Then, for $0 < t < t+h < T$,

$$\mathbb{E}_{\mathbb{Q}_\pi}[M_\pi(t+h) - M_\pi(t)|\mathcal{G}_t] = 0,$$

or, equivalently, that for all bounded $\mathcal{G}_t$ -measurable random variables α we have

$$0 = \mathbb{E}_{\mathbb{Q}_\pi}\left[\alpha(M_\pi(t+h) - M_\pi(t))\middle|\mathcal{G}_t\right] = \mathbb{E}_{\mathbb{Q}_\pi}\left[\int_0^T \alpha \mathbf{1}_{(t, t+h]}(s) d^- M_\pi(s) \middle|\mathcal{G}_t\right].$$

Taking linear combinations we get that

$$0 = \mathbb{E}\left[U'(X_\pi(T))X_\pi(T) \int_0^T \beta(s) d^- M_\pi(s)\right] \quad (5.3.14)$$

for any $\beta \in \mathcal{A}_\mathbb{G}^e$. Since the mapping $y \rightarrow \mathbb{E}[U(X_{\pi+y\beta}(T))]$ is concave on $|y| < \delta_{\pi, \beta}$ then π is a weak local maximum.

Part iii). The conditions of ii) are satisfied so (5.3.14) holds for all $\beta \in \mathcal{A}_\mathbb{G}^e$. Let $\beta \in \mathcal{A}_\mathbb{G}$, β be bounded, and β_j , $j = 1, \dots$, be a sequence of elementary stochastic processes $\beta_j \in \mathcal{A}_\mathbb{G}^e$ such β_j converges pointwise in ω and uniformly in t to β .

Since M_π is adapted and has the martingale property, it is a local martingale and

$$\int_0^T \beta_j(s) dM_\pi(s) \longrightarrow \int_0^T \beta(s) dM_\pi(s) \quad \text{in probability as } j \rightarrow \infty.$$

By assumption $A_{u.i.}$, the random variable $\int_0^T \beta(s) dM_\pi(s)$ is $\mathbb{Q}_\pi$ -integrable so that

$$\mathbb{E}_{\mathbb{Q}_\pi}\left[\int_0^T \beta(s) dM_\pi(s)\right] = 0. \quad (5.3.15)$$

Since the mapping $y \rightarrow \mathbb{E}[U(X_{\pi+y\beta}(T))]$ is concave, from the computations in part i) we see that (5.3.15) can only be zero if π is a local maximum.

□

With the introduction of the forthcoming assumption A_{d^2} we can detail additional results on the convexity of (5.3.9) and the uniqueness of local maximums.

Assumption A_{d^2} : The utility function U is twice differentiable, strictly increasing and concave. For $\pi \in \mathcal{A}_{\mathbb{G}}$, we assume that for all $\beta \in \mathcal{A}_{\mathbb{G}}$ bounded, there exists a $\delta_{\pi,\beta} > 0$, that may depend on β , such that the family

$$\left\{ U''(X_{\pi+y\beta}(T))X_{\pi+y\beta}^2(T)\Psi^2(y, \beta, \pi) + U'(X_{\pi+y\beta}(T))X_{\pi+y\beta}(T)[\Psi(y, \beta, \pi) + \Psi_y(y, \beta, \pi)] \right\}_{|y| < \delta_{\pi,\beta}}$$

is uniformly integrable where $\Psi(y, \beta, \pi)$ is defined in (5.3.6) and

$$\begin{aligned} \Psi_y(y, \beta, \pi) &:= \frac{d}{dy} \Psi(y, \beta, \pi) \\ &= - \int_0^\tau \beta^2(s) \sigma^2(s) ds - \int_0^\tau \int_{\mathbb{R}_0} \frac{\beta^2(s) \theta^2(s, z)}{[1 + (\pi(s) + y\beta(s))\theta(s, z)]^2} N(d^-s, dz) \\ &\quad - \int_0^\tau \frac{\beta^2(s) \kappa^2(s)}{[1 + (\pi(s) + y\beta(s))\kappa(s)]^2} dH(s). \end{aligned} \quad (5.3.16)$$

Since it is reasonable to assume that the coefficients σ , θ and κ are not zero on the same time intervals, then $\Psi_y(y, \beta, \pi) < 0$ for $|y| < \delta_{\pi,\beta}$ and $\beta \neq 0$.

Lemma 5.3.1 will give us a sufficient condition for the concavity of (5.3.9) in Theorem 5.3.3.

Lemma 5.3.1. Suppose $A_{u,i}$ and A_{d^2} hold with $|y| < \delta_{\pi,\beta}$, $\beta \in \mathcal{A}_{\mathbb{G}}$ bounded, and that the utility function U satisfies

$$xU''(x) + U'(x) \leq 0 \quad \text{for all } x > 0. \quad (5.3.17)$$

Then for $\pi \in \mathcal{A}_{\mathbb{G}}$ the mappings (5.3.9), $y \rightarrow \mathbb{E}[U(X_{\pi+y\beta}(T))]$, $y \in (-\delta_{\pi,\beta}, \delta_{\pi,\beta})$, $\delta_{\pi,\beta} > 0$, are concave for all bounded controls $\beta \in \mathcal{A}_{\mathbb{G}}$.

Proof. By assumptions $A_{u,i}$ and A_{d^2} the following equations hold true:

$$\begin{aligned} \frac{d^2}{dy^2} \mathbb{E}[U(X_{\pi+y\beta}(T))] &= \\ &= \frac{d}{dy} \mathbb{E}\left[\left(U'(X_{\pi+y\beta}(T))X_{\pi+y\beta}(T)\Psi(y, \beta, \pi)\right)\right] \\ &= \mathbb{E}\left[X_{\pi+y\beta}(T)\Psi^2(y, \beta, \pi)\left(U''(X_{\pi+y\beta}(T))X_{\pi+y\beta}(T) + U'(X_{\pi+y\beta}(T))\right)\right. \\ &\quad \left.+ U'(X_{\pi+y\beta}(T))X_{\pi+y\beta}(T)\Psi_y(y, \beta, \pi)\right], \quad |y| < \delta_{\pi,\beta} \end{aligned} \quad (5.3.18)$$

Thanks to (5.3.17) and the observation that $\Psi_y(y, \beta, \pi) < 0$ for all $|y| < \delta_{\pi,\beta}$, both summands are negative and the mapping (5.3.9) is locally concave. $\square$

Remark 5.3.2. Examples of utility functions satisfying (5.3.17) are the power utility $U(x) = \frac{1}{1-c}x^{1-c}$ when $c > 1$, and logarithmic utility $U(x) = \log(x)$, while the exponential utility, $U(x) = \frac{-1}{\gamma}e^{-\gamma x}$, does not.

Remark 5.3.3. Condition (5.3.17) can also be discussed in terms of the *Arrow Pratt measure of relative risk aversion*. This measure is defined by

$$R_u(x) = \frac{-xU''(x)}{U'(x)},$$

so an equivalent way of stating condition (5.3.17) would be to require the $R_u(x) \geq 1$.

