About the Pricing Equations in Finance

Stéphane Crépey*
Département de Mathématiques
Université d'Évry Val d'Essonne
91025 Évry Cedex, France

November 12, 2009

Abstract

In this article we study a decoupled *Forward Backward Stochastic Differential Equation* (FBSDE) and the associated *system of partial integro-differential obstacle problems*, in a flexible Markovian set-up made of a jump–diffusion with regimes.

These equations are motivated by numerous applications in financial modeling, whence the title of the paper. This financial motivation is developed in the first part of the paper, which provides a synthetic view of the theory of pricing and hedging financial derivatives, using Backward Stochastic Differential Equations (BSDEs) as main tool.

In the second part of the paper, we establish the well-posedness of reflected BSDEs with jumps coming out of the pricing and hedging problems exposed in the first part. We first provide a construction for our Markovian model made of a jump diffusion like component X interacting with a continuous time Markov chain like component N (so this model includes in particular Markovian jump-diffusions and continuous time Markov chain as special cases). The jump process N defines the so-called *regime* of the coefficients of X, whence the name of *Jump–Diffusion with Regimes* for this model. Motivated by *optimal stopping* and *optimal stopping game* problems (pricing equations of *American or game contingent claims*), we introduce the related *reflected and doubly reflected Markovian BSDEs*, showing that they are *well-posed* in the sense that they have *unique solutions, which depend continuously on their input data.* As an aside, we recover the *Markov property* of our model.

In the third part of the paper we derive the related *variational inequality approach.* We first introduce the systems of partial integro-differential variational inequalities formally associated to our reflected BSDEs, and we state suitable definitions of viscosity solutions for these problems (accounting for jumps and/or systems of equations). We then show that the state-processes (first components Y) of the solutions to our reflected BSDEs can be characterized in terms of the *value functions* of related optimal stopping or game problems, given as *viscosity solutions with polynomial growth* to the related obstacle problems. We further establish a *comparison principle* for semi-continuous viscosity solutions to our problems, which implies in particular *uniqueness* of viscosity solutions for these problems. This comparison principle is subsequently used for proving the convergence of *stable, monotone and consistent* approximation schemes to our value functions.

Finally in the last part of the paper we provide various extensions of the results needed for applications in finance to pricing problems involving *discrete dividends* (on

*This research benefited from the support of the "Chaire Risque de crédit" (Fédération Bancaire Française), of the Europlace Institute of Finance and of Ito33.

a financial derivative or on an underlying asset), as well as various forms of *discrete path-dependence*.

Contents

1 Introduction

In this article, we establish the well-posedness of a decoupled *Forward Backward stochastic differential equation* and of the associated *system of partial integro-differential obstacle problems*, in a rather flexible Markovian set-up made of a jump–diffusion model with regimes.

These equations are motivated by numerous applications in financial modeling, whence the title of the paper. This financial motivation is developed in Part I, where we essentially reduce the problem of pricing and hedging financial derivatives to that of solving (typically reflected) backward stochastic differential equations (BSDEs), or, equivalently in the Markovian case, associated partial integro-differential equations or integro-differential variational inequalities (PIDEs or PDEs for short).

In Parts II to IV, we tackle the resulting Markovian BSDE and PDE problems. In Crépey and Matoussi [42], a priori estimates and comparison principles were derived for reflected or doubly reflected BSDEs in the general (non-Markovian) set-up of a model driven by a continuous local martingale and an integer-valued random measure. In Part II we use these results to establish the well-posedness of a *Markovian doubly reflected BSDE*, which is used in Part III for studying the associated partial integro-differential system of double obstacle problems, in a rather flexible Markovian set-up made of a jump–diffusion model with regimes. As an aside we prove the convergence of any *stable, monotone and consistent* approximation scheme to the resulting parabolic system. Part IV provides various extensions of the previous results needed for applications in finance to pricing problems involving *discrete dividends* (on a financial derivative or on an underlying asset), as well as various forms of *discrete path-dependence*.
The main results are summed-up in Propositions 9.4 and 12.4, which synthesize the major findings of Part II and III, respectively.

The paper has above all a unifying perspective. However, even if rather expected in their statement for most of them, some of the mathematical results derived in Parts II to IV are, to some extent, innovative. For instance, doubly reflected BSDEs with a delayed or an even more general intermittent upper barrier (RDBSDEs and RIBSDEs, see Definitions 5.4(ii) and 16.3), have not been considered elsewhere in the literature (if not for the preliminary RDBSDE results of Crépey and Matoussi [42]).

As for Part I, we believe that, beyond providing the motivation for the mathematical results of Parts II to IV, it also has the merit of giving a unified, cross market perspective (see Sections 3.3.3 and 6.6) on the theory of pricing and hedging financial derivatives, via the use of BSDEs as a main tool.

Part I on one hand, and Parts II to IV on the other hand, can be read essentially independently. The reader who would be mainly interested in the financial applications can thus read Part I first, taking for granted the results of Parts II to IV whenever they are used therein (see Propositions 3.2, 3.3, 4.1, 4.12 and 4.14 in particular). Likewise readers mainly interested by the mathematical results of Parts II to IV can skip Part I at first reading.

1.1 Detailed Outline

Section 2 develops the theory of risk-neutral pricing and hedging of financial derivatives, using BSDEs as a main tool (see El Karoui et al. [51] for a general reference on BSDEs in finance). The central result, Proposition 2.3, can be informally stated as follows: Under the assumption, thoroughly investigated in Part II, that a reflected *Backward Stochastic Differential Equation (BSDE)* related to a financial derivative, relative to a risk-neutral probability measure $\mathbb{P}$ over a primary market of hedging instruments, admits a solution Π, then Π is the minimal *superhedging price up to a* $\mathbb{P}$ – *local martingale cost process* for the derivative at hand, this cost being equal to 0 in the case of complete markets. This notion of hedge *with local martingale cost* thus establishes a connection between arbitrage prices and hedging, in a rather general, possibly incomplete, market.

In Section 3, we consider the specification of these results to the *Markovian set-up*. Using the results of Part III, a complementary *variational inequality* approach may then be developed, and more *explicit and constructive hedging strategies* may be given (see Section 3.5 in particular).

Section 4 presents various extensions of the previous results. Section 4.1 generalizes the previous risk-neutral approach to a martingale modeling approach relative to an arbitrary *numeraire* B (positive primary asset price process) which may be used for discounting other price processes, rather than implicitly the savings account $B = \beta^{-1}$ in the risk-neutral approach. This extension is particularly important for dealing with interest-rate derivatives. Section 4.2, which is based on Bielecki et al. [17], refines the risk-neutral martingale modeling approach of Sections 2 and 3 to the specific case, important for equity-to-credit applications, of *defaultable derivatives*, with all cash flows killed at the default time θ of a reference entity. Finally in Section 4.3 we deal with the issue of callability and *call protection* (*intermittent* call protection versus call protection *before a stopping time*).

Up to this point (Part I), well-posedness of the resulting pricing reflected BSDEs and PDEs was taken for granted. The following sections of the paper (Parts II to IV) are devoted to the rigorous mathematics of these pricing equations.

In Section 5 we recall the general set-up of [42] and the general form of the BSDEs we are interested in.

In Section 6, we present a versatile Markovian specification of this general set-up, made of a jump diffusion X interacting with a pure jump process N (which in the simplest case reduces to a Markov chain in continuous time). The interaction between X and N is materialized by the fact that the coefficients of the dynamics of X depend on N, and also, by a mutual dependence of the jump intensity of either process on the other one. Such coupled dependence is motivated by applications like modeling *frailty* and *contagion* in the field of *portfolio credit risk* in finance (see [19]).

A related model was already considered and some of the results of this paper were announced in [19, 42, 17]. But the possibility to construct a model with all the required properties was taken for granted there. Indeed the construction of a model with such mutual dependence is a non-trivial issue, and we treat it in detail in Section 7, resorting to a suitable *Markovian change of probability measure*.

This model may also be viewed as a generalization of the interacting (continuous) Itô process and point process model considered by Becherer and Schweizer in [10]. Yet as opposed to the set-up of [10] where linear reaction-diffusion systems of parabolic equations (pricing equations of *European contingent claims*, from the point of view of the financial interpretation) are considered from the point of view of *classical solutions*, here the application one has in mind consists of more general *optimal stopping* or *optimal stopping game* problems (pricing equations of *American or game contingent claims*, see Part I) for which the related reaction-diffusion systems typically do not have classical solutions (and even less so, that there are also jumps in the component X of our model). This leads us to study in Section 8 the related *reflected and doubly reflected Markovian BSDEs* (see [51, 50, 17]), showing that they are *well-posed* in the sense that they have *unique solutions, which depend continuously on their input data.*

In Section 9 we derive the associated *Markov and flow properties.*

In Section 10 we introduce the systems of partial integro-differential variational inequalities formally associated to our reflected BSDEs, and we state suitable definitions of semi-continuous viscosity solutions and solutions for these problems.

In Section 11 we show that the state-processes (first components Y) of the solutions to our reflected BSDEs can be characterized in terms of the *value functions* to related optimal stopping or game problems, given as *viscosity solutions with polynomial growth* to the related obstacle problems.

We establish in Section 12 a *semi-continuous viscosity solutions comparison principle* for our problems, which implies in particular *uniqueness* of viscosity solutions for these problems.

This comparison principle is subsequently used in Section 13 for proving the convergence of *stable, monotone and consistent* approximation schemes (cf. Barles and Souganidis; see also [9] Briani, La Chioma and Natalini [30], Cont and Voltchkova [38] or Jakobsen et al. [69]) to the viscosity solutions of the equations. These results thus extend to models with regimes (whence *systems* of PDEs [65, 6]) the results of [9, 30], among others.

In Sections 14 to 16 we provide extensions of the previous results to a factor process model $\mathcal{X} = (X, N)$ possibly involving further *deterministic jumps* at some fixed times T_ls. This is required for applications to pricing problems involving *discrete dividends* (on a financial derivative or on an underlying asset), and also, to be able to deal with the issue of *discrete path-dependence.*

Part I

Martingale Modeling in Finance

In this part (see Section 1 for a detailed outline), we show how the task of pricing and hedging financial derivatives can generically be reduced to that of solving (typically reflected) BSDEs, or, equivalently in the Markovian case, P(I)DEs. These equations are called *pricing equations* in this paper. Well-posedness of these equations in suitable spaces of solutions will be taken for granted whenever needed in this part, and will then be thoroughly studied in the remaining three parts of the paper.

2 General Set-Up

The evolution of a financial market model is given throughout this part in terms of stochastic processes defined on a continuous time stochastic basis $(\Omega, \mathbb{F}, \widehat{\mathbb{P}})$, where $\widehat{\mathbb{P}}$ denotes the *objective* (also called statistical, historical, physical..) probability measure. We may and do assume that the filtration $\mathbb{F}$ satisfies the usual completeness and right-continuity conditions, and that all semimartingales are càdlàg (i.e., almost surely right continuous with left limits). Finally, since we are always in the context of pricing contingent claims with a fixed maturity T, we further assume that $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ with $\mathcal{F}_0$ trivial and $\mathcal{F}_T = \mathcal{F}$, for simplicity. Moreover, we declare that a *process* on $[0, T]$ (resp. *a random variable*) has to be $\mathbb{F}$-adapted (resp. $\mathcal{F}$-measurable), by definition.

Since our aim in this article is mainly computational (derivation and study of the pricing equations within a given class of pricing models), we shall typically work under a *risk-neutral* probability measure $\mathbb{P} \sim \widehat{\mathbb{P}}$, or more generally, under a martingale probability measure $\mathbb{P}$ relative to a suitable *numeraire* (see Section 4.1), such that the prices of primary assets, once properly discounted and adjusted for dividends, are $\mathbb{P}$ − local martingales.
As we shall see in Section 2.1, under mild technical conditions, existence of such a martingale measure $\mathbb{P}$ is equivalent to a suitable notion of no-arbitrage. In practical applications, it is convenient to think of $\mathbb{P}$ as 'the pricing measure chosen by the market' to price a contingent claim.

2.1 Pricing by Arbitrage

2.1.1 Primary Market Model

To model a financial derivative with maturity T, we consider a primary market composed of the savings account B and of d primary risky assets. The discount factor β is supposed to be absolutely continuous with respect to the Lebesgue measure, and given by

$$\beta_t = \exp(-\int_0^t r_u \, du) \tag{1}$$

(so $\beta_0 = 1$), for a bounded from below *short-term interest rate* process r .
The primary risky assets, with $\mathbb{R}^d$-valued price process P, may pay dividends, whose cumulative value process, denoted by $\mathcal{D}$, is assumed to be an $\mathbb{R}^d$-valued process of finite variation.

Given the price process P, we define the *cumulative price* $\widehat{P}$ of the asset as

$$\widehat{P}_t = P_t + \beta_t^{-1} \int_{[0,t]} \beta_u \, d\mathcal{D}_u. \tag{2}$$

In the financial interpretation, the last term in (2) represents the current value at time t of all dividend payments of the asset over the period $[0,t]$, under the assumption that all dividends are immediately reinvested in the savings account.

For technical reasons we assume that $\widehat{P}$ is a locally bounded semimartingale.

We assume that the primary market model is free of arbitrage opportunities (though presumably incomplete), in the sense that the so-called *No Free Lunch with Vanishing Risk* (NFLVR) condition is satisfied. This NFLVR condition is a specific no arbitrage condition involving wealth processes of admissible self-financing primary trading strategies (see Delbaen and Schachermayer [45]). We do not reproduce here the full definition of arbitrage price, since it is rather technical and will not be explicitly used in the sequel. It will be enough for us to recall the related notions of trading strategies in the primary market.

Definition 2.1 A *primary trading strategy* (ζ^0, ζ) built on the primary market is an $\mathbb{R} \times \mathbb{R}^{1\otimes d}$-valued process, with ζ predictable and locally bounded, representing the number of units held in the savings account and in the primary risky assets, respectively. The related *wealth process* $\mathcal{W}$ is thus given by:

$$\mathcal{W}_t = \zeta_t^0 B_t + \zeta_t P_t \ , \tag{3}$$

for $t \in [0,T]$. Accounting for dividends, we say that the strategy is *self-financing* if

$$d\mathcal{W}_t = \zeta_t^0 dB_t + \zeta_t \left(dP_t + d\mathcal{D}_t \right)$$

or, equivalently[1]

$$d(\beta_t \mathcal{W}_t) = \zeta_t \, d(\beta_t \widehat{P}_t) \ . \tag{4}$$

If, moreover, the discounted wealth process $\beta\mathcal{W}$ is bounded from below, the strategy is said to be *admissible*.

Given the initial wealth w of a self-financing primary trading strategy and the strategy ζ in the primary risky assets, the related wealth process is thus given by, for $t \in [0,T]$:

$$\beta_t \mathcal{W}_t = w + \textstyle\int_0^t \zeta_u \, d(\beta_u \widehat{P}_u) \tag{5}$$

and the process ζ^0 (number of units held in the savings account) is then uniquely determined as

$$\zeta_t^0 = \beta_t \left(\mathcal{W}_t - \zeta_t P_t \right) \ .$$

In the sequel we restrict ourselves to self-financing trading strategies. We thus ***redefine*** a (self-financing) primary trading strategy as a pair (w, ζ), of an initial wealth $w \in \mathbb{R}$ and a $\mathbb{R}^{1\otimes d}$-valued predictable locally bounded primary strategy in the risky assets ζ, with related wealth process $\mathcal{W}$ defined by (5).

By classic arbitrage theory (see, e.g., [45, 34, 15]), the NFLVR condition in a perfect market (without transaction costs, in particular) is equivalent to the existence of a *risk-neutral measure* $\mathbb{P} \in \mathcal{M}$, where $\mathcal{M}$ denotes the set of probability measures $\mathbb{P} \sim \widehat{\mathbb{P}}$ such that $\beta\widehat{P}$ is a $\mathbb{P}$ – local martingale.

[1] This equivalence is very general (cf. Section 4.1), and it is an easy exercise in the present context where β, given by (1), is a finite variation and continuous process.

2.1.2 Financial Derivatives

In the sequel we are going to extend the financial market by introducing a financial *derivative* relative to the primary market. A derivative is a financial claim between an investor (or *holder* of a claim) and a financial institution (or *issuer*), involving in a sense made precise in Definition 2.3 below, some or all of the following cash flows (or payoffs):
• a bounded variation *dividend process* $D = (D_t)_{t\in[0,T]}$,
• terminal cash flows, consisting of:
– a *payment* ξ *at maturity* T, where ξ denotes a bounded from below real-valued random variable,
– and, in the case of (American or game) products with early exercice features, *put and/or call payment processes* $L = (L_t)_{t\in[0,T]}$ and $U = (U_t)_{t\in[0,T]}$, given as real-valued, bounded from below, càdlàg processes such that $L \le U$ and $L_T \le \xi \le U_T$.

The put payment L_t corresponds to a payment made by the issuer to the holder of the claim at time t, in case the holder of the claim would decide to terminate ('put') the contract at time t. Likewise, the call payment U_t corresponds to a payment made by the issuer to the holder of the claim at time t, in case the issuer of the claim would decide to terminate ('call') the contract at time t.

Of course, there is also the initial cash flow (null in the case of a swapped derivative with initial value equal to zero, by construction), namely the purchasing price of the contract paid at the initiation time by the holder and received by the issuer.

The terminology 'derivative' comes from the fact that all the above cash flows are typically given as functions of the 'primary' asset price processes P. More generally, the price Π of a derivative and the prices P of the primary assets may be given as functions of a common set of *factors* (traded or not) X (cf. Section 3). One may then consider the issue of factor hedging the claim with price process Π by the primary assets with price process P, via the common dependence of Π and P on X.

Here and henceforth all the financial cash flows are seen from the point of view of the *holder* of the claim. In this perspective, the implicit assumption above that all the cash flows are bounded from below, which from the mathematical point of view ensures their integrability in $\mathbb{R} \cup \{+\infty\}$, is indeed satisfied by a vast majority of real-life financial derivatives.

Remark 2.2 Usually in the derivative pricing and hedging literature, dividends are implicitly set to zero, or equivalently, implicitly amalgamated with the terminal cash flows L, U and ξ. The related notion of price thus effectively corresponds to a *cum-dividend price* (present value of future cash flows plus already perceived dividends reinvested in the savings account), as opposed to the market notion of *ex-dividend price*. Since an important proportion of financial derivatives (starting with all swapped derivatives) only entails dividends (terminal cash flows $L = U = \xi = 0$), it is our opinion that it is better to make the dividends appear explicitly. This is in fact a necessity for the study of defaultable derivatives in Section 4.2, where we shall see that the specific structure of the products' cash flows and their distribution between dividends (in the sense of coupons and recovery) and terminal payoffs, is fruitfully exploited in the so-called reduced form approach to these problems.

We are now in a position to introduce the formal definition of a financial derivative, distinguishing more specifically European claims, American claims and game claims. It will soon become apparent that European claims can be considered as special cases of American claims, which are themselves included in game claims, so that we shall eventually be able to reduce attention to game claims.

In the following definitions, the put time (or *put or maturity* time, more precisely) τ and the call (or maturity) time σ represent stopping times at the holder's and at the issuer's convenience, respectively.

Definition 2.3 **(i)** An *European claim* is a financial claim with dividend process D, and with payment ξ at maturity T.
(ii) An *American claim* is a financial claim with dividend process D, and with payment at the terminal (put or maturity) time τ given by,

$$\mathbb{1}_{\{\tau<T\}}L_\tau + \mathbb{1}_{\{\tau=T\}}\xi \ . \tag{6}$$

(iii) A *game claim* is a financial claim with dividend process D, and with payment at the terminal (call, put or maturity) time $\nu = \tau \wedge \sigma$ given by,[2]

$$\mathbb{1}_{\{\nu=\tau<T\}}L_\tau + \mathbb{1}_{\{\sigma<\tau\}}U_\sigma + \mathbb{1}_{\{\nu=T\}}\xi \ . \tag{7}$$

Moreover, there may be a *call protection* modeled in the form of a stopping time $\bar{\sigma}$ such that calls are not allowed to occur before $\bar{\sigma}$.

Example 2.4 In the simplest case of an European vanilla call/put option with maturity T and strike K on $S = P^1$, the first primary risky asset, one has $D = 0$ and $\xi = (S_T - K)^{\pm}$.

Comments 2.1 **(i)** The above classification, which is good enough for the purpose of this article, is by no means exhaustive. For instance Bermudan products corresponding to constrained put policies might also be introduced. Note however that Bermudan products can be included in the above set-up by considering a suitably adjusted put payoff process L. This is indeed a consequence of Proposition 2.1(ii) below, in conjunction with our boundedness from below assumption on all the cash flows at hand.
On the opposite the explicit introduction of call protections appears to be a useful modeling ingredient. Such protections are actually quite typical in the case of real-life callable products like, for instance, convertible bonds (see Section 4.2.1.1), with the effect of making the product cheaper to the investor (holder of the claim). The introduction of such call protections also allows one to consider an American claim as a game claim with call protection $\bar{\sigma} = T$.
(ii) In Section 4.3, building on the mathematical results of Section 16, we consider products with more general, hence potentially more realistic forms of *intermittent* call protection, namely call protection *whenever a certain condition* is satisfied, rather than more specifically call protection *before a stopping time* above.

Let $\mathcal{T}_t$ and $\bar{\mathcal{T}}_t$ (or simply $\mathcal{T}$ and $\bar{\mathcal{T}}$, in case $t = 0$) denote the set of $[t, T]$-valued and $[t \vee \bar{\sigma}, T]$-valued stopping times. Let also ν stand for $\sigma \wedge \tau$, for any $(\sigma, \tau) \in \bar{\mathcal{T}}_t \times \mathcal{T}_t$.

In the sequel, the statement '$(\Pi_t)_{t\in[0,T]}$ *is an arbitrage price for a derivative*' is to be understood as '$(P_t, \Pi_t)_{t\in[0,T]}$ *is an arbitrage price for the extended market consisting of the*

[2]With priority of a put over a call, here, though this happens to be rather immaterial in terms of pricing and hedging the claim.

primary market and the derivative'. The notion of arbitrage price process of a financial derivative referred to in the next result is thus the classical notion of No Free Lunch with Vanishing Risk condition of Delbaen and Schachermayer [45] in the case of European claims, subsequently extended to game (including American) claims by Kallsen and Kühn [72] (see Bielecki et al. [15]). The proof of this result is based on a rather straightforward application of Theorem 2.9 in Kallsen and Kühn [72] (see Bielecki et al. [15] for the details).

Proposition 2.1 **(i)** *For any $\mathbb{P} \in \mathcal{M}$, the process $\Pi = (\Pi_t)_{t\in[0,T]}$ defined by*

$$\beta_t \Pi_t = \mathbb{E}_{\mathbb{P}}\Big\{ \int_t^T \beta_u \, dD_u + \beta_T \xi \,\Big|\, \mathcal{F}_t \Big\}, \quad t \in [0,T] \tag{8}$$

is an arbitrage price of the related European claim. Moreover, any arbitrage price of the claim is of this form provided

$$\sup_{\mathbb{P}\in\mathcal{M}} \mathbb{E}_{\mathbb{P}}\Big\{ \int_{[0,T]} \beta_u \, dD_u + \beta_T \xi \Big\} < \infty \, ; \tag{9}$$

(ii) *For any $\mathbb{P} \in \mathcal{M}$, the process $\Pi = (\Pi_t)_{t\in[0,T]}$ defined by*

$$\beta_t \Pi_t = \mathrm{esssup}_{\tau\in\mathcal{T}_t} \mathbb{E}_{\mathbb{P}}\Big\{ \int_t^\tau \beta_u \, dD_u + \beta_\tau \big(\mathbb{1}_{\{\tau<T\}} L_\tau + \mathbb{1}_{\{\tau=T\}} \xi \,\Big|\, \mathcal{F}_t \Big\}, \quad t \in [0,T] \tag{10}$$

is an arbitrage price of the related American claim as soon as it is a semimartingale. Moreover, any arbitrage price of the claim is of this form provided

$$\sup_{\mathbb{P}\in\mathcal{M}} \mathbb{E}_{\mathbb{P}} \sup_{t\in[0,T]} \Big\{ \int_{[0,t]} \beta_u \, dD_u + \beta_t \big(\mathbb{1}_{\{t<T\}} L_t + \mathbb{1}_{\{t=T\}} \xi \big) \Big\} < \infty \, ; \tag{11}$$

(iii) *For any $\mathbb{P} \in \mathcal{M}$, the process $\Pi = (\Pi_t)_{t\in[0,T]}$ defined by*

$$\begin{aligned} \mathrm{esssup}_{\tau\in\mathcal{T}_t} \mathrm{essinf}_{\sigma\in\bar{\mathcal{T}}_t} \mathbb{E}_{\mathbb{P}}\Big\{ \int_t^\nu \beta_u \, dD_u + \beta_\nu \big(\mathbb{1}_{\{\nu=\tau<T\}} L_\tau + \mathbb{1}_{\{\sigma<\tau\}} U_\sigma + \mathbb{1}_{\{\nu=T\}} \xi \big) \,\Big|\, \mathcal{F}_t \Big\} = \beta_t \Pi_t \\ = \mathrm{essinf}_{\sigma\in\bar{\mathcal{T}}_t} \mathrm{esssup}_{\tau\in\mathcal{T}_t} \mathbb{E}_{\mathbb{P}}\Big\{ \int_t^\nu \beta_u \, dD_u + \beta_\nu \big(\mathbb{1}_{\{\nu=\tau<T\}} L_\tau + \mathbb{1}_{\{\sigma<\tau\}} U_\sigma + \mathbb{1}_{\{\nu=T\}} \xi \big) \,\Big|\, \mathcal{F}_t \Big\}, \quad t \in [0,T] \end{aligned} \tag{12}$$

is an arbitrage price of the related game claim as soon as it is a well-defined semimartingale (which supposes in particular that equality indeed holds between the left hand side and the right hand side in (12)). Moreover, any arbitrage price of the claim is of this form assuming (11).

In view of these results, one may interpret an European claim as an American claim with a (fictitious) put payment process L defined by $\beta L = -(c+1)$, where $-c$ is a minorant of $\int_t^T \beta_u \, dD_u + \beta_T \xi$. Indeed, in view of Propositions 2.1(ii), for this specification of L, exercise of the put before maturity is always sub-optimal to the holder of the claim. It is thus equivalent for a process Π to be an arbitrage price of the European claim with the cash flows D, ξ, or to be an arbitrage price of the American claim with the cash flows D, L, ξ, with L thus specified.

Henceforth by default, by 'financial derivative' or 'game option', we shall mean game claim, possibly with a call protection $\bar{\sigma}$, including American claim (case $\bar{\sigma} = T$, in particular European claim with L as specified above) as a special case. Arbitrage prices of the form (8), (10) or (12) will be called $\mathbb{P}$-*prices* in the sequel.

2.2 Connection with Hedging

We adopt a definition of hedging of a game option stemming from successive developments, starting from the hedging of American options examined by Karatzas [73], and subsequently followed by El Karoui and Quenez [52], Kifer [74], Ma and Cvitanić [83], Hamadène [60], and (in the context of defaultable derivatives examined in Section 4.2) Bielecki et al. [16, 17] (see also Schweizer [93]). This definition will be later shown to be consistent with the concept of arbitrage pricing of Proposition 2.1(iii) for a game option (which encompasses American and European options as special cases).

We first introduce a (very large, to be specified later) class of hedges *with semimartingale cost process* Q. The issuer of a financial derivative immediately sets up a primary hedging strategy such that the corresponding wealth process $\mathcal{W}$ reduces to a *cost* or *hedging error* Q, after accounting for the 'dividend cost' $-D$ and for the 'terminal loss' given by $-L$, $-U$ or $-\xi$. The initial wealth w may then be used as a safe issuer price, up to the hedging error Q, for the derivative at hand. Recall that we denote $\nu = \tau \wedge \sigma$.

Definition 2.5 An *hedge with semimartingale cost process* Q (*issuer hedge starting at time 0*) for a game option is represented by a triplet (w, ζ, σ) such that:
• (w, ζ) is a (self-financing) primary trading strategy,
• the call time σ belongs to $\bar{\mathcal{T}}$,
• the wealth process $\mathcal{W}$ of the strategy (w, ζ) satisfies for every put time τ in $\mathcal{T}$, almost surely,

$$\beta_\nu \mathcal{W}_\nu + \int_0^\nu \beta_u dQ_u \geq \int_0^\nu \beta_u dD_u \; + \; \beta_\nu \Big(\mathbb{1}_{\{\nu=\tau<T\}} L_\tau + \mathbb{1}_{\{\sigma<\tau\}} U_\sigma + \mathbb{1}_{\{\tau=\sigma=T\}} \xi \Big) \, . \tag{13}$$

In the special case of European derivatives, in which case $\bar{\sigma} = T$, and if moreover equality holds in (13) at $t = T$, then, almost surely,

$$\beta_T \mathcal{W}_T + \int_0^T \beta_u dQ_u = \int_0^T \beta_u dD_u \; + \; \beta_T \xi \, . \tag{14}$$

In this case one effectively deals with a *replicating strategy with cost* Q.

Comments 2.2 **(i)** The process Q is to be interpreted as the (running) *financing cost*, that is, the amount of cash added to (if $dQ_t \geq 0$) or withdrawn from (if $dQ_t \leq 0$) the hedging portfolio in order to get a perfect, but no longer self-financing, hedge.
(ii) Hedges *at no cost* (that is, with $Q = 0$) are thus in effect *superhedges*.
(iii) In relation with admissibility issues (see the end of Definition 2.1), note that the left hand side of (13) (discounted wealth process with financing costs included) is bounded from below, for any hedge (w, ζ, σ) with cost Q.

This class of hedges with cost Q is obviously too large for any practical purpose, so we will restrict our attention to hedges with a *local martingale* cost Q under a particular risk-neutral measure $\mathbb{P}$ (cf. the related notions of *risk-minimizing strategy* in Föllmer and Sondermann [55] and *mean self-financing hedge* in Schweizer [93]). ***Henceforth in this part, we thus work under a fixed but arbitrary risk-neutral measure*** $\mathbb{P}$**,** with $\mathbb{P}$-expectation denoted by $\mathbb{E}$. All the measure-dependent notions like (local) martingale, compensator, etc., implicitly refer to this probability measure $\mathbb{P}$. In practical applications, it is convenient to think of $\mathbb{P}$ as 'the pricing measure chosen by the market' to price a contingent claim. For pricing and hedging purposes this measure is typically estimated by calibration of a model to market data.

2.2.1 BSDE Modeling

We shall now postulate suitable integrability and regularity conditions embedded in the standing assumption that a related reflected Backward Stochastic Differential Equation (BSDE, see El Karoui et al. [51] for a general reference in connection with finance) has a solution. We shall thus introduce a reflected BSDE (15) under the probability measure $\mathbb{P}$, with data defined in terms of those of a derivative. Assuming that (15) has a solution (for which various sets of sufficient regularity and integrability conditions are known in the literature, see Part II and [42, 62, 61]), we shall be in a position to deduce explicit hedging strategies with minimal initial wealth for the related derivative.

We assume further for the sake of simplicity that $dD_t = C_t dt$ for some progressively measurable time-integrable coupon rate process C.

Remark 2.6 It is important to note for applications that it is also possible to deal with discrete dividends: see [17] and Section 14 in Part IV.

We then consider the following *reflected BSDE* with data $\beta, C, \xi, L, U, \bar{\sigma}$:

$$\begin{cases} \beta_t \Pi_t = \beta_T \xi + \int_t^T \beta_u C_u du + \int_t^T \beta_u (dK_u - dM_u), \quad t \in [0,T] \\ L_t \leq \Pi_t \leq \bar{U}_t, \quad t \in [0,T] \\ \int_0^T (\Pi_u - L_u)\, dK_u^+ = \int_0^T (\bar{U}_u - \Pi_u)\, dK_u^- = 0 \end{cases} \tag{15}$$

where, with the convention that $0 \times \pm\infty = 0$ in the last two lines above,

$$\bar{U}_t = \mathbb{1}_{\{t<\bar{\sigma}\}} \infty + \mathbb{1}_{\{t \geq \bar{\sigma}\}} U_t \ . \tag{16}$$

Definition 2.7 *(See Part II for more formal definitions, including in particular the specification of* spaces *for the inputs and outputs to (15)).* By a $\mathbb{P}$-*solution* to (15), we mean a triplet (Π, M, K) such that all conditions in (15) are satisfied, where:
• the *state-process* Π is a real valued, càdlàg process,
• M is a $\mathbb{P}$-martingale vanishing at time 0,
• K is a non-decreasing (null at time 0) continuous process, and $K^{\pm}$ denote the components of the Jordan decomposition of K.

By *the Jordan decomposition* of K in the last bullet point, we mean the unique decomposition $K = K^+ - K^-$ into the difference of non-decreasing null at 0 processes $K^{\pm}$ defining mutually singular random measures on $\mathbb{R}^+$.