We can use a concavity argument from the derivatives to get some form of uniqueness. A similar argument occurs in [72], where it is proven that local maximums are unique in the case of logarithmic utility under some restriction on admissible controls. In our case we have the following result:

Theorem 5.3.4. Suppose A is a convex set in $\mathcal{A}_{\mathbb{G}}$ such that all $\pi \in A$ are bounded. If for all $\pi \in A$ assumptions $A_{u.i.}$, A_{d^2} with $|y| < \delta_{\pi,\beta}$, $\beta \in \mathcal{A}_{\mathbb{G}}$ bounded, and (5.3.17) is satisfied, then there can at most be one local maximum in A .

Proof. Suppose $\pi_1, \pi_2 \in A$ are two local maximums. Let $\pi_2 - \pi_1 = \bar{\pi}$. Since A is convex, we have $\pi_1 + y\bar{\pi} \in A$ for $y \in [0, 1]$. We note that

$$\frac{d}{dy}\mathbb{E}[U(X_{\pi_1+y\bar{\pi}}(T))]|_{y=a} = \frac{d}{d\zeta}\mathbb{E}[U(X_{(\pi_1+a\bar{\pi})+\zeta\bar{\pi}}(T))]|_{\zeta=0} \quad \text{for } a \in [0, 1].$$

Indeed $A_{u.i.}$ and A_{d^2} hold for $(\pi_1 + a\bar{\pi})$ as it is an element of A . In particular we can apply Lemma 5.3.1 to conclude that $\frac{d}{dy}\mathbb{E}[U(X_{\pi_1+y\bar{\pi}}(T))]$ is strictly monotone for $y \in [0, 1]$.

We show that there cannot exist two local maximums by contradiction. Consider

$$\frac{d}{dy}\mathbb{E}[U(X_{\pi_1+y\bar{\pi}}(T))]|_{y=1} = \frac{d}{d\zeta}\mathbb{E}[U(X_{\pi_1+\bar{\pi}+\zeta\bar{\pi}}(T))]|_{\zeta=0} = \frac{d}{d\zeta}\mathbb{E}[U(X_{\pi_2+\zeta\bar{\pi}}(T))]|_{\zeta=0} = 0 \quad (5.3.19)$$

since $\pi_1 + \bar{\pi} = \pi_2$, and π_2 is a local maximum. On the other hand, we also have that π is a local maximum, hence

$$\frac{d}{dy}\mathbb{E}[U(X_{\pi_1+y\bar{\pi}}(T))]|_{y=0} = 0.$$

Consequently $\frac{d}{dy}\mathbb{E}[U(X_{\pi_1+y\bar{\pi}}(T))]$ is strictly monotone and zero at two different points, which is absurd. $\square$

5.3.1 Some examples with logarithmic utility

We concentrate on the logarithmic utility to reduce computation and highlight some interesting aspects of the analysis. Note that if $U(x) = \ln(x)$ then $F_{\pi}(T) = 1$ in (5.3.8). By application of

Theorem 5.3.3 the following equation plays a crucial role:

$$\begin{aligned}
0 &= \mathbb{E}\left[F_\pi(T)(M_\pi(s) - M_\pi(t)) \middle| \mathcal{G}_t\right] \\
&= \mathbb{E}\left[M_\pi(s) - M_\pi(t) \middle| \mathcal{G}_t\right] \\
&= \mathbb{E}\left[\int_{t \wedge \tau}^{s \wedge \tau} [\mu(r) - \rho(r) - \sigma^2(r)\pi(r) - \int_{\mathbb{R}_0} \frac{\pi(r)\theta^2(r, z)}{1 + \pi(r)\theta(r, z)} \nu(dz)] dr \right. \\
&\quad + \int_{t \wedge \tau}^{s \wedge \tau} \sigma(r) d^-W(r) + \int_{t \wedge \tau}^{s \wedge \tau} \int_{\mathbb{R}_0} \frac{\theta(r, z)}{1 + \pi(r)\theta(r, z)} \tilde{N}(d^-r, dz) \\
&\quad \left. + \int_{t \wedge \tau}^{s \wedge \tau} \frac{\kappa(r)}{1 + \kappa(r)\pi(r)} dH(r) \middle| \mathcal{G}_t\right] \quad s \geq t. \tag{5.3.20}
\end{aligned}$$

Example 5.3.5. Assume that H is independent of W and N and that all the coefficients are $\mathbb{F}$ -adapted, as in classical market modeling. Further we assume that $\Lambda(ds) = \lambda(s)ds$, for some positive stochastic process λ .

We consider the case of an investor having access to an information flow $\mathbb{G}$ with $\mathcal{G}_t \subseteq \mathcal{F}_t \vee \mathcal{F}_t^H$, for all $t \in [0, T]$. We call this a case of partial information. The expectation of the forward integrals in (5.3.20) are zero in this setup, so the equation can be written

$$\begin{aligned}
0 &= \mathbb{E}\left[\int_{t \wedge \tau}^{s \wedge \tau} [\mu(r) - \rho(r) - \sigma^2(r)\pi(r) - \int_{\mathbb{R}_0} \frac{\pi(r)\theta^2(r, z)}{1 + \pi(r)\theta(r, z)} \nu(dz)] dr \right. \\
&\quad \left. + \int_{t \wedge \tau}^{s \wedge \tau} \frac{\kappa(r)}{1 + \kappa(r)\pi(r)} dH(r) \middle| \mathcal{G}_t\right], \quad s \geq t.
\end{aligned}$$

Dividing by $(s - t)$ and letting $s \rightarrow t$, we find that the locally optimal $\pi(t)$ in this case must satisfy

$$\begin{aligned}
0 &= \mathbf{1}_{\{\tau > t\}} \mathbb{E}\left[\mu(t) - \rho(t) - \sigma^2(t)\pi(t) \right. \\
&\quad \left. - \int_{\mathbb{R}_0} \frac{\pi(r)\theta^2(r, z)}{1 + \pi(r)\theta(r, z)} \nu(dz) + \frac{\kappa(t)}{1 + \kappa(t)\pi(t)} \lambda(t) \middle| \mathcal{G}_t\right]. \tag{5.3.21}
\end{aligned}$$

For illustration, assume $\theta = 0$. Then (5.3.21) yields a polynomial equation in $\pi(t)$ of degree 2:

$$\begin{aligned}
0 &= \mathbf{1}_{\tau > t} \left(\mathbb{E}\left[\mu(t) - \rho(t) + \kappa(t)\lambda(t) \middle| \mathcal{G}_t\right] + \pi(t) \mathbb{E}\left[\mu(t)\kappa(t) - \rho(t)\kappa(t) - \sigma^2(t)\lambda(t) \middle| \mathcal{G}_t\right] \right. \\
&\quad \left. - \pi^2(t) \mathbb{E}\left[\sigma^2(t)\kappa(t)\lambda(t) \middle| \mathcal{G}_t\right] \right)
\end{aligned}$$

Example 5.3.6. Assume that $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{F}_t^H$ and that H contains no anticipating information on $\mathbb{F}$. In this case we say that the investor has full information. If $\Lambda(ds) = \lambda(s)ds$ and the

(a) Default risk *not* compensated by higher drift

(b) Default risk compensated by higher drift

Figure 5.1: Optimal investment π as a function of λ .

coefficients $\mu, \sigma, \theta, \kappa$ are adapted to $\mathbb{G}$, equation (5.3.21) reduces to

$$0 = \mathbf{1}_{\{\tau > t\}} \left(\mu(t) - \rho(t) - \sigma^2(t)\pi(t) - \int_{\mathbb{R}_0} \frac{\pi(r)\theta^2(r, z)}{1 + \pi(r)\theta(r, z)} \nu(dz) + \frac{\kappa(t)}{1 + \kappa(t)\pi(t)} \lambda(t) \right) \quad (5.3.22)$$

If we assume $\theta = 0$, the explicit solution of (5.3.22) is given by

$$\pi = \frac{1}{2\kappa} \left(\frac{\kappa(\mu - \rho)}{\sigma^2} - 1 + \sqrt{\left(1 - \frac{\kappa(\mu - \rho)}{\sigma^2}\right)^2 + 4\kappa \left(\frac{\mu - \rho + \lambda\kappa}{\sigma^2}\right)} \right), \quad (5.3.23)$$

where we used (5.3.1) to exclude one of the two solutions of the quadratic equation. Remark that equation (5.3.22) gives us Merton ratio when $\kappa = \theta = 0$.

Two explicit examples with full information can be found in the figures. In Figure 5.1a the

stock price is modeled as

$$dS_1(t) = S(t-)\left(\mu_o dt + \sigma dW(t) + \kappa dH(t)\right), \quad (5.3.24)$$

with μ_o, σ, κ fixed, and $\rho = 0$. We see that with higher default risk the agent invests less and the asset is also shorted when the overall return becomes negative at the point $\lambda\kappa = -\mu_o$.

In Figure 5.1b the stock price is modeled as

$$dS_1(t) = S(t-)\left((\mu_o - \lambda\kappa)dt + \sigma dW(t) + \kappa dH(t)\right) \quad (5.3.25)$$

The assumptions in (5.3.25) are similar to (5.3.24). But with the term $-\lambda\kappa$ in the drift, the expected return of the asset is invariant to the value of λ . So the agent invests less due to risk aversion and not due to changes in the asset returns.

Next we explicitly detail how after-default and/or multiple defaults are easily treated in our framework.

Example 5.3.7. Here we discuss a model with default time ζ . After default the asset has a recovery process with different dynamics than before default. We assume that it is possible to invest both before and after default and set $\tau = T$. Set $H(t) = \mathbf{1}_{\{\zeta \leq t\}}(t)$ and

$$\begin{aligned} \mu(s) &= \mu_1 \mathbf{1}_{\{H(s-)=0\}}(s) + \mu_2 \mathbf{1}_{\{H(s-)>0\}}(s) \\ \theta(s, z) &= z\theta_1 \mathbf{1}_{\{H(s-)=0\}}(s) + z\theta_2 \mathbf{1}_{\{H(s-)>0\}}(s) \\ \sigma(s) &= \sigma_1 \mathbf{1}_{\{H(s-)=0\}}(s) + \sigma_2 \mathbf{1}_{\{H(s-)>0\}}(s) \end{aligned}$$

where $\mu_1, \mu_2, \theta_1, \theta_2, \sigma_1, \sigma_2 \in \mathbb{R}$. Here θ_1 and σ_1 are the coefficients of the noises of the asset dynamics pre-default and θ_2 and σ_2 the coefficients after default while μ_1 and μ_2 are the drift coefficients before and after default respectively.

In the case of full information as above, $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{F}_t^H$, the optimization scheme separates into pre-default and after default. The optimal portfolio π satisfies:

$$\begin{aligned} 0 &= \mathbf{1}_{\{\zeta > t\}} \left(\mu_1 - \rho(t) - \sigma_1^2(t)\pi(t) \right. \\ &\quad \left. - \int_{\mathbb{R}_0} \frac{\pi(r)z\theta_1^2}{1 + \pi(r)z\theta_1} \nu(dz) + \frac{\kappa(t)}{1 + \kappa(t)\pi(t)} \lambda(t) \right) \\ &\quad + \mathbf{1}_{\{\zeta \leq t\}} \left(\mu_2 - \rho(t) - \sigma_2^2(t)\pi(t) \right. \\ &\quad \left. - \int_{\mathbb{R}_0} \frac{\pi(r)z\theta_2^2}{1 + \pi(r)z\theta_2} \nu(dz) \right). \end{aligned}$$

The cases of anticipating information, i.e. $\mathcal{G}_t \supseteq \mathcal{F}_t \vee \mathcal{F}_t^H$, are more subtle than partial or full information, with various approaches being possible depending on the specific conditions. The main challenge with anticipating information is to evaluate the terms $\mathbb{E}\left[\int_{t \wedge \tau}^{s \wedge \tau} \sigma(r) d^-W(r) | \mathcal{G}_t\right]$ and $\mathbb{E}\left[\int_{t \wedge \tau}^{s \wedge \tau} \int_{\mathbb{R}_0} \frac{\theta(r, z)}{1 + \pi(r)\theta(r, z)} \tilde{N}(dz, d^-r) | \mathcal{G}_t\right]$ in (5.3.20).

One possible way to compute the expectations of the forward integrals above is to exploit Malliavin calculus, see [43, Chapter 8 and Chapter 15] for the theoretical framework. However

we must stress that this general approach cannot always be taken here. In fact the application of Malliavin calculus requires that the integrands are measurable with respect to $\mathcal{F}_T$, which is not, in general, the case when considering default risk.

See also [40] on how the $\tilde{N}(d^-t, dz)$ integral can be evaluated using predictable compensators of the measure with respect to $\mathbb{G}$ and [12, 75, 72] for other examples on the d^-W integral in insider models without default risk.

Hereafter we show an example where the process H contains anticipating information on the jumps of N . In this example the process $\int_0^t \int_{\mathbb{R}_0} z \tilde{N}(ds, dz)$, $t \in [0, T]$, is martingale under $(\mathbb{P}, \mathbb{F})$ but not under $(\mathbb{P}, \mathbb{G})$. While we claim no particular market model related to this, we show how the suggested framework enables solutions for optimization problems with anticipating information. In particular the process H generalizes the optimization problem not only because of it's precense, but also because it adds knowledge on the other noises driving S_1 . Nevertheless a solution is obtained from Theorem 5.3.3.

Example 5.3.8. Assume $\nu(dz) = \gamma \mathbf{1}_{\{1\}}(dz)$, i.e. the Poisson random measure $N(dt, dz)$ is actually $dN(t)$ where $N(t)$, $t \geq 0$ is a Poisson process with intensity γ . Define $\tau = \inf_t \{N(t + \epsilon) - N(t - \epsilon) \geq 2\}$, $\epsilon > 0$, and $H(t) = \mathbf{1}_{\{\tau \leq t\}}(t)$. Thus H is independent of W but is dependent on N . Note that τ contains anticipating information with respect to $\mathbb{F}$ since $(\tau > t)$ implies $N(t + \epsilon) - N(t - \epsilon) < 2$. Set $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{F}_t^H$.