Remark 2.8 The first line of (15) can be interpreted as giving the Doob-Meyer decomposition $\int_0^t \beta_u (dK_u - dM_u)$ of the special semimartingale

$$\beta_t \widehat{\Pi}_t := \beta_t \Pi_t + \int_0^t \beta_u C_u du \ . \tag{17}$$

So an equivalent definition of a solution to (15) would be that of a special semimartingale Π (rather than a triplet of processes (Π, M, K)) such that all conditions in (15) are satisfied, where M and K therein are to be understood as the canonical local martingale and finite variation predictable components of $\int_{[0,\cdot]} \beta_t^{-1} d(\beta_t \widehat{\Pi}_t)$.

Note that the first line of (15) is equivalent to

$$\Pi_t = \xi + \int_t^T (C_u - r_u\Pi_u)du + (K_T - K_t) - (M_T - M_t), \quad t \in [0,T] \,. \tag{18}$$

As established in [61, 62, 42], existence and uniqueness of a solution to (15) (under suitable L_2-integrability conditions on the data and the solution) are essentially equivalent to the so-called *Mokobodski condition*, namely, the existence of a *quasimartingale* Y (special semimartingale with additional integrability properties, Section 16.2.2) such that $L \le Y \le U$ on $[0,T]$. Existence and uniqueness for (15) thus hold when one of the barriers is a quasimartingale and, in particular, when one of the barriers is given as $S \vee c$ where S is a square-integrable Itô process and c is a constant in $\mathbb{R} \cup \{-\infty\}$ (see [42] as well as Comments 5.1(v) and Proposition 9.4 in Part II). This covers, for instance, the put payment process L of an American vanilla option, or of a convertible bond (see Definition 4.3 and Bielecki et al. [15, 18]). Moreover one typically has $K = 0$ in the case of an European derivative.

We thus work henceforth in this part under the following hypothesis.

Assumption 2.9 Equation (15) admits a solution (Π, M, K), with K equal to zero in the special case of an European derivative.

Proposition 2.2 Π *is the* $\mathbb{P}$*-price process of the derivative.*

Proof. If (Π, M, K) is a solution to (15), then Π is a (special) semimartingale (see (18)), and, by a standard verification principle (cf. Proposition 5.2 in Part II), Π satisfies (12), which in the special cases of American (resp. European) options reduces to (10) (resp. (8)). One thus concludes by an application of Proposition 2.1. □

We are now ready to interpret the $\mathbb{P}$-price Π, thus defined via (15), in terms of the notion of hedging introduced in Section 2.2. Let us set

$$\sigma^* = \inf \left\{ u \in [t,T] \,;\, \Pi_u \ge U_u \right\} \wedge T \,. \tag{19}$$

Using the minimality condition (third line) in (15) and the continuity of $K^{\pm}$, one thus has,

$$K^- = 0 \text{ and } K = K^+ \ge 0 \text{ on } [0,\sigma^*] \,,\; \Pi_{\sigma^*} = U_{\sigma^*} \,. \tag{20}$$

Note that for any primary strategy ζ, the issuer's cumulative discounted *Profit and Loss* (or *Tracking Error*) process $(e_t)_{t\in[0,T]}$ relative to the price process Π of Proposition 2.2 is given for $t \in [0,T]$ by:

$$\beta_t e_t = \Pi_0 - \int_0^t \beta_u C_u du \, + \int_0^t \zeta_u d(\beta_u \widehat{P}_u) - \beta_t \Pi_t = \int_0^t \Big(-d(\beta_u \widehat{\Pi}_u) + \zeta_u d(\beta_u \widehat{P}_u) \Big) \tag{21}$$

where $\widehat{\Pi}$ is defined by (17), so that, in view of Proposition 2.2, $\beta\widehat{\Pi}$ can be interpreted as the $\mathbb{P}$ – *cumulative price* of the option (cf. (2)). Observe in view of (18) that βe is a special semimartingale. Let the $\mathbb{P}$ – local martingale $\rho = \rho(\zeta)$ be such that $\rho_0 = 0$ and $\int_0^\cdot \beta_t d\rho_t$ is the local martingale component of the special semimartingale βe, so (cf. (21)–(18))

$$d\rho_t = \; dM_t - \zeta_t \, \beta_t^{-1} d(\beta_t \widehat{P}_t) \tag{22}$$

$$\beta_t e_t = \int_0^t \beta_u dK_u - \int_0^t \beta_u d\rho_u \,. \tag{23}$$

The arguments underlying the following result are classical, and already present for instance in Lepeltier and Maingueneau [82] (in the specific contexts of the Cox–Ross–Rubinstein or Black–Scholes models, analogous results can also be found in Kifer [74]).

Proposition 2.3 **(i)** *For any primary strategy* ζ, (Π_0, ζ, σ^*), *is an hedge with* $\mathbb{P}$ *– local martingale cost* $\rho(\zeta)$*;*
(ii) Π_0 *is the minimal initial wealth of an hedge with* $\mathbb{P}$ *– local martingale cost;*
(iii) *In the special case of an European derivative with* $K = 0$, *then* (Π_0, ζ) *is a replicating strategy with* $\mathbb{P}$ *– local martingale cost* ρ. Π_0 *is thus also the minimal initial wealth of a replicating strategy with* $\mathbb{P}$ *– local martingale cost.*

Proof. **(i)** One must show that for any $\tau \in \mathcal{T}$, almost surely:

$$\Pi_0 + \int_0^{\sigma^* \wedge \tau} \zeta_u d(\beta_u \widehat{P}_u) + \int_0^{\sigma^* \wedge \tau} \beta_u d\rho_u \geq \tag{24}$$
$$\int_0^{\sigma^* \wedge \tau} \beta_u C_u du \; + \; \beta_{\sigma^* \wedge \tau} \Big(\mathbb{1}_{\{\sigma^* \wedge \tau = \tau < T\}} L_t + \mathbb{1}_{\{\sigma^* < \tau\}} U_{\sigma^*} + \mathbb{1}_{\{\sigma^* = \tau = T\}} \xi \Big)$$

or equivalently, using (22):

$$\Pi_0 + \int_0^{\sigma^* \wedge \tau} \beta_u dM_u \geq \tag{25}$$
$$\int_0^{\sigma^* \wedge \tau} \beta_u C_u du \; + \; \beta_{\sigma^* \wedge \tau} \Big(\mathbb{1}_{\{\sigma^* \wedge \tau = \tau < T\}} L_\tau + \mathbb{1}_{\{\sigma^* < \tau\}} U_{\sigma^*} + \mathbb{1}_{\{\sigma^* = \tau = T\}} \xi \Big)$$

where by the first line in (15):

$$\Pi_0 + \int_0^{\sigma^* \wedge \tau} \beta_u dM_u = \beta_{\sigma^* \wedge \tau} \Pi_{\sigma^* \wedge \tau} + \int_0^{\sigma^* \wedge \tau} \beta_u C_u du + \int_0^{\sigma^* \wedge \tau} \beta_u dK_u \; .$$

Inequality (25) then follows by non-negativity of K on $[0, \sigma^*]$ (cf. (20)) and by the following relations, which are valid by the terminal and put conditions in (15) and by (20):

$$\Pi_T = \xi \; , \; \Pi_\tau \geq L_\tau \; , \; \Pi_{\sigma^*} \geq U_{\sigma^*} \; .$$

(ii) There exists an hedge with initial wealth Π_0 and $\mathbb{P}$ – local martingale cost, by (i) applied with, for instance, $\zeta = 0$. Moreover, for any hedge (w, ζ, σ) with $\mathbb{P}$ – local martingale cost Q, one has for every $t \in [0, T]$:

$$w + \int_0^{\sigma \wedge t} \zeta_u d(\beta_u \widehat{P}_u) + \int_0^{\sigma \wedge t} \beta_u dQ_u \geq \tag{26}$$
$$\int_0^{\sigma \wedge t} \beta_u C_u du \; + \; \beta_{\sigma \wedge t} \Big(\mathbb{1}_{\{\sigma \wedge t = t < T\}} L_t + \mathbb{1}_{\{\sigma < t\}} U_\sigma + \mathbb{1}_{\{\sigma = t = T\}} \xi \Big)$$

The left hand side is thus a bounded from below local martingale, hence it is a supermartingale. Moreover, (26) also holds with a stopping time $\tau \in \mathcal{T}$ instead of t therein. So, by taking expectations in (26) with τ instead of t therein:

$$w \geq \mathbb{E} \left\{ \int_0^{\sigma \wedge \tau} \beta_u C_u du + \beta_{\sigma \wedge \tau} \Big(\mathbb{1}_{\{\sigma \wedge = \tau < T\}} L_\tau + \mathbb{1}_{\{\sigma < \tau\}} U_\sigma + \mathbb{1}_{\{\sigma = \tau = T\}} \xi \Big) \right\} \; .$$

Hence $w \geq \Pi_0$ follows, by (12).
(iii) In the special case of an European derivative, the stated results follow by setting $K = 0$ in the previous points of the proof. □

Comments 2.3 **(i)** Proposition 2.3 thus *characterizes* the $\mathbb{P}$-price (*arbitrage price* relative to the risk-neutral measure $\mathbb{P}$) of a derivative as the *smallest initial wealth of a hedge* with $\mathbb{P}$ –local martingale cost, under the assumption that the related reflected BSDE (15) has a solution. For related results, see also Föllmer and Sondermann [55] or Schweizer [93].
(ii) The special case $\rho = 0$ in the previous results corresponds to a suitable form of model completeness (replicability of European options, cf. point (iii) of the proposition), in which the issuer of the option *wishes to* hedge all the risks embedded in the option.
The case $\rho \neq 0$ corresponds to either model incompleteness, or a situation of model completeness in which the issuer *wishes not to* hedge all the risks embedded in the product at hand, for instance because she wants to limit transaction costs, or because she *wishes to take some bets* in specific risk directions.
(iii) In case where ρ may be taken equal to 0 in Proposition 2.3, the minimality statements in this proposition can be used to prove uniqueness of the related arbitrage prices.
(iv) Analogous definitions and results hold for holder hedges.
(v) It is also easy to see that one could state analogous definitions and results regarding hedging a defaultable game option starting at any date $t \in [0, T]$, rather than at time 0 above.

3 Markovian Set-Up

3.1 Markovian FBSDE Approach

In order to be usable in practice, a dynamic pricing model needs to be constructive, or *Markovian* in some sense, relatively to a given derivative. This will be achieved by assuming that the related BSDE (15) is *Markovian* (see Section 4 of [51] and Part II).

Definition 3.1 We say that the BSDE (15) is a *Markovian Backward Stochastic Differential Equation* if the input data r, C, ξ, L and U of (15) are given by Borel-measurable functions of some $\mathbb{R}^q$-valued $(\mathbb{F}, \mathbb{P})$-Markov *factor* process X, so

$$r_t = r(t, X_t)\,,\ C_t = C(t, X_t)\,,\ \xi = \xi(X_T)\,,\ L_t = L(t, X_t)\,,\ U_t = U(t, X_t)\,, \tag{27}$$

and if $\bar{\sigma}$ is the first time of entry (capped at T) by the process (t, X) into a given closed subset of $[0, T] \times \mathbb{R}^q$.

Remark 3.2 By a slight abuse of notation, the related functions are thus denoted in (27) by the same symbols as the corresponding processes or random variables.

In particular, the system made of the specification of a forward dynamics for X, together with the BSDE (15), constitutes a decoupled *Markovian forward-backward system of equations* in (X, Π, M, K). The system is decoupled in the sense that the forward component of the system serves as an input for the backward component (X is an input to (15), cf. (27)), but not the other way round. See Definition 6.4 in Part II for more complete and formal statements.

From the point of view of interpretation, the components of X are observable *factors*. These are typically, though non-trivially, connected with the primary risky asset price process P,

as follows:
• Most factors are typically given as primary price processes. The components of X that are not included in P (if any) are to be understood as simple factors that may be required to 'Markovianize' the payoffs of the derivative at hand, such as factors accounting for path dependence in the derivative's payoff, and/or non-traded factors such as stochastic volatility in the dynamics of the assets underlying the derivative;
• Some of the primary price processes may not be needed as factors, but are used for hedging purposes.
Note that observability of the factor process X in the mathematical sense of $\mathbb{F}$-adaptedness is not sufficient in practice. In order for a factor process model to be usable in practice, a constructive *mapping* from a collection of meaningful and directly observable economic variables to X is really needed. Otherwise, the model will be useless.

3.2 Factor Process Dynamics

Under a rather generic specification for the Markov factor process X, we now derive a *variational inequality approach* for pricing and hedging a financial derivative.

We thus assume that the factor process X is an $(\mathbb{F} = \mathbb{F}^{W,N}, \mathbb{P})$-solution of the following Markovian (forward) stochastic differential equation in $\mathbb{R}^q$:

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dW_t + \delta(t, X_{t-})dN_t \ , \tag{28}$$

where:
• W is a q-dimensional Brownian motion, and
• N is a compensated integer-valued random measure with finite jump intensity measure $\lambda(t, X_t, dx)$, for some deterministic function λ.
So $\delta(t, X_{t-})dN_t$ in (28) is a short-hand for $\int_{\mathbb{R}^q} \delta(t, X_{t-}, x)N(dt, dx)$, where the integration is with respect to the x variable, and where the *response jump size function* δ and the *intensity measure* λ, like the other model coefficients b and σ of X, are to be specified depending on the application at hand: see Section 3.3 for specific examples and Definition 6.2 in Part II for more precise statements.

Remark 3.3 The generic and 'abstract' jump–diffusion (28) will be made precise and specified in Part II in the form of a process $\mathcal{X} = (X, N)$ in which a jump diffusion like component X interacts with a continuous time Markov chain like component N; so the process $\mathcal{X}$ in Part II corresponds to X here.

Let us introduce the following additional notation:
• J_t, a random variable on $\mathbb{R}^q$ with law $\frac{\lambda(t, X_{t-}, dx)}{\lambda(t, X_{t-}, \mathbb{R}^q)}$ conditional on X_{t-}, where x represents the size of a jump in $\delta(t, X_{t-}, x)$,
• (t_l), the ordered sequence of the times of jumps of N (note that we deal with a *finite* jump measure λ, so (t_l) is well defined),
• For any (possibly vector-valued) function u on $\mathbb{R}^q$,

$$\delta u(t, x, y) = u(t, x + \delta(t, x, y)) - u(t, x)\ , \ \bar{\delta} u(t, x) = \int_{\mathbb{R}^q} \delta u(t, x, y)\lambda(t, x, dy) \tag{29}$$
$$\delta u_t = \delta u(t, X_{t-}, J_t)\ , \ \bar{\delta} u_t = \bar{\delta} u(t, X_{t-})\ ,$$

We apologize to the reader for this admittedly heavy notation, which is motivated by the wish to give intuitive and compact forms to various expressions of the model's dynamics, generator and Itô formula. Denoting further

$$\bar{\delta}(t,x) := \bar{\delta}\mathrm{Id}_{\mathbb{R}^q}(t,x) = \int_{\mathbb{R}^q} \delta(t,x,y)\lambda(t,x,dy)\ ,\ \delta_t = \delta(t,X_{t-},J_t)\ ,\ \bar{\delta}_t = \bar{\delta}(t,X_{t-})\ ,$$

one thus has for instance:

$$\delta(t,X_{t-})dN_t = d\left(\sum_{t_l\le t}\delta_{t_l}\right) - \bar{\delta}_t dt \tag{30}$$

and the dynamics (28) of X may be rewritten as follows:

$$dX_t = \widetilde{b}(t,X_t)dt + \sigma(t,X_t)\,dW_t + d\left(\sum_{t_l\le t}\delta_{t_l}\right) \tag{31}$$

where we set $\widetilde{b}(t,x) = b(t,x) - \bar{\delta}(t,x)$.

3.2.1 Itô Formula and Model Generator

In view of (31), the following variant of the Itô formula holds, for any real-valued function u of class $\mathcal{C}^{1,2}$ on $[0,T]\times\mathbb{R}^q$:

$$du(t,X_t) = \widetilde{\mathcal{G}}u(t,X_t)\,dt + \partial u(t,X_t)\,\sigma(t,X_t)\,dW_t + d\left(\sum_{t_l\le t}\delta u_{t_l}\right) \tag{32}$$

with

$$\widetilde{\mathcal{G}}u(t,x) = \partial_t u(t,x) + \partial u(t,x)\widetilde{b}(t,x) + \frac{1}{2}\mathrm{Tr}[a(t,x)\mathcal{H}u(t,x)] \tag{33}$$

where $a(t,x) = \sigma(t,x)\sigma(t,x)^{\mathsf{T}}$, and where ∂u and $\mathcal{H}u$ denote the *row-gradient* and the *Hessian* of u with respect to x — so in particular

$$\mathrm{Tr}[a(t,x)\mathcal{H}u(t,x)] = \sum_{1\le i,j,k\le q}\sigma_{i,k}(t,x)\sigma_{j,k}(t,x)\partial^2_{x_i,x_j}u(t,x)\ .$$

Moreover (cf. (30)),

$$\delta u(t,X_{t-})dN_t = d\left(\sum_{t_l\le t}\delta u_{t_l}\right) - \bar{\delta}u_t dt\ . \tag{34}$$

The Itô formula (32) may thus be rewritten as, with the short-hand $\delta u(t,X_{t-})dN_t = \int_{x\in\mathbb{R}^q}\delta u(t,X_{t-},x)N(dt,dx)$,

$$du(t,X_t) = \mathcal{G}u(t,X_t)\,dt + \partial u(t,X_t)\,\sigma(t,X_t)\,dW_t + \delta u(t,X_{t-})dN_t \tag{35}$$

where we set

$$\begin{aligned}\mathcal{G}u(t,x) &= \widetilde{\mathcal{G}}u(t,x) + \bar{\delta}u(t,x)\\ &= \partial_t u(t,x) + \partial u(t,x)b(t,x) + \frac{1}{2}\mathrm{Tr}[a(t,x)\mathcal{H}u(t,x)] + \bar{\delta}u(t,x) - \partial u(t,x)\bar{\delta}(t,x)\ .\end{aligned} \tag{36}$$

The process X is thus a Markov process with generator $\mathcal{G}$ (see Proposition 9.2 in Part III for a more formal derivation).

Remark 3.4 By a convenient abuse of terminology we call here and henceforth $\mathcal{G}$ the generator of X, whereas strictly speaking $\mathcal{G}$ is the generator of the time-extended process (t,X) (the generator of X does not contain the ∂_t term).

3.2.2 Brackets

Let Π^c and Θ^c, resp. $\Delta\Pi$ and $\Delta\Theta$, denote the continuous local martingale components, resp. the jump processes, of two given real-valued semimartingales Π and Θ. Recall that the quadratic covariation or *bracket* $[\Pi, \Theta]$ is given by

$$d[\Pi,\Theta]_t = d(\Pi_t\Theta_t) - \Pi_{t-}d\Theta_t - \Theta_{t-}d\Pi_t \tag{37}$$

$$= d\langle\Pi^c,\Theta^c\rangle_t + d\left(\sum_{s\le t}\Delta\Pi_s\Delta\Theta_s\right) \tag{38}$$

with the initial condition $[\Pi,\Theta]_0 = 0$. The *sharp bracket* $\langle\Pi,\Theta\rangle$ corresponds to the *compensator* of $[\Pi,\Theta]$, which is well defined provided $[\Pi,\Theta]$ is of locally integrable variation (see, e.g., Protter [91]).

Assuming Π and Θ to be defined in terms of the process X of (28) by $\Pi_t = u(t, X_t)$ and $\Theta_t = v(t, X_t)$ for deterministic and 'smooth enough' functions u and v, an application of (38) yields:

$$d[\Pi,\Theta]_t = \partial ua(\partial v)^{\mathsf{T}}\,(t,X_t)\,dt + d\left(\textstyle\sum_{t_l\le t}\delta u_{t_l}\delta v_{t_l}\right) .$$

The bracket $[\Pi,\Theta]$ thus admits a compensator $< \Pi,\Theta >$ given as a time-differentiable process with the following Lebesgue-density:

$$\tfrac{d<\Pi,\Theta>_t}{dt} = (u,v)\,(t,X_t) \tag{39}$$

where we denote, for any vector-valued functions u and v on $\mathbb{R}^q$ such that the matrix-product uv^{T} makes sense:

$$(u,v)(t,x) = \partial ua(\partial v)^{\mathsf{T}}(t,x) + \int_{y\in\mathbb{R}^q}\delta u(\delta v)^{\mathsf{T}}(t,x,y)\lambda(t,x,dy) . \tag{40}$$

Remark 3.5 In the vector-valued case ∂u and ∂v are defined component by component, and can thus be identified to the Jacobian matrices of u and v.

Besides, it comes by application of the Itô formula (35) to the functions u, v and uv, '$\triangleq$' standing for 'equality up to a local martingale term':

$$d[\Pi,\Theta]_t = d(\Pi_t\Theta_t) - \Pi_{t-}d\Theta_t - \Theta_{t-}d\Pi_t$$
$$\triangleq \{\mathcal{G}(uv) - u\mathcal{G}v - v\mathcal{G}u\}\,(t,X_t)\,dt .$$

This yields the following alternative expression for $\frac{d<\Pi,\Theta>_t}{dt}$ (cf. (39)):

$$\tfrac{d<\Pi,\Theta>_t}{dt} = \{\mathcal{G}(uv) - u\mathcal{G}v - v\mathcal{G}u\}\,(t,X_t) . \tag{41}$$

We are now ready to prove the following,

Proposition 3.1 *For processes Π and Θ given as $\Pi_t = u(t,X_t)$ and $\Theta_t = v(t,X_t)$, where u and v are 'smooth enough' and in the domain of the operator $\mathcal{G}$, one has in probability, for almost every t,*

$$\tfrac{d\langle\Pi,\Theta\rangle_t}{dt} = \lim_{h\to 0} h^{-1}\mathbb{C}\mathrm{ov}_t(\Pi_{t+h}-\Pi_t,\Theta_{t+h}-\Theta_t) \tag{42}$$

where the subscript t stands for 'conditional on $\mathcal{F}_t$'.

Proof. For any fixed $h > 0$, one has,

$$\mathbb{C}\mathrm{ov}_t(\Pi_{t+h} - \Pi_t, \Theta_{t+h} - \Theta_t) + \mathbb{E}_t(\Pi_{t+h} - \Pi_t)\mathbb{E}_t(\Theta_{t+h} - \Theta_t) = \\ \mathbb{E}_t\left(\Pi_{t+h}\Theta_{t+h} - \Pi_t\Theta_t\right) - \Pi_t\mathbb{E}_t\left(\Theta_{t+h} - \Theta_t\right) - \Theta_t\mathbb{E}_t\left(\Pi_{t+h} - \Pi_t\right) \ . \tag{43}$$

Now one has by the Itô formula (35) applied to u, v and uv, respectively:

$$\lim_{h\to 0} h^{-1}\mathbb{E}_t(\Pi_{t+h} - \Pi_t) = \mathcal{G}u(t, X_t)$$
$$\lim_{h\to 0} h^{-1}\mathbb{E}_t(\Theta_{t+h} - \Theta_t) = \mathcal{G}v(t, X_t)$$
$$\lim_{h\to 0} h^{-1}\mathbb{E}_t\left(\Pi_{t+h}\Theta_{t+h} - \Pi_t\Theta_t\right) = \mathcal{G}(uv)(t, X_t)$$

Hence, by (43):

$$\lim_{h\to 0} h^{-1}\mathbb{C}\mathrm{ov}_t(\Pi_{t+h} - \Pi_t, \Theta_{t+h} - \Theta_t) = \\ \left\{\mathcal{G}(uv) - u\mathcal{G}v - v\mathcal{G}u\right\}(t, X_t) = \frac{d\langle \Pi, \Theta\rangle}{dt} \ ,$$

by (41). □

3.3 Examples

3.3.1 Model Specifications

In case $\lambda = 0$, the jump component of the generic jump-diffusion (28) vanishes, and we are left with a diffusion X.

In case $b = \bar{\delta}$ (so $\widetilde{b} = 0$ in (31)) and $\sigma = 0$, the general jump-diffusion X reduces to a pure jump process.
Under a more specific structure on δ and λ (see Section 6 in Part II for the related mathematics), the jump process X is supported by a finite set which can be identified with $E = \{1, \dots, n\}$, without loss of generality, and X is a continuous-time E-valued Markov chain X such that (cf. (31))

$$dX_t = d\left(\sum_{t_l \le t} \delta_{t_l}\right) \ , \tag{44}$$

with generator $\mathcal{G}$ such that, for any time-differentiable function u over $[0, T] \times E$ (or, equivalently, any system $u = (u^i)_{1\le i\le n}$ of time-differentiable functions u^i over $[0, T]$):

$$\mathcal{G}u^i(t) = \partial_t u^i(t) + \bar{\delta}u^i(t) = \partial_t u^i(t) + \sum_{j\neq i} \lambda^{i,j}(t)\big(u^j(t) - u^i(t)\big) \ . \tag{45}$$

3.3.2 Unbounded Jump Measures

For simplicity we did not consider above the 'infinite activity' case of possibly unbounded jump intensity measures $\lambda(t, x, \cdot)$. Note however that reinforcing our local boundedness assumption on the response jump size function δ into

$$|\delta(t, x, y)| < C(1 \wedge |y|) \tag{46}$$

for some constant C locally uniform in (t,x)[3], then most statements in this part (and the related developments in Parts II to IV as well) can be extended to more general *Lévy jump measures* $\lambda(t,x,\cdot)$ on $\mathbb{R}^q$ such that, locally uniformly in (t,x),

$$\int_{\mathbb{R}^q}(1\wedge|y|^2)\lambda(t,x,dy)<C\,. \tag{47}$$

The stochastic differential equation (28) then defines a Markov process X with generator (to be compared with (36))

$$\begin{aligned}\mathcal{G}u(t,x)=\partial_t u(t,x)+\partial u(t,x)b(t,x)+\frac{1}{2}\mathrm{Tr}[a(t,x)\mathcal{H}u(t,x)]+\\ \int_{\mathbb{R}^q}\Big(\delta u(t,x,y)-\partial u(t,x)\delta(t,x,y)\Big)\lambda(t,x,dy)\end{aligned} \tag{48}$$

where the integral converges for functions $u=u(t,x)$ of class $\mathcal{C}^2$ in x, under (46)–(47).

Remark 3.6 In the context of Lévy jump measures λ on $\mathbb{R}^q$, the process X is typically defined via its Lévy triplet $(\bar{b},\sigma,\lambda)$ in the following form (see, e.g., Cont and Tankov [37]):

$$dX_t=\bar{b}(t,X_t)\,dt+\sigma(t,X_t)\,dW_t+d\left(\sum_{\bar{t}_l\le t}\delta(t,X_{\bar{t}_l-},J_{\bar{t}_l})\right)+\int_{|x|<1}\delta(t,X_{t-},x)\,N(dt,dx) \tag{49}$$

where the $\bar{t}_l$s stand for the successive jump times of $N(\bar{B}\times[0,t])$, in which $\bar{B}$ denotes the complement of the unit ball in $\mathbb{R}^q$ (note that the ordered sequence $(\bar{t}_l)$ is well defined, in the case of Lévy jump measures $\lambda(t,x,\cdot)$). By identification with (28), it comes:

$$b(t,x)=\bar{b}(t,x)+\int_{|y|\ge 1}\delta(t,x,y)\lambda(t,x,dy)\,.$$

The following equivalent form of the generator $\mathcal{G}$ in terms of $\bar{b}$ follows (cf. (48)):

$$\begin{aligned}\mathcal{G}u(t,x)=\partial_t u(t,x)+\partial u(t,x)\bar{b}(t,x)+\frac{1}{2}\mathrm{Tr}[a(t,x)\mathcal{H}u(t,x)]\\ +\textstyle\int_{\mathbb{R}^q}\Big(\delta u(t,x,y)-\partial u(t,x)\delta(t,x,y)\mathbb{1}_{|y|<1}\Big)\lambda(t,x,dy)\,.\end{aligned} \tag{50}$$

3.3.3 Applications

With such versatile specifications ranging from pure diffusions, or (resorting to unbounded jump measures as explained in Section 3.3.2) Lévy processes, to continuous-time Markov chains, the jump–diffusion model factor process model (28) offers a flexible setting which is sufficient for most applications in financial derivatives modeling.

This set-up includes in particular the most common forms of stochastic volatility and/or jump *equity derivatives* models, like the *Black–Scholes model*, *local volatility models*, the *Merton model*, the *Heston model*, the *Bates model*, the most common forms of *Lévy models* used in finance for pricing purposes, etc.

[3] In the sense that for every compact set in the (t,x) variables there exists a constant C such that (46) holds for every (t,x) in this set and $y\in\mathbb{R}^q$.

Moreover, as we shall see in Section 4.2, one can easily accommodate in this risk-neutral martingale modeling approach *defaultable derivatives* with terminal payoffs of the form $\mathbb{1}_{T<\theta}\phi(X_T)$ (or $\mathbb{1}_{\nu<\theta}\phi(X_\nu)$ upon exercise at a stopping time ν, in case of American or game claims), where θ represents the *default-time* of a reference entity. This allows one to deal with *equity-to-credit* derivatives, like, for instance, convertible bonds (see Section 4.2.1.1). A model X as of (28) is then typically used in the mode of a *pre-default factor process model* (see Section 4.2 and [17]).

As will be explained in Section 4.1, the risk-neutral martingale modeling approach can also be readily extended to a martingale modeling approach relative to an arbitrary *numeraire*, rather than the savings account implicit in the risk-neutral approach. This allows one to extend the previous models to *interest-rates* and *foreign exchange* derivatives, yielding for instance the *Black model* or the *SABR model*, to quote but a few.

Finally continuous-time Markov chains, or continuous-time Markov chains modulated by diffusions, which, as illustrated in Section 3.3.1 and made precise in Part II (see Sections 6 and 7 therein), can all be considered as specific instances of the general jump-diffusion framework (28), cover most of the dynamic models used in the field of *portfolio credit derivatives.* Let us thus quote:

- The so called *local intensity model*, or pure birth process, which is used for modeling a credit portfolio cumulative loss process in Laurent, Cousin and Fermanian [81], Cont and Minca [36] or Herbertsson [64],
- A more general *homogeneous groups model* considered for different purposes by various authors in [58, 14, 41], among others,
- An even more general *basket credit migrations model* of Bielecki et al. [19, 21] in which the dynamics of the credit ratings of reference entities are modulated by the evolution of macro-economic factors, or a recent generation of *Markovian copula models* of Bielecki et al. [22] with model marginals automatically calibrated to the individual CDS curves.

3.4 Markovian Reflected BSDEs and PDEs with obstacles

3.4.1 No Protection Price

With the jump-diffusion factor process X defined by (28) and in the special case of a game option with no call protection ($\bar{\sigma} = 0$), the partial integro-differential equation formally related to the pricing BSDE (15) writes,

$$\begin{aligned}\min\Big(\max\Big(&\mathcal{G}u(t,x) + C(t,x) - r(t,x)u(t,x),\\ &L(t,x) - u(t,x)\Big), U(t,x) - u(t,x)\Big) = 0, \quad t < T,\ x \in \mathbb{R}^q,\end{aligned} \tag{51}$$

with terminal condition $u(T,x) = \xi(x)$.

An application of the results of Part III (see Proposition 12.4(i) therein) yields,

Proposition 3.2 *Under mild conditions, problem (51) is well-posed in the sense of viscosity solutions, and its solution $u(t,x)$ is related to the solution (Π, M, K) of (15) as follows, for $t \in [0,T]$:*

$$\Pi_t = u(t, X_t) \ . \tag{52}$$

In view of Proposition 2.3(ii), $u(0, X_0) = \Pi_0$ is therefore the minimal initial wealth of a (super-)hedge with $\mathbb{P}$ – local martingale cost process for the option.

Remark 3.7 When the pricing function u is sufficiently regular for an Itô formula to be applicable, one has further, for $t \in [0, T]$ (see, e.g., [11, 12, 8, 5, 4]),

$$dM_t = \partial u \sigma(t, X_t) dW_t + \delta u(t, X_{t-}) dN_t \ . \tag{53}$$

3.4.2 Protection Price

We now consider a call protection of the form

$$\bar{\sigma} = \inf\{ t > 0 \, ; \, X_t \notin \mathcal{O} \} \wedge \overline{T} \tag{54}$$

for a constant $\overline{T} \in [0, T]$ and an open subset $\mathcal{O} \subseteq \mathbb{R}^q$ satisfying suitable regularity properties (see, e.g., Example 8.2 in Part II).