Our ad hoc interpretation is that too many bad events (represented by N) in a limited time span (2ϵ) will cause the firm to default (with a loss κH and the asset is no longer tradeable τ).

We assume $\rho, \mu, \sigma, \theta$ and κ are constants. Starting from (5.3.20)

$$\begin{aligned} 0 = \mathbb{E} \Bigg[& \int_{t \wedge \tau}^{s \wedge \tau} \left[\mu - \rho - \sigma^2 \pi(r) - \frac{\pi(r) \theta^2}{1 + \pi(r) \theta^2} \gamma \right] dr + \int_{t \wedge \tau}^{s \wedge \tau} \sigma(r) d^-W(r) \\ & + \int_{t \wedge \tau}^{s \wedge \tau} \frac{\theta}{1 + \pi(r) \theta} (N(dr) - \gamma dr) + \int_{t \wedge \tau}^{s \wedge \tau} \frac{\kappa}{1 + \kappa \pi(r)} dH(r) \Big| \mathcal{G}_t \Bigg]. \end{aligned}$$

Computing the predictable compensators of H and N (sketched below), and dividing by $(t - s)$ we find that the optimal π is a solution of

$$\begin{aligned} 0 = & \mu - \sigma^2 \pi(r) - \frac{\pi(r) \theta^2}{1 + \pi(r) \theta^2} \gamma + \frac{\theta(r, z)}{1 + \pi(r) \theta} \frac{\gamma}{1 + \gamma \epsilon} \mathbf{1}_{\{N(t) - N(t - \epsilon) = 0\}} \\ & + \mathbf{1}_{\{N(t) - N(t - \epsilon) = 0\}} \frac{\gamma^2 \epsilon}{1 + \gamma \epsilon} + \mathbf{1}_{\{N(t) - N(t - \epsilon) = 1\}} \gamma. \end{aligned}$$

To compute the $\mathbb{G}$ -predictable compensators of H and N we investigate the intensities (see, e.g. [21, Section 3.2]) on the set $\{t < \tau\}$

$$\begin{aligned} \lambda_t^N &:= \lim_{\Delta t \rightarrow 0^+} \frac{1}{\Delta t} \mathbb{E}[N(t + \Delta t) - N(t) | \mathcal{G}_t, t < \tau], \\ \lambda_t &:= \lim_{\Delta t \rightarrow 0^+} \frac{1}{\Delta t} \mathbb{E}[(H(t + \Delta t) - H(t)) | \mathcal{G}_t, t < \tau]. \end{aligned}$$

The $\mathbb{G}$ -predictable compensators of H and N are then given by $\Lambda(t) = \int_0^t \lambda(s)ds$ and $\Lambda^N(t) = \int_0^t \lambda^N(s)ds$. We consider the case of N , the computations for H are similar. First note that

$$\lim_{\Delta t \rightarrow 0^+} \frac{1}{\Delta t} \mathbb{P}(N(t + \Delta t) - N(t) > 1 | \mathcal{G}_t, t \leq \tau) = 0.$$

Recall that $(\tau > t)$ implies $N(t + \epsilon) - N(t - \epsilon) < 2$ and that $(\tau > t, N(t) - N(t - \epsilon) = 1)$ implies $N(t + \epsilon) - N(t) = 0$. For $\Delta t < \epsilon$,

$$\begin{aligned} \mathbb{P}(N(t + \Delta t) - N(t) = 1 | \mathcal{G}_t, t \leq \tau) &= \\ \mathbf{1}_{\{N(t) - N(t - \epsilon) = 0\}} \mathbb{P}(N(t + \Delta t) - N(t) = 1 | N(t) - N(t - \epsilon) = 0, t \leq \tau) &+ \\ + \mathbf{1}_{\{N(t) - N(t - \epsilon) = 1\}} \mathbb{P}(N(t + \Delta t) - N(t) = 1 | N(t) - N(t - \epsilon) = 1, t \leq \tau). \end{aligned}$$

We have

$$\begin{aligned} \mathbb{P}(N(t + \Delta t) - N(t) = 1 | t \leq \tau, N(t) - N(t - \epsilon) = 0) &= \\ = \mathbb{P}(N(t + \Delta t) - N(t) = 1 | N(t + \epsilon) - N(t) < 2) &= \\ = \frac{\mathbb{P}(N(t + \Delta t) - N(t) = 1, N(t + \epsilon) - N(t + \Delta t) = 0)}{\mathbb{P}(N(t + \epsilon) - N(t) < 2)} &= \\ = \frac{\gamma \Delta t e^{-\gamma \Delta t} e^{-\gamma(\epsilon - \Delta t)}}{e^{-\gamma \epsilon} + \gamma \epsilon e^{-\gamma \epsilon}} \end{aligned}$$

and

$$\mathbb{P}(N(t + \Delta t) - N(t) = 1 | t \leq \tau, N(t) - N(t - \epsilon) = 1) = 0.$$

Thus

$$\lambda^N(t) = \frac{\gamma}{1 + \gamma \epsilon} \mathbf{1}_{\{N(t) - N(t - \epsilon) = 0\}} \quad \text{for } t \leq \tau.$$

Similarly we find

$$\lambda_t = \mathbf{1}_{\{N(t) - N(t - \epsilon) = 0\}} \frac{\gamma^2 \epsilon}{1 + \gamma \epsilon} + \mathbf{1}_{\{N(t) - N(t - \epsilon) = 1\}} \gamma \quad \text{for } t \leq \tau.$$

5.4 On the driving processes as semi-martingales

The results of Theorem 5.3.3 take a more specific form when M_π is $\mathbb{G}$ -adapted. This will be our standing assumption throughout the section, implying that $\mathcal{F}_t, \mathcal{F}_t^H \subset \mathcal{G}_t$ for all $t \in [0, T]$ and that the integrands $\rho, \mu, \sigma, \theta$ and κ are $\mathbb{G}$ -adapted.

Theorem 5.4.1. Suppose that M_π is $\mathbb{G}$ -adapted and that for $\pi \in \mathcal{A}_{\mathbb{G}}$ assumption $A_{u.i.}$ holds.

- i) If π is a local maximum, then $M_\pi(t), t \in [0, T]$, is a martingale under $(\mathbb{Q}_\pi, \mathbb{G})$.

ii) If π is a local maximum, then the stochastic process

$$\hat{M}_\pi(t) = M_\pi(t) - \int_0^t \frac{1}{Z(s)} d[M_\pi, Z](s), \quad t \in [0, T],$$

is a martingale under $(\mathbb{P}, \mathbb{G})$. Here, we have set

$$Z(t) := \mathbb{E}_{\mathbb{Q}_\pi} \left[\frac{d\mathbb{P}}{d\mathbb{Q}_\pi} \middle| \mathcal{G}_t \right] = \left(\mathbb{E}[F_\pi(T) | \mathcal{G}_t] \right)^{-1}.$$

Assume that the mapping $y \rightarrow \mathbb{E}[U(X_{\pi+y\beta}(T))]$ is concave for all bounded controls $\beta \in \mathcal{A}_\mathbb{G}$. Then we also have the converse conclusions

iii) If M_π is a martingale under $(\mathbb{Q}_\pi, \mathbb{G})$, then π is a local maximum.

iv) If the stochastic process

$$\hat{M}_\pi(t) = M_\pi(t) - \int_0^t \frac{1}{Z(s)} d[M_\pi, Z](s), \quad t \in [0, T],$$

is a martingale under $(\mathbb{P}, \mathbb{G})$, then π is a local maximum.