A further application of the results of Part III (Proposition 12.4 therein) then yields,

Proposition 3.3 (i) (No-protection price). *On $[\bar{\sigma}, T]$, the $\mathbb{P}$-price process Π can be represented as $\Pi_t = u(t, X_t)$, where u is the unique viscosity solution of (51);*
(ii)(Protection price). *On $[0, \bar{\sigma}]$, the the $\mathbb{P}$-price process Π can be represented as $\bar{u}(t, X_t)$, where the function $\bar{u}$ is the unique viscosity solution of*

$$\max\Big(\mathcal{G}\bar{u}(t, x) + C(t, x) - r(t, x)\bar{u}(t, x), L(t, x) - \bar{u}(t, x) \Big) = 0, \quad t < \overline{T}, \ x \in \mathcal{O}, \tag{55}$$

with boundary condition $\bar{u} = u$ on $\big([0, T] \times \mathbb{R}^q\big) \setminus \big([0, \overline{T}) \times \mathcal{O}\big)$.

Remark 3.8 Because of the jumps in X, one needs to deal with the 'thick' parabolic boundary $\big([0, T] \times \mathbb{R}^q\big) \setminus \big([0, \overline{T}) \times \mathcal{O}\big)$.

Moreover (cf. Remark 3.7), in case the pricing functions u and $\bar{u}$ are sufficiently regular for an Itô formula to be applicable, one has further, for $t \in [0, T]$,

$$dM_t = \partial \nu \sigma(t, X_t) dW_t + \delta \nu(t, X_{t-}) dN_t \ , \tag{56}$$

where the random function ν therein is to be understood as u for $t > \bar{\sigma}$ and as $\bar{u}$ for $t \le \bar{\sigma}$.

Remark 3.9 Under more specific assumptions on the structure of the jump component of the model (see, e.g., Section 6 in Part II), the generic cascade of two PDEs (51), (55) can be amended in various ways. For instance, in the case of a continuous time Markov chain X over $E = \{1, \dots, n\}$ and for $\bar{\sigma}$ defined by (54) with $\mathcal{O}$ therein given as a subset of E, equations (51), (55) on $\mathbb{R}^q$ in fact reduce to a cascade of two systems of ODEs to be solved in $(u, \bar{u}) = (u^i(t), \bar{u}^i(t))_{1 \le i \le n}$, namely,

$$\begin{cases} u^i(T) = \xi^i(T), \quad 1 \le i \le n \\ \min\Big(\max\Big(\mathcal{G}u^i(t) + C^i(t) - r^i(t)u^i(t), \\ \qquad L^i(t) - u^i(t) \Big), U^i(t) - u^i(t) \Big) = 0, \quad t < T, \ 1 \le i \le n \\ \bar{u} = u \text{ on } \big([0, T] \times E\big) \setminus \big([0, \overline{T}) \times \mathcal{O}\big) \\ \max\Big(\mathcal{G}\bar{u}^i(t) + C^i(t) - r^i(t)\bar{u}^i(t), L^i(t) - \bar{u}^i(t) \Big) = 0, \quad t < \overline{T}, \ i \in \mathcal{O} \end{cases} \tag{57}$$

where the generator $\mathcal{G}$ therein assumes the form (45).

Note that we find it convenient in this article to refer to a decoupled system of partial integro-differential equations or obstacle problems as a *cascade* of PDEs, and particularly so (but not only, cf. above), when this system consists of equations defined over successive time intervals $[T_{l-1}, T_l]$, so that the solution of the equation which is posed over the next (in 'backward time') time interval is used as a terminal condition for the equation over the previous time interval.

3.5 Discussion of Various Hedging Schemes

In view of Proposition 3.3, the first line of (15) takes the following form (cf. (18)):

$$-d\nu(t, X_t) = (C - r\nu)(t, X_t)dt + dK_t - \partial\nu\sigma(t, X_t)dB_t - \delta\nu(t, X_{t-})dN_t \tag{58}$$

where the function ν therein is to be understood as u for $t > \bar{\sigma}$ and as $\bar{u}$ for $t \leq \bar{\sigma}$.

Let us assume the same structure (except for the barriers) on the primary market price process P, so $P_t = v(t, X_t)$ for a deterministic function $v(t, x)$, and

$$-dv(t, X_t) = (\mathcal{C} - rv)(t, X_t)dt - \partial v\sigma(t, X_t)dB_t - \delta v(t, X_{t-})dN_t\ , \tag{59}$$

where $\mathcal{C}(t, X_t)$ represents a primary market coupon rate process.

Note that v is an $\mathbb{R}^d$-valued function, so in particular ∂v lives in $\mathbb{R}^{d\otimes q}$, and identity (59) holds in $\mathbb{R}^d$.

The cost ρ relative to the strategy ζ (cf. (22)) can in turn be expressed in terms of the pricing functions u and v and the related delta functions.

Proposition 3.4 *Under the previous conditions in the Markovian jump-diffusion set-up (28), the dynamics (22) for the cost process ρ relative to the strategy ζ (and thus the related tracking error e in (23)) may be rewritten as (using the notation introduced in (29)):*

$$\begin{aligned} d\rho_t = &\Big(\partial\nu\sigma(t, X_t) - \zeta_t\partial v\sigma(t, X_t)\Big)dW_t \\ &+ \Big(\delta\nu(t, X_{t-}) - \zeta_t\delta v(t, X_{t-})\Big)dN_t \end{aligned} \tag{60}$$

It is thus possible to hedge completely the market risk W by setting, provided $\partial v\sigma$ is left-invertible,

$$\zeta_t = \partial\nu\sigma(\partial v\sigma)^{-1}(t, X_t) \tag{61}$$

In the simplest case where $q = d$ and ∂v and σ are invertible this formula further reduces to

$$\zeta_t = \partial\nu(\partial v)^{-1}(t, X_t) \tag{62}$$

Plugging this strategy into (60), one is left with the cost process

$$\rho_t = \int_0^{\cdot} \Big(\delta\nu(t, X_{t-}) - \zeta_t\delta v(t, X_{t-})\Big)dN_t \tag{63}$$

with ζ defined by (61) (or (62)). It is thus interesting to note that this strategy, which is perfect on one hand from the point of view of hedging the market risk W, potentially *creates some jump risk* on the other hand via the dependence on ζ of the integrand in (63).

At the other extreme, in case the jump measure has finite support (like in the case of a continuous-time Markov chain X with state-space reducible to a finite set E, cf. Remark 3.9), it is alternatively possible to hedge completely the jump risk N by setting, provided $\delta v(t, X_{t-})$ is left-invertible,

$$\zeta_t = \delta\nu(\delta v)^{-1}(t, X_{t-}) \ . \tag{64}$$

Plugging this strategy into (60), one is left with the cost process

$$\rho_t = \int_0^{\cdot} \Big(\partial\nu\sigma(t, X_t) - \zeta_t \partial v\sigma(t, X_t)\Big) dW_t \tag{65}$$

with ζ defined by (64). Note however that this strategy potentially *creates market risk* via the dependence in ζ of the integrand in (65).

Remark 3.10 In the context of credit derivatives (see also Section 4.2 in this regard), hedging the source risk W typically amounts to *hedging the spread risk*, whereas hedging the source risk N typically amounts to *hedging default risk*. We thus see that hedging the spread risk without caring about default risk, which has been the tendency in the practical risk management of credit derivatives in the last years (to spare the high cost of hedging default risk), can lead to leveraged default risk.

3.5.1 Min-Variance Hedging

Again a perfect hedge ($\rho = 0$) is hopeless unless the jump measure of X has finite support. In the context of incomplete markets the choice of a hedging strategy is up to one's *optimality criterion*, relative to the hedging cost (22)–(60). For instance, a trader may wish to minimize the (objective, $\widehat{\mathbb{P}}$ –) variance of $\int_0^T \beta_t d\rho_t$. Yet the related strategy $\widehat{\zeta}^{va}$ is hardly accessible in practice (in particular it typically depends on the objective model drift, a quantity notoriously difficult to estimate from financial data). As a proxy to this strategy, traders commonly use the strategy ζ^{va} which minimizes the *risk-neutral* variance of the error. Note that under mild conditions $\int_0^{\cdot} \beta dM$ and $\beta\widehat{P}$ are square integrable martingales, as they can typically be defined in terms of the martingales components of the solutions to related BSDEs. The risk-neutral min-variance hedging strategy ζ^{va} is then given by the following *Galtchouk-Kunita-Watanabe decomposition* of $\int_0^{\cdot} \beta dM$ with respect to $\beta\widehat{P}$ (see, e.g., Protter [91, IV.3, Corollary 1]):

$$\beta_t dM_t = \zeta_t^{va} d(\beta_t \widehat{P}_t) \ + \ \beta_t d\rho_t^{va} \tag{66}$$

for some $\mathbb{R}^d$-valued $\beta\widehat{P}$-integrable process ζ^{va} and a real-valued square integrable martingale $\beta_t d\rho_t^{va}$ strongly orthogonal to $\beta\widehat{P}$. Denoting in vector-matrix form

$$< A, B >= (< A^i, B^j >)_i^j \ , \ < A >=< A, A > \ ,$$

one thus has by (66) and (39):

$$\zeta_t^{va} = \tfrac{d<\Pi, P>_t}{d<P>_t} = \tfrac{(\nu, v)}{(v, v)}\big(t, X_{t-}\big) \ . \tag{67}$$

Comments 3.1 **(i)** For every fixed $t \in [0, T]$ and $h > 0$, it follows from (66) that $(\zeta_u^{va})_{u\in[t,t+h]}$ minimizes

$$\mathbb{V}\mathrm{ar}_t\Big(\int_t^{t+h} \beta_u dM_u - \int_t^{t+h} \zeta_u d(\beta_u d\widehat{P}_u) \Big) \ ,$$

where the subscript t stands for 'conditional on $\mathcal{F}_t$', over the set of all (self-financing) primary strategies (ζ_u) on the time interval $[t, t+h]$. Let likewise $\zeta_t^{va,h}$ minimize

$$\mathbb{V}\mathrm{ar}_t\Big(\int_t^{t+h} \beta_u dM_u - \zeta_t^h \int_t^{t+h} d(\beta_u d\widehat{P}_u)\Big)$$

over the set of all *buy-and-hold* constant strategies ζ_t^h on the time interval $[t, t+h]$. The strategy $\zeta_t^{va,h}$ is given as the solution of the linear regression problem of $\int_t^{t+h} \beta_u dM_u$ against $\int_t^{t+h} d(\beta_u d\widehat{P}_u)$, so:

$$\zeta_t^{va,h} = \mathbb{C}\mathrm{ov}_t\big(\int_t^{t+h} \beta_u dM_u, \int_t^{t+h} d(\beta_u d\widehat{P}_u)\big)\mathbb{V}\mathrm{ar}_t\big(\int_t^{t+h} d(\beta_u d\widehat{P}_u)\big)^{-1} .$$

In view of (43) we deduce that $\zeta_t^{va} = \lim_{h\to 0} \zeta_t^{va,h}$, as it was natural to expect.
(ii) In case of a diffusion X (without jumps), sharp brackets coincide with square brackets and are independent of the equivalent probability measure under consideration. It follows that the risk-neutral min-variance hedging strategy ζ^{va} defined by (67) satisfies $\zeta_t^{va} = \lim_{h\to 0} \widehat{\zeta}_t^{va,h}$, where the strategies $\widehat{\zeta}_t^{va,h}$ are the counterpart relative to the objective probability measure $\widehat{\mathbb{P}}$ of the strategies $\zeta_t^{va,h}$ introduced in part (i). In the no jumps case the risk-neutral min-variance hedging strategy ζ^{va} is thus also an objective locally (but possibly not globally) minimal variance strategy.

4 Extensions

4.1 More General Numeraires

Up to this point, we implicitly chose the savings account β^{-1}, assumed to be a positive finite variation process, as a *numeraire*, namely a primary asset with positive price process, used for discounting other price processes. However for certain applications, like dealing with stochastic interest rates in the field of *interest rate derivatives*, this choice may not be available (inasmuch as there may not be a riskless asset in the primary market), or it may not be the most appropriate (even if there is a riskless asset, the choice of another asset as a numeraire may be more convenient). This motivates the extension of the previous developments to the case where B is a general locally bounded positive semimartingale, not necessarily of finite variation. The interpretation of B as savings account and of $\beta = B^{-1}$ as a riskless discount factor is now replaced by the interpretation of B as a simple numeraire, referring to the fact that other price processes will be typically expressed as relative (rather than discounted) prices βP.

Understanding a *discounted price* as a *relative price*, a *risk-neutral model* as a *martingale model relative to the numeraire* B, etc., the risk-neutral modeling approach developed in the previous sections holds mutatis mutandis under this relaxed assumption on B. Note in particular that the self-financing condition still assumes the form of equation (4) (see, e.g., Protter [92]), though this is not as obvious as in the special case where B was a finite variation and continuous process. Also note that the concept of arbitrage is now to be understood relatively to the numeraire B (the set of admissible strategies being a numeraire dependent notion).

In this more general situation, we define a formal correspondence between triplets of processes (Π, M, K) and (π, m, k) by setting

$$\pi_t = \beta_t \Pi_t \,,\; dm_t = \beta_t \, dM_t \,,\; dk_t = \beta_t \, dK_t \text{ with } m_0 = 0 \text{ and } k_0 = 0 \tag{68}$$

where β now refers to the discount factor relative to an arbitrarily fixed numeraire. The pricing BSDE (15) (with β therein just mentioned above) to be solved in (Π, M, K), is then equivalent to the following BSDE with data $(c, \chi, \ell, \bar{h}) := (\beta C, \beta_T \xi, \; \beta L, \; \beta \bar{U})$, to be solved in (π, m, k) (cf. (18)):

$$\begin{array}{c} \pi_t = \chi + c_T - c_t + k_T - k_t - (m_T - m_t), \quad t \in [0, T] \\ \ell_t \leq \pi_t \leq \bar{h}_t, \quad t \in [0, T] \\ \int_0^T (\pi_u - \ell_u) \, dk_u^+ = \int_0^T (\bar{h}_u - \pi_u) \, dk_u^- = 0 \,. \end{array} \tag{69}$$

Note that equation (69) is but equation (15) with input data $r, C, \xi, L, \bar{U}$ defined as $0, c, \chi, \ell, \bar{h}$.

The conclusions of Propositions 2.2, 2.3 are still valid in this context, provided that 'a solution (Π, M, K) to (15)' therein is understood as the process (Π, M, K) defined via (68) in terms of a solution (π, m, k) to (69).

The Markovian case now corresponds to the situation where (cf. (27)):

$$c_t = c(t, X_t) \,,\; \chi = \chi(X_T) \,,\; \ell_t = \ell(t, X_t) \,,\; h_t = h(t, X_t) \tag{70}$$

for a suitable $\mathbb{R}^q$-valued $(\mathbb{F}, \mathbb{P})$-Markov factor process X.

In the generic jump-diffusion model X defined by (28) under a valuation measure $\mathbb{P}$ corresponding to the numeraire under consideration, with generator $\mathcal{G}$ given by (36), and for $\bar{\sigma}$ given by (54), the cascade of two PDEs to be solved in the no-protection and protection pricing functions u, $\bar{u}$ formally related to the BSDE (69) writes:

$$\begin{cases} u(T, x) = \chi(x) \,,\; x \in \mathbb{R}^q \\ \min\left(\max\left(\mathcal{G}u + c, \ell - u\right), h - u\right) = 0 \text{ on } [0, T) \times \mathbb{R}^q \\ \bar{u} = u \text{ on } \big([0, T] \times \mathbb{R}^q\big) \setminus \big([0, \overline{T}) \times \mathcal{O}\big) \\ \max\Big(\mathcal{G}\bar{u} + c, \ell - \bar{u}\Big) \text{ on } [0, \overline{T}) \times \mathcal{O} \end{cases} \tag{71}$$

We then have the following analog to Propositions 3.2–3.3.

Proposition 4.1 *Under suitable conditions, the BSDE (69) admits a unique solution (π, m, k), and the cascade of PDEs (71) admits a unique viscosity solution $(u, \bar{u})$. The connection between (π, m, k) and $(u, \bar{u})$ writes, for $t \in [0, T]$:*

$$\pi_t = \nu(t, X_t)$$

where ν is to be understood as u for $t > \bar{\sigma}$ and $\bar{u}$ for $t \leq \bar{\sigma}$.

Moreover, in case the pricing functions u, $\bar{u}$ are sufficiently regular for an Itô formula to be applicable, one has further, for $t \in [0, T]$,

$$dm_t = \partial \nu \sigma(t, X_t) dW_t + \delta \nu(t, X_{t-}) dN_t \,.$$

Let us further assume that the primary risky price process P satisfies likewise $p = \beta P = v(t, X_t)$ for a function v such that

$$d(\beta_t \widehat{P}_t) = \partial v \sigma(t, X_t) dW_t \, + \, \delta v(t, X_{t-}) dN_t \, . \tag{72}$$

One then has the following analog to Proposition 3.4.

Proposition 4.2 $\Pi_0 = B_0 \nu(t, X_0)$ *is the minimal initial wealth of a (super-)hedge with* $\mathbb{P}$ *– local martingale cost process. Moreover the cost process* $\rho = \rho(\zeta)$ *and the tracking error process* $e = e(\zeta)$ *in (21), (22), (23) may be rewritten as, respectively (with* $\rho_0 = 0$*):*

$$\begin{aligned} d\rho_t = & \Big(\partial \nu \sigma(t, X_t) - \zeta_t \partial v \sigma(t, X_t)\Big) dW_t \\ & + \Big(\delta \nu(t, X_{t-}) - \zeta_t \delta v(t, X_{t-})\Big) dN_t \end{aligned} \tag{73}$$

$$\beta_t e_t = \pi_0 - \int_0^t c_u du \, + \int_0^t \zeta_u d(\beta_u \widehat{P}_u) - \pi_t = \int_0^t dk_u - \int_0^t \beta_u d\rho_u \, . \tag{74}$$

It is thus possible to hedge completely the market risk represented by W by setting, provided $\partial v \sigma$ is left-invertible,

$$\zeta_t = \partial \nu \sigma (\partial v \sigma)^{-1}(t, X_t) \tag{75}$$

In the simplest case where $q = d$ and ∂v and σ are invertible this formula further reduces to

$$\zeta_t = \partial \nu (\partial v)^{-1}(t, X_t) \tag{76}$$

Alternatively, it is possible to hedge completely the jump risk N by setting, provided $\delta v(t, X_{t-})$ is left-invertible (assuming a jump measure with finite support, here),

$$\zeta_t = \delta \nu (\delta v)^{-1}(t, X_{t-}) \tag{77}$$

Still another possibility is to use the strategy ζ^{va} which minimizes the risk-neutral variance of the error, and which is given by

$$\zeta_t^{va} = \frac{d<\pi, p>_t}{d<p>_t} = \frac{(\nu, v)}{(v, v)}(t, X_{t-}) \, . \tag{78}$$

4.2 Defaultable Derivatives

To illustrate the flexibility of the above martingale modeling approach to pricing and hedging problems in finance, we now consider an extension of the previous developments to *defaultable derivatives*. This class of assets, including convertible bonds in particular (see Definition 4.2.1.1), plays an important role in the sphere of equity-to-credit / credit-to-equity capital structure arbitrage strategies.

Back to risk-neutral modeling with respect to a numeraire B given as a savings account and for a riskless discount factor $\beta = B^{-1}$ as of (1), we thus consider defaultable derivatives with terminal payoffs of the form $\mathbb{1}_{T<\theta}\phi(S_T)$ (or $\mathbb{1}_{\nu<\theta}\phi(S_\nu)$ upon exercise at a stopping time ν, in case of American or game claims), where θ represents the *default-time* of a reference entity. We shall follow the reduced-form intensity approach originally introduced by Lando [79] or Jarrow and Turnbull [70], subsequently generalized in many ways in the credit risk literature (see for instance Bielecki and Rutkowski [20]), and extended in particular to American and game claims in Bielecki et al. [15, 16, 17, 18], on which the material of this section is based. We shall give hardly no proofs in this section, referring the interested reader to [15, 16, 17, 18].

The main message here is that defaultable claims can be handled in essentially the same way as default-free claims, provided the default-free discount factor process β is replaced by a *credit-risk adjusted* discount factor α, and a fictitious dividend continuously paid at rate γ (the *default intensity*) is introduced to account for recovery on the claim upon default.

Incidentally note that the 'original default-free' discount factor β can itself be regarded as a default probability, at the *killing rate* r in (1).

4.2.1 Cash Flows

Given a $[0, +\infty]$-valued stopping time θ representing the default time of a reference entity (firm), let us set

$$I_t = \mathbb{1}_{\{\theta \le t\}} \,,\ J_t = 1 - I_t \,.$$

We shall directly consider the case of defaultable game options with call protection $\bar{\sigma}$. For reasons analogous to those developed above, these encompass as a special case defaultable American options (case $\bar{\sigma} = T$), themselves including defaultable European options.

In few words, a defaultable game option is a game option in the sense of Definition 2.3(iii), with all cash flows killed at the default time θ.

Given a call protection $\bar{\sigma} \in \mathcal{T}$ and a pricing time $t \in [0, T]$, let ν stand for $\sigma \wedge \tau \wedge \theta$, for any $(\sigma, \tau) \in \bar{\mathcal{T}}_t \times \mathcal{T}_t$.

Definition 4.1 A *defaultable game option* is a game option with the *ex-dividend cumulative discounted cash flows* $\beta_t \pi(t; \sigma, \tau)$, where the $\mathcal{F}_\nu$-measurable random variable $\pi(t; \sigma, \tau)$ is given by the formula, for any pricing time $t \in [0, T]$, holder call time $\sigma \in \bar{\mathcal{T}}_t$ and issuer put time $\tau \in \mathcal{T}_t$,

$$\beta_t \pi(t; \sigma, \tau) = \int_t^\nu \beta_u \, dD_u + \beta_\nu J_\nu \Big(\mathbb{1}_{\{\nu = \tau < T\}} L_\tau + \mathbb{1}_{\{\nu < \tau\}} U_\sigma + \mathbb{1}_{\{\nu = T\}} \xi \Big), \tag{79}$$

where:
- the *dividend process* $D = (D_t)_{t \in [0,T]}$ equals

$$D_t = \int_{[0,t]} J_u C_u du + R_u \, dI_u \,,$$

for some *coupon rate process* $C = (C_t)_{t \in [0,T]}$, and some predictable locally bounded *recovery process* $R = (R_t)_{t \in [0,T]}$;
- the *put payment* $L = (L_t)_{t \in [0,T]}$ and the *call payment* $U = (U_t)_{t \in [0,T]}$ are càdlàg processes, and the *payment at maturity* ξ is a random variable such that

$$L \le U \text{ on } [0, T] \,,\ L_T \le \xi \le U_T \,.$$

We further assume that R, L and ξ are bounded from below, so that the cumulative discounted payoff is bounded from below. Specifically, there exists a constant c such that

$$\int_{[0,t]} \beta_u \, dD_u + \beta_t J_t \Big(\mathbb{1}_{\{t<T\}} L_t + \mathbb{1}_{\{t=T\}} \xi \Big) \geq -c \,, \qquad t \in [0,T] \,. \tag{80}$$

Remark 4.2 One can also cope with the case of *discrete coupons* (see [15, 16, 17, 18] and Section 14 in Part IV).

4.2.1.1 Convertible Bonds

The standing example of a defaultable game option is a (defaultable) *convertible bond.* Convertible bonds have two important and distinguishing features:
• early put and call clauses at the holder's and issuer's convenience, respectively;
• defaultability, since they are corporate bonds, and one of the main vehicles of the so called *equity to credit* and *credit to equity* strategies.

To describe the covenants of a (stylized) convertible bond, we need to introduce some additional notation:

$\bar{N}$: the par (or nominal) value,

S: the price process of the asset underlying the bond,

$\bar{R}$: the recovery process on the bond upon default of the issuer,

κ : the bond's conversion factor,

$\bar{P} \leq \bar{C}$: the put and call nominal payments, respectively; by assumption $\bar{P} \leq \bar{N} \leq \bar{C}$.

Definition 4.3 A convertible bond is a (defaultable) game option with coupon rate process C, recovery process R^{cb} and payoffs L^{cb}, U^{cb}, ξ^{cb} such that

$$R^{cb}_t = (1-\eta)\kappa S_{t-} \vee \bar{R}_t \,, \; \xi^{cb} = \bar{N} \vee \kappa S_T \tag{81}$$

$$L^{cb}_t = \bar{P} \vee \kappa S_t \,, \; U^{cb}_t = \bar{C} \vee \kappa S_t \,. \tag{82}$$

See [15] for a more detailed description of covenants of convertible bonds, with further important real-life features like discrete coupons or call protection.

4.2.2 Reduction of Filtration in the Hazard Intensity Set-Up

An application of Proposition 2.1 yields (see Bielecki et al. [16]),

Proposition 4.3 *Assume that a semimartingale* Π *is the value of the* Dynkin game *related to a defaultable game option under some risk-neutral measure* $\mathbb{P}$ *on the primary market, that is, for* $t \in [0,T]$:

$$\begin{aligned} \mathrm{esssup}_{\tau \in \mathcal{T}_t} \mathrm{essinf}_{\sigma \in \bar{\mathcal{T}}_t} \mathbb{E}_{\mathbb{P}} \big(\pi(t;\sigma,\tau) \,\big|\, \mathcal{F}_t \big) &= \Pi_t \\ &= \mathrm{essinf}_{\sigma \in \bar{\mathcal{T}}_t} \mathrm{esssup}_{\tau \in \mathcal{T}_t} \mathbb{E}_{\mathbb{P}} \big(\pi(t;\sigma,\tau) \,\big|\, \mathcal{F}_t \big) \,. \end{aligned} \tag{83}$$

Then Π *is an arbitrage price process for the defaultable game option. Moreover, a converse to this result holds under a suitable integrability assumption.*

We work henceforth under a given risk-neutral measure $\mathbb{P} \in \mathcal{M}$, with $\mathbb{P}$-expectation denoted by $\mathbb{E}$.

In view of applying the so-called *reduced-form approach* in (single-name) credit risk (see, e.g., [20]), we assume further that $\mathbb{F} = \mathbb{H} \vee \widetilde{\mathbb{F}}$, where the filtration $\mathbb{H}$ is generated by the *default indicator process* $I_t = \mathbb{1}_{\{\theta \le t\}}$ and $\widetilde{\mathbb{F}}$ is some *reference filtration*. Moreover, we assume that the optional projection of J, defined by, for $t \in [0, T]$,

$$^oJ_t = \mathbb{P}(\theta > t \,|\, \widetilde{\mathcal{F}}_t) =: Q_t$$

(the so-called *Azema's supermartingale*), is a (strictly) positive, continuous and non-increasing process.

Comments 4.1 **(i)** If Q is continuous, θ is a *totally inaccessible* $\mathbb{F}$-stopping time (see, e.g., Dellacherie and Meyer [46]). Moreover, θ *avoids* $\widetilde{\mathbb{F}}$*-stopping times*, in the sense that $\mathbb{P}(\theta = \tau) = 0$, for any $\widetilde{\mathbb{F}}$-stopping time τ (see Coculescu et al. [35]).
(ii) Assuming Q continuous, the further assumption that Q has a finite variation in fact implies that Q is non-increasing. This further assumption lies somewhere between assuming further the (stronger) $(\mathcal{H})$ (or *immersion*) Hypothesis and assuming further that θ is an $\widetilde{\mathbb{F}}$-*pseudo-stopping time*. Recall that the $(\mathcal{H})$ Hypothesis means that all $\widetilde{\mathbb{F}}$-local martingales are $\mathbb{F}$-local martingales; θ being an $\widetilde{\mathbb{F}}$*-pseudo-stopping time* means that all $\widetilde{\mathbb{F}}$-local martingales *stopped at* θ are $\mathbb{F}$-local martingales (see Nikeghbali and Yor [85]).

We assume for simplicity of presentation in this article that Q is time-differentiable, and we define the *default (hazard) intensity* γ, the *credit-risk adjusted interest rate* μ and the *credit-risk adjusted discount factor* α by, respectively;

$$\gamma_t = -\frac{d \ln Q_t}{dt}\,,\ \mu_t = r_t + \gamma_t\,,\ \alpha_t = \beta_t \exp(-\int_0^t \gamma_u du) = \exp(-\int_0^t \mu_u du)$$

Under the previous assumptions, the *compensated jump-to-default process* $H_t = I_t - \int_0^t J_u \gamma_u du$, $t \in [0, T]$, is known to be an $\mathbb{F}$-martingale. Also note that the process α is time-differentiable and bounded, like β.

The quantities $\widetilde{\tau}$ and $\widetilde{\Pi}$ introduced in the next lemma are called the *pre-default values* of τ and Π, respectively.

Lemma 4.4 (see, e.g., Bielecki et al. [16]) **(i)** *For any $\mathbb{F}$-adapted, resp. $\mathbb{F}$-predictable process Π over $[0, T]$, there exists an (unique) $\widetilde{\mathbb{F}}$-adapted, resp. $\widetilde{\mathbb{F}}$-predictable process $\widetilde{\Pi}$ over $[0, T]$ such that $J\Pi = J\widetilde{\Pi}$, resp. $J_{\cdot-}\Pi = J_{\cdot-}\widetilde{\Pi}$ over $[0, T]$.*
(ii) *For any $\tau \in \mathcal{T}$, there exists a $[0, T]$-valued $\widetilde{\mathbb{F}}$-stopping time $\widetilde{\tau}$ such that $\tau \wedge \theta = \widetilde{\tau} \wedge \theta$.*

In view of the structure of the payoffs π in (79), we thus may assume without loss of generality that the data C, R, L, U, ξ, the call protection $\bar{\sigma}$ and the stopping policies σ, τ are defined relative to the filtration $\widetilde{\mathbb{F}}$, rather than $\mathbb{F}$ above. More precisely, ***we assume in the sequel that*** C, L, U ***are*** $\widetilde{\mathbb{F}}$***−adapted,*** $\xi \in \widetilde{\mathcal{F}}_T$***,*** R ***is*** $\widetilde{\mathbb{F}}$***-predictable and*** $\bar{\sigma}, \sigma, \tau$ ***are*** $\widetilde{\mathbb{F}}$***-stopping times. For any*** $t \in [0, T]$***,*** $\mathcal{T}_t$ ***(or*** $\mathcal{T}$***, in case*** $t = 0$***) henceforth denotes the set of*** $[t, T]$***-valued*** $\widetilde{\mathbb{F}}$***- (rather than*** $\mathbb{F}$***- before) stopping times;*** ν ***denotes*** $\sigma \wedge \tau$ ***(rather than*** $\sigma \wedge \tau \wedge \theta$ ***before), for any*** $t \in [0, T]$ ***and*** $\sigma, \tau \in \mathcal{T}_t$***.***

The next lemma (which is rather standard, if not for the presence of the stopping policies σ, τ therein) shows that the computation of conditional expectations of cash flows $\pi(t;\sigma,\tau)$ with respect to $\mathcal{F}_t$, can then be reduced to the computation of conditional expectations of $\widetilde{\mathbb{F}}$-equivalent cash flows $\widetilde{\pi}(t;\sigma,\tau)$ with respect to $\widetilde{\mathcal{F}}_t$.

Lemma 4.5 (see Bielecki et al. [16]) *For any stopping times $(\sigma,\tau) \in \bar{\mathcal{T}}_t \times \mathcal{T}_t$, one has,*

$$\mathbb{E}\big(\pi(t;\sigma,\tau)\,\big|\,\mathcal{F}_t\big) = J_t\,\mathbb{E}\big(\widetilde{\pi}(t;\sigma,\tau)\,\big|\,\widetilde{\mathcal{F}}_t\big),$$

where $\widetilde{\pi}(t;\sigma,\tau)$ is given by, with $\nu = \tau \wedge \sigma$,

$$\alpha_t \widetilde{\pi}(t;\sigma,\tau) = \int_t^\nu \alpha_u f_u du + \alpha_\nu \big(\mathbb{1}_{\{\nu=\tau<T\}} L_\tau + \mathbb{1}_{\{\nu<\tau\}} U_\sigma + \mathbb{1}_{\{\nu=T\}}\xi\big) \tag{84}$$

where we set $f = C + \gamma R$.

As a corollary to the previous results, we have,

Proposition 4.6 (see Bielecki et al. [16]) *If an $\widetilde{\mathbb{F}}$-semimartingale $\widetilde{\Pi}$ solves the $\widetilde{\mathbb{F}}$-Dynkin game with payoff $\widetilde{\pi}$, in the sense that, for any $t \in [0,T]$,*

$$\begin{aligned}\mathrm{esssup}_{\tau\in\mathcal{T}_t}\mathrm{essinf}_{\sigma\in\bar{\mathcal{T}}_t}\mathbb{E}\big(\widetilde{\pi}(t;\sigma,\tau)\,\big|\,\widetilde{\mathcal{F}}_t\big) &= \widetilde{\Pi}_t \\ = \mathrm{essinf}_{\sigma\in\bar{\mathcal{T}}_t}\mathrm{esssup}_{\tau\in\mathcal{T}_t}\mathbb{E}\big(\widetilde{\pi}(t;\sigma,\tau)\big|\,\widetilde{\mathcal{F}}_t\big)&\ ,\end{aligned}$$

then $\Pi := J\widetilde{\Pi}$ is an $\mathbb{F}$-semimartingale solving the $\mathbb{F}$-Dynkin game with payoff π.

Hence, by Proposition 4.3, Π is an arbitrage price for the option, with related pre-default price process $\widetilde{\Pi}$. A converse to this result may be established under a suitable integrability assumption.

We thus effectively moved our considerations from the original market subject to the default risk, in which cash flows are discounted according to the discount factor β, to the fictitious default-free market, in which cash flows are discounted according to the credit risk adjusted discount factor α.

4.2.3 Backward Stochastic Differential Equations Pre-default Modeling

The next step consists in modeling $\widetilde{\Pi}$ as the state-process of a solution $(\widetilde{\Pi}, \widetilde{M}, \widetilde{K})$, assumed to exist, to the following doubly reflected BSDE with data $\alpha, f = C+\gamma R, \xi, L, \bar{U} = \mathbb{1}_{\{\cdot<\bar{\sigma}\}}\infty + \mathbb{1}_{\{\cdot\geq\bar{\sigma}\}}U$ (cf. Definition 2.7 for the definition of a solution to (85)):

$$\begin{cases}\alpha_t\widetilde{\Pi}_t = \alpha_T\xi + \int_t^T \alpha_u\left(f_u du + d\widetilde{K}_u - d\widetilde{M}_u\right),\ t\in[0,T], \\ L_t \leq \widetilde{\Pi}_t \leq \bar{U}_t, \quad t\in[0,T], \\ \int_0^T (\widetilde{\Pi}_u - L_u)\,d\widetilde{K}_u^+ = \int_0^T (\bar{U}_u - \widetilde{\Pi}_u)\,d\widetilde{K}_u^- = 0\ .\end{cases} \tag{85}$$

Hence, by Proposition 2.2, the $\widetilde{\mathbb{F}}$-semimartingale $\widetilde{\Pi}$ solves the $\widetilde{\mathbb{F}}$-Dynkin game with payoff $\widetilde{\pi}$. Thus, by Proposition 4.6, $\Pi := J\widetilde{\Pi}$ is an arbitrage price for the option, with related pre-default price process $\widetilde{\Pi}$.

Let us set further, for $t \in [0,T]$ (cf. (17)),

$$\Pi_t = \mathbb{1}_{\{t<\theta\}}\widetilde{\Pi}_t \,,\ \beta_t\widehat{\Pi}_t = \beta_t\Pi_t + \int_{[0,t]} \beta_u \, dD_u \tag{86}$$

where we recall that $D_t = \int_{[0,t]} J_u C_u du + R_u \, dI_u$. We define M by $M_0 = 0$ and, for $t \in [0,T]$,

$$\int_{[0,t]} \beta_u dM_u = \beta_t\widehat{\Pi}_t + \int_0^t \beta_u J_u \, dK_u. \tag{87}$$

The following lemma is key in this section. It allows one in particular to interpret (87) as the canonical decomposition of the $\mathbb{F}$ – special semimartingale $\beta\widehat{\Pi}$. In particular M is but the canonical $\mathbb{F}$-local martingale component of $\int_{[0,\cdot]} \beta_t^{-1} d(\beta_t\widehat{\Pi}_t)$ (cf. Remark 2.8).

Lemma 4.7 *The process M defined by (87) is an $\mathbb{F}$-local martingale (stopped at θ).*

Proof. One has by (85), for every $t \in [0,T]$,

$$\int_0^t \alpha_u \, d\widetilde{M}_u = \alpha_t\widetilde{\Pi}_t - \widetilde{\Pi}_0 + \int_0^t \alpha_u \, d\widetilde{K}_u + \int_0^t \alpha_u \, (C_u + \gamma_u R_u) du$$

So by standard computations (cf. Lemma 4.5), for any $0 \le t \le u \le T$,

$$\mathbb{E}\Big(\beta_t^{-1} \int_t^u \beta_v \, dM_v \,\Big|\, \mathcal{F}_t\Big) = J_t \, \mathbb{E}\Big(\alpha_t^{-1} \int_t^u \alpha_v \, d\widetilde{M}_v \,\Big|\, \widetilde{\mathcal{F}}_t\Big) = 0 \,.$$

□

Let

$$\sigma^* = \inf\big\{\, u \in [\bar{\sigma}, T] \, ; \, \widetilde{\Pi}_u \ge U_u \big\} \wedge T \,. \tag{88}$$

For any primary strategy ζ, let the $\mathbb{F}$-*local martingale* $\rho(\zeta) = \rho$ be given by $\rho_0 = 0$ and

$$d\rho_t = dM_t - \zeta_t \, \beta_t^{-1} d(\beta_t\widehat{P}_t). \tag{89}$$

Proposition 4.8 can be seen as an extension of Proposition 2.3 to the defaultable case, in which two filtrations are involved. Note that our assumptions here are made relative to the filtration $\widetilde{\mathbb{F}}$ (the one with respect to which the BSDE (85) is defined), whereas conclusions are drawn relative to the filtration $\mathbb{F}$.

Proposition 4.8 (see Bielecki et al. [17, 16]) (i) *For any hedging strategy ζ, (Π_0, ζ, σ^*), is an hedge with $(\mathbb{F}, \mathbb{P})$ – local martingale cost ρ;*
(ii) *Π_0 is the minimal initial wealth of an hedge with $(\mathbb{F}, \mathbb{P})$ – local martingale cost;*
(iii) *In the special case of an European derivative with $\widetilde{K} = 0$, then (Π_0, ζ) is a replicating strategy with $(\mathbb{F}, \mathbb{P})$ – local martingale cost ρ. Π_0 is thus also the minimal initial wealth of a replicating strategy with $(\mathbb{F}, \mathbb{P})$ – local martingale cost.*

4.2.3.1 Analysis of Hedging Strategies

Let $H_t = I_t - \int_0^t J_u \gamma_u du$ stand for the compensated jump-to-default $\mathbb{F}$-martingale. Our analysis of hedging strategies will rely on the following lemma, which yields the dynamics of the price process $\widehat{\Pi}$ of a game option or, more precisely, of the $\mathbb{F}$ − local martingale component M of $\int_{[0,\cdot]} \beta_t^{-1} d(\beta_t \widehat{\Pi}_t)$.

Lemma 4.9 *The $\mathbb{F}$-local martingale M defined in (87) satisfies, for $t \in [0, T \wedge \theta]$:*

$$dM_t = d\widetilde{M}_t + \Delta\widehat{\Pi}_t \, dH_t \tag{90}$$

with $\Delta\widehat{\Pi}_t := R_t - \widetilde{\Pi}_{t-}$.

Sketch of Proof (see Bielecki et al. [16] for the detail). This follows by computations similar to those of the proof of Kusuoka's Theorem 2.3 in [78] (where the $(\mathcal{H})$ hypothesis and a more specific Brownian reference filtration $\widetilde{\mathbb{F}} = \widetilde{\mathbb{F}}^W$ are assumed), using in particular the avoidance property viewed at Comment 4.1(i) that $\mathbb{P}(\theta = \tau) = 0$ for any $\widetilde{\mathbb{F}}$-stopping time τ. □

In analogy with the structure of the payoffs of a defaultable derivative, we assume henceforth that the dividend (vector) process $\mathcal{D}$ of the primary market price process P is given as

$$\mathcal{D}_t = \int_{[0,t]} J_u \, \mathcal{C}_u du + \mathcal{R}_u \, dH_u$$

for suitable coupon rate and recovery processes $\mathcal{C}$ and $\mathcal{R}$. We also assume that $P = J\widetilde{P}$, without loss of generality with respect to the application of hedging a defaultable derivative (in particular any value of the primary market at θ is embedded in the recovery part of the dividend process $\mathcal{D}$ for P). We further define, along with the cumulative price $\widehat{P}$ as usual, the *pre-default cumulative price*, by, for $t \in [0, T]$:

$$\bar{P}_t = \widetilde{P}_t + \alpha_t^{-1} \int_0^t \alpha_u \, g_u du$$

where we set $g = \mathcal{C} + \gamma \mathcal{R}$.

The following result is the analog, relative to the primary market, of identity (90) for a game option.

Lemma 4.10 (see Bielecki et al. [17]) *$\alpha \bar{P}$ is an $\widetilde{\mathbb{F}}$-local martingale and one has, for $t \in [0, T \wedge \theta]$:*

$$\beta_t^{-1} d(\beta_t \widehat{P}_t) = \alpha_t^{-1} d(\alpha_t \bar{P}_t) + \Delta\widehat{P}_t \, dH_t \tag{91}$$

with $\Delta\widehat{P}_t := \mathcal{R}_t - \widetilde{P}_{t-}$.

Plugging (91) and (90) into (89), one gets the following *decomposition of the hedging cost ρ* of the strategy (Π_0, ζ, σ^*).

Proposition 4.11 *Under the previous assumptions, for any primary strategy ζ, the related cost $\rho = \rho(\zeta)$ in Proposition 4.8 satisfies, for every $t \in [0, T \wedge \theta]$,*

$$d\rho_t = dM_t - \zeta_t \, \beta_t^{-1} d(\beta_t \widehat{P}_t) = \Big[d\widetilde{M}_t - \zeta_t \, \alpha_t^{-1} d(\alpha_t \bar{P}_t) \Big] + \Big[\Delta\widehat{\Pi}_t - \zeta_t \Delta\widehat{P}_t \Big] dH_t \, . \tag{92}$$

4.2.4 Pre-default Markovian Set-Up

We now assume that the pre-default pricing BSDE (85) is Markovian, in the sense that the pre-default input data $\mu = r+\gamma$, $f = C+\gamma R$, ξ, L, U of (85) are given by Borel-measurable functions of an $(\widetilde{\mathbb{F}}, \mathbb{P})$-Markov process X, so

$$\mu_t = \mu(t, X_t)\,,\ f_t = f(t, X_t)\,,\ \xi = \xi(X_T)\,,\ L_t = L(t, X_t)\,,\ U_t = U(t, X_t)\,.$$

We assume more specifically that the pre-default factor process X is defined by (28) with respect to $\widetilde{\mathbb{F}} = \mathbb{F}^{W,N}$, with related generator $\mathcal{G}$, and that $\bar{\sigma}$ is defined by (54).

One can then introduce the pre-default pricing PDE cascade formally related to the pre-default pricing BSDE (85), to be solved in the (no protection pricing function, protection pricing function) pair $(u, \bar{u})$, namely, with $f = C + \gamma R$, (cf. equations (51)-(55) or (71) above; see also [17]):

$$\begin{cases} u(T,x) = \xi(x)\,,\ x \in \mathbb{R}^q \\ \min\left(\max\left(\mathcal{G}u + f - \mu u, L - u\right), U - u\right) = 0 \text{ on } [0,T) \times \mathbb{R}^q \\ \bar{u} = u \text{ on } \left([0,T] \times \mathbb{R}^q\right) \setminus \left([0,\overline{T}) \times \mathcal{O}\right) \\ \max\left(\mathcal{G}\bar{u} + f - \mu u, L - \bar{u}\right) \text{ on } [0,\overline{T}) \times \mathcal{O} \end{cases} \tag{93}$$

One then has as before, by application of the results of Parts II and III,

Proposition 4.12 *Problem (93) is well-posed in the sense of viscosity solutions under mild conditions, and its solution $(u, \bar{u})$ is related to the solution $(\widetilde{\Pi}, \widetilde{M}, \widetilde{K})$ of (85) as follows, for $t \in [0,T]$:*

$$\widetilde{\Pi}_t = \nu(t, X_t) \tag{94}$$

where ν is to be understood as u for $t > \bar{\sigma}$ and $\bar{u}$ for $t \le \bar{\sigma}$.

Moreover, in case the pricing functions u and $\bar{u}$ are sufficiently regular for an Itô formula to be applicable, one has further, for $t \in [0,T]$,

$$d\widetilde{M}_t = \partial\nu\sigma(t, X_t)dW_t + \delta\nu(t, X_{t-})dN_t\,. \tag{95}$$

Accordingly, the first line of (85) takes the following form:

$$-\,d\nu(t, X_t) = (f - \mu\nu)(t, X_t)dt + d\widetilde{K}_t - \partial\nu\sigma(t, X_t)dB_t - \delta\nu(t, X_{t-})dN_t\,. \tag{96}$$

Let us assume the same structure (without barriers) on the primary market price process P, thus $P_t = v(t, X_t)$, where, setting $g(t,z) = \mathcal{C}(t,z) + \gamma(t,z)\mathcal{R}(t,z)$,

$$-\,dv(t, X_t) = (g - \mu v)(t, X_t)dt - \partial v\sigma(t, X_t)dB_t - \delta v(t, X_{t-})dN_t\,. \tag{97}$$

Exploiting (96) and (97) in (92), one gets,

Proposition 4.13 *For* $t \in [0, T \wedge \theta]$,

$$d\rho_t = \Big[(\partial\nu\sigma(t, X_t), \delta\nu(t, X_{t-}), \Delta\nu(t, X_{t-})) - \zeta_t(\partial v\sigma(t, X_t), \delta v(t, X_{t-}), \Delta v(t, X_{t-}))\Big] d\begin{pmatrix} B_t \\ N_t \\ H_t \end{pmatrix} . \quad (98)$$

where we set $\Delta\nu(t, x) = (R - \nu)(t, x), \Delta v(t, x) = (\mathcal{R} - v)(t, x)$.

As in Section 3.5 (see also Bielecki et al. [18]), this decomposition of the hedging cost ρ can then be used for devising practical hedging schemes of a defaultable game option, like superhedging ($\rho = 0$), hedging only the market (spread) risk B, hedging only the default risk H, or min-variance hedging.

Comments 4.2 **(i)** Under more specific assumptions on the structure of the jump component of the model, the cascade of PDEs (93) can assume various forms, like, for instance, being reducible to a cascade of systems of ODEs, cf. Remark 3.9 and Part III.
(ii) Analogous developments regarding defaultable derivatives can also be made relatively to a more general numeraire, cf. Section 4.1.

4.3 Intermittent Call Protection

We now want to consider callable products with more general, hence potentially more realistic forms of *intermittent* call protection, namely call protection *whenever a certain condition* is satisfied, rather than more specifically call protection *before a stopping time* earlier in this part. This leads us to introduce financial derivatives with an effective call payoff process $\bar{U}$ of the following form:

$$\bar{U}_t = \Omega^c_t \infty + \Omega_t U_t \; , \quad (99)$$

for given càdlàg event-processes[4] $\Omega_t, \Omega^c_t = 1 - \Omega_t$. The interpretation of (99) is that call is possible whenever $\Omega_t = 1$, otherwise call protection is in force. This is thus a generalization of (16), which corresponds to the special case for which $\Omega_t = \mathbb{1}_{\{t \geq \bar{\sigma}\}}$ in (99).

The identification between the *arbitrage*, or *infimal super-hedging*, $\mathbb{P}$*-price process* of the derivative at hand, and the state-process Π of a solution (Π, M, K), *assumed to exist*, to the generalization of the BSDE (15) with $\bar{U}$ given by (99) therein, can be established by a straightforward adaptation of the arguments developed in Section 2 (See Remark 16.2 in Part IV) .

In the Markovian jump-diffusion model X defined by (28), and assuming

$$\Omega_t = \Omega(t, X_t, N_t) \quad (100)$$

for a suitably extended finite-dimensional Markovian factor process (X_t, N_t) and a related Boolean function Ω of (t, X, N), it is expected that one should then have $\Pi_t = v(t, X_t, N_t)$ on $[0, T]$ for a suitable pricing function v.

[4]In the sense of Boolean-valued processes.

Under suitable technical conditions (including U being given as a Lipschitz function of (t,x)), this is precisely what comes out from the results of Section 16, in case of a call protection *discretely monitored* at the dates of a finite time grid $\mathcal{T} = \{T_0, T_1 \dots, T_m\}$.

As standing examples of such discretely monitored forms of call protection, we shall consider the following clauses, which are commonly found in convertible bonds contracts on an underlying stock S.

Let $S_t = X_t^1$ denote the first component of the vector-process X_t.

Example 4.4 Given a constant trigger level $\bar{S}$ and a constant integer $\imath$:
(i) Call possible whenever $S_t \geq \bar{S}$ at the last $\imath$ monitoring times T_ls, Call protection otherwise,
Or more generally, given a further integer $\jmath \geq \imath$,
(ii) Call possible whenever $S_t \geq \bar{S}$ on at least $\imath$ of the last $\jmath$ monitoring times T_ls, Call protection otherwise.

Let $S = x_1$ denote the first component of the mute vector-variable x, and let $u(T_l-, x)$ be a notation for the formal limit, given a function $u = u(t,y)$,

$$\lim_{(t,y)\to(T_l,x) \text{ with } t\leq T_l} u(t,y) \ . \tag{101}$$

Let finally, in the situation of Example 4.4(ii),

$$|k| = \sum_{1\leq j\leq \jmath} k_j \ , \ k_+ = k_+(k,x) = (\mathbb{1}_{S\geq\bar{S}}, k_1, \dots, k_{\jmath-1}) \ .$$

One thus has by application of the results of Section 16 (cf. in particular (269)–(270)),

Proposition 4.14 ***In the situation of Example 4.4(i)****, the BSDE (15) with $\bar{U}$ given by (99) admits a unique solution (Π, M, K), and one has $\Pi_t = v(t, X_t, N_t)$ on $[0,T]$, for a suitable pricing function $v = v(t,x,k) = v_k(t,x)$ with $k \in \mathbb{N}_\imath$, and where N_t represents the number of consecutive monitoring dates T_ls with $S_{T_l} \geq \bar{S}$ from time t backwards, capped at $\imath$. The restrictions of the v_ks to every set $[T_{l-1}\, T_l) \times [0, +\infty)$ are continuous, and $v_k(T_l-, x)$ as formally defined by (101) exists for every $k \in \mathbb{N}_\imath$, $l \geq 1$ and x in the hyperplane $\{S \neq \bar{S}\}$ of $\mathbb{R}^q$. Moreover v solves the following cascade of variational inequalities,*

For l decreasing from N to 1:
- *At $t = T_l$, for $k \in \mathbb{N}_\imath$,*

$$v_k(T_l-, x) = \begin{cases} v_{k+1}(T_l, x) \ (\textit{or } v_k(T_l,x) \textit{ if } k = \imath) \ \ \textit{on } \{S > \bar{S}\} \times \mathbb{R}^{q-1} \\ v_0(T_l, x) \ (\textit{or } \min(v_0(T_l,x), U(T_l,x)) \textit{ if } k = \imath) \textit{ on } \{S < \bar{S}\} \times \mathbb{R}^{q-1} \end{cases}, \tag{102}$$

Or, in case $l = m$, $v_k(T_l-, x) = \xi(x)$ on $\mathbb{R}^q$,
- *On the time interval $[T_{l-1}, T_l)$,*

$$\max\left(\mathcal{G}v_k + C - rv_k \ , \ L - v_k\right) = 0 \ , \ k = 0 \dots \imath - 1$$
$$\min\left(\max\left(\mathcal{G}v_\imath + C - rv_\imath \ , \ L - v_\imath\right), U - v_\imath\right) = 0 \ .$$

In the situation of Example 4.4(ii)*, the BSDE (15) with* $\bar{U}$ *given by (99) admits a unique solution* (Π, M, K)*, and one has* $\Pi_t = v(t, X_t, N_t)$ *on* $[0, T]$*, for a suitable pricing function* $v = v(t, S, k) = v_k(t, S)$ *with* $k \in \{0, 1\}^\jmath$*, and where* N_t *represents the vector of the indicator functions of the events* $S_{T_l} \geq \bar{S}$ *at the last* $\jmath$ *monitoring dates preceding time* t*. The restrictions of the* v_k*s to every set* $[T_{l-1}\, T_l) \times [0, +\infty)$ *are continuous, and the limit* $v_k(T_l-, x)$ *as defined by (101) exists for every* $k \in \{0, 1\}^\jmath$*,* $l \geq 1$ *and* x *in the hyperplane* $\{S \neq \bar{S}\}$ *of* $\mathbb{R}^q$*. Moreover* v *solves the following cascade of equations:*

For l *decreasing from* N *to 1:*
- *At* $t = T_l$*, for* $k \in \{0, 1\}^\jmath$*,*

$$v_k(T_l-, x) = v_{k_+}(T_l, x) \ \text{(or } \min(v_{k_+}(T_l, x), U(T_l, x)) \text{ if } |k| \geq \imath \text{ and } |k_+| < \imath) \quad \text{on } \{S \neq \bar{S}\}\ , \tag{103}$$

Or, in case $l = m$*,* $v_k(T_l-, x) = \xi(x)$ *on* $\mathbb{R}^q$*,*
- *On the time interval* $[T_{l-1}, T_l)$*, for* $k \in \{0, 1\}^\jmath$*,*

$$\begin{aligned} &\max\left(\mathcal{G}v_k + C - rv_k\ ,\ L - v_k\right) = 0\ ,\ |k| < \imath \\ &\min\left(\max\left(\mathcal{G}v_\imath + C - rv_\imath\ ,\ L - v_\imath\right), U - v_\imath\right) = 0\ ,\ |k| \geq \imath\ . \end{aligned} \tag{104}$$

Comments 4.3 **(i)** Existence of the limits $v_k(T_l-, x)$ in (102) or (103) for x in the hyperplane $\{S \neq \bar{S}\}$ of $\mathbb{R}^q$ follows in view of Remark 16.12.
(ii) Note that (103)-(104) is a cascade of $2^\jmath$ equations, which precludes the practical use of deterministic schemes for solving it numerically as soon as $\jmath$ is greater than a few units. Simulation methods on the opposite can be a fruitful alternative (see [32, 33]).

Moreover, in case the pricing functions v_ks are sufficiently regular for an Itô formula to be applicable, one has further, for $t \in [0, T]$,

$$dM_t = \partial v(t, X_t, N_t)\sigma(t, X_t)dW_t + \delta v(t, X_{t-}, N_{t-})dN_t\ .$$

Part II

Main BSDE Results

As opposed to Part I which was mainly focused on the financial interpretation and use of the results, Parts II to IV will be mainly mathematical.

In this part (see Section 1 for a detailed outline), we construct a rather generic Markovian model *(jump–diffusion with regimes)* $\mathcal{X}$ which gives a precise and rigorous mathematical content to the factor process X underlying a financial derivative in Part I, informally defined by equation (28) therein.

Using the general results of Crépey and Matoussi [42], we then show that related Markovian reflected and doubly reflected BSDEs, covering the ones considered in Part I (see Comments 5.1(v), Definition 6.4 and Section 6.6), are *well-posed*, in the sense that they have unique solutions, which depend continuously on their input data.

This part can thus be seen as a justification of the fact that we were legitimate in assuming well-posedness of the Markovian BSDEs that arose from the derivatives pricing problems considered in Part I.

5 General Set-Up

We first recall the general set-up of [42]. Let us thus be given an initial time conventionally taken in this section as 0, a finite time horizon $T > 0$, a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ with $\mathcal{F}_T = \mathcal{F}$. By default henceforth one considers the right-continuous and completed versions of all filtrations, a *random variable* has to be $\mathcal{F}$-measurable, and a *process* is defined on the time interval $[0, T]$ and $\mathbb{F}$-adapted. All semimartingales are assumed to be càdlàg, without restriction.

Let $B = (B_t)_{t\in[0,T]}$ be a d-dimensional Brownian motion. Given an auxiliary measured space $(E, \mathcal{B}_E, \rho)$, where ρ is a non-negative σ-finite measure on $(E, \mathcal{B}_E)$, let $\mu = (\mu(dt, de))_{t\in[0,T],e\in E}$ be an *integer valued random measure* on $\big([0, T] \times E, \mathcal{B}([0, T]) \otimes \mathcal{B}_E\big)$ (see Jacod–Shiryaev [67, Definition II.1.13 page 68]). Denoting by $\mathcal{P}$ the predictable sigma field on $\Omega \times [0, T]$, we assume that the compensator of μ is defined by $dt \otimes \zeta\rho(de) := \zeta_t(\omega, e)\rho(de)dt$, for a $\mathcal{P} \otimes \mathcal{B}_E$-measurable non-negative bounded random *intensity function* ζ. We refer the reader to the literature [67, 13] regarding the definition of the integral process of $\mathcal{P} \otimes \mathcal{B}_E$-measurable integrands with respect to random measures such as $\mu(dt, de)$ or its compensated form $\widetilde{\mu}(dt, de) = \mu(dt, de) - \zeta_t(\omega, e)\rho(de)dt$. By default, all (in)equalities between random quantities are to be understood $d\mathbb{P}$ – almost surely, $d\mathbb{P} \otimes dt$ – almost everywhere or $d\mathbb{P} \otimes dt \otimes \zeta\rho(de)$ – almost everywhere, as suitable in the situation at hand. For simplicity we omit all dependences in ω of any process or random function in the notation.

We denote by:

- $|X|$, the (d-dimensional) Euclidean norm of a vector or row vector X in $\mathbb{R}^d$ or $\mathbb{R}^{1\otimes d}$;
- $|M|$, the supremum of $|MX|$ over the unit ball of $\mathbb{R}^d$, for M in $\mathbb{R}^{d\otimes d}$;
- $\mathcal{M}_\rho = \mathcal{M}(E, \mathcal{B}_E, \rho; \mathbb{R})$, the set of measurable functions from $(E, \mathcal{B}_E, \rho)$ to $\mathbb{R}$ endowed with

the topology of convergence in measure, and for $v \in \mathcal{M}_\rho$ and $t \in [0,T]$:

$$|v|_t = \Big[\int_E v(e)^2 \zeta_t(e)\rho(de) \Big]^{\frac{1}{2}} \in \mathbb{R}_+ \cup \{+\infty\} \; ; \tag{105}$$

• $\mathcal{B}(\mathcal{O})$, the Borel sigma field on $\mathcal{O}$, for any topological space $\mathcal{O}$.

Let us now introduce some Banach (or Hilbert, in case of $\mathcal{L}^2$, $\mathcal{H}^2_d$ or $\mathcal{H}^2_\mu$) spaces of random variables or processes, where p denotes here and henceforth a real number in $[2,\infty)$:
• $\mathcal{L}^p$, the space of real valued ($\mathcal{F}_T$-measurable) random variables ξ such that

$$\|\xi\|_{\mathcal{L}^p} := \Big(\mathbb{E}\Big[\xi^p \Big] \Big)^{\frac{1}{p}} < +\infty \, ;$$

• $\mathcal{S}^p_d$ (or $\mathcal{S}^p$, in case $d=1$), the space of $\mathbb{R}^d$-valued càdlàg processes X such that

$$\|X\|_{\mathcal{S}^p_d} := \Big(\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t|^p \Big] \Big)^{\frac{1}{p}} < +\infty \, ;$$

• $\mathcal{H}^p_d$ (or $\mathcal{H}^p$, in case $d=1$), the space of $\mathbb{R}^{1\otimes d}$-valued predictable processes Z such that

$$\|Z\|_{\mathcal{H}^p_d} := \Big(\mathbb{E}\Big[\int_0^T |Z_t|^2 \, dt \Big]^{\frac{p}{2}} \Big)^{\frac{1}{p}} < +\infty \; ;$$

• $\mathcal{H}^p_\mu$, the space of $\mathcal{P} \otimes \mathcal{B}_E$-measurable functions $V : \Omega \times [0,T] \times E \to \mathbb{R}$ such that

$$\|V\|_{\mathcal{H}^p_\mu} := \Big(\mathbb{E}\Big[\int_0^T \int_E |V_t(e)|^p \zeta_t(e)\rho(de)dt \Big] \Big)^{\frac{1}{p}} < +\infty \; ,$$

so in particular (cf. (105))

$$\|V\|_{\mathcal{H}^2_\mu} = \Big(\mathbb{E}\Big[\int_0^T |V_t|_t^2 dt \Big] \Big)^{\frac{1}{2}} \; ;$$

• $\mathcal{A}^2$, the space of finite variation continuous processes K with (continuous and non-decreasing) Jordan components $K^\pm \in \mathcal{S}^2$ null at time 0;
• $\mathcal{A}^2_i$, the space of non-decreasing processes in $\mathcal{A}^2$.

By *the Jordan decomposition* of $K \in \mathcal{A}^2$, we mean the unique decomposition $K = K^+ - K^-$ as the difference of two non-decreasing processes $K^\pm$ null at 0 and defining mutually singular random measures on $\mathbb{R}^+$.

Remark 5.1 By a slight abuse of notation we we shall also write $\|X\|_{\mathcal{H}^p}$ for $\Big(\mathbb{E}\Big[\int_0^T X_t^2 \, dt \Big]^{\frac{p}{2}} \Big)^{\frac{1}{p}}$ in the case of merely progressively measurable (not necessarily *predictable*) real-valued processes X.

For the reader's convenience we recall the following well known facts which will be used implicitly throughout (Regarding (ii) see e.g. Bouchard and Elie [27]).

Proposition 5.1
(i) $\int_0^{\cdot} Z_t dB_t$ *and* $\int_0^{\cdot}\int_E V_t(e)\widetilde{\mu}(dt,de)$ *are genuine martingales, for any* $Z \in \mathcal{H}^p_d$ *and* $V \in \mathcal{H}^p_\mu$;
(ii) *Assuming that the jump measure* ρ *is* finite, *then there exist positive constants* c_p *and* C_p *depending only on* p, $\rho(E)$, T *and a bound on* ζ, *such that:*

$$c_p\|V\|_{\mathcal{H}^p_\mu} \le \|\int_0^{\cdot}\int_E V_t(e)\widetilde{\mu}(dt,de)\|_{\mathcal{S}^p_d} \le C_p\|V\|_{\mathcal{H}^p_\mu} \tag{106}$$

for any $V \in \mathcal{H}^p_\mu$. □

5.1 General Reflected and Doubly Reflected BSDEs

Let us now be given a *terminal condition* ξ, and a *driver coefficient* $g : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^{1\otimes d} \times \mathcal{M}_\rho \to \mathbb{R}$, such that:
(**H.0**) $\xi \in \mathcal{L}^2$;
(**H.1.i**) $g_\cdot(y,z,v)$ is a progressively measurable process, and $\|g_\cdot(y,z,v)\|_{\mathcal{H}^2} < \infty$, for any $y \in \mathbb{R}$, $z \in \mathbb{R}^{1\otimes d}$, $v \in \mathcal{M}_\rho$;
(**H.1.ii**) g is uniformly Λ − Lipschitz continuous with respect to (y,z,v), in the sense that Λ is a constant such that for every $t \in [0,T]$, $y,y' \in \mathbb{R}$, $z,z' \in \mathbb{R}^{1\otimes d}$, $v,v' \in \mathcal{M}_\rho$, one has:

$$|g_t(y,z,v) - g_t(y',z',v')| \le \Lambda(|y-y'| + |z-z'| + |v-v'|_t)\ .$$

Remark 5.2 Given the Lipschitz continuity property (H.1.ii) of g, the requirement that

$$\|g_\cdot(y,z,v)\|_{\mathcal{H}^2} < \infty \text{ for any } y \in \mathbb{R},\ z \in \mathbb{R}^{1\otimes d},\ v \in \mathcal{M}_\rho$$

in (H.1.i) reduces of course to $\|g_\cdot(0,0,0)\|_{\mathcal{H}^2} < \infty$.

We also introduce the *barriers* (or *obstacles*) L and U such that:
(**H.2.i**) L and U are càdlàg processes in $\mathcal{S}^2$;
(**H.2.ii**) $L_t \le U_t$, $t \in [0,T)$ and $L_T \le \xi \le U_T$, $\mathbb{P}$-a.s.

Definition 5.3 **(a)** An $(\Omega,\mathbb{F},\mathbb{P}),(B,\mu)$-*solution* $\mathcal{Y}$ *to the doubly reflected backward stochastic differential equation (R2BSDE, for short) with data* (g,ξ,L,U) is a quadruple $\mathcal{Y} = (Y,Z,V,K)$, such that:

(i) $Y \in \mathcal{S}^2, Z \in \mathcal{H}^2_d, V \in \mathcal{H}^2_\mu, K \in \mathcal{A}^2$,
(ii) $$Y_t = \xi + \int_t^T g_s(Y_s,Z_s,V_s)ds + K_T - K_t - \int_t^T Z_s dB_s - \int_t^T\int_E V_s(e)\widetilde{\mu}(ds,de) \ \text{ for any } \ t \in [0,T],\ \mathbb{P}\text{-a.s.}$$
(iii) $L_t \le Y_t \le U_t$ for any $t \in [0,T]$, $\mathbb{P}$-a.s., and $\int_0^T (Y_t - L_t)dK^+_t = \int_0^T (U_t - Y_t)dK^-_t = 0$, $\mathbb{P}$-a.s.