Proof.

Part i) Being M_π $\mathbb{G}$ -adapted, it is a $(\mathbb{P}, \mathbb{G})$ -martingale.

Part ii) is obtained by application of the Girsanov theorem (see in particular [92, Part III, Theorem 39]).

Part iii) is a direct application of Theorem 5.3.3.

Part iv) is again an application of the Girsanov theorem. $\square$

The existence of a local maximum also has other implications.

Theorem 5.4.2. If a local maximum exists, M_π is $\mathbb{G}$ -adapted and $A_{u.i.}$ holds, then W and $\int_0^{t \wedge \tau} \theta(s, z) \tilde{N}(ds, dz)$, $t \in [0, T]$, are semi-martingales under $(\mathbb{P}, \mathbb{G})$.

Proof. Assume a local maximum $\pi \in \mathcal{A}_\mathbb{G}$ exists. By Theorem 5.3.3 this implies that M_π is a $(\mathbb{Q}_\pi, \mathbb{G})$ -martingale. Define

$$M_\pi^H(t) := \int_0^{t \wedge \tau} \frac{\kappa(s)}{1 + \pi(s)\kappa(s)} dH(s) - \int_0^{t \wedge \tau} \frac{\kappa(s)}{1 + \pi(s)\kappa(s)} \Lambda^{\mathbb{Q}_\pi}(ds), \quad t \in [0, T],$$

where $\Lambda^{\mathbb{Q}_\pi}$ is the $(\mathbb{Q}_\pi, \mathbb{G})$ -predictable compensator of H . Note that M_π^H is a $(\mathbb{Q}_\pi, \mathbb{G})$ -martingale and thus $M_\pi - M_\pi^H$ is also a $(\mathbb{Q}_\pi, \mathbb{G})$ -martingale. We can (uniquely) decompose $M_\pi - M_\pi^H$

into a continuous martingale and pure jump martingale [58, Theorem 1.4.18] which we denote by M_π^W and M_π^N respectively. By the definition of M_π it follows that we can write

$$\begin{aligned} M_\pi^W(t) &:= \int_0^{t \wedge \tau} \sigma(s) d^-W(s) + \int_0^{t \wedge \tau} \sigma(s) a(s) ds, \quad t \in [0, T], \\ M_\pi^N(t) &:= \int_0^{t \wedge \tau} \int_{\mathbb{R}_0} \frac{\theta(s, z)}{1 + \pi(s)\theta(s, z)} \tilde{N}(d^-s, dz) + \int_0^{t \wedge \tau} \gamma(s) ds, \quad t \in [0, T], \end{aligned}$$

where a and γ is such that

$$\begin{aligned} \int_0^{t \wedge \tau} \sigma(s) a(s) ds + \int_0^{t \wedge \tau} \gamma(s) ds &= \int_0^{t \wedge \tau} \mu(s) - \rho(s) - \sigma^2(s) \pi(s) \\ &\quad - \int_{\mathbb{R}_0} \frac{\theta^2(s, z) \pi(s)}{1 + \pi(s)\theta(s, z)} \nu(dz) ds + \int_0^{t \wedge \tau} \frac{\kappa(s)}{1 + \kappa(s) \pi(s)} \Lambda^{\mathbb{Q}_\pi}(ds). \end{aligned} \quad (5.4.1)$$

The right hand side of (5.4.1) have a finite $\mathbb{P}$ expectation by the assumptions (5.1.4) and Definition 5.3.1 so that $\int_0^t \sigma(s) a(s) ds$ and $\int_0^t \gamma(s) ds$ are processes of finite variation.

As in Theorem 5.4.1, $M_\pi^W - \int_0^{t \wedge \tau} \frac{1}{Z(s)} d[M_\pi^W, Z]$ is a $(\mathbb{P}, \mathbb{G})$ -martingale. We note that $[M_\pi^W, Z]$ is absolutely continuous with respect to Lebesgue by the Kunita-Watanabe-inequality (see for instance [92, Theorem 25]) since the quadratic variation of M_π^W is absolutely continuous with respect to Lebesgue. Thus the quadratic variation of $\int_0^{t \wedge \tau} \frac{1}{\sigma(s)} M_\pi^W(ds)$ is t , making $\frac{1}{\sigma} M_\pi^W$ a $(\mathbb{P}, \mathbb{G})$ -Brownian motion. Hence W has the $(\mathbb{P}, \mathbb{G})$ semi-martingale decomposition

$$W(t) = \tilde{W}(t) + \int_0^{t \wedge \tau} a(s) ds - \int_0^{t \wedge \tau} \frac{1}{Z(s) \sigma(s)} [M_\pi^W, Z] ds$$

where $\tilde{W}$ is a $(\mathbb{P}, \mathbb{G})$ -Brownian motion.

Similarly,

$$\begin{aligned} M_\pi^N(t) &- \int_0^{t \wedge \tau} \frac{1}{Z(s)} d[M_\pi^N, Z](s) \\ &= \int_0^{t \wedge \tau} \int_{\mathbb{R}_0} \frac{\theta(s, z)}{1 + \theta(s, z) \pi(s)} \tilde{N}(d^-s, dz) + \int_0^{t \wedge \tau} \gamma(s) ds - \int_0^t \frac{1}{Z(s)} d[M_\pi^N, Z](s), \quad t \in [0, T], \end{aligned} \quad (5.4.2)$$

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

$$\int_0^{t \wedge \tau} \frac{1}{Z(s)} d[M_\pi^N, Z](s) = \int_0^{t \wedge \tau} \frac{\mathbb{E}[U(X_\pi(T))X_\pi(T)|\mathcal{G}_s]}{\mathbb{E}[U(X_\pi(T))X_\pi(T)]} d[M_\pi^N, Z](s), \quad t \in [0, T],$$

is a $\mathbb{G}$ -adapted process of finite variation and thus a $(\mathbb{P}, \mathbb{G})$ semi-martingale (recall that $[M_\pi^N, Z]$ is of finite variation [92, p. 67]). Since also M_π^N and $\int_0^t \gamma(s)ds$ are $(\mathbb{P}, \mathbb{G})$ semi-martingales we must have that

$$\int_0^{t \wedge \tau} \int_{\mathbb{R}_0} \frac{\theta(s, z)}{1 + \theta(s, z)\pi(s)} \tilde{N}(d^-s, dz), \quad t \in [0, T]$$

is a $(\mathbb{P}, \mathbb{G})$ semi-martingale by (5.4.2). Since the $1 + \theta\pi$ is bounded away from zero (recall Definition 5.3.1) we must also have that $\int_0^{t \wedge \tau} \int_{\mathbb{R}_0} \theta(s, z) \tilde{N}(d^-s, dz)$ is a semi-martingale. $\square$

Finally we do an analysis on the jumps of H .