(b) An $(\Omega, \mathbb{F}, \mathbb{P}), (B, \mu)$-*solution* $\mathcal{Y}$ *to the reflected BSDE (RBSDE, for short) with data* (g, ξ, L) is a quadruple $\mathcal{Y} = (Y, Z, V, K)$ such that:

(i) $Y \in \mathcal{S}^2, Z \in \mathcal{H}^2_d, V \in \mathcal{H}^2_\mu, K \in \mathcal{A}^2_i$

(ii) $$Y_t = \xi + \int_t^T g_s(Y_s, Z_s, V_s)ds + K_T - K_t - \int_t^T Z_s dB_s - \int_t^T \int_E V_s(e)\widetilde{\mu}(ds, de) \text{ for any } t \in [0, T], \ \mathbb{P}\text{-a.s.}$$

(iii) $L_t \le Y_t$ for any $t \in [0, T]$, $\mathbb{P}$-a.s., and $\int_0^T (Y_t - L_t)dK_t = 0$, $\mathbb{P}$-a.s.

(c) When there is no barrier, we define likewise solutions to the BSDE with data (g, ξ).

5.1.1 Extensions with Stopping Times

Motivated by applications (cf. Part I), we now consider two variants of the above problems involving a further $[0, T]$-valued stopping time τ. Note that $(\mathbb{1}_{\cdot \le \tau} g, \xi, L_{\cdot \wedge \tau}, , U_{\cdot \wedge \tau})$ satisfies (H.0), (H.1) and (H.2), like (g, ξ, L, U). One can thus state the following

Definition 5.4 Assuming that ξ is $\mathcal{F}_\tau$-measurable,
(i) A solution to the *stopped R2BSDE with data* (g, ξ, L, U, τ) is a quadruple (Y, Z, V, K) which solves the R2BSDE with data $(\mathbb{1}_{\cdot \le \tau} g, \xi, L_{\cdot \wedge \tau}, U_{\cdot \wedge \tau})$, and such that $Y = Y_\tau$, $K = K_\tau$ and $Z = V = 0$ on $[\tau, T]$.
A solution to the *stopped RBSDE with data* (g, ξ, L, τ) is a quadruple (Y, Z, V, K) which solves the RBSDE with data $(\mathbb{1}_{\cdot \le \tau} g, \xi, L_{\cdot \wedge \tau})$, and such that $Y = Y_\tau$, $K = K_\tau$ and $Z = V = 0$ on $[\tau, T]$.
(ii) The *RDBSDE with data* (g, ξ, L, U, τ) (where 'D' stands for 'delayed') is the generalization of an R2BSDE in which the upper barrier U is inactive before τ. Formally, we replace U by

$$\bar{U}_t := \mathbb{1}_{\{t<\tau\}}\infty + \mathbb{1}_{\{t \ge \tau\}} U_t \tag{107}$$

in Definition 5.3(a)(iii), with the convention that $0 \times \pm\infty = 0$.

Comments 5.1 **(i)** All these definitions admit obvious extensions to problems in which the driving term contains a further finite variation process A (not necessarily absolutely continuous).
(ii) In [42], reflected BSDEs stopped at a random time were introduced and presented as reflected BSDEs with random terminal time as of Darling and Pardoux [44] (only defined over the time interval $[0, \tau]$). Such (possibly doubly) reflected BSDEs stopped at a random time and the above stopped R(2)BSDEs are in fact equivalent notions. We refer the reader to [42] for preliminary general results on stopped RBSDEs and on RDBSDEs.
(iii) In the special case when $\tau = 0$, resp. $\tau = T$, then the RDBSDE with data (g, ξ, L, U, τ) reduces to the R2BSDE with data (g, ξ, L, U), resp. to the RBSDE with data (g, ξ, L).
(iv) If (Y, Z, V, K) is a solution to the RDBSDE with data (g, ξ, L, U, τ), then the process

$$(Y_{\cdot \wedge \tau}, \mathbb{1}_{\cdot \le \tau} Z, \mathbb{1}_{\cdot \le \tau} V, K_{\cdot \wedge \tau})$$

is a solution to the stopped RBSDE with data $(g, Y_{\cdot\wedge\tau}, L, \tau)$.
(v) The problem we are most interested in from the point of view of applications consists of the general RDBSDE (cf. equations (15), (69) or (85) in Part I). However it will come out from the results of this part (Theorem 8.4; see also [42]) that the solution of a RDBSDE is essentially given as the solution of a stopped RBSDE before τ, appropriately pasted at τ with the solution of a (standard) R2BSDE after τ. So the results of this part effectively reduce the study of RDBSDEs to those of RBSDEs and R2BSDEs. In Part III of this paper we shall not deal explicitly with RDBSDEs. Yet, given the results of this part, the results of Part III are applicable to RDBSDEs, giving a way to compute their solutions in two pieces, before and after τ (cf. the related *cascades of two PDEs* in Part I).
(vi) In Section 16 in Part IV we shall consider doubly reflected BSDEs with intermittent upper barrier, or RIBSDEs, generalizing RDBSDEs to an effective upper barrier $\bar{U}$ of the form (to be compared with (107))

$$\bar{U}_t = \Omega_t^c \infty + \Omega_t U_t \ , \tag{108}$$

for a larger class of càdlàg event-processes[5] Ω_t, $\Omega_t^c = 1 - \Omega_t$.

5.1.2 Verification Principle

Originally, R2BSDEs have been developed in connection with *Dynkin games*, or optimal stopping game problems (see, e.g., Lepeltier and Maingueneau [82], Cvitanić and Karatzas [43]). Given a $[0,T]$-valued stopping time θ, let $\mathcal{T}_\theta$ (or simply $\mathcal{T}$, in case $\theta = 0$) denote the set of $[\theta, T]$-valued stopping times. We thus have the following *Verification Principle*, which was used in the proof of Proposition 2.2 in Part I. We state it for an RDBSDE as of Definition 5.4(ii), which in view of Comments 5.1(iii) covers RBSDEs and R2BSDEs as special cases. Note that in the case of RBSDEs (special case where $\tau = T$) the related Dynkin game reduces to an optimal stopping problem.

Proposition 5.2 (Verification Principle) *If $\mathcal{Y} = (Y, Z, V, K)$ solves the RDBSDE with data (g, ξ, L, U, τ), then the state process Y is the conditional value process of the Dynkin game with payoff functional given by, for any $t \in [0,T]$ and $(\rho, \theta) \in \mathcal{T}_\tau \times \mathcal{T}_t$:*

$$J(t;\rho,\theta) = \int_t^{\rho\wedge\theta} g_s(Y_s, Z_s, V_s)ds + L_\theta \mathbb{1}_{\{\rho\wedge\theta=\theta<T\}} + U_\rho \mathbb{1}_{\{\rho<\theta\}} + \xi \mathbb{1}_{[\rho\wedge\theta=T]} \ .$$

More precisely, a saddle-point of the game at time t is given by:

$$\rho_t = \inf\Big\{ s \in [t \vee \tau, T] \,;\, Y_s = U_s \Big\} \wedge T \ , \ \theta_t = \inf\Big\{ s \in [t, T] \,;\, Y_s = L_s \Big\} \wedge T \ .$$

So, for any $t \in [0,T]$:

$$\mathbb{E}\big[J(t;\rho_t,\theta)\big|\mathcal{F}_t\big] \ \le \ Y_t \ = \ \mathbb{E}\big[J(t;\rho_t,\theta_t)\big|\mathcal{F}_t\big] \ \le \ \mathbb{E}\big[J(t;\rho,\theta_t)\big|\mathcal{F}_t\big] \ \text{for any } (\rho,\theta) \in \mathcal{T}_\tau \times \mathcal{T}_t \tag{109}$$

Proof. Except for the presence of τ, the result is standard (see, e.g., Lepeltier and Maingueneau [82]; or see also Bielecki et al. [16] for a proof of an analogous result in a context of

[5] Boolean-valued processes.

mathematical finance). We nevertheless give a self-contained proof for the reader's convenience.
The result of course reduces to showing (109). Let us first check that the right-hand side inequality in (109) is valid for any $\rho \in \mathcal{T}_\tau$. Let θ denote $\theta_t \wedge \rho$. By definition of θ_t, we see that K^+ equals 0 on $[t,\theta]$. Since K^- is non-decreasing, taking conditional expectations in the RDBSDE, and using also the facts that $Y_{\theta_t} \le L_{\theta_t}$ if $\theta_t < T$, $Y_\rho \le U_\rho$ if $\rho < T$ (recall that $\rho \in \mathcal{T}_\tau$, so that $\rho \ge \tau$ and $\bar{U}_\rho = U_\rho$), and $Y_T = \xi$, we obtain:

$$\begin{aligned} Y_t &\le \mathbb{E}\Big(\int_t^\theta g_s(Y_s,Z_s,V_s)ds + Y_\theta \,|\, \mathcal{F}_t\Big) \\ &\le \mathbb{E}\Big(\int_t^\theta g_s(Y_s,Z_s,V_s)ds + \big(\mathbb{1}_{\{\theta=\theta_t<T\}}L_{\theta_t} + \mathbb{1}_{\{\rho<\theta_t\}}U_\rho + \mathbb{1}_{\{\theta=T\}}\xi\big)\,\Big|\, \mathcal{F}_t\Big). \end{aligned}$$

We conclude that $Y_t \le \mathbb{E}\big(J(t;\theta_t,\rho)\,\big|\,\mathcal{F}_t\big)$ for any $\rho \in \mathcal{T}_\tau$. This completes the proof of the right-hand side inequality in (109). The left-hand side inequality can be shown similarly. It is in fact standard, since it does not involve τ, and thus we leave the details to the reader.
□

Remark 5.5 For general well-posedness (in the sense of existence, uniqueness and a priori estimates) and comparison results on the different variants of reflected BSDEs (specifically: RBSDEs, R2BSDEs and RDBSDEs) above, we refer the reader to Crépey and Matoussi [42]. We do not reproduce explicitly these results here, since we will state in Section 16.2 extensions of these results to more general RIBSDEs (see Comment 5.1(vi)).

5.2 General Forward SDE

To conclude this section we consider the (forward) stochastic differential equation

$$d\widetilde{X}_s = \widetilde{b}_s(\widetilde{X}_s)ds + \widetilde{\sigma}_s(\widetilde{X}_s)\,dB_s + \int_E \widetilde{\delta}_s(\widetilde{X}_s,e)\,\zeta_s(e)\widetilde{\mu}(ds,de)\;, \tag{110}$$

where $\widetilde{b}_s(x)$, $\widetilde{\sigma}_s(x)$ and $\widetilde{\delta}_s(x,e)$ are d-dimensional *drift* vector, *dispersion* matrix and *jump size* vector **random** coefficients such that:
- $\widetilde{b}_s(x)$, $\widetilde{\sigma}_s(x)$ and $\widetilde{\delta}_s(x,e)$ are Lipschitz continuous in x uniformly in $s \ge 0$ and $e \in E$;
- $\widetilde{b}_s(0)$, $\widetilde{\sigma}_s(0)$ and $\widetilde{\delta}_s(0,e)$ are bounded in $s \ge 0$ and $e \in E$.

The following proposition can be shown by standard applications of Burkholder's inequality used in conjunction with (106) and Gronwall's lemma (see for instance Fujiwara–Kunita [59, Lemma 2.1 page 84] for analogous results with proofs).

Proposition 5.3 *Assuming that the jump measure ρ is* finite*, then for any strong solution $\widetilde{X}$ to the stochastic differential equation (110) with initial condition $\widetilde{X}_0 \in \mathcal{F}_0 \cap \mathcal{L}^p$, the following bound and error estimates are available:*

$$\|\widetilde{X}\|^p_{\mathcal{S}^p_d} \le C_p\mathbb{E}\Big[|\widetilde{X}_0|^p + \int_0^T |\widetilde{b}_s(0)|^p ds + \int_0^T |\widetilde{\sigma}_s(0)|^p ds + \int_0^T\int_E |\widetilde{\delta}_s(0,e)|^p \zeta_s(e)\rho(de)ds\Big] \tag{111}$$

$$\begin{aligned} \|\widetilde{X}-\widetilde{X}'\|^p_{\mathcal{S}^p_d} \le C_p\mathbb{E}\Big[&|\widetilde{X}_0-\widetilde{X}'_0|^p + \int_0^T |\widetilde{b}_s(\widetilde{X}_s)-\widetilde{b}'_s(\widetilde{X}_s)|^p ds + \int_0^T |\widetilde{\sigma}_s(\widetilde{X}_s)-\widetilde{\sigma}'_s(\widetilde{X}_s)|^p ds + \\ &\int_0^T\int_E |\widetilde{\delta}_s(\widetilde{X}_s,e)-\widetilde{\delta}'_s(\widetilde{X}_s,e)|^p \zeta_s(e)\rho(de)ds\Big] \end{aligned} \tag{112}$$

where, in (112), $\widetilde{X}'$ is the solution of a stochastic differential equation of the form (110) with coefficients $\widetilde{b}'$, $\widetilde{\sigma}'$, $\widetilde{\delta}'$ and initial condition $\widetilde{X}'_0 \in \mathcal{F}_0 \cap \mathcal{L}^p$. □

6 A Markovian Decoupled Forward Backward SDE

We now present a versatile Markovian specification of the general set-up of the previous section. This model was already considered and used in applications in [17, 19, 42], but the construction of the model has been deferred to the present work.

6.1 Infinitesimal Generator

Given integers d and k, we define the following linear operator $\mathcal{G}$ acting on regular functions $u = u^i(t,x)$ for $(t,x,i) \in \mathcal{E} = [0,T] \times \mathbb{R}^d \times I$ with $I = \{1,\ldots,k\}$, and where ∂u (resp. $\mathcal{H}u$) denotes the row-gradient (resp. Hessian) of $u(t,x,i) = u^i(t,x)$ with respect to x :

$$\begin{aligned}\mathcal{G}u^i(t,x) &= \partial_t u^i(t,x) + \frac{1}{2}\mathrm{Tr}[a^i(t,x)\mathcal{H}u^i(t,x)] + \partial u^i(t,x)\widetilde{b}^i(t,x) \\ &+ \int_{\mathbb{R}^d} \Big(u^i(t,x+\delta^i(t,x,y)) - u^i(t,x)\Big) f^i(t,x,y)m(dy) \\ &+ \sum_{j\in I} n^{i,j}(t,x)\Big(u^j(t,x) - u^i(t,x)\Big)\end{aligned} \tag{113}$$

with

$$\widetilde{b}^i(t,x) = b^i(t,x) - \int_{\mathbb{R}^d} \delta^i(t,x,y) f^i(t,x,y)m(dy)\ . \tag{114}$$

Assumption 6.1 In (113)–(114), $m(dy)$ is a *finite jump measure* (not charging the origin 0_d in $\mathbb{R}^d$), and all the coefficients are Borel–measurable functions such that:
• the $a^i(t,x)$ are d-dimensional *covariance* matrices, with $a^i(t,x) = \sigma^i(t,x)\sigma^i(t,x)^{\mathsf{T}}$, for some d-dimensional *dispersion* matrices $\sigma^i(t,x)$;
• the $b^i(t,x)$ are d-dimensional *drift* vector coefficients;
• the *jump intensity* functions $f^i(t,x,y)$ are bounded, and the *jump size* functions $\delta^i(t,x,y)$ are bounded with respect to y at fixed (t,x), locally uniformly in (t,x)[6];
• the $n^{i,j}(t,x)_{i,j\in I}$ are *regime switching intensities* such that the functions $n^{i,j}(t,x)$ are non-negative and bounded for $i \neq j$, and $n^{i,i}(t,x) = 0$.

We shall often find convenient to denote $v(t,x,i,\ldots)$ rather than $v^i(t,x,\ldots)$ for a function v of $(t,x,i,\ldots)$, and $n(t,x,i,j)$, for $n^{i,j}(t,x)$. For instance, the notation $f(t,X_t,N_t,y)$, or $f(t,\mathcal{X}_t,y)$ with $\mathcal{X}_t = (X_t,N_t)$ below, will typically be used rather than $f^{N_t}(t,X_t,y)$. Also note that a function u on $[0,T]\times\mathbb{R}^d\times I$ is equivalently referred to in this paper as a *system* $u = (u^i)_{i\in I}$ of functions $u^i = u^i(t,x)$ on $[0,T]\times\mathbb{R}^d$.

[6] In the sense that the bound with respect to y may be chosen uniformly as (t,x) varies in a compact set.

6.2 Model Dynamics

Definition 6.2 A *model with generator* $\mathcal{G}$ *and initial condition* (t,x,i) is a triple

$$(\Omega, \mathbb{F}^t, \mathbb{P}^t),\ (B^t, \chi^t, \nu^t),\ \mathcal{X}^t = (X^t, N^t)\,,$$

where the superscript ^{t} stands in reference to the initial condition $(t,x,i) \in \mathcal{E}$, such that $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$ is a stochastic basis on $[t,T]$, relative to which the following processes and random measures are defined:
(i) A d-dimensional standard Brownian motion B^t starting at t, and integer-valued random measures χ^t on $[t,T]\times\mathbb{R}^d$ and ν^t on $[t,T]\times I$, such that χ^t and ν^t cannot jump together at stopping times;
(ii) An $\mathbb{R}^d \times I$-valued process $\mathcal{X}^t = (X^t, N^t)$ on $[t,T]$ with initial condition (x,i) at t and such that for $s \in [t,T]$:

$$\left\{\begin{array}{ll} dN_s^t & = \sum_{j\in I} (j - N_{s-}^t)\, d\nu_s^t(j) \\ dX_s^t & = b(s,\mathcal{X}_s^t)ds + \sigma(s,\mathcal{X}_s^t)\, dB_s^t + \int_{\mathbb{R}^d} \delta(s,\mathcal{X}_{s-}^t,,y)\, \widetilde{\chi}^t(ds,dy) \end{array}\right. \tag{115}$$

(thus in particular $\nu_s^t(j)$ counts the number of transitions of N^t to state j between times t and s), and the $\mathbb{P}^t$-compensatrices $\widetilde{\nu}^t$ and $\widetilde{\chi}^t$ of ν^t and χ^t are such that

$$\left\{\begin{array}{l} d\widetilde{\nu}_s^t(j) = d\nu_s^t(j) - n(s,\mathcal{X}_s^t,j)\, ds \\ \widetilde{\chi}^t(ds,dy) = \chi^t(ds,dy) - f(s,\mathcal{X}_s^t,y)m(dy)ds \end{array}\right. \tag{116}$$

with $n(s,\mathcal{X}_s^t,j) = n^{N_s^t,j}(s,X_s^t)$, $f(s,\mathcal{X}_s^t,y) = f^{N_s^t}(s,X_s^t,y)$.

By an application of Jacod [66, Theorem 3.89 page 109], the following variant of the *Itô formula* holds (cf. formula (35) in Part I).

Proposition 6.1 *Given a model* $(\Omega,\mathbb{F}^t,\mathbb{P}^t)$, (B^t,χ^t,ν^t), $\mathcal{X}^t = (X^t,N^t)$ *with generator* $\mathcal{G}$, *one has for any system* $u = (u^i)_{i\in I}$ *of functions* $u^i = u^i(t,x)$ *of class* $\mathcal{C}^{1,2}$ *on* $[0,T]\times\mathbb{R}^d$, *for* $s \in [t,T]$,

$$\begin{aligned} du(s,\mathcal{X}_s^t) &= \mathcal{G}u(s,\mathcal{X}_s^t)ds + (\partial u\sigma)(s,\mathcal{X}_s^t)dB_s^t \\ &\quad + \int_{y\in\mathbb{R}^d} \big(u(s,X_{s-}^t + \delta(s,\mathcal{X}_{s-}^t,y),N_{s-}^t) - u(s,\mathcal{X}_{s-}^t)\big)\widetilde{\chi}^t(ds,dy) \\ &\quad + \sum_{j\in I} (u(s,X_{s-}^t,j) - u(s,\mathcal{X}_{s-}^t))d\widetilde{\nu}_s^t(j)\,. \end{aligned} \tag{117}$$

In particular $(\Omega,\mathbb{F}^t,\mathbb{P}^t,\mathcal{X}^t)$ is a solution to the *time-dependent local martingale problem* with generator $\mathcal{G}$ and initial condition (t,x,i) (see Ethier–Kurtz [53, sections 7.A and 7.B]).

Comments 6.1 **(i)** If we suppose that the coefficients b,σ,δ and f do not depend on i, then X is a 'standard' jump-diffusion. Alternatively, if n does not depend on x, then N is an inhomogeneous continuous time Markov chain with finite state space I. In general the above model defines a rather generic class of Markovian factor processes $\mathcal{X} = (X,N)$, in the form of an N-modulated jump-diffusion component X and of an X-modulated I-valued component N. The pure jump process N may be interpreted as defining the so-called *regime*

of the coefficients b, σ, δ and f, whence the name of *jump–diffusion with regimes* for this model.
For simplicity we do not consider the 'infinite activity' case of an infinite jump measure m. Note however that our approach could be extended to Lévy jump measures without major changes if wished (see in this respect Section 3.3.2 in Part I). Yet this would be at the cost of a significantly heavier formalism, regarding in particular the viscosity solutions approach of Part III (see the seminal paper by Barles et al. [6], complemented by Barles and Imbert [7]).
(ii) The general construction of such a model with mutual dependence between N and X, is a non-trivial issue. It will be treated in detail in Section 7, resorting to a suitable *Markovian change of probability* approach. It should be noted that more specific sub-cases or related models were frequently considered in the literature. So (see also Section 6.6 for more comments about financial applications of this model):
• Barles et al. [6] consider jumps in X without regimes N, for a Lévy jump measure m (cf. point (i) above);
• Pardoux et al. [87] consider a diffusion model with regimes, which corresponds to the special case of our model in which f is equal to 0, and the regimes are driven by a standard Poisson process with constant intensity (instead of a family of independent Poisson processes with intensities $n^{l,j}$ in our case, cf. Remark 7.3);
• Becherer and Schweizer consider in [10] a diffusion model with regimes which corresponds to the special case of our model in which f is equal to 0.

6.3 Mapping with the General Set-Up

The model $\mathcal{X}^t = (X^t, N^t)$ is thus a rather generic Markovian specification of the general set-up of Section 5, with (note that the initial time is t here instead of 0 therein; superscripts ^{t} are therefore added below to the notation of Section 5 where need be):

• E, the subset $\big(\mathbb{R}^d \times \{0\}\big) \cup \big(\{0_d\} \times I\big)$ of $\mathbb{R}^{d+1}$;

• $\mathcal{B}_E$, the sigma field generated by $\mathcal{B}(\mathbb{R}^d) \times \{0\}$ and $\{0_d\} \times \mathcal{B}_I$ on E, where $\mathcal{B}(\mathbb{R}^d)$ and $\mathcal{B}_I$ stand for the Borel sigma field on $\mathbb{R}^d$ and the sigma field of all parts of I, respectively;

• $\rho(de)$ and $\zeta^t_s(e)$ respectively given by, for any $e = (y, j) \in E$ and $s \in [t, T]$:

$$\rho(de) = \begin{cases} m(dy) & \text{if} \quad j = 0 \\ 1 & \text{if} \quad y = 0_d \end{cases} , \quad \zeta^t_s(e) = \begin{cases} f(t, \mathcal{X}^t_s, y) & \text{if} \quad j = 0 \\ n(t, \mathcal{X}^t_s, j) & \text{if} \quad y = 0_d \ ; \end{cases}$$

• μ^t, the integer-valued random measure on $\big([t, T] \times E, \mathcal{B}([t, T]) \otimes \mathcal{B}_E\big)$ counting the jumps of X of size $y \in A$ and the jumps of N to state j between t and s, for any $s \geq t$, $A \in \mathcal{B}(\mathbb{R}^d)$, $j \in I$.

We denote for short:

$$(E, \mathcal{B}_E, \rho) = (\mathbb{R}^d \oplus I, \mathcal{B}(\mathbb{R}^d) \oplus \mathcal{B}_I, m(dy) \oplus \mathbb{1}) ,$$

and $\mu^t = \chi^t \oplus \nu^t$ on $\big([t, T] \times E, \mathcal{B}([t, T]) \otimes \mathcal{B}_E\big)$. So the compensator of the random measure μ^t is given by, for any $s \geq t$, $A \in \mathcal{B}(\mathbb{R}^d)$, $j \in I$, with $A \oplus \{j\} := \big(A \times \{0\}\big) \cup \big(\{0_d\} \times \{j\}\big)$:

$$\int_t^s \int_{A \oplus \{j\}} \zeta^t_r(e)\rho(de)dr = \int_t^s \int_A f(r, \mathcal{X}^t_r, y)m(dy)dr + \int_t^s n(r, \mathcal{X}^t_r, j)\, dr \ .$$

Note that $\mathcal{H}^2_{\mu^t}$ can be identified with the product space $\mathcal{H}^2_{\chi^t} \times \mathcal{H}^2_{\nu^t}$, and that $\mathcal{M}_\rho = \mathcal{M}(E, \mathcal{B}_E, \rho; \mathbb{R})$ can be identified with the product space $\mathcal{M}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m(dy); \mathbb{R}) \times \mathbb{R}^k$. These identifications will be used freely in the sequel. Let $\widetilde{v}$ denote a generic pair $(v, w) \in \mathcal{M}_\rho \equiv \mathcal{M}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m(dy); \mathbb{R}) \times \mathbb{R}^k$. We denote accordingly, for $s \geq t$ (cf. (105)):

$$|\widetilde{v}|_s^2 = \textstyle\int_{\mathbb{R}^d} v(y)^2 f(s, \mathcal{X}_s^t, y) m(dy) + \sum_{j \in I} w(j)^2 n(s, \mathcal{X}_s^t, j) \tag{118}$$

(with the slight abuse of notation that $|\widetilde{v}|_s$ implicitly depends on t, x, i in (118)).

Remark 6.3 Of course ultimately the related semi-group and Markov properties will be formally established (see in particular Proposition 8.3, 8.6 and 9.2 as well as Theorems 9.1 and 9.3), so that in applications one shall be able to restrict attention to a 'single' process $\mathcal{X}$, corresponding in practice to the 'true' initial condition (t, x, i) of interest. In the context of pricing in finance this 'true initial condition of interest' corresponds to values of the model parameters calibrated to the current market data, see Part I (cf. also the last section of [42] in which some of the results of this part were announced without proof). Yet at the stage of *deriving* these results in the present paper, it is necessary to consider *families* of processes $\mathcal{X}^t$ parameterized by their initial condition $(t, x, i) \in \mathcal{E}$.

6.4 Cost Functionals

We denote by $\mathcal{P}_q$ the class of functions u on $\mathcal{E}$ such that u^i is Borel-measurable with polynomial growth of exponent $q \geq 0$ in x, for any $i \in I$. Here by *polynomial growth of exponent p in x* we mean the existence of a constant C (which may depend on u) such that for any $(t, x, i) \in \mathcal{E}$:

$$\left|u^i(t, x)\right| \leq C(1 + |x|^q) .$$

Let also $\mathcal{P} = \cup \mathcal{P}_q$ denote the class of functions u on $\mathcal{E}$ such that u^i is Borel-measurable with polynomial growth in x for any $i \in I$.

Let us further be given a system $\mathcal{C}$ of real-valued continuous *cost functions*, namely a *running cost function* $g^i(t, x, u, z, r)$ (where $(u, z, r) \in \mathbb{R}^k \times \mathbb{R}^{1 \otimes d} \times \mathbb{R}$), a *terminal cost function* $\Phi^i(x)$, and *lower and upper cost functions* $\ell^i(t, x)$ and $h^i(t, x)$, such that:
(**M.0**) Φ lies in $\mathcal{P}_q$;
(**M.1.i**) $(t, x, i) \mapsto g^i(t, x, u, z, r)$ lies in $\mathcal{P}_q$, for any $(u, z, r) \in \mathbb{R}^k \times \mathbb{R}^{1 \otimes d} \times \mathbb{R}$;
(**M.1.ii**) g is uniformly Λ – Lipschitz continuous with respect to (u, z, r), in the sense that Λ is a constant such that for every $(t, x, i) \in \mathcal{E}$ and $(u, z, r), (u', z', r') \in \mathbb{R}^k \times \mathbb{R}^{1 \otimes d} \times \mathbb{R}$:

$$\left|g^i(t, x, u, z, r) - g^i(t, x, u', z', r')\right| \leq \Lambda \left(\left|u - u'\right| + \left|z - z'\right| + \left|r - r'\right| \right) ;$$

(**M.1.iii**) g is non-decreasing with respect to r ;
(**M.2.i**) ℓ and h lie in $\mathcal{P}_q$;
(**M.2.ii**) $\ell \leq h$, $\ell(T, \cdot) \leq \Phi \leq h(T, \cdot)$.

Fixing an initial condition $(t, x, i) \in \mathcal{E}$ for $\mathcal{X} = (X, N)$, we define for any $(s, y, z, \widetilde{v}) \in [t, T] \times \mathbb{R} \times \mathbb{R}^{1 \otimes d} \times \mathcal{M}_\rho$, with $\widetilde{v} = (v, w) \in \mathcal{M}_\rho \equiv \mathcal{M}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m(dy); \mathbb{R}) \times \mathbb{R}^k$:

$$\widetilde{g}(s, \mathcal{X}_s^t, y, z, \widetilde{v}) = g(s, \mathcal{X}_s^t, \widetilde{u}_s^t, z, \widetilde{r}_s^t) - \sum_{j \in I} w_j n(s, \mathcal{X}_s^t, j) , \tag{119}$$

where $\widetilde{u}^t_s = \widetilde{u}^t_s(y,w)$ and $\widetilde{r}^t_s = \widetilde{r}^t_s(v)$ are defined by

$$(\widetilde{u}^t_s)^j = \left\{ \begin{array}{ll} y, & j = N^t_s \\ y + w_j, & j \neq N^t_s \end{array} \right. , \; \widetilde{r}^t_s = \int_{\mathbb{R}^d} v(y) f(s, \mathcal{X}^t_s, y) m(dy) \; . \tag{120}$$

Given the previous ingredients, we now define the main decoupled Forward Backward stochastic differential equation (FBSDE, for short) in this work, encapsulating all the SDEs and BSDEs of interest for us in this article. Recall that $\widetilde{g}$ is defined by (119) and that $\widetilde{v}$ denotes a generic pair $(v,w) \in \mathcal{M}_\rho$.

Definition 6.4 **(a)** A *solution to the Markovian decoupled Forward Backward stochastic differential equation with data* $\mathcal{G}$, $\mathcal{C}$ *and* τ is a parameterized family of triples $\mathcal{Z}^t = (\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, χ^t, ν^t), $(\mathcal{X}^t, \mathcal{Y}^t, \bar{\mathcal{Y}}^t)$, where the superscript ^{t} stands in reference to the *initial condition* $(t,x,i) \in \mathcal{E}$, such that:
(i) $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, χ^t, ν^t), $\mathcal{X}^t = (X^t, N^t)$ is a model with generator $\mathcal{G}$ and initial condition (t,x,i);
(ii) $\mathcal{Y}^t = (Y^t, Z^t, \mathcal{V}^t, K^t)$, with $\mathcal{V}^t = (V^t, W^t) \in \mathcal{H}^2_{\mu^t} = \mathcal{H}^2_{\chi^t} \times \mathcal{H}^2_{\nu^t}$, is an $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t) – solution to the R2BSDE on $[t,T]$ with data

$$\widetilde{g}(s, \mathcal{X}^t_s, y, z, \widetilde{v}) \; , \; \Phi(\mathcal{X}^t_T) \; , \; \ell(s, \mathcal{X}^t_s) \; , \; h(s, \mathcal{X}^t_s) \; ; \tag{121}$$

(iii) $\bar{\mathcal{Y}}^t = (\bar{Y}^t, \bar{Z}^t, \bar{\mathcal{V}}^t, \bar{K}^t)$, with $\bar{\mathcal{V}}^t = (\bar{V}^t, \bar{W}^t) \in \mathcal{H}^2_{\mu^t} = \mathcal{H}^2_{\chi^t} \times \mathcal{H}^2_{\nu^t}$, is an $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t) – solution to the stopped RBSDE on $[t,T]$ with data

$$\widetilde{g}(s, \mathcal{X}^t_s, y, z, \widetilde{v}) \; , \; Y^t_{\tau^t} \; , \; \ell(s, \mathcal{X}^t_s) \; , \; \tau^t \tag{122}$$

where Y^t is the state-process of $\mathcal{Y}^t$ in (ii).

(b) The solution is said to be *consistent*, if:
(i) $Y^t_t =: u^i(t,x)$ defines as (t,x,i) varies in $\mathcal{E}$, a continuous *value function* of class $\mathcal{P}$ on $\mathcal{E}$, and one has for every $t \in [0,T]$, $\mathbb{P}^t$-a.s.:

$$Y^t_s = u(s, \mathcal{X}^t_s) \; , \; s \in [t,T] \tag{123}$$

$$\text{For any } j \in I: \; W^t_s(j) = u^j(s, X^t_{s-}) - u(s, \mathcal{X}^t_{s-}) \; , \; s \in [t,T] \tag{124}$$

$$\begin{aligned} \int_t^s \widetilde{g}(\zeta, \mathcal{X}^t_\zeta, Y^t_\zeta, Z^t_\zeta, \mathcal{V}^t_\zeta) d\zeta = \int_t^s \Big[& g(\zeta, \mathcal{X}^t_\zeta, u(\zeta, X^t_\zeta), Z^t_\zeta, \widetilde{r}^t_\zeta) \\ & - \textstyle\sum_{j \in I} n(\zeta, \mathcal{X}^t_\zeta, j) \big(u^j(\zeta, X^t_\zeta) - u(\zeta, \mathcal{X}^t_\zeta) \big) \Big] d\zeta \; , \; s \in [t,T] \end{aligned} \tag{125}$$

with in (125):

$$u(\zeta, X^t_\zeta) := (u^j(\zeta, X^t_\zeta))_{j \in I} \; , \; \widetilde{r}^t_\zeta = \int_{\mathbb{R}^d} V_\zeta(y) f(\zeta, \mathcal{X}^t_\zeta, y) m(dy)$$

(cf. (120));
(ii) $\bar{Y}^t_t =: v^i(t,x)$ defines as (t,x,i) varies in $\mathcal{E}$, a continuous *value function* of class $\mathcal{P}$ on $\mathcal{E}$, and one has for every $t \in [0,T]$, $\mathbb{P}^t$-a.s.:

$$\bar{Y}^t_s = v(s, \mathcal{X}^t_s) \; , \; s \in [t, \tau^t] \tag{126}$$

$$\text{For any } j \in I: \; \bar{W}^t_s(j) = v^j(s, X^t_{s-}) - v(s, \mathcal{X}^t_{s-}) \;\; , \; s \in [t, \tau^t] \tag{127}$$

$$\begin{aligned} \int_t^s \widetilde{g}(\zeta, \mathcal{X}^t_\zeta, \bar{Y}^t_\zeta, \bar{Z}^t_\zeta, \bar{\mathcal{V}}^t_\zeta) d\zeta = \int_t^s \Big[& g(\zeta, \mathcal{X}^t_\zeta, v(\zeta, X^t_\zeta), \bar{Z}^t_\zeta, \bar{r}^t_\zeta) \\ & - \textstyle\sum_{j \in I} n(\zeta, \mathcal{X}^t_\zeta, j) \big(v^j(\zeta, X^t_{\zeta-}) - v(\zeta, \mathcal{X}^t_{\zeta-}) \big) \Big] d\zeta \; , \; s \in [t, \tau^t] \end{aligned} \tag{128}$$

with in (128):

$$v(\zeta, X^t_\zeta) := (v^j(\zeta, X^t_\zeta))_{j\in I} \ , \ \bar{r}^t_\zeta := \tilde{r}^t_\zeta(\bar{V}^t_\zeta) = \int_{\mathbb{R}^d} \bar{V}^t_\zeta(y) f(\zeta, \mathcal{X}^t_\zeta, y) m(dy) \tag{129}$$

(cf. (120)).

6.5 Markovian Verification Principle

The following proposition is a Markovian counterpart to the general verification principle of Proposition 5.2 in Section 5.1.2.

Proposition 6.2 *If $\mathcal{Z}^t = (\Omega, \mathbb{F}^t, \mathbb{P}^t)$, $(B^t, \chi^t, \nu^t), (\mathcal{X}^t, \mathcal{Y}^t, \bar{\mathcal{Y}}^t)$, is a consistent solution to the Markovian decoupled Forward Backward stochastic differential equation with data $\mathcal{G}$, $\mathcal{C}$, τ, with related value functions u and v, then:*
(i) *A saddle-point (ρ_t, θ_t) of the Dynkin game related to $\mathcal{Y}^t$ is given by:*

$$\rho_t = \inf\{s \in [t, T]\,;\, (s, \mathcal{X}^t_s) \in \mathcal{E}_-\} \wedge T \ , \ \theta_t = \inf\{s \in [t, T]\,;\, (s, \mathcal{X}^t_s) \in \mathcal{E}_+\} \wedge T,$$

with

$$\begin{aligned} \mathcal{E}_- &= \{(t, x, i) \in [0, T] \times \mathbb{R}^d \times I\,;\, u^i(t, x) = h^i(t, x)\} \\ \mathcal{E}_+ &= \{(t, x, i) \in [0, T] \times \mathbb{R}^d \times I\,;\, u^i(t, x) = \ell^i(t, x)\}\ ; \end{aligned}$$

(ii) *An optimal stopping time θ_t of the optimal stopping problem related to $\bar{\mathcal{Y}}^t$ is given by:*

$$\theta_t = \inf\{s \in [t, \tau^t]\,;\, (s, \mathcal{X}^t_s) \in \mathcal{E}^+\} \wedge T, \tag{130}$$

with

$$\mathcal{E}^+ = \{(t, x, i) \in [0, T] \times \mathbb{R}^d \times I\,;\, v^i(t, x) = \ell^i(t, x)\}\ .$$

Proof. **(i)** This follows immediately from identity (123) and the definition of the barriers in (121), given the general verification principle of Proposition 5.2.
(ii) By (126) and the fact that $\bar{\mathcal{Y}}^t$ is stopped at τ^t, it comes,

$$\bar{Y}^t_s = v(s \wedge \tau^t, \mathcal{X}^t_{s\wedge\tau^t}) \ , \ s \in [t, T]\ .$$

Using also the definition of the barrier in (122), θ_t defined by (130) is hence an optimal stopping time of the related optimal stopping problem, by application of the general verification principle of Proposition 5.2 (special case $\tau = T$ therein). □

6.6 Financial Application

Jump-diffusions, respectively continuous time Markov chains, are the major ingredients of most dynamic financial pricing models in the field of equity and interest- rates derivatives, respectively credit portfolio derivatives. The above jump–diffusion with regimes $\mathcal{X} = (X, N)$ can thus be fit to virtually any situation one may think of in the context of pricing and hedging financial derivatives (see Section 3.3.3 in Part I, where this model is represented, denoted by X, in the formalism of the abstract jump-diffusion (28)).

Let us give a few comments about more specific applications illustrating the fact that the generality of the set-up of this model is indeed required in order to cover the variety of situations encountered in financial modeling. So:

• In Bielecki et al. [17], this model is presented as a flexible risk-neutral pricing model in finance, for *equity and equity-to-credit (defaultable, cf. Section 4.2 in Part I) derivatives.* In this case the main component of the model, that is, the one in which the *payoffs* of the product under consideration are expressed, is X, while N represents *implied pricing regimes* which may be viewed as a simple form of *stochastic volatility.* More standard, diffusive, forms of stochastic volatility, may be accounted for in the diffusive component of X, whereas the jumps in X are motivated by the empirical evidence of the short-term volatility smile on financial derivatives markets.

In the context of single-name credit derivatives, N may also represent the credit rating of the reference obligor. So, in the area of *structural arbitrage*, *credit–to–equity* models and/or *equity–to–credit* interactions are studied. For example, if one of the factors is the price process of the equity issued by a credit name, and if credit migration intensities depend on this factor, then one has an equity–to–credit type interaction. On the other hand, if the credit rating of the obligor impacts the equity dynamics, then we deal with a credit–to–equity type interaction. The model $\mathcal{X}$ can nest both types of interactions.

• In Bielecki et al. [19], this model is used in the context of *portfolio credit risk* for the valuation and hedging of basket credit derivatives. The main component in the model is then the 'Markov chain like' component N, representing the vector of (implied) credit ratings of the reference obligors, which is modulated by the 'jump-diffusion like' component X, representing the evolution of economic variables which impact the likelihood of credit rating migrations. *Frailty* and *default contagion* are accounted for in the model by the coupled interaction between N and X.

Now, in the case of risk-neutral pricing problems in finance (see Part I), the driver coefficient function g is typically given as $g_1^i(t,x) - g_2^i(t,x)y$, for *dividend and interest-rate* related functions g_1 and g_2 (or dividends and interest-rates *adjusted for credit spread* in a more general context of defaultable contingent claims, cf. Section 4.2). Observe that in order for a consistent solution $\mathcal{Z}^t$ to our main FBSDE to satisfy

$$\int_t^s \widetilde{g}(\zeta, \mathcal{X}_\zeta^t, Y_\zeta^t, Z_\zeta^t, \mathcal{V}_\zeta^t)d\zeta = \int_t^s \left(g_1(\zeta, \mathcal{X}_\zeta^t) - g_2(\zeta, \mathcal{X}_\zeta^t)Y_\zeta^t\right)d\zeta\ ,\ s \in [t,T]$$
$$\int_t^s \widetilde{g}(\zeta, \mathcal{X}_\zeta^t, \bar{Y}_\zeta^t, \bar{Z}_\zeta^t, \bar{\mathcal{V}}_\zeta^t)d\zeta = \int_t^s \left(g_1(\zeta, \mathcal{X}_\zeta^t) - g_2(\zeta, \mathcal{X}_\zeta^t)\bar{Y}_\zeta^t\right)d\zeta\ ,\ s \in [t,\tau^t]$$

for given functions g_1 and g_2 on $\mathcal{E}$, it suffices in view of identities (125)-(128) to set

$$g^i(t,x,u,z,r) = g_1^i(t,x) - g_2^i(t,x)u^i + \sum_{j \in I} n^{i,j}(t,x)(u^j - u^i) \tag{131}$$

Note that g in (131) does not depend on z nor r, so $g^i(t,x,u,z,r) = g^i(t,x,u)$ here. However, modeling the pricing problem under the historical probability (as opposed to directly under the risk-neutral probability in Part I) would lead to a '(z,r)-*dependent*' driver coefficient function g.

Moreover we tacitly assumed in Part I a perfect, frictionless financial market. Accounting for market imperfections would lead to a *nonlinear* coefficient g.

Also, in the financial interpretation (see Part I):
• $\Phi(\mathcal{X}_T^t)$ corresponds to a *terminal payoff* that is paid by the issuer to the holder at time T if the contract was not exercised before T;
• $\ell(\mathcal{X}_s^t)$, resp. $h(\mathcal{X}_s^t)$, corresponds to a *lower*, resp. *upper payoff* that is paid by the issuer to the holder of the claim in the event of early termination of the contract at the initiative of the holder, resp. issuer;
• The stopping time τ^t (corresponding to $\bar{\sigma}$ in Part 4.2) is interpreted as the *time of lifting of a call protection*. This call protection prevents the issuer of the claim from calling it back (enforcing early exercise) before time τ^t. For instance, one has $\tau^t = T$ in the case of American contingent claims, which may only be exercised at the convenience of the *holder* of the claim.
The contingent claims under consideration are thus general *game contingent claims*, covering American claims and European claims as special cases;
• $\mathcal{X}$ (alias X in Part I) corresponds to a vector of observable *factors* (cf. Section 3.1).

Recall finally from Section 4.2 that in a context of *vulnerable claims* (or *defaultable derivatives*), it is enough, to account for credit-risk, to work with suitably *credit-spread adjusted interest-rates* μ and *recovery-adjusted dividend-yields* c in (131).

Remark 6.5 In Section 16 in Part IV (see also Section 4.3 in Part I), we consider products with more general, hence potentially more realistic forms of *intermittent* call protection, namely call protection *whenever a certain condition* is satisfied, rather than more specifically call protection *before a stopping time*.

7 Study of the Markovian Forward SDE

In few words, Sections 7 to 9, which culminate in Proposition 9.4 below, are devoted to finding explicit and general enough, even if admittedly technical and involved, conditions on the data $\mathcal{G}$, $\mathcal{C}$ and τ, under which existence of a consistent solution

$$\mathcal{Z}^t = (\Omega, \mathbb{F}^t, \mathbb{P}^t),\ (B^t, \chi^t, \nu^t), (\mathcal{X}^t, \mathcal{Y}^t, \bar{\mathcal{Y}}^t)$$

to the related Markovian FBSDE can be established.

Our approach for constructing a Markovian model $\mathcal{X} = (X, N)$ with mutual dependence between X and N is to start from a model with independent components (Section 7.1). We shall then apply a suitable Markovian change of probability measure in order to get a model with mutual dependence under the changed measure (Section 7.2).

7.1 Homogeneous Case

In this section we consider a first set of data with coefficients $n, f, b = \widehat{n}, \widehat{f}, \widehat{b}$ and related generator $\widehat{\mathcal{G}}$ such that

Assumption 7.1 **(i)** $\widehat{f} = 1$, and $\widehat{n}^{i,j}(t,x) = \widehat{n}^{i,j} \geq 0$ for any $i, j \in I$;
(ii) $\widehat{b}^i(t,x)$, $\sigma^i(t,x)$ and $\delta^i(t,x,y)$ are Lipschitz continuous in x uniformly in t, y, i;
(iii) $\widehat{b}^i(t,0)$, $\sigma^i(t,0)$ and $\int_{\mathbb{R}^d} \delta^i(t,0,y)m(dy)$ are bounded in t, y, i.

Let us be given a stochastic basis $(\Omega, \mathbb{F}, \mathbb{P})$, with $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$, assumed to support the following processes, independent of each other:
• a d-dimensional standard Brownian motion B;
• a compound Poisson process P with jump measure $m(dy)$;
• a continuous time Markov chain Q on $E = I^2 \times \{0,1\}$ with jump intensity from (l,j,ε) to (l',j',ε') given by $\mathbb{1}_{\varepsilon\neq\varepsilon'}\widehat{n}^{l',j'}$, for any $(l,j,\varepsilon) \neq (l',j',\varepsilon')$ (and a given law at time 0, the nature of which plays no role in the sequel).

Remark 7.2 Since P and Q are independent of each other and the jumping times of P are totally inaccessible, thus P and Q cannot jump together.

We denote by χ the random measure $\chi(ds,dy)$ on $[0,T] \times \mathbb{R}^d$ counting the jumps of P of size y between times 0 and s, and by ν the random measure $d\nu_s(l,j)$ on $[0,T] \times I^2$ counting the jumps of Q to the set $\{(l,j,0),(l,j,1)\}$ between times 0 and s. The $\mathbb{P}$-compensatrices $\widetilde{\chi}$ of χ and $\widetilde{\nu}$ of ν are thus respectively given by

$$\widetilde{\chi}(ds,dy) = \chi(ds,dy) - \widehat{f}(s,\mathcal{X}^t_{s-},y)m(dy)ds\ ,\ d\widetilde{\nu}_s(l,j) = d\nu_s(l,j) - \widehat{n}^{l,j}\,ds\ . \tag{132}$$

Remark 7.3 Conditionally on being in state (l,j,ε) at time s, Q jumps into the set $\{(l',j',0),(l',j',1)\}$ with probability $\widehat{n}^{l',j'}$ on the time interval $(s,s+ds)$. The $\nu(l,j)$s for l and j varying in I thus define independent Poisson processes with intensities $\widehat{n}^{l,j}$. In other words, $(\nu(l,j))_{(l,j)\in I^2}$ is a *multivariate point process* with intensity $(\widehat{n}^{l,j})_{(l,j)\in I^2}$ (see, e.g., Brémaud [29]).

We now consider the following stochastic differential equation, for $s \in [t,T]$:

$$\begin{cases} dN^t_s &= \sum_{j\in I}(j - N^t_{s-})\,d\nu_s(N^t_{s-},j) \\ dX^t_s &= \widehat{b}(s,\mathcal{X}^t_s)ds + \sigma(s,\mathcal{X}^t_s)\,dB_s + \int_{\mathbb{R}^d}\delta(s,X^t_{s-},N^t_{s-},y)\,\widetilde{\chi}(ds,dy)\ . \end{cases} \tag{133}$$

Remark 7.4 The reason why we introduce N^t indirectly via Q through (133) is that we need not only to define a process N^t for every initial condition (t,x,i), but also to ensure some kind of *consistency* between the family of processes thus defined, in order to enjoy error estimates like (136) below, where it should be noted that (136) ultimately relies on (144), which is valid by construction of N^t in (133). In particular, we define Q on the product state space $E = I^2 \times \{0,1\}$ as described above to ensure that N^t jumps from N^t_{s-} to $j \neq N^t_{s-}$ with probability $\widehat{n}(N^t_{s-},j)ds$ in the time interval $(s,s+ds)$, as must be the case for a Markov chain on I with intensities $\widehat{n}^{l,j}$, for every possible state of the underlying Markov chain Q at time s.

Proposition 7.1 *The stochastic differential equation (133) on $[t,T]$ with initial condition (x,i) on $[0,t]$ has a unique $(\Omega,\mathbb{F},\mathbb{P})$ – solution $\mathcal{X}^t = (X^t,N^t)$. For any $p \in [2,+\infty)$, one has:*

$$\|X^t\|^p_{\mathcal{S}^p_d} \le C_p\,(1 + |x|^p) \tag{134}$$

$$\|\mathbb{1}_{(s,r)}(X^t_\cdot - X^t_s)\|^p_{\mathcal{S}^p_d} \le C_p\,(1+|x|^p)\,(r-s) \tag{135}$$

Moreover, $^{t'}$ referring to a perturbed initial condition (t', x', i), one has:

$$\mathbb{P}(N^t \neq N^{t'}) \leq C|t - t'| \tag{136}$$

$$\|X^t - X^{t'}\|^p_{\mathcal{S}^p_d} \leq C_p \left(|x - x'|^p + (1 + \bar{x}^p)|t - t'|^{\frac{1}{2}} \right) \tag{137}$$

with $\bar{x} = |x| \vee |x'|$.

Proof. Note that the first line of (133) can be rewritten as

$$dN^t_s = \sum_{(l,j)\in I^2} (j - l)\, \mathbb{1}_{\{l = N^t_{s-}\}} d\nu_s(l, j) \tag{138}$$

$$= \sum_{(l,j)\in I^2} (j - l)\, \mathbb{1}_{\{l = N^t_{s-}\}} \widehat{n}^{l,j} ds + \sum_{(l,j)\in I^2} (j - l)\, \mathbb{1}_{\{l = N^t_{s-}\}} d\widetilde{\nu}_s(l, j)\ . \tag{139}$$

The last formulation corresponds to the special semimartingale canonical decomposition of N^t. One thus has the following equivalent form of (133),

$$\begin{cases} dN^t_s = \sum_{(l,j)\in I^2} (j - l)\, \mathbb{1}_{s<t} \mathbb{1}_{\{l = N^t_{s-}\}} \widehat{n}^{l,j} ds + \sum_{(l,j)\in I^2} (j - l)\, \mathbb{1}_{s<t} \mathbb{1}_{\{l = N^t_{s-}\}} d\widetilde{\nu}_s(l, j) \\ dX^t_s = \mathbb{1}_{s>t} \widehat{b}(s, \mathcal{X}^t_s) ds + \mathbb{1}_{s>t} \sigma(s, \mathcal{X}^t_s)\, dB_s + \int_{\mathbb{R}^d} \mathbb{1}_{s>t} \delta(s, \mathcal{X}^t_{s-}, y)\, \widetilde{\chi}(ds, dy) \end{cases} \tag{140}$$

with initial condition (x, i) at time 0. Any square integrable martingale or martingale measure is an L_2-*integrator* in the sense of Bichteler [13] (see Theorem 2.5.24 and its proof page 78 therein). Therefore by application of [13, Proposition 5.2.25 page 297], the stochastic differential equation (140) with initial condition (x, i) at time 0, or, equivalently, the stochastic differential equation (133) with initial condition (x, i) on $[0, t]$, has a unique $(\Omega, \mathbb{F}, \mathbb{P})$ – solution $\mathcal{X}^t = (X^t, N^t)$.

The general estimates (111)–(112) then yield, under Assumption 7.1:

$$\|X^t\|^p_{\mathcal{S}^p_d} \leq C_p C^t_p \tag{141}$$

$$\|\mathbb{1}_{(s,r)}(X^t_\cdot - X^t_s)\|^p_{\mathcal{S}^p_d} \leq C_p C^t_p (r - s) \tag{142}$$

$$\|X^t - X^{t'}\|^p_{\mathcal{S}^p_d} \leq C_p \left(C^t_p |t - t'| + C^{t,t'}_p \right) \tag{143}$$

with

$$\begin{aligned} C^t_p &= |x|^p + \mathbb{E}\left[\sup_{[t,T]} |\widehat{b}(\cdot, 0, N^t_\cdot)|^p + \sup_{[t,T]} |\sigma(\cdot, 0, N^t_\cdot)|^p + \sup_{[t,T]} \int_{\mathbb{R}^d} |\delta(\cdot, 0, N^t_\cdot, y)|^p m(dy) \right] \\ C^{t,t'}_p &= |x - x'|^p + \mathbb{E}\Big[\int_{t\wedge t'}^T |\widehat{b}(s, X^t_s, N^t_s) - \widehat{b}(s, X^t_s, N^{t'}_s)|^p ds \\ &\quad + \int_{t\wedge t'}^T |\sigma(s, X^t_s, N^t_s) - \sigma(s, X^t_s, N^{t'}_s)|^p ds \\ &\quad + \int_{t\wedge t'}^T \int_{\mathbb{R}^d} |\delta(s, X^t_s, N^t_s, y) - \delta(s, X^t_s, N^{t'}_s, y)|^p m(dy) ds \Big] \end{aligned}$$

The bound estimates (134)–(135) result from (141)–(142) by the boundedness Assumption 7.1(iii) on the coefficients. As for the error estimates (136)–(137), note that by construction of N via Q in (133), one has (assuming $t \leq t'$, w.l.o.g.):

$$N^t_{t'} = i \ \Rightarrow\ N^{t'}_s = N^t_s,\ s \in [t', T] \tag{144}$$

Now

$$\mathbb{P}(\{N^t_{t'} \neq i\}) \leq \sum_{j\in I} 1 - e^{-\widehat{n}^{i,j}|t-t'|} \leq \sum_{j\in I} \widehat{n}^{i,j} |t - t'| = |\widehat{n}^{i,i}||t - t'|\ ,$$

whence (136). Therefore

$$\mathbb{E}\int_t^T |\widehat{b}(s,X_s^t,N_s^t)-\widehat{b}(s,X_s^t,N_s^{t'})|^p ds \le \\ C|t-t'|^{\frac{1}{2}}\Big(\mathbb{E}\int_t^T \big(|\widehat{b}(s,X_s^t,N_s^t)|^{2p}+|\widehat{b}(s,X_s^t,N_s^{t'})|^{2p}\big)ds\Big)^{\frac{1}{2}}$$

where by (134) and the properties of b:

$$\mathbb{E}\int_t^T |\widehat{b}(s,X_s^t,N_s^t)|^{2p}ds \le C\mathbb{E}\int_t^T \big(|\widehat{b}(s,0,N_s^t)|^{2p}+|X_s^t|^{2p}\big)ds \le C_{2p}\big(1+\bar{x}^{2p}\big)$$

and likewise for $\mathbb{E}\int_t^T |\widehat{b}(s,X_s^t,N_s^{t'})|^{2p}ds$. So

$$\mathbb{E}\int_t^T |\widehat{b}(s,X_s^t,N_s^t)-\widehat{b}(s,X_s^t,N_s^{t'})|^p ds \le C_p(1+\bar{x}^p)|t-t'|^{\frac{1}{2}}$$

and by similar estimates regarding the terms in σ and δ of $C_p^{t,t'}$:

$$C_p^{t,t'} \le |x-x'|^p + C_p(1+\bar{x}^p)|t-t'|^{\frac{1}{2}} .$$

Hence (137) follows, by (143). □

Remark 7.5 In case where there are no regimes in the model (case $k=1$), one can see by inspection of the above proof that $|t-t'|^{\frac{1}{2}}$ can be improved into $|t-t'|$ in (137).

Let us define further on $[t,T]$:

$$B_s^t = B_s - B_t \ , \ \chi^t = \chi - \chi_{t-} \ , \ \nu^t = \nu(N_{\cdot-}^t, j) - \nu(N_{t-}^t, j) \ . \tag{145}$$

Note that ν^t is a random measure on $[0,T]\times I$, whereas ν is a random measure on $[0,T]\times I^2$. Let $\mathbb{F}_{B^t}$, $\mathbb{F}_{\chi^t}$, $\mathbb{F}_{\nu^t}$ and $\mathbb{F}^t$ stand for the filtrations on $[t,T]$ generated by B^t, χ^t, ν^t, and the three processes together, respectively. Given a further *initial condition at time t* ($\mathcal{F}$-measurable random variable) denoted by $\widetilde{M}_t$, with generated sigma field denoted by $\Sigma(\widetilde{M}_t)$, let in turn $\widetilde{\mathbb{F}}_{B^t}$, $\widetilde{\mathbb{F}}_{\chi^t}$, $\widetilde{\mathbb{F}}_{\nu^t}$ and $\widetilde{\mathbb{F}}^t$ stand for the filtrations on $[t,T]$ generated by $\Sigma(\widetilde{M}_t)$ and, respectively, $\mathbb{F}_{B^t}$, $\mathbb{F}_{\chi^t}$, $\mathbb{F}_{\nu^t}$ and $\mathbb{F}^t$.

Proposition 7.2 **(i)** *Let $\mathcal{X}^t$ be defined as in Proposition 7.1. The stochastic differential equation (133)–(140) on $[t,T]$ with initial condition (x,i) at t admits a unique strong $(\Omega,\mathbb{F}^t,\mathbb{P})$ – solution, which is given by the restriction of $\mathcal{X}^t$ to $[t,T]$. In particular, $(\Omega,\mathbb{F}^t,\mathbb{P}),(B^t,\chi^t,\nu^t),\mathcal{X}^t$ is a solution to the time-dependent local martingale problem with generator $\widehat{\mathcal{G}}$ and initial condition (t,x,i).*
(ii) $(\mathbb{F}^t,\mathbb{P};B^t,\chi^t,\nu^t)$ has the local martingale predictable representation property*, in the sense that for any random variable $\widetilde{M}_t$, any $(\widetilde{\mathbb{F}}^t,\mathbb{P})$ – local martingale M with initial condition $\widetilde{M}_t$ at time t admits a representation*

$$M_s = M_t + \int_t^s Z_r\,dB_r + \int_t^s\int_{\mathbb{R}^d} V_r(dx)\widetilde{\chi}(dx,dr) + \sum_{j\in I}\int_t^s W_r(j)d\widetilde{\nu}(N_{s-}^t,j)\,, \quad s\in[t,T] \tag{146}$$

for processes Z,V,W in the related spaces of predictable integrands. □

Proof. **(i)** is straightforward, given Proposition 7.1 and the fact that the restriction of $\mathcal{X}^t$ to $[t,T]$ is $\mathbb{F}^t$-adapted. The fact that $(\Omega, \mathbb{F}^t, \mathbb{P}), (B^t, \chi^t, \nu^t), \mathcal{X}^t$ is a model with generator $\widehat{\mathcal{G}}$ immediately follows by application of the Itô formula (117).
(ii) One has the following local martingale predictable representation properties for $(\mathbb{F}_{B^t}, \mathbb{P}; B^t)$, $(\mathbb{F}_{\chi^t}, \mathbb{P}; \chi^t)$ and $(\mathbb{F}_{\nu^t}, \mathbb{P}; \nu^t)$, respectively (see, e.g., Jacod–Shiryaev [67, Theorem 4.34(a) Chaper III page 189] for the two former and Boel et al. [24, 25] for the latter):
• Every $(\widetilde{\mathbb{F}}_{B^t}, \mathbb{P}; B^t)$ – local martingale M with initial condition $\widetilde{M}_t$ at time t admits a representation

$$M_s = M_t + \int_t^s Z_r \, dB_r \,, \quad s \in [t,T] \,;$$

• Every $(\widetilde{\mathbb{F}}_{\chi^t}, \mathbb{P}; \chi^t)$ – local martingale M with initial condition $\widetilde{M}_t$ at time t admits a representation

$$M_s = M_t + \int_t^s \int_{\mathbb{R}^d} V_r(dx) \widetilde{\chi}(dx, dr) \,, \quad s \in [t,T] \,;$$

• Every $(\widetilde{\mathbb{F}}_{\nu^t}, \mathbb{P}; \nu^t)$ – local martingale M with initial condition $\widetilde{M}_t$ at time t admits a representation

$$M_s = M_t + \sum_{j \in I} \int_t^s W_r(j) d\widetilde{\nu}(N^t_{s-}, j) \,, \quad s \in [t,T] \,,$$

for processes Z, V, W in the related spaces of predictable integrands.
By independence of B, P and Q, added to the fact that the related square brackets are null (see, e.g., Jeanblanc et al. [71]), this implies the local martingale predictable representation property (146) for $(\mathbb{F}^t, \mathbb{P}; B^t, \chi^t, \nu^t)$. □

7.2 Inhomogeneous Case

Our next goal is to show how to construct a model with generator of a more general form (113) (if not of the completely general form (113): see Assumption 7.6 below), under less restrictive conditions than in the previous section, with *state-dependent intensities*. Towards this end we shall apply to the model of Section 7.1 a suitable *Markovian change of probability measures* (see Kunita–Watanabe [76], Palmowski and Rolski [86]; cf. also Bielecki et al. [19] and Becherer–Schweizer [10]).

Let thus the *change of measure function* γ be defined as the exponential of a function of class $\mathcal{C}^{1,2}$ with compact support on $\mathcal{E}$. Starting from $\widehat{\mathcal{G}}$, we define the operator $\mathcal{G}$ of the form (113) with data n, f and b as follows (and other data as before), for $(t,x,i) \in \mathcal{E}$:

$$\left\{ \begin{array}{l} n^{i,j}(t,x) = \frac{\gamma^j(t,x)}{\gamma^i(t,x)} \widehat{n}^{i,j} \\ f^i(t,x,y) = \frac{\gamma^i(t,x+\delta^i(t,x,y))}{\gamma^i(t,x)} \, \widehat{f}^i(t,x,y), \\ b^i(t,x) = \widehat{b}^i(t,x) + \int_{\mathbb{R}^d} \delta^i(t,x,y)(f^i(t,x,y) - \widehat{f}^i(t,x,y)) m(dy) \,. \end{array} \right. \tag{147}$$

Lemma 7.3 ***(i)*** *The function n is bounded, and the function f is positively bounded and Lipschitz continuous with respect to x uniformly in t, y, i.*

(ii) *The* $(\mathbb{F}^t, \mathbb{P})$ *– local martingale* Γ^t *defined by, for* $s \in [t, T]$,

$$\frac{d\Gamma_s^t}{\Gamma_{s-}^t} = \int_{\mathbb{R}^d} \left(\frac{f(s, \mathcal{X}_{s-}^t, y)}{\widehat{f}(s, \mathcal{X}_{s-}^t, y)} - 1 \right) \widetilde{\chi}(ds, dy) + \sum_{j \in I} \left(\frac{n^j(s, \mathcal{X}_{s-}^t)}{\widehat{n}^j(N_{s-}^t)} - 1 \right) d\widetilde{\nu}_s(N_{s-}^t, j) \tag{148}$$

is a positive $(\mathbb{F}^t, \mathbb{P})$*-martingale with* $\mathbb{E}\Gamma_s^t = 1$ *and such that (with* Γ^t *extended by one on* $[0, t]$*):*

$$\|\Gamma^t\|_{\mathcal{S}_d^p}^p \leq C_p \ . \tag{149}$$

Proof. **(i)** is straightforward, given Assumptions 7.1(ii) and the regularity assumptions on γ.
(ii) By application of Bichteler [13, Proposition 5.2.25 page 297], the stochastic differential equation (148) with initial condition 1 on $[0, t]$, has a unique $(\Omega, \mathbb{F}, \mathbb{P})$ – solution Γ^t. Estimate (149) follows by application of the general estimate (111) to Γ^t. So the local martingale Γ^t is a genuine martingale. □

We thus define an equivalent probability measure $\mathbb{P}^t$ on $(\Omega, \mathcal{F}_T^t)$ by setting, for every $s \in [t, T]$:

$$\frac{d\mathbb{P}_s^t}{d\mathbb{P}_s} = \Gamma_s^t, \quad \mathbb{P}\text{-a.s.} \tag{150}$$

Γ_s^t is then the $\mathcal{F}_s^t$-measurable version of the Radon-Nikodym density of $\mathbb{P}^t$ with respect to $\mathbb{P}$ on $\mathcal{F}_s^t$.