Theorem 5.4.3. Assume that a local maximum exists, $N = 0$, M_π is $\mathbb{G}$ -adapted and Assumption $A_{u.i.}$ holds. Then the jumps of H are totally inaccessible stopping times (for the filtration $\mathbb{G}$) and the $(\mathbb{Q}_\pi, \mathbb{G})$ -predictable compensators of H is absolutely continuous with respect to Lebesgue.

Proof. With the above assumptions, M_π is a $(\mathbb{Q}_\pi, \mathbb{G})$ -martingale by Theorem 5.3.3. Denote

$$A_1(t) = M_\pi(t) - \int_0^t \frac{\kappa(s)}{1 + \kappa(s)\pi(s)} H(ds)$$

$$A_2(t) = \int_0^t \frac{\kappa(s)}{1 + \kappa(s)\pi(s)} H(ds),$$

for $t \in [0, T]$. Remark that $A_1 + A_2 = M_\pi$ is a martingale. Since A_2 is discontinuous, A_1 must be the sum of the predictable compensator of A_2 and a martingale. Hence, since A_1 is continuous the compensator of A_2 is continuous. It immediately follows that the jump times of H are totally inaccessible (see, e.g. [65, Corollary 22.18]). $\square$

Bibliography

- [1] K. AASE, B. ØKSENDAL, N. PRIVAULT, AND J. UBØE, *White noise generalizations of the Clark-Haussmann-Ocone theorem with application to mathematical finance*, Finance and Stochastics, 4 (2000), pp. 465–496.
- [2] T. T. K. AN AND B. ØKSENDAL, *Maximum principle for stochastic differential games with partial information*, Journal of Optimization Theory and Applications, 139 (2008), pp. 463–483.
- [3] —, *A maximum principle for stochastic differential games with g -expectations and partial information*, Stochastics An International Journal of Probability and Stochastic Processes, 84 (2012), pp. 137–155.
- [4] T. T. K. AN, F. PROSKE, AND M. RUBTSOV, *A SPDE maximum principle for stochastic differential games under partial information with application to optimal portfolios on fixed income markets*, Stochastics, An International Journal of Probability and Stochastic Processes, 82 (2010), pp. 3–23.
- [5] D. APPLEBAUM, *Lévy Processes and Stochastic Calculus - second edition*, Cambridge university press, 2009.
- [6] F. BAGHERY AND B. ØKSENDAL, *A maximum principle for stochastic control with partial information*, Stochastic Analysis and Applications, 25 (2007), pp. 705–717.
- [7] O. BARNDORFF-NIELSEN, M. MAEJIMA, AND K.-I. SATO, *Infinite divisibility for stochastic processes and time change*, Journal of Theoretical Probability, 19 (2006), pp. 411–446.
- [8] O. E. BARNDORFF-NIELSEN, E. NICOLATO, AND N. SHEPHARD, *Some recent developments in stochastic volatility modelling*, Quantitative Finance, 2 (2002), pp. 11–23.
- [9] N. BÄUERLE AND U. RIEDER, *Portfolio optimization with jumps and unobservable intensity process*, Mathematical Finance, 17 (2007), pp. 205–224.
- [10] A. BENSOUSSAN, *Stochastic control of partially observable systems*, Cambridge University Press Cambridge, 1992.
- [11] F. E. BENTH, G. DI NUNNO, A. LØKKA, B. ØKSENDAL, AND F. PROSKE, *Explicit representation of the minimal variance portfolio in markets driven by Lévy processes*, Mathematical Finance, 13 (2003), pp. 55–72.

- [12] F. BIAGINI AND B. ØKSENDAL, *A general stochastic calculus approach to insider trading*, Applied Mathematics Optimization, 52 (2005), pp. 167–181.
- [13] T. BIELECKI AND I. JANG, *Portfolio optimization with a defaultable security*, Asia-Pacific Financial Markets, 13 (2006), pp. 113–127.
- [14] T. BIELECKI, M. JEANBLANC, AND M. RUTKOWSKI, *Hedging of Defaultable Claims*. Paris-Princeton Lectures on Mathematical Finance, 2004.
- [15] —, *PDE approach to valuation and hedging of credit derivatives*, Quantitative Finance, 5 (2007), pp. 257–270.
- [16] P. BILLINGSLEY, *Probability and Measure*, Wiley, 1995.
- [17] C. BLANCHET-SCALLIET, N. E. KAROUI, M. JEANBLANC, AND L. MARTELLINI, *Optimal investment decisions when time-horizon is uncertain*, Journal of Mathematical Economics, 44 (2008), pp. 1100–1113.
- [18] O. BOBROVNYTSKA AND M. SCHWEIZER, *Mean-variance hedging and stochastic control: beyond the brownian setting*, Automatic Control, IEEE Transactions on, 49 (2004), pp. 396–408.
- [19] R. BOEL, P. VARAIYA, AND E. WONG, *Martingales on jump processes I. Representation results*, SIAM Journal on control, 13 (1975).
- [20] B. BOUCHARD AND H. PHAM, *Wealth-path dependent utility maximization in incomplete markets*, Finance and Stochastics, 8 (2004), pp. 579–603. 10.1007/s00780-004-0125-8.
- [21] P. BRÉMAUD, *Point Processes and Queues - Martingale dynamics*, Springer, 1981.
- [22] P. BREMAUD AND M. YOR, *Changes of filtrations and of probability measures*, Probability Theory and Related Fields, 45 (1978), pp. 269–295.
- [23] R. CAIROLI AND J. WALSH, *Stochastic integrals in the plane*, Acta Mathematica, 134 (1975), pp. 111–183.
- [24] R. CARBONE, B. FERRARIO, AND M. SANTACROCE, *Backward stochastic differential equations driven by càdlàg martingales*, Theory of Probability and its Applications, 52 (2008), pp. 304–314.
- [25] P. CARR, H. GERMAN, D. MADAN, AND M. YOR, *Stochastic volatility for Lévy processes*, Mathematical Finance, 13 (2003), pp. 345–282.
- [26] E. ÇINLAR, *Measure and integration*, in Probability and Stochastics, vol. 261 of Graduate Texts in Mathematics, Springer New York, 2011, pp. 1–47.
- [27] D. COCULESCU, M. JEANBLANC, AND A. NIKEGHBALI, *Default times, no-arbitrage conditions and changes of probability measures*, Finance and Stochastics, 16 (2012), pp. 513–535.