Let us define, for $s \in [t, T]$:

$$\left\{ \begin{array}{l} \widetilde{\chi}^t(ds, dy) = \chi^t(ds, dy) - f(s, \mathcal{X}_s^t, y) m(dy) ds \\ d\widetilde{\nu}_s^t(j) = d\nu_s^t(j) - n(s, \mathcal{X}_s^t, j)\, ds \ . \end{array} \right. \tag{151}$$

The proof of the following lemma is classical and therefore deferred to Appendix A.1. Note that this result does not depend on the special form of b in (147). Recall (145) for the definition of B^t.

Lemma 7.4 B^t *is an* $(\mathbb{F}^t, \mathbb{P}^t)$ *– Brownian motion starting at time* t*, and* $\widetilde{\chi}^t$ *and* $\widetilde{\nu}^t$ *are the* $\mathbb{P}^t$*-compensatrices of* χ^t *and* ν^t*.*

Proposition 7.5 **(i)** *The restriction to* $[t, T]$ *of* $\mathcal{X}^t = (X^t, N^t)$ *in Propositions 7.1-7.2(i) is the unique* $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$ *– solution of the following SDE on* $[t, T]$ *with initial condition* (x, i) *at time* t *:*

$$\left\{ \begin{array}{ll} dN_s^t & = \sum_{j \in I} (j - N_{s-}^t)\, d\nu_s^t(j) = \sum_{j \in I} (j - N_{s-}^t)\, n(s, \mathcal{X}_{s-}^t, j)\, ds + \sum_{j \in I} (j - N_{s-}^t)\, d\widetilde{\nu}_s^t(j) \\ dX_s^t & = b(s, \mathcal{X}_s^t) ds + \sigma(s, \mathcal{X}_s^t)\, dB_s^t + \int_{\mathbb{R}^d} \delta(s, \mathcal{X}_{s-}^t, y)\, \widetilde{\chi}^t(ds, dy) \ . \end{array} \right. \tag{152}$$

In particular $(\Omega, \mathbb{F}^t, \mathbb{P}^t), (B^t, \chi^t, \nu^t), \mathcal{X}^t$ *is a solution to the time-dependent local martingale problem with generator* $\mathcal{G}$ *and initial condition* (t, x, i)*.*
(ii) $(\mathbb{F}^t, \mathbb{P}^t; B^t, \chi^t, \nu^t)$ *has the local martingale predictable representation property, in the*

sense that for any random variable $\widetilde{M}_t$, any $(\widetilde{\mathbb{F}}^t, \mathbb{P}^t)$ – local martingale M with initial condition $\widetilde{M}_t$ at time t, where $\widetilde{\mathbb{F}}^t$ denotes the filtration on $[t, T]$ generated by $\mathbb{F}^t$ and $\Sigma(\widetilde{M}_t)$, admits a representation

$$M_s = M_t + \int_t^s Z_r\, dB_r^t + \int_t^s \int_{\mathbb{R}^d} V_r(dx)\widetilde{\chi}^t(dx, dr) + \sum_{j\in I} \int_t^s W_r(j) d\widetilde{\nu}_s^t(j)\,, \quad s \in [t, T] \tag{153}$$

for processes Z, V, W in the related spaces of predictable integrands.

Proof. **(i)** Given (151)–(147), $\mathcal{X}^t$ is a strong $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$ – solution of the stochastic differential equation (152) with initial condition (x, i) at time t if and only if it is a strong $(\Omega, \mathbb{F}^t, \mathbb{P})$ – solution of the stochastic differential equation (133) with initial condition (x, i) at time t. The result hence follows from Proposition 7.2(i).
(ii) The local martingale predictable representation property is preserved by equivalent changes of probability measures (see, e.g., Jacod–Shiryaev [67, Theorem 5.24 page 196]), so the result follows from Proposition 7.2(ii). □

Comments 7.1 **(i)** As an alternative to (148), one might consider the following variant:

$$\begin{aligned}\frac{d\Gamma_s^t}{\Gamma_{s-}^t} = \frac{\partial\gamma\sigma}{\gamma}(s, \mathcal{X}_s^t)dB_s + \\ \int_{\mathbb{R}^d} \left(\frac{f(s, \mathcal{X}_{s-}^t, y)}{\widehat{f}(s, \mathcal{X}_{s-}^t, y)} - 1\right) \widetilde{\chi}(ds, dy) + \sum_{j\in I} \left(\frac{n^j(s, \mathcal{X}_{s-}^t)}{\widehat{n}(N_{s-}^t, j)} - 1\right) d\widetilde{\nu}_s(N_{s-}^t, j)\end{aligned} \tag{154}$$

As compared with (148), the change of probability measure defined by (154), which is used for instance in [19], would have the additional effect to further change the Brownian motion into

$$d\widetilde{B}_s^t = dB_s^t - \tfrac{(\partial\gamma\sigma)^{\mathsf{T}}}{\gamma}(s, \mathcal{X}_s^t)ds \tag{155}$$

in (151), and to modify accordingly the coefficient of the first-order term in the generator.
(ii) From the point of view of financial interpretation (see Part I):
• The changed measure $\mathbb{P}^t$ with associated generator $\mathcal{G}$ of $\mathcal{X}^t$ may be thought of as representing the *risk-neutral pricing measure* chosen by the market to value financial instruments (or, in the case of defaultable single-name credit instruments as of Section 4.2, the *pre-default pricing measure*).
In the risk-neutral pricing context, this imposes a specific *arbitrage consistency condition* that must be satisfied by the risk-neutral drift coefficient b of $\mathcal{G}$ in (147). Namely, in the simplest, default-free case, and for those components x_l of X which correspond to price processes of primary risky assets, in an economy with constant riskless interest-rate r and dividend yields q_l, arbitrage requirements imply that

$$b_l^i(t, x) = (r - q_l)x_l\ ,$$

for $(t, x, i) \in \mathcal{E}$. An analogous *pre-default arbitrage drift condition* may also be derived in the case of a pre-default factor process $\mathcal{X}$ in the case of defaultable derivatives, see Section 4.2 and [17]. The corresponding components b_l of b are thus pre-determined in (147). The change of measure (147) must then be understood in the reverse-engineering mode, for

deducing $\widehat{b}_l$ from b_l rather than the other way round. The change of measure function γ in (147), possibly parameterized in some relevant way depending on the application at hand, may be determined along with other model parameters at the stage of the *calibration* of the model to market data;
• Another possible interpretation and use of the change of measure (as in Bielecki et al. [19], where the variant (154) of (148) is used therein), is that of a *change of numeraire* (cf. Section 4.1).

7.3 Synthesis

Assumption 7.6 In the remaining sections in this part (Sections 8 and 9), we shall work with the models $(\Omega, \mathbb{F}^t, \mathbb{P}^t), (B^t, \chi^t, \nu^t), \mathcal{X}^t = (X^t, N^t)$ with generator $\mathcal{G}$ thus constructed, for initial conditions (t, x, i) varying in $\mathcal{E}$. We thus effectively reduce attention from the general case (113) to the case of a generator with data n, f, b deduced from one with 'independent ingredients' $\widehat{n}, \widehat{f} = 1, \widehat{b}$ by the formulas (147).

$\mathbb{P}^t$-expectation and $\mathbb{P}$-expectation will be denoted henceforth by $\mathbb{E}^t$ and $\mathbb{E}$, respectively. The original stochastic basis $(\Omega, \mathbb{F}, \mathbb{P})$ and generator $\widehat{\mathcal{G}}$ will be used for deriving error estimates in Sections 8 and 9, where we shall express with respect to this common basis differences between $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$-solutions corresponding to different initial conditions (t, x, i).

Towards this view, in addition to the notation already introduced in section 6.3, and applied to the model $\mathcal{X}^t$ specified as above relatively to $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, we define likewise, in relation with the process $\mathcal{X}^t$ considered relatively to $(\Omega, \mathbb{F}, \mathbb{P})$:
• F, the subset $\big(\mathbb{R}^d \times \{0_2\}\big) \cup \big(\{0_d\} \times I^2\big)$ of $\mathbb{R}^d \times \mathbb{R}^2$;
• $\mathcal{B}_F$, the sigma field generated by $\mathcal{B}(\mathbb{R}^d) \times \{0_2\}$ and $\{0_d\} \times \mathcal{B}_{I^2}$ on F, where $\mathcal{B}(\mathbb{R}^d)$ and $\mathcal{B}_{I^2}$ stand for the Borel sigma field on $\mathbb{R}^d$ and the sigma field of all parts of I^2, respectively;
• $\pi(de)$ and $\zeta_t(e)$ respectively given by, for any $e = (y, (l, j)) \in E$ and $t \in [0, T]$:

$$\pi(de) = \begin{cases} m(dy) \text{ if } (l, j) = 0_2 \\ 1 \text{ if } y = 0_d \end{cases} \quad , \quad \zeta_t(e) = \begin{cases} 1 \text{ if } (l, j) = 0_2 \\ \widehat{n}^{l,j} \text{ if } y = 0_d \end{cases}$$

• μ, the integer-valued random measure on $\big([0, T] \times E, \mathcal{B}([0, T]) \otimes \mathcal{B}_E\big)$ counting the jumps of χ of size $y \in A$ and the jumps of ν into the set $\{(l, j, 0), (l, j, 1)\}$ between 0 and t, for any $t \geq 0$, $A \in \mathcal{B}(\mathbb{R}^d), (l, j) \in I^2$, viewed as a random measure relative to the stochastic basis $(\Omega, \mathbb{F}, \mathbb{P})$.

We denote for short (cf. section 6.3):

$$(F, \mathcal{B}_F, \pi) = (\mathbb{R}^d \oplus I^2, \mathcal{B}(\mathbb{R}^d) \oplus \mathcal{B}_{I^2}, m(dy) \oplus \mathbb{1})$$

and $\mu = \chi \oplus \nu$. The $(\Omega, \mathbb{F}, \mathbb{P})$-compensator of μ is thus given by, for any $t \geq 0, A \in \mathcal{B}(\mathbb{R}^d), (l, j) \in I^2$, with $A \oplus \{(l, j)\} := \big(A \times \{0_2\}\big) \cup \big(\{0_d\} \times \{(l, j)\}\big)$:

$$\int_0^t \int_{A \oplus \{(l,j)\}} \zeta_t(e) \rho(de) ds = \int_0^t \int_A m(dy) ds + \int_0^t \widehat{n}^{l,j} \, ds \, .$$

Note that $\mathcal{H}^2_\mu$ can be identified with the product space $\mathcal{H}^2_\chi \times \mathcal{H}^2_\nu$, and that $\mathcal{M}_\pi = \mathcal{M}(F, \mathcal{B}_F, \pi; \mathbb{R})$ can be identified with the product space $\mathcal{M}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m(dy); \mathbb{R}) \times \mathbb{R}^{k^2}$. For

$$\widehat{v} = (v, w) \in \mathcal{M}_\pi \equiv \mathcal{M}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m(dy); \mathbb{R}) \times \mathbb{R}^{k^2} \ ,$$

we denote accordingly (cf. (105)):

$$|\widehat{v}|^2 = \int_{\mathbb{R}^d} v(y)^2 m(dy) + \sum_{(l,j)\in I^2} w(l,j)^2 \widehat{n}^{l,j} \,. \tag{156}$$

In the sequel $\widetilde{v}$ and $\widehat{v}$ denote generic pairs (v,w) in $\mathcal{M}_\rho$ and $\mathcal{M}_\pi$, respectively.

8 Study of the Markovian BSDEs

We assume that the cost functions $\mathcal{C}$ satisfy the Markovian assumptions (M.0) to (M.2) introduced in Section 6.4, as well as
(**M.3**) $\ell = \phi \vee c$ for a $\mathcal{C}^{1,2}$-function ϕ on $\mathcal{E}$ such that

$$\phi,\, \mathcal{G}\phi,\, \partial\phi\sigma,\, (t,x,i) \mapsto \int_{\mathbb{R}^d} |\phi^i(t, x+\delta^i(t,x,y))| m(dy) \;\in\; \mathcal{P} \tag{157}$$

and for a constant $c \in \mathbb{R} \cup \{-\infty\}$.

Comments 8.1 **(i)** The standing example for ϕ in (M.3) (see [42]) is $\phi = x_1$, the first component of $x \in \mathbb{R}^d$ (assuming $d \geq 1$ in our model), whence $\mathcal{G}\phi = b_1$. In this case (157) reduces to

$$b_1,\, \sigma_1,\, (t,x,i) \mapsto \int_{\mathbb{R}^d} |\delta_1^i(t,x,y)| m(dy) \;\in\; \mathcal{P} \,.$$

(ii) Alternatively to (M.3), one might work with the symmetric assumptions regarding h, namely $h = \phi \wedge C$ where ϕ satisfies (157). However it turns out that this kind of call payoff does not correspond to any known applications, at least in finance.

Theorem 8.1 **(i)** *The R2BSDE on* $[t,T]$ *with data (cf. (121))*

$$\widetilde{g}(s, \mathcal{X}_s^t, y, z, \widetilde{v}) \,,\; \Phi(\mathcal{X}_T^t) \,,\; \ell(s, \mathcal{X}_s^t) \,,\; h(s, \mathcal{X}_s^t) \tag{158}$$

has a unique $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t) *– solution* $\mathcal{Y}^t = (Y^t, Z^t, \mathcal{V}^t, K^t)$.
(ii) *Denoting* $\mathcal{V}^t = (V^t, W^t)$ *with* $V^t \in \mathcal{H}^2_{\chi^t}, W^t \in \mathcal{H}^2_{\nu^t}$, *we extend* Y^t *by* Y_t^t *and* K^t, Z^t *and* $\mathcal{V}^t$ *by* 0 *on* $[0,t]$, *and we define on* $[0,T]$:

$$\widetilde{W}_s^t(l,j) = \mathbb{1}_{\{l = N_{s-}^t\}} W_s^t(j) \text{ for } l,j \in I \,,\; \widetilde{\mathcal{V}} = (V, \widetilde{W}) \,.$$

Then $\widetilde{\mathcal{Y}}^t = (Y^t, Z^t, \widetilde{\mathcal{V}}^t, K^t)$ *is an* $(\Omega, \mathbb{F}, \mathbb{P}), (B, \mu)$ *– solution to the R2BSDE on* $[0,T]$ *with data*

$$\mathbb{1}_{\{s>t\}} \widehat{g}(s, \mathcal{X}_s^t, y, z, \widehat{v}) \,,\; \Phi(\mathcal{X}_T^t) \,,\; \ell(s \vee t, \mathcal{X}_{s\vee t}^t) \,,\; h(s \vee t, \mathcal{X}_{s\vee t}^t) \,, \tag{159}$$

where

$$\widehat{g}(s, \mathcal{X}_s^t, y, z, \widehat{v}) := g(s, \mathcal{X}_s^t, \widehat{u}_s^t, z, \widetilde{r}_s^t) + (\widetilde{r}_s^t - \widehat{r}) - \sum_{(l,j)\in I^2, l\neq j} w_{l,j} \widehat{n}^{l,j} \tag{160}$$

with

$$(\widehat{u}_s^t)_j(y,w) = \begin{cases} y \,, & j = N_s^t \\ y + \sum_{l\in I} w_{l,j} \,, & j \neq N_s^t \end{cases} \,,\; \widehat{r}(v) = \int_{\mathbb{R}^d} v(y) m(dy) \,.$$

Proof. **(i)** Given (M.0)–(M.1)–(M.2) and the bound estimates (134) on X^t and (149) on Γ^t, the following conditions are satisfied:
(H.0)' $\Phi(\mathcal{X}^t_T) \in \mathcal{L}^2$;
(H.1.i)' $\widetilde{g}(\cdot, \mathcal{X}^t_\cdot, y, z, \widetilde{v})$ is a progressively measurable process on $[t, T]$ with

$$\mathbb{E}^t\Big[\int_t^T \widetilde{g}(\cdot, \mathcal{X}^t_\cdot, y, z, \widetilde{v})^2\, dt\Big] \, < \, +\infty \,,$$

for any $y \in \mathbb{R}$, $z \in \mathbb{R}^{1\otimes d}$, $\widetilde{v} \in \mathcal{M}_\rho$ (where $\mathbb{E}^t$ denotes $\mathbb{P}^t$-expectation);
(H.1.ii)' $\widetilde{g}(\cdot, \mathcal{X}^t_\cdot, y, z, \widetilde{v})$ is uniformly Λ – Lipschitz continuous with respect to $(y, z, \widetilde{v})$, in the sense that for every $s \in [t, T], y, y' \in \mathbb{R},\ z, z' \in \mathbb{R}^{1\otimes d},\ \widetilde{v}, \widetilde{v}' \in \mathcal{M}_\rho$:

$$|\widetilde{g}(s, \mathcal{X}^t_s, y, z, \widetilde{v}) - \widetilde{g}(s, \mathcal{X}^t_s, y', z', \widetilde{v}')| \le \Lambda\big(|y - y'| + |z - z'| + |\widetilde{v} - \widetilde{v}'|_s\big)$$

(cf. (118) for the definition of $|\widetilde{v} - \widetilde{v}'|_s$);
(H.2.i)' $\ell(s, \mathcal{X}^t_s)$ and $h(s, \mathcal{X}^t_s)$ are càdlàg quasi-left continuous processes in $\mathcal{S}^2$;
(H.2.ii)' $\ell(\cdot, \mathcal{X}^t_\cdot) \le h(\cdot, \mathcal{X}^t_\cdot)$ on $[t, T)$, and $\ell(T, \mathcal{X}^t_T) \le \Phi(\mathcal{X}^t_T) \le h(T, \mathcal{X}^t_T)$.

Therefore the general assumptions (H.0)–(H.1)–(H.2) are satisfied by the data (158) relative to $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t). Given the local martingale predictable representation property of Proposition 7.5(ii) and the form postulated in (M.3) for ℓ, existence and uniqueness of an $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t) – solution $\mathcal{Y}^t = (Y^t, Z^t, \mathcal{V}^t, K^t)$ to the R2BSDE with data (158) on $[t, T]$ follows by application of the general results of [42].

(ii) By the previous R2BSDE, one thus has for $s \in [t, T]$:

$$\begin{aligned}
-dY^t_s &= \widetilde{g}(s, \mathcal{X}^t_s, Y^t_s, Z^t_s, \mathcal{V}^t_s)ds + dK^t_s - Z^t_s dB_s - \int_{\mathbb{R}^d} V^t_s(y)\widetilde{\chi}^t(ds, dy) - \sum_{j\in I} W^t_s(j) d\widetilde{\nu}^t_s(j) \\
&= g(s, \mathcal{X}^t_s, \widetilde{u}^t_s, Z^t_s, \widetilde{r}^t_s)ds + dK^t_s - Z^t_s dB_s + \int_{\mathbb{R}^d} V^t_s(y)(\widetilde{\chi} - \widetilde{\chi}^t)(ds, dy) \\
&\quad - \int_{\mathbb{R}^d} V^t_s(y)\widetilde{\chi}(ds, dy) - \sum_{j\in I} W^t_s(j) d\nu^t_s(j) \ .
\end{aligned}$$

Given (151), (147) (where $\widehat{f} = 1$) and the facts that for $s \ge t$:

$$\textstyle\sum_{j\in I} W^t_s(j) d\nu^t_s(j) = \sum_{(l,j)\in I^2} \widetilde{W}^t_s(l, j) d\nu_s(l, j) \ , \ \widetilde{u}^t_s(Y^t_s, W^t_s) = \widehat{u}^t_s(Y^t_s, \widetilde{W}^t_s) \ ,$$

one gets that for $s \ge t$:

$$\begin{aligned}
-dY^t_s = g(s, \mathcal{X}^t_s, \widehat{u}^t_s, Z^t_s, \widetilde{r}^t_s)ds + dK^t_s - Z^t_s dB_s + \int_{\mathbb{R}^d} V^t_s(y)(f(s, \mathcal{X}^t_s, y) - 1)m(dy)ds \\
- \int_{\mathbb{R}^d} V^t_s(y)\widetilde{\chi}(ds, dy) - \sum_{(l,j)\in I^2} \widetilde{W}^t_s(l, j) d\nu_s(l, j) \ .
\end{aligned}$$

It is then immediate to check that $\widetilde{\mathcal{Y}}^t$ is an $(\Omega, \mathbb{F}, \mathbb{P}), (B, \mu)$ – solution of the R2BSDE with data (159) on $[0, T]$. □

By application of the general estimates of [42] to $\widetilde{\mathcal{Y}}^t$ (where the $\widetilde{\mathcal{Y}}^t$s for varying t, x, i are defined with respect to the common stochastic basis $(\Omega, \mathbb{F}, \mathbb{P})$, cf. section 7.3), one can derive the following stability result. Since this derivation is rather standard but tedious and lengthy, it is deferred to Appendix A.2.

Proposition 8.2 **(i)** *One has the following estimate on $\widetilde{\mathcal{Y}}^t$ in Theorem 8.1:*

$$\|Y^t\|^2_{\mathcal{S}^2} + \|Z^t\|^2_{\mathcal{H}^2_d} + \|\widetilde{\mathcal{V}}^t\|^2_{\mathcal{H}^2_\mu} + \|K^{t,+}\|^2_{\mathcal{S}^2} + \|K^{t,-}\|^2_{\mathcal{S}^2} \le C(1+|x|^{2q}) \; . \tag{161}$$

(ii) *Moreover, ^{t_n} referring to a perturbed initial condition (t_n, x_n, i) with $(t_n, x_n) \to (t, x)$ as $n \to \infty$, one has that $\widetilde{\mathcal{Y}}^{t_n}$ $\mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu \times \mathcal{A}^2$ – converges to $\widetilde{\mathcal{Y}}^t$ as $n \to \infty$.*

8.1 Semi-Group Properties

Let ^{t} refer to the constant initial condition (t, x, i) as usual. Let $\mathcal{X}^t = (X^t, N^t)$ and $\mathcal{Y}^t$ be defined as in Proposition 7.1 and Theorem 8.1, respectively. Given $t' \ge t$, let $\widetilde{\mathbb{F}}^{t'}$ stand for $(\widetilde{\mathcal{F}}^{t'}_r)_{r \ge t'}$ with for $r \ge t'$

$$\widetilde{\mathcal{F}}^{t'}_r = \sigma(\mathcal{X}^t_{t'}) \bigvee \mathcal{F}^{t'}_r \; .$$

As for $\mathbb{F}^{t'} = (\mathcal{F}^{t'}_r)_{r \ge t'}$, $\mathbb{P}^{t'}$, $B^{t'}$ and $\mu^{t'}$, they are still defined as in Sections 7.1-7.2, with t' instead of t therein. Note in particular that $\widetilde{\mathbb{F}}^{t'}$ is smaller than or equal to the restriction $\mathbb{F}^t_{|[t',T]}$ of $\mathbb{F}^t$ to $[t', T]$.

Proposition 8.3 **(i)** *Let $\mathcal{X}^t$ be defined as in Proposition 7.1. The stochastic differential equation (133)-(140) on $[t', T]$ with initial condition $\mathcal{X}^t_{t'}$ at t' admits a unique strong $(\Omega, \widetilde{\mathbb{F}}^{t'}, \mathbb{P})$ – solution $\mathcal{X}^{t'} = (X^{t'}, N^{t'})$, which coincides with the restriction of $\mathcal{X}^t$ to $[t', T]$, so:*

$$\mathcal{X}^{t'} = (X^{t'}_r, N^{t'}_r)_{t' \le r \le T} = (\mathcal{X}^t_r)_{t' \le r \le T} \; .$$

(ii) *The R2BSDE on $[t', T]$ with data*

$$\widetilde{g}(s, \mathcal{X}^{t'}_s, y, z, \widetilde{v}) \; , \; \Phi(\mathcal{X}^{t'}_T) \; , \; \ell(s, \mathcal{X}^{t'}_s) \; , \; h(s, \mathcal{X}^{t'}_s) \tag{162}$$

has a unique $(\Omega, \widetilde{\mathbb{F}}^{t'}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$ – solution $\mathcal{Y}^{t'} = (Y^{t'}, Z^{t'}, \mathcal{V}^{t'}, K^{t'})$ such that, with $\mathcal{Y}^t = (Y^t_r, Z^t_r, \mathcal{V}^t_r, K^t_r)_{t \le r \le T}$ defined as in Theorem 8.1:

$$\mathcal{Y}^{t'} = (Y^{t'}_r, Z^{t'}_r, \mathcal{V}^{t'}_r, K^{t'}_r)_{t' \le r \le T} = (Y^t_r, Z^t_r, \mathcal{V}^t_r, K^t_r - K^t_{t'})_{t' \le r \le T} \; . \tag{163}$$

Proof. **(i)** By Bichteler [13, Proposition 5.2.25 page 297], the stochastic differential equation (133) with initial condition $(t', \mathcal{X}^t_{t'})$ admits a unique $(\Omega, \widetilde{\mathbb{F}}^{t'}, \mathbb{P})$ – solution $\mathcal{X}^{t'} = (X^{t'}, N^{t'})$, and it also admits a unique $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P})$ – solution, which by uniqueness is given by $\mathcal{X}^{t'}$ as well, since $\widetilde{\mathbb{F}}^{t'}$ is smaller than or equal to $\mathbb{F}^t_{|[t',T]}$. Now, $(N^t_r)_{t' \le r \le T}$ is an $\mathbb{F}^t_{|[t',T]}$-adapted process satisfying the first line of (133) on $[t', T]$. $(X^t_r)_{t' \le r \le T}$ is then in turn an $\mathbb{F}^t_{|[t',T]}$-adapted process satisfying the second line of (133) on $[t', T]$. Therefore $\mathcal{X}^{t'} = (\mathcal{X}^t_r)_{t' \le r \le T}$, by uniqueness relative to $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P})$.

(ii) Note that the bound estimate (134) on X^t is in fact valid for more general solutions of stochastic differential equations with random initial condition like $X^{t'}$ in part (i) above, by application of Proposition 5.3 (cf. proof of Proposition 7.1). We thus have for any $p \in [2, +\infty)$, with $X^{t'}$ extended by $X^{t'} = X^t_{t'}$ on $[0, t']$:

$$\|X^{t'}\|^p_{\mathcal{S}^p_d} \le C_p \left(1 + \mathbb{E}|X^t_{t'}|^p\right) \le C'_p \left(1 + |x|^p\right)$$

where the last inequality comes from (134). Consequently, (H.0)'–(H.1)'–(H.2)' in the proof of Theorem 8.1(i) still hold with $^{t'}$ (in the sense of the initial condition $(t', \mathcal{X}^t_{t'})$ for $\mathcal{X}$) instead of ^{t} therein. Given the local martingale predictable representation property of Proposition 7.5(ii) applied with t and $\widetilde{M}_t$ therein equal to t' and $\mathcal{X}^t_{t'}$ here, and in view of the form postulated in (M.3) for ℓ, existence and uniqueness of an $(\Omega, \widetilde{\mathbb{F}}^{t'}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$ – solution $\mathcal{Y}^{t'} = (Y^{t'}, Z^{t'}, \mathcal{V}^{t'}, K^{t'})$ to the R2BSDE with data (162) on $[t', T]$ follows by application of the general results of [42]. These results also imply uniqueness of an $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$ – solution to the R2BSDE with data (162) on $[t', T]$, by (H.0)'–(H.1)'–(H.2)' as above. Since $\widetilde{\mathbb{F}}^{t'}$ is smaller than or equal to $\mathbb{F}^t_{|[t',T]}$, $\mathcal{Y}^{t'} = (Y^{t'}, Z^{t'}, \mathcal{V}^{t'}, K^{t'})$ is thus the unique $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$ – solution to the R2BSDE with data (162) on $[t', T]$. Finally given part (i) it is immediate to check that $(Y^t_r, Z^t_r, \mathcal{V}^t_r, K^t_r - K^t_{t'})_{t' \le r \le T}$ is an $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$ – solution to the R2BSDE with data (162) on $[t', T]$. We conclude by uniqueness relative to $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$. □

8.2 Stopped Problem

Let τ^t denote a stopping time in $\mathcal{T}_t$, parameterized by the initial condition (t, x, i) of $\mathcal{X}$.

Theorem 8.4 **(i)** *The RDBSDE on* $[t, T]$ *with data (cf. (122))*

$$\widetilde{g}(s, \mathcal{X}^t_s, y, z, \widetilde{v}) \ , \ \Phi(\mathcal{X}^t_T) \ , \ \ell(s, \mathcal{X}^t_s) \ , \ h(s, \mathcal{X}^t_s) \ , \ \tau^t \tag{164}$$

has a unique $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t) *– solution* $\widehat{\mathcal{Y}}^t = (\widehat{Y}^t, \widehat{Z}^t, \widehat{\mathcal{V}}^t, \widehat{K}^t)$. *Moreover,* $\widehat{Y}^t = Y^t$ *on* $[\tau^t, T]$, *where* Y^t *is the state-process of the solution* $\mathcal{Y}^t$ *defined at Theorem 8.1.*
(ii) *Let us denote* $\widehat{\mathcal{V}}^t = (\widehat{V}^t, \widehat{W}^t)$ *with* $\widehat{V}^t \in \mathcal{H}^2_{\chi^t}, \widehat{W}^t \in \mathcal{H}^2_{\nu^t}$. *We extend* $\widehat{Y}^t$ *by* $\widehat{Y}^t_t$ *and* $\widehat{K}^t$, $\widehat{Z}^t$ *and* $\widehat{\mathcal{V}}^t$ *by* 0 *on* $[0, t]$, *and we define on* $[0, T]$:

$$\begin{gathered}\bar{Y}^t = \widehat{Y}^t_{\cdot \wedge \tau^t} \ , \ \bar{Z}^t = \mathbb{1}_{\cdot \le \tau^t} \widehat{Z}^t \ , \ \bar{\mathcal{V}}^t = \mathbb{1}_{\cdot \le \tau^t} \widehat{\mathcal{V}}^t, \bar{K}^t = \widehat{K}^t_{\cdot \wedge \tau^t} \\ \overline{W}^t(l, j) = \mathbb{1}_{\{l = N^t_{\cdot -}\}} \widehat{W}^t(j) \text{ for } l, j \in I \ , \ \overline{\mathcal{V}}^t = \mathbb{1}_{\cdot \le \tau^t} (\widehat{V}^t, \overline{W}^t) \\ \bar{\mathcal{Y}}^t = (\bar{Y}^t, \bar{Z}^t, \bar{\mathcal{V}}^t, \bar{K}^t) \ , \ \overline{\mathcal{Y}}^t = (\bar{Y}^t, \bar{Z}^t, \overline{\mathcal{V}}^t, \bar{K}^t) \ .\end{gathered}$$

Then (cf. (119) and (160) for the definitions of $\widetilde{g}$ *and* $\widehat{g}$*):*
- $\bar{\mathcal{Y}}^t$ *is an* $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t) *– solution to the stopped RBSDE on* $[t, T]$ *with data*

$$\widetilde{g}(s, \mathcal{X}^t_s, y, z, \widetilde{v}) \ , \ \widehat{Y}^t_{\tau^t} = Y^t_{\tau^t} \ , \ \ell(s, \mathcal{X}^t_s) \ , \ \tau^t \ , \tag{165}$$

- $\overline{\mathcal{Y}}^t$ *is an* $(\Omega, \mathbb{F}, \mathbb{P}), (B, \mu)$ *– solution to the stopped RBSDE on* $[0, T]$ *with data*

$$\mathbb{1}_{\{s > t\}} \widehat{g}(s, \mathcal{X}^t_s, y, z, \widehat{v}) \ , \ Y^t_{\tau^t} \ , \ \ell(s \vee t, \mathcal{X}^t_{s \vee t}) \ , \ \tau^t \ . \tag{166}$$

Proof. **(i)** By the general results of [42], existence of an $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t) – solution $\mathcal{Y}^t$ to the R2BSDE on $[t, T]$ with data (158) in Theorem 8.1(i) implies existence of an $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t) – solution $(\widehat{Y}^t, \widehat{Z}^t, \widehat{\mathcal{V}}^t, \widehat{K}^t)$ to the RDBSDE on $[t, T]$ with data (164), such that $\widehat{Y}^t = Y^t$ on $[\tau^t, T]$.
(ii) This implies as in the proof of Theorem 8.1(ii) that $\widehat{\mathcal{Y}}^t = (\widehat{Y}^t, \widehat{Z}^t, (V^t, \overline{W}^t), \widehat{K}^t)$, defined on $[0, T]$ as described in the statement of the theorem, is an $(\Omega, \mathbb{F}, \mathbb{P}), (B, \mu)$ – solution to the RDBSDE on $[0, T]$ with data

$$\mathbb{1}_{\{s > t\}} \widehat{g}(s, \mathcal{X}^t_s, y, z, \widehat{v}) \ , \ \Phi(\mathcal{X}^t_T) \ , \ \ell(t \vee s, \mathcal{X}^t_{s \vee t}), h(t \vee s, \mathcal{X}^t_{s \vee t}), \tau^t \ .$$

The results of part (ii) follow in view of Comments 5.1(iv). □

We work henceforth under the following standing assumption on τ^t.