- [28] D. R. COX, *Some statistical methods connected with series of events*, Journal of the Royal Statistical Society. Series B (Methodological), 17 (1955), pp. 129–164.
- [29] D. DALEY AND D. VERE-JONES, *An Introduction to the Theory of Point Processes*, Springer, 2003.
- [30] D. DAVID, Y. Y. OKUR, ET AL., *Optimal consumption and portfolio for an insider in a market with jumps*, Communications on Stochastic Analysis, 3 (2009), pp. 101–117.
- [31] M. H. A. DAVIS, *The representation of martingales of jump processes*, SIAM Journal on control and optimization, 14 (1976), pp. 623–638.
- [32] M. H. A. DAVIS, *Martingale representation and all that*, in Advances in Control, Communication Networks, and Transportation Systems, E. H. Abed, ed., Systems & Control: Foundations & Applications, Birkhauser Boston, 2005, pp. 57–68.
- [33] L. DELONG, *Optimal investment and consumption in the present of default in a financial market driven by a Lévy process*, Annales Universitatis Mariae Curie-Sklodowska, LX (2006).
- [34] G. DI NUNNO, *Stochastic integral representation, stochastic derivatives and minimal variance hedging*, Stochastics and Stochastics Reports, 73 (2002), pp. 181–198.
- [35] ———, *Random Fields Evolution: non-anticipating integration and differentiation*, Theory of Probability and Mathematical Statistics, AMS, 66 (2003), pp. 91–104.
- [36] ———, *On orthogonal polynomials and the Malliavin derivative for Lévy stochastic measures*, SMF Seminaires et Congres, 16 (2007), pp. 55–69.
- [37] ———, *Random Fields: non-anticipating derivative and differentiation formulas*, Infin. Dimens. Anal. Quantum Probab. Relat. Top., 10 (2007), pp. 465–481.
- [38] G. DI NUNNO AND I. B. EIDE, *Minimal-variance hedging in large financial markets: random fields approach*, Stochastic Analysis and Applications, 28 (2010).
- [39] G. DI NUNNO, T. MEYER-BRANDIS, B. ØKSENDAL, AND F. PROSKE, *Malliavin calculus and anticipative Itô formulae for Lévy processes*, Infinite Dimensional Analysis, Quantum Probability and Related Topics, 8 (2005), pp. 235–258.
- [40] ———, *Optimal portfolio for an insider in a market driven by Lévy processes*, Quantitative Finance, 6 (2006), pp. 83–94.
- [41] G. DI NUNNO AND B. ØKSENDAL, *Optimal portfolio, partial information and malliavin calculus*, Stochastics: An International Journal of Probability and Stochastics Processes, 81 (2009), pp. 303–322.
- [42] G. DI NUNNO, B. ØKSENDAL, AND F. PROSKE, *White noise analysis for Lévy processes*, Journal of Functional Analysis, 206 (2004), pp. 109 – 148.

- [43] —, *Malliavin Calculus for Lévy Processes with Applications to Finance*, Springer, 2009.
- [44] G. DI NUNNO AND Y. ROZANOV, *Stochastic integrals and adjoint derivatives*, in *Stochastic Analysis and Applications*, F. E. Benth, G. Di Nunno, T. Lindstrøm, B. Øksendal, and T. Zhang, eds., vol. 2 of *Abel Symposia*, Springer Berlin Heidelberg, 2007, pp. 265–307.
- [45] R. DUDLEY, *Uniform Central Limit Theorems*, Cambridge University Press, 1999.
- [46] D. DUFFIE, *Credit risk modeling with affine processes*, *Journal of Banking & Finance*, 25 (2005).
- [47] D. DUFFIE AND K. SINGLETON, *Credit Risk: Pricing, Measurement, and Management*, Princetown University Press, 2003.
- [48] N. C. FRAMSTAD, B. ØKSENDAL, AND A. SULEM, *Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance*, *Journal of Optimization Theory and Applications*, 121 (2004), pp. 77–98.
- [49] E. R. GIANIN, *Risk measures via g-expectations*, *Insurance: Mathematics and Economics*, 39 (2006), pp. 19 – 34.
- [50] J. GRANDELL, *Doubly Stochastic Poisson Processes*, vol. 529 of *Lecture Notes in Mathematics*, Springer-Verlag, 1976.
- [51] B. GRIGELIONIS, *Characterization of stochastic processes with conditionally independent increments*, *Lithuanian Mathematical Journal*, 15 (1975), pp. 562–567.
- [52] S. L. HESTON, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, *Review of financial studies*, 6 (1993), pp. 327–343.
- [53] Y. HOU AND X. JIN, *Optimal investment with default risk*. FAME Research Paper No. 46, 2002.
- [54] J. HULL AND A. WHITE, *The pricing of options on assets with stochastic volatilities*, *The journal of finance*, 42 (1987), pp. 281–300.
- [55] J. JACOD, *Multivariate point processes: Predictable projection, Radon-Nikodym derivatives, representation of martingales*, *Probability Theory and Related Fields*, 31 (1975), pp. 235–253.
- [56] J. JACOD, *Calcul Stochastique et Problèmes de Martingales*, vol. 714 of *Lecture Notes in Mathematics*, Springer Berlin / Heidelberg, 1979.
- [57] —, *Existence and uniqueness for stochastic differential equations*, in *Stochastic Control Theory and Stochastic Differential Systems*, M. Kohlmann and W. Vogel, eds., vol. 16 of *Lecture Notes in Control and Information Sciences*, Springer Berlin / Heidelberg, 1979, pp. 435–446.

- [58] J. JACOD AND A. N. SHIRYAEV, *Limit Theorems for Stochastic Processes*, Springer, 2003.
- [59] R. A. JARROW AND F. YU, *Counterparty risk and the pricing of defaultable securities*, The Journal of Finance, 56 (2001), pp. 1765–1799.
- [60] M. JEANBLANC AND Y. LE CAM, *Immersion property and credit risk modelling*, in Optimality and Risk - Modern Trends in Mathematical Finance, Springer Berlin Heidelberg, 2010, pp. 99–132.
- [61] M. JEANBLANC, M. MANIA, M. SANTACROCE, AND M. SCHWEIZER, *Mean-variance hedging via stochastic control and BSDE's for general semimartingales*. Preprint, to appear in Annals of Applied probability, 2011.
- [62] M. JEANBLANC, M. YOR, M. CHESNEY, M. JEANBLANC, M. YOR, AND M. CHESNEY, *Default risk: An enlargement of filtration approach*, in Mathematical Methods for Financial Markets, Springer Finance, Springer London, 2009, pp. 407–456.
- [63] X. JIANMING, *Backward stochastic differential equation with random measures*, Acta Mathematicae Applicatae Sinica (English Series), 16 (2000), pp. 225–234.
- [64] O. KALLENBERG, *Random measures*, Berlin: Akademie-Verlag, 1986.
- [65] —, *Foundations of Modern Probability*, Springer, 1997.
- [66] J. KALLSEN AND J. MUHLE-KARBE, *Utility maximization in affine stochastic volatility models*, International Journal of Theoretical and Applied Finance, 13 (2010), pp. 459–477.
- [67] —, *Utility maximization in models with conditionally independent increments*, The Annals of Applied Probability, 20 (2010), pp. 2162–2177.
- [68] J. KALLSEN AND A. PAUWELS, *Variance-optimal hedging for time-changed Lévy processes*, Applied Mathematical Finance, 18 (2011), pp. 1–28.
- [69] J. KALLSEN AND R. VIERTHAUER, *Quadratic hedging in affine stochastic volatility models*, Review of Derivatives Research, 12 (2009), pp. 3–27.
- [70] I. KARATZAS AND X.-X. XUE, *A note on utility maximization under partial observations I*, Mathematical Finance, 1 (1991), pp. 57–70.
- [71] N. E. KAROUI, M. JEANBLANC, AND Y. JIAO, *What happens after a default: The conditional density approach*, Stochastic Processes and their Applications, 120 (2010), pp. 1011 – 1032.
- [72] A. KOHATSU-HIGA AND A. SULEM, *Utility maximization in an insider influenced market*, Mathematical Finance, 16 (2006), pp. 153–179.
- [73] M. KOHLMANN, D. XIONG, AND Z. YE, *Mean variance hedging in a general jump model*, Applied Mathematical Finance, 17 (2010), pp. 29–57.

- [74] D. LANDO, *On Cox processes and credit risky securities*, Review of Derivatives Research, 2 (1998), pp. 99–120.
- [75] J. A. LEÓN, R. NAVARRO, AND D. NUALART, *An anticipating calculus approach to the utility maximization of an insider*, Mathematical Finance, 13 (2003), pp. 171–185.
- [76] F. C. LEONE, L. S. NELSON, AND R. B. NOTTINGHAM, *The Folded Normal Distribution*, Technometrics, 3 (1961).
- [77] A. LIM, *Mean-variance hedging when there are jumps*, Siam J. Control Optim., 44 (2005), pp. 1893–1922.
- [78] T. LIM AND M.-C. QUENEZ, *Exponential utility maximization in an incomplete market with defaults*, Electronic Journal of Probability, 16 (2011), pp. 1434–1464.
- [79] A. LØKKA, *Martingale Representation of Lévy Processes*, Stochastic Analysis and Applications, 22 (2005), pp. 867–892.
- [80] S. MATARAMVURA AND B. ØKSENDAL, *Risk minimizing portfolios and hjbi equations for stochastic differential games*, Stochastics An International Journal of Probability and Stochastic Processes, 80 (2008), pp. 317–337.
- [81] Q. MENG, *A maximum principle for optimal control problem of fully coupled forward-backward stochastic systems with partial information*, Science in China Series A: Mathematics, 52 (2009), pp. 1579–1588.
- [82] T. MEYER-BRANDIS, B. ØKSENDAL, AND X. Y. ZHOU, *A mean-field stochastic maximum principle via Malliavin calculus*, Stochastics An International Journal of Probability and Stochastic Processes, 84 (2012), pp. 643–666.
- [83] D. NUALART, *The Malliavin Calculus and Related Topics*, Springer, 1995.
- [84] D. NUALART AND W. SCHOUTENS, *Chaotic and predictable representations for lévy processes*, Stochastic Processes and their Applications, 90 (2000), pp. 109 – 122.
- [85] B. ØKSENDAL, *Stochastic Differential Equations*, Springer, 2005.
- [86] B. ØKSENDAL AND T. ZHANG, *Backward stochastic differential equations with respect to general filtrations and applications to insider finance*. Preprint, <http://urn.nb.no/URN:NBN:no-23398>, 2009.
- [87] O. M. PAMEN, F. PROSKE, AND H. B. SALLEH, *Stochastic differential games in insider markets via Malliavin calculus*, Journal of Optimization Theory and Applications, (2009), pp. 1–42.
- [88] E. PARDOUX AND S. PENG, *Adapted solution of a backward stochastic differential equation*, Systems & Control Letters, 14 (1990), pp. 55–61.
- [89] H. PHAM, *Continuous-time Stochastic Control and Optimization with Financial Applications*, vol. 61 of Stochastic Modelling and Applied Probability, Springer, 2009.

- [90] H. PHAM, *Stochastic control under progressive enlargement of filtrations and applications to multiple defaults risk management*, Stochastic Processes and their Applications, 120 (2010), pp. 1795–1820.
- [91] N. PRIVAULT, *Generalized Bell polynomials and the combinatorics of Poisson central moments*, The electronic journal of combinatorics, 18 (2011).
- [92] P. PROTTER, *Stochastic Integration and Differential Equations*, Springer, 2005. Version 2.1.
- [93] D. REVUZ AND M. YOR, *Continuous Martingales and Brownian Motion*, vol. 293 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, 1991.
- [94] M. ROYER, *Backward stochastic differential equations with jumps and related non-linear expectations*, Stochastic Processes and their Applications, 116 (2006), pp. 1358 – 1376.
- [95] F. RUSSO AND P. VALLOIS, *Elements of stochastic calculus via regularization*, in Séminaire de Probabilités XL, C. Donati-Martin, M. Émery, A. Rouault, and C. Stricker, eds., vol. 1899 of Lecture Notes in Mathematics, Springer Berlin Heidelberg, 2007, pp. 147–185.
- [96] F. RUSSO AND P. VALLOIS, *Forward, backward and symmetric stochastic integration*, Probability Theory and Related Fields, 97 (1993), pp. 403–421.
- [97] —, *The generalized covariation process and Itô formula*, Stochastic Processes and their Applications, 59 (1995), pp. 81–104.
- [98] —, *Stochastic calculus with respect to continuous finite quadratic variation processes*, Stochastics and Stochastics Reports, 70 (2000), pp. 1–40.
- [99] A. SEIERSTAD AND K. SYDSÆTER, *Optimal control theory with economic applications*, Elsevier, 1987.
- [100] R. F. SERFOZO, *Processes with conditional stationary independent increments*, Journal of applied probability, 9 (1972).
- [101] J. L. SOLÉ, F. UTZET, AND J. VIVES, *Canonical Lévy process and Malliavin calculus*, Stochastic Processes and their Applications, 117 (2007), pp. 165 – 187.
- [102] E. M. STEIN AND J. C. STEIN, *Stock price distributions with stochastic volatility: an analytic approach*, Review of financial Studies, 4 (1991), pp. 727–752.
- [103] S. TANG, *The maximum principle for partially observed optimal control of stochastic differential equations*, SIAM Journal on Control and Optimization, 36 (1998), pp. 1596–1617.
- [104] S. TANG AND X. LI, *Necessary conditions for optimal control of stochastic systems with random jumps*, SIAM J. Control Optim., 32 (1994), pp. 1447–1475.

- [105] M. WINKEL, *The recovery problem for time-changed Lévy processes*. Research Report MaPhySto 2001-37, October 2001.
- [106] E. WONG AND M. ZAKAI, *Martingales and stochastic integrals for processes with a multi-dimensional parameter*, Probability Theory and Related Fields, 29 (1974), pp. 109–122.
- [107] A. L. YABLONSKI, *The Malliavin calculus for processes with conditionally independent increments*, in Stochastic Analysis and Applications, F. E. Benth, G. Nunno, T. Lindstrøm, B. Øksendal, and T. Zhang, eds., vol. 2 of Abel Symposia, Springer Berlin Heidelberg, 2007, pp. 641–678.