Assumption 8.1 τ^t is an almost surely continuous random function of (t, x, i) on $\mathcal{E}$.

Example 8.2 Let τ^t denote the minimum of T and of the first exit time by $\mathcal{X}^t$ of a given domain $D \subseteq \mathbb{R}^d \times I$, that is:

$$\tau^t = \inf\{s \geq t\,;\, \mathcal{X}_s^t \notin D\} \wedge T \tag{167}$$

where for every $i \in I$:

$$D \cap \big(\mathbb{R}^d \times \{i\}\big) = \{\psi^i > 0\} \text{ for some } \psi^i \in \mathcal{C}^2(\mathbb{R}^d) \text{ with } |\nabla\psi^i| > 0 \text{ on } \{\psi^i = 0\}\,. \tag{168}$$

Then Assumption 8.1 is typically satisfied on a suitable *uniform ellipticity condition* on the data. For classical results in this direction, see, e.g., Darling and Pardoux [44], Dynkin [48, Theorem 13.8], Freidlin [56], or Assumption A2.2 and the related discussion in Kushner–Dupuis [77, page 281]. See also [32] for a precise statement and proof in a diffusion set-up.

Under Assumption 8.1, one has the following stability result on $\overline{\mathcal{Y}}^t$. The proof is deferred to Appendix A.3.

Proposition 8.5 **(i)** *One has the following estimate on $\overline{\mathcal{Y}}^t$ in Theorem 8.4(ii):*

$$\|\bar{Y}^t\|_{\mathcal{S}^2}^2 + \|\bar{Z}^t\|_{\mathcal{H}_d^2}^2 + \|\overline{\mathcal{V}}^t\|_{\mathcal{H}_\mu^2}^2 + \|\bar{K}^t\|_{\mathcal{S}^2}^2 \leq C(1 + |x|^{2q})\,. \tag{169}$$

(ii) *Moreover, ^{t_n} referring to a perturbed initial condition (t_n, x_n, i) with $(t_n, x_n) \to (t, x)$ as $n \to \infty$, then $\overline{\mathcal{Y}}^{t_n}$ $\mathcal{S}^2 \times \mathcal{H}_d^2 \times \mathcal{H}_\mu^2 \times \mathcal{S}^2$ – converges to $\overline{\mathcal{Y}}^t$ as $n \to \infty$.*

8.2.1 Semi-Group Properties

Let $\mathcal{X}^t = (X^t, N^t)$ and $\mathcal{Y}^t$ be defined as in Section 8.1, $\overline{\mathcal{Y}}^t = (\bar{Y}^t, \bar{Z}^t, \overline{\mathcal{V}}^t, \bar{K}^t)$ and $\widehat{\mathcal{V}}^t$ be defined as in Theorem 8.4(ii), and let $\bar{\mathcal{X}}^t = (\bar{X}^t, \bar{N}^t)$ stand for $\mathcal{X}^t_{\cdot\wedge\tau^t}$. Given $t' \geq t$, let $\bar{\mathbb{F}}^{t'} = (\bar{\mathcal{F}}_r^{t'})_{r \geq t'}$ be defined by, for $r \in [t', T]$:

$$\bar{\mathcal{F}}_r^{t'} = \sigma(\bar{\mathcal{X}}_{t'}^t) \bigvee \mathcal{F}_r^{t'}\,,$$

and let $\tau' := t' \vee \tau^t$. As for $\mathbb{F}^{t'} = (\mathcal{F}_r^{t'})_{r \geq t'}$, $\mathbb{P}^{t'}$, $B^{t'}$ and $\mu^{t'}$, they are still defined as in Sections 7.1-7.2, with t' instead of t therein. Note in particular that $\bar{\mathbb{F}}^{t'}$ is smaller than or equal to the restriction $\mathbb{F}^t_{|[t',T]}$ of $\mathbb{F}^t$ to $[t', T]$.

To proceed we make the following

Assumption 8.3 τ' is an $\bar{\mathbb{F}}^{t'}$-stopping time.

Note that Assumption 8.3 is satisfied in the case of Example 8.2.

Remark 8.4 This would not be the case if the domain D had been taken closed instead of open in (168), for instance with $\{\psi^i \geq 0\}$ instead of $\{\psi^i > 0\}$ therein.

Proposition 8.6 **(i)** *The following stochastic differential equation on* $[t', T]$:

$$\left\{\begin{array}{ll} d\bar{N}^{t'}_s & = \mathbb{1}_{s<\tau^t}\Big(\sum_{(l,j)\in I^2}(j-l)\,\mathbb{1}_{\{l=\bar{N}^{t'}_{s-}\}}\widehat{n}^{l,j}ds + \sum_{(l,j)\in I^2}(j-l)\,\mathbb{1}_{\{l=\bar{N}^{t'}_{s-}\}}d\widetilde{\nu}_s(l,j)\Big) \\ d\bar{X}^{t'}_s & = \mathbb{1}_{s<\tau^t}\Big(\widehat{b}(s,\bar{\mathcal{X}}^{t'}_s)ds + \sigma(s,\bar{\mathcal{X}}^{t'}_s)\,dB_s + \int_{\mathbb{R}^d}\delta(s,\bar{\mathcal{X}}^{t'}_{s-},,y)\,\widetilde{\chi}(ds,dy)\Big) \end{array}\right. \tag{170}$$

with initial condition $\bar{\mathcal{X}}^t_{t'}$ *at* t' *admits a unique strong* $(\Omega, \bar{\mathbb{F}}^{t'}, \mathbb{P})$ *– solution, which is given by the restriction of* $\bar{\mathcal{X}}^t$ *to* $[t', T]$*, so:*

$$\bar{\mathcal{X}}^{t'} = (\bar{X}^{t'}, \bar{N}^{t'}) = (\bar{X}^{t'}_{\cdot\wedge\tau^t}, \bar{N}^{t'}_{\cdot\wedge\tau^t}) = (\bar{\mathcal{X}}^t_r)_{t'\leq r\leq T}\ . \tag{171}$$

(ii) *The stopped RBSDE on* $[t', T]$ *with data*

$$\widetilde{g}(s, \bar{\mathcal{X}}^{t'}_s, y, z, \widetilde{v})\ ,\ Y^t_{\tau^t}\ ,\ \ell(s, \bar{\mathcal{X}}^{t'}_s)\ ,\ \tau' \tag{172}$$

has a unique $(\Omega, \bar{\mathbb{F}}^{t'}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$ *– solution* $\bar{\mathcal{Y}}^{t'} = (\bar{Y}^{t'}_r, \bar{Z}^{t'}_r, \bar{\mathcal{V}}^{t'}_r, \bar{K}^{t'}_r)_{t'\leq r\leq T}$*, given by:*

$$(\bar{Y}^{t'}_r, \bar{Z}^{t'}_r, \bar{\mathcal{V}}^{t'}_r, \bar{K}^{t'}_r)_{t'\leq r\leq T} = (\bar{Y}^t_r, \bar{Z}^t_r, \bar{\mathcal{V}}^t_r, \bar{K}^t_r - \bar{K}^t_{t'})_{t'\leq r\leq T}\ . \tag{173}$$

Proof. **(i)** By Bichteler [13, Proposition 5.2.25 page 297], the stochastic differential equation (170) with initial condition $(t', \bar{\mathcal{X}}^t_{t'})$ admits a unique $(\Omega, \bar{\mathbb{F}}^{t'}, \mathbb{P})$ – solution $\bar{\mathcal{X}}^{t'} = (\bar{X}^{t'}, \bar{N}^{t'})$, and it also admits a unique $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P})$ – solution, which by uniqueness is given by $\bar{\mathcal{X}}^{t'}$ as well, given that $\bar{\mathbb{F}}^{t'}$ is smaller than or equal to $\mathbb{F}^t_{|[t',T]}$. Now, $(\bar{N}^t_r)_{t'\leq r\leq T}$ is an $\mathbb{F}^t_{|[t',T]}$-adapted process satisfying the first line of (170) on $[t', T]$. $(\bar{X}^t_r)_{t'\leq r\leq T}$ is then in turn an $\mathbb{F}^t_{|[t',T]}$-adapted process satisfying the second line of (170) on $[t', T]$. Therefore $\bar{\mathcal{X}}^{t'} = (\bar{\mathcal{X}}^t_r)_{t'\leq r\leq T}$, by uniqueness relative to $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P})$.
(ii) One has as in the proof of Proposition 8.3(ii):

$$\|\bar{X}^{t'}\|^p_{\mathcal{S}^p_d} \leq C_p\left(1 + \mathbb{E}|\bar{X}^t_{t'}|^p\right) \leq C'_p\left(1 + |x|^p\right)\ .$$

Consequently the data

$$\mathbb{1}_{\{s<\tau'\}}\widetilde{g}(s, \bar{\mathcal{X}}^{t'}_s, y, z, \widetilde{v})\ ,\ Y^t_{\tau^t}\ ,\ \ell(s\wedge\tau', \bar{\mathcal{X}}^{t'}_{s\wedge\tau'}) \tag{174}$$

satisfy the general assumptions (H.0)–(H.1)–[assumptions regarding L in](H.2) relative to $(\Omega, \bar{\mathbb{F}}^{t'}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$ or $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$. Given the local martingale predictable representation property of $(\mathbb{F}^t, \mathbb{P}^t; B^t, \chi^t, \nu^t)$ (cf. Proposition 7.5(ii)) and the form postulated in (M.3) for ℓ, this implies existence and uniqueness of an $(\Omega, \bar{\mathbb{F}}^{t'}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$ – solution $\bar{\mathcal{Y}}^{t'} = (\bar{Y}^{t'}, \bar{Z}^{t'}, \bar{\mathcal{V}}^{t'}, \bar{K}^{t'})$ to the stopped RBSDE with data (172) on $[t', T]$, which is also the unique $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$ – solution to the stopped RBSDE with data (172) on $[t', T]$, by application of the general results of [42]. Besides, by Theorem 8.4(ii), $(\bar{Y}^t_r, \bar{Z}^t_r, \bar{\mathcal{V}}^t_r, \bar{K}^t_r)_{t\leq r\leq T}$ is an $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t) – solution to the stopped RBSDE on $[t, T]$ with data (165), where in particular $\bar{\mathcal{V}}^t = \mathbb{1}_{\cdot\leq\tau^t}\widehat{\mathcal{V}}^t$ and $\widehat{\mathcal{V}}^t = (\widehat{V}^t, \widehat{W}^t)$ for some

$\widehat{V}^t \in \mathcal{H}^2_{\chi^t}, \widehat{W}^t \in \mathcal{H}^2_{\nu^t}$. So by Definition 5.4(i):

$$\begin{cases} \bar{Y}^t_s = Y^t_{\tau^t} + \int_{s\wedge\tau^t}^{\tau^t} \widetilde{g}(r, \bar{\mathcal{X}}^t_r, \bar{Y}^t_r, \bar{Z}^t_r, \bar{\mathcal{V}}^t_r)dr + \bar{K}^t_{\tau^t} - \bar{K}^t_{s\wedge\tau^t} \\ \qquad - \int_{s\wedge\tau^t}^{\tau^t} \bar{Z}^t_r dB_r - \int_{s\wedge\tau^t}^{\tau^t}\int_{\mathbb{R}^d} \widehat{V}^t_r \widetilde{\chi}^t(dy, dr) - \sum_{j\in I}\int_{s\wedge\tau^t}^{\tau^t} \widehat{W}^t_r d\widetilde{\nu}^t_r(j)\ ,\ s \in [t, T] \\ \ell(s, \bar{\mathcal{X}}^t_s) \le \bar{Y}^t_s \text{ for } s \in [t, \tau^t]\ ,\ \text{ and } \int_t^{\tau^t} (\bar{Y}^t_s - \ell(s, \bar{\mathcal{X}}^t_s))d\bar{K}^t_s = 0 \\ \bar{Y}^t, \bar{K}^t \text{ constant on } [\tau^t, T]\ . \end{cases}$$

Therefore, given in particular (171) in part (i):

$$\begin{cases} \bar{Y}^t_s = Y^t_{\tau^t} + \int_{s\wedge\tau'}^{\tau'} \widetilde{g}(r, \bar{\mathcal{X}}^{t'}_r, \bar{Y}^t_r, \bar{Z}^t_r, \bar{\mathcal{V}}^t_r)dr + \bar{K}^t_{\tau'} - \bar{K}^t_{s\wedge\tau'} \\ \qquad - \int_{s\wedge\tau'}^{\tau'} \bar{Z}^t_r dB_r - \int_{s\wedge\tau'}^{\tau'}\int_{\mathbb{R}^d} \widehat{V}^t_r \widetilde{\chi}^t(dy, dr) - \sum_{j\in I}\int_{s\wedge\tau'}^{\tau'} \widehat{W}^t_r d\widetilde{\nu}^t_r(j)\ ,\ s \in [t', T] \\ \ell(s, \bar{\mathcal{X}}^{t'}_s) \le \bar{Y}^t_s \text{ for } s \text{ in } (t', \tau']\ ,\ \text{ and } \int_{t'}^{\tau'} (\bar{Y}^t_s - \ell(s, \bar{\mathcal{X}}^{t'}_s))d(\bar{K}^t_s - \bar{K}^t_{t'}) = 0 \\ \bar{Y}^t, \bar{K}^t - \bar{K}^t_{t'} \text{ constant on } [\tau', T]\ . \end{cases}$$

where $\ell(s, \bar{\mathcal{X}}^{t'}_s) \le \bar{Y}^t_s$ for s in $(t', \tau']$ in the third line implies that the last inequality also holds at $s = t'$, by right-continuity. So $(\bar{Y}^t_r, \bar{Z}^t_r, \bar{\mathcal{V}}^t_r, \bar{K}^t_r - \bar{K}^t_{t'})_{t'\le r\le T}$ is an $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$ – solution to the stopped RBSDE with data (172) on $[t', T]$ (cf. Definition 5.4(i)). This implies (173), by uniqueness, established above, of an $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$ – solution to the stopped RBSDE with data (172) on $[t', T]$. □

9 Markov Properties

Our next goal is to establish the Markov properties which are expected for the solutions $\mathcal{X}$ of our forward Markovian SDE and $\mathcal{Y}$, $\bar{\mathcal{Y}}$ of our backward Markovian SDEs.

Theorem 9.1 *For any initial condition $(t, x, i) \in \mathcal{E}$, let $\mathcal{Y}^t = (Y^t, Z^t, \mathcal{V}^t, K^t)$ with $\mathcal{V}^t = (V^t, W^t) \in (\mathcal{H}^2_{\chi^t}, \mathcal{H}^2_{\nu^t})$ be the $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t) – solution to the R2BSDE on $[t, T]$ with data (158) of Theorem 8.1.*
(i) *Y^t_t defines as (t, x, i) varies in $\mathcal{E}$ a continuous function u of class $\mathcal{P}$ on $\mathcal{E}$.*
(ii) *One has, $\mathbb{P}^t$-a.s. (cf. (123)–(124)–(125)):*

$$Y^t_s = u(s, \mathcal{X}^t_s)\ ,\ s \in [t, T] \tag{175}$$

$$\text{For any } j \in I:\ W^t_s(j) = u^j(s, X^t_{s-}) - u(s, \mathcal{X}^t_{s-})\ ,\ s \in [t, T] \tag{176}$$

$$\int_t^s \widetilde{g}(\zeta, \mathcal{X}^t_\zeta, Y^t_\zeta, Z^t_\zeta, \mathcal{V}^t_\zeta)d\zeta = \int_t^s \Big[g(\zeta, \mathcal{X}^t_\zeta, u(\zeta, X^t_\zeta), Z^t_\zeta, \widetilde{r}^t_\zeta) - \textstyle\sum_{j\in I} n(\zeta, \mathcal{X}^t_\zeta, j)\big(u^j(\zeta, X^t_\zeta) - u(\zeta, \mathcal{X}^t_\zeta)\big)\Big]d\zeta\ ,\ s \in [t, T] \tag{177}$$

with in (177):

$$u(\zeta, X^t_\zeta) := (u^j(\zeta, X^t_\zeta))_{j\in I}\ ,\ \widetilde{r}^t_\zeta = \int_{\mathbb{R}^d} V_\zeta(y) f(\zeta, \mathcal{X}^t_\zeta, y)m(dy)$$

(cf. (120)).

Proof. By taking $r = t' = s$ in the semi-group property (163) of $\mathcal{Y}$, one gets:

$$Y_s^t = u(s, \mathcal{X}_s^t)\ ,\ \mathbb{P}^t\!-\!a.s. \tag{178}$$

for a deterministic function u on $\mathcal{E}$. In particular,

$$Y_t^t = u^i(t,x)\ , \text{ for any } (t,x,i) \in \mathcal{E}\ . \tag{179}$$

The fact that u is of class $\mathcal{P}$ then directly follows from (179) by the bound estimate (161) on $\widetilde{\mathcal{Y}}^t$. Let $\mathcal{E} \ni (t_n, x_n) \to (t,x)$ as $n \to \infty$. We decompose

$$|u^i(t,x) - u^i(t_n,x_n)| = |Y_t^t - Y_{t_n}^{t_n}| \le |\mathbb{E}(Y_t^t - Y_{t_n}^t)| + \mathbb{E}|Y_{t_n}^t - Y_{t_n}^{t_n}|,$$

where the second term goes to 0 as $n \to \infty$ by Proposition 8.2(ii). As for the first term, one has by the R2BSDE with data (159) solved by $\widetilde{\mathcal{Y}}^t$:

$$|\mathbb{E}(Y_t^t - Y_{t_n}^t)| \le \mathbb{E}\int_{t\wedge t_n}^{t\vee t_n} |\widehat{g}(s, \mathcal{X}_s^t, Y_s^t, Z_s^t, \widetilde{\mathcal{V}}_s^t)|ds + \mathbb{E}|K_{t\vee t_n}^t - K_{t\wedge t_n}^t|$$

in which the second term goes to 0 as $n \to \infty$ by Proposition 8.2(i), and

$$\mathbb{E}\int_{t\wedge t_n}^{t\vee t_n} |\widehat{g}(s, \mathcal{X}_s^t, Y_s^t, Z_s^t, \widetilde{\mathcal{V}}_s^t)|ds \le \|\widehat{g}(\cdot, \mathcal{X}_\cdot^t, Y_\cdot^t, Z_\cdot^t, \widetilde{\mathcal{V}}_\cdot^t)\|_{\mathcal{H}^2}|t-t_n|^{\frac{1}{2}}\ ,$$

which also goes to 0 as $n \to \infty$, by the properties of g and the bound estimate (8.2) on $\widetilde{\mathcal{Y}}^t$. Finally $u^i(t_n, x_n) \to u^i(t,x)$ whenever $\mathcal{E} \ni (t_n, x_n) \to (t,x)$ as $n \to \infty$, which establishes the continuity of u on $\mathcal{E}$. Identity (175) then follows from (178) by the fact that Y^t and (given the continuity of u) $u(\cdot, \mathcal{X}_\cdot^t)$ are càdlàg processes. We then have on $\{(\omega, s);\ N_s^t \neq N_{s-}^t\}$ (whence $\chi(\mathbb{R}^d, ds) = 0$), using also the continuity of u :

$$\Delta Y_s^t = u(s, \mathcal{X}_s^t) - u(s, \mathcal{X}_{s-}^t) = \sum_{j\in I}\big(u^j(s, X_{s-}^t) - u(s, \mathcal{X}_{s-}^t)\big)\Delta\nu_s^t(j) = \sum_{j\in I} W_s(j)\Delta\nu_s^t(j)\ ,$$

where the last equality comes the R2BSDE with data (121) satisfied by $\mathcal{Y}^t$. The last equality also trivially holds on $\{(\omega, s);\ N_s^t = N_{s-}^t\}$. Denoting $\mathcal{W}_s^t(j) = u^j(s, X_{s-}^t) - u(s, \mathcal{X}_{s-}^t)$, one thus has on $[t,T]$:

$$\begin{aligned} 0 &= \sum_{j\in I}\big(\mathcal{W}_s^t(j) - W_s^t(j)\big)\Delta\nu_s^t(j) \\ &= \sum_{j\in I}\big(\mathcal{W}_s^t(j) - W_s^t(j)\big)\Delta\widetilde{\nu}_s^t(j) + \sum_{j\in I}\big(\mathcal{W}_s^t(j) - W_s(j)\big)n(s, \mathcal{X}_s^t, j)ds \end{aligned}$$

(recall (151)for the definition of $\widetilde{\nu}^t$), $\mathbb{P}^t$ – almost surely. Therefore $\mathcal{W}_s^t(j) = W_s^t(j)$ on $[t,T]$, $\mathbb{P}^t$ – almost surely, by uniqueness of the canonical decomposition of a special semimartingale. This proves (176). Now note that for $(y, z, \widetilde{v}) = (Y_s^t, Z_s^t, \mathcal{V}_s^t)$ in (120):

$$\widetilde{u}_s^t(N_s^t) = Y_s^t = u(s, \mathcal{X}_s^t)\ ,$$

by (175), and then for $j \neq N_s^t$:

$$(\widetilde{u}_s^t)^j = Y_s^t + W_s^t(j) = u(s, \mathcal{X}_s^t) + \big(u^j(s, X_{s-}^t) - u(s, \mathcal{X}_{s-}^t)\big)\ ,$$

by (176). Therefore $\widetilde{u}^t_{s-} = u(s, X^t_{s-})$, so that by definition (119) of $\widetilde{g}$:

$$\int_t^s \widetilde{g}(\zeta, \mathcal{X}^t_\zeta, Y^t_\zeta, Z^t_\zeta, \mathcal{V}^t_\zeta) d\zeta = \int_t^s \Big[g(\zeta, \mathcal{X}^t_\zeta, \widetilde{u}^t_\zeta, z, \widetilde{r}^t_\zeta) - \sum_{j\in I} W^t_\zeta(j) n(\zeta, \mathcal{X}^t_\zeta, j) \Big] d\zeta$$
$$= \int_t^s \Big[g(\zeta, \mathcal{X}^t_\zeta, \widetilde{u}^t_{\zeta-}, z, \widetilde{r}^t_\zeta) - \sum_{j\in I} (u^j(s, X^t_{\zeta-}) - u(\zeta, \mathcal{X}^t_{\zeta-})) n(\zeta, \mathcal{X}^t_\zeta, j) \Big] d\zeta$$
$$= \int_t^s \Big[g(\zeta, \mathcal{X}^t_\zeta, u(\zeta, X^t_{\zeta-}), z, \widetilde{r}^t_\zeta) - \sum_{j\in I} (u^j(\zeta, X^t_\zeta) - u(\zeta, \mathcal{X}^t_\zeta)) n(\zeta, \mathcal{X}^t_\zeta, j) \Big] d\zeta$$
$$= \int_t^s \Big[g(\zeta, \mathcal{X}^t_\zeta, u(\zeta, X^t_\zeta), z, \widetilde{r}^t_\zeta) - \sum_{j\in I} (u^j(\zeta, X^t_\zeta) - u(\zeta, \mathcal{X}^t_\zeta)) n(\zeta, \mathcal{X}^t_\zeta, j) \Big] d\zeta$$

which gives (177). □

As a by-product of Theorem 9.1, one has the following

Proposition 9.2 *$\mathcal{X}^t$ is an $(\mathbb{F}^t, \mathbb{P}^t)$-Markov process.*

Proof. In the case of a classical BSDE (without barriers) with

$$g^i(t, x, u, z, r) = \sum_{j\in I} n^{i,j}(t, x)(u_j - u_i),$$

and using also the Verification Principle of Proposition 5.2, identities (175) and (177) give:

$$Y^t_s = \mathbb{E}^t\big[\Phi(\mathcal{X}^t_T)\big|\mathcal{F}^t_s\big] = u(s, \mathcal{X}^t_s) \ ,$$

for a continuous bounded function u in $\mathcal{P}$. Therefore

$$\mathbb{E}^t\big[\Phi(\mathcal{X}^t_T)\big|\mathcal{F}^t_s\big] = \mathbb{E}^t\big[\Phi(\mathcal{X}^t_T)\big|\mathcal{X}^t_s \Sigma(\mathcal{X}^t_s)\big] \ , \tag{180}$$

where $\Sigma(\mathcal{X}^t_s)$ denotes the sigma-field generated by $\mathcal{X}^t_s$. By the monotone class theorem, (180) then holds for any Borel-measurable bounded function Φ on $\mathcal{E}$, which proves that $\mathcal{X}^t$ is an $(\mathbb{F}^t, \mathbb{P}^t)$-Markov process. □

9.1 Stopped BSDE

Theorem 9.3 *For any initial condition $(t, x, i) \in \mathcal{E}$, let $\bar{\mathcal{Y}} = (\bar{Y}^t, \bar{Z}^t, \bar{\mathcal{V}}^t, \bar{K}^t)_{t\le r\le T}$, with in particular $\bar{\mathcal{V}}^t = \mathbb{1}_{\cdot\le\tau^t}\widehat{\mathcal{V}}^t$ and $\widehat{\mathcal{V}}^t = (\widehat{V}^t, \widehat{W}^t) \in \mathcal{H}^2_{\chi^t} \times \mathcal{H}^2_{\nu^t}$, be the unique $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t) – solution to the stopped RBSDE on $[t, T]$ with data (165) of Theorem 8.4(ii).*
(i) *$\bar{Y}^t_t$ defines as (t, x, i) varies in $\mathcal{E}$ a continuous function v of class $\mathcal{P}$ on $\mathcal{E}$.*
(ii) *One has, $\mathbb{P}^t$-a.s. (cf. (126)–(127)–(128)):*

$$\bar{Y}^t_s = v(s, \mathcal{X}^t_s) \ , \ s \in [t, \tau^t] \tag{181}$$
$$v(\tau^t, \mathcal{X}^t_{\tau^t}) = u(\tau^t, \mathcal{X}^t_{\tau^t}) \tag{182}$$
$$\text{For any } j \in I : \ \widehat{W}^t_s(j) = v^j(s, X^t_{s-}) - v(s, \mathcal{X}^t_{s-}) \ \ , \ s \in [t, \tau^t] \tag{183}$$
$$\int_t^s \widetilde{g}(\zeta, \mathcal{X}^t_\zeta, \bar{Y}^t_\zeta, \bar{Z}^t_\zeta, \bar{\mathcal{V}}^t_\zeta) d\zeta = \int_t^s \Big[g(\zeta, \mathcal{X}^t_\zeta, v(\zeta, X^t_\zeta), \bar{Z}^t_\zeta, \bar{r}^t_\zeta) \tag{184}$$
$$- \textstyle\sum_{j\in I} n(\zeta, \mathcal{X}^t_\zeta, j)\big(v^j(\zeta, X^t_{\zeta-}) - v(\zeta, \mathcal{X}^t_{\zeta-})\big)\Big] d\zeta \ , \ s \in [t, \tau^t]$$

with in (184):

$$v(\zeta, X^t_\zeta) := (v^j(\zeta, X^t_\zeta))_{j\in I}\ ,\ \tilde{r}^t_\zeta := \tilde{r}^t_\zeta(\widehat{V}^t_\zeta) = \int_{\mathbb{R}^d} \widehat{V}^t_\zeta(y) f(\zeta, \mathcal{X}^t_\zeta, y) m(dy)$$

(cf. (120) for the definition of $\tilde{r}^t$*).*

Proof. By taking $r = t' = s$ in the semi-group property (173) of $\bar{\mathcal{Y}}$, one gets:

$$\bar{Y}^t_s = v(s, \bar{\mathcal{X}}^t_s)\ ,\ \mathbb{P}^t\text{−}a.s. \tag{185}$$

for a deterministic function v on $\mathcal{E}$. In particular,

$$\bar{Y}^t_t = v^i(t, x)\ ,\ \text{for any } (t, x, i) \in \mathcal{E}\ . \tag{186}$$

The fact that v is of class $\mathcal{P}$ then directly follows from the bound estimate (169) on $\overline{\mathcal{Y}}^t$. Moreover, given $\mathcal{E} \ni (t_n, x_n) \to (t, x)$ as $n \to \infty$, we decompose

$$|v^i(t, x) - v^i(t_n, x_n)| = |\bar{Y}^t_t - \bar{Y}^{t_n}_{t_n}| \le |\mathbb{E}(\bar{Y}^t_t - \bar{Y}^t_{t_n})| + \mathbb{E}|\bar{Y}^t_{t_n} - \bar{Y}^{t_n}_{t_n}|,$$

where the second term goes to 0 as $n \to \infty$ by Proposition 8.5(ii). As for the first term, one has by the stopped RBSDE with data (166) solved by $\overline{\mathcal{Y}}^t$:

$$|\mathbb{E}(\bar{Y}^t_t - \bar{Y}^t_{t_n})| \le \mathbb{E}\int_{t\wedge t_n}^{t\vee t_n} |\widehat{g}(s, \bar{\mathcal{X}}^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s)| ds + \mathbb{E}|\bar{K}^t_{t\vee t_n} - \bar{K}^t_{t\wedge t_n}|$$

in which the second term goes to 0 as $n \to \infty$ by Proposition 8.5(i), and:

$$\mathbb{E}\int_{t\wedge t_n}^{t\vee t_n} |\widehat{g}(s, \bar{\mathcal{X}}^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s)| ds \le \|\widehat{g}(\cdot, \bar{\mathcal{X}}^t_\cdot, \bar{Y}^t_\cdot, \bar{Z}^t_\cdot, \overline{\mathcal{V}}^t_\cdot)\|_{\mathcal{H}^2} |t - t_n|^{\frac{1}{2}}$$

which also goes to 0 as $n \to \infty$, by the properties of g and the bound estimate (169) on $\overline{\mathcal{Y}}^t$. So $v^i(t_n, x_n) \to v^i(t, x)$ whenever $\mathcal{E} \ni (t_n, x_n) \to (t, x)$ as $n \to \infty$, which establishes the continuity of v on $\mathcal{E}$. Identity (181) then follows from (185) by the fact that $\bar{Y}^t$ and (given the continuity of v) $v(\cdot, \mathcal{X}^t_\cdot)$, are càdlàg processes. Given that $\bar{Y}^t_{\tau^t} = Y^t_{\tau^t}$ (cf. Theorem 8.4(ii)), (181) and (175) in turn imply (182). One has further on $\{(\omega, s)\,;\ s \in [t, \tau^t],\, N^t_s \neq N^t_{s-}\}$ (whence $\chi(\mathbb{R}^d, ds) = 0$), using also the continuity of v :

$$\Delta\bar{Y}^t_s = v(s, \mathcal{X}^t_s) - v(s, \mathcal{X}^t_{s-}) = \sum_{j\in I}\big(v^j(s, X^t_{s-}) - v(s, \mathcal{X}^t_{s-})\big)\Delta\nu^t_s(j) = \sum_{j\in I}\widehat{W}_s(j)\Delta\nu^t_s(j)$$

where the last equality comes the stopped RBSDE on $[t, T]$ with data (165) solved by $\bar{\mathcal{Y}}^t$. The last equality also trivially holds on $\{(\omega, s)\,;\ s \in [t, \tau^t],\, N^t_s = N^t_{s-}\}$. Denoting $\mathcal{W}^t_s(j) := v^j(s, X^t_{s-}) - v(s, \mathcal{X}^t_{s-})$, one thus has, on $[t, \tau^t]$:

$$\begin{aligned} 0 &= \sum_{j\in I}\big(\mathcal{W}^t_s(j) - \widehat{W}^t_s(j)\big)\Delta\nu^t_s(j) \\ &= \sum_{j\in I}\big(\mathcal{W}^t_s(j) - \widehat{W}^t_s(j)\big)\Delta\tilde{\nu}^t_s(j) + \sum_{j\in I}\big(\mathcal{W}^t_s(j) - \widehat{W}_s(j)\big) n(s, \mathcal{X}^t_s, j) ds \end{aligned}$$

(recall (151)for the definition of $\tilde{\nu}^t$), $\mathbb{P}^t$ − almost surely. Therefore $\mathcal{W}^t_s(j) = \widehat{W}_s(j)$ on $[t, \tau^t]$, by uniqueness of the canonical decomposition of a special semimartingale, whence (183).

Finally (184) follows from (181) and (183) like (177) from (175) and (176) (cf. proof of (177)). □

In summary, one has established in Sections 7 to 9 the following proposition relative to the main FBSDE of Section 6 (cf. Definition 6.4).

Proposition 9.4 *Under the assumptions of Sections 7 to 9, the Markovian FBSDE with generator $\mathcal{G}$, cost functions $\mathcal{C}$ and (parameterized) stopping time τ has a consistent solution* $\mathcal{Z}^t = (\Omega, \mathbb{F}^t, \mathbb{P}^t)$, $(B^t, \chi^t, \nu^t), (\mathcal{X}^t, \mathcal{Y}^t, \bar{\mathcal{Y}}^t)$.

The related assumptions, based on the Markovian change of probability measure defined by (148)–(147) (see Assumption 7.6), are admittedly technical and involved, and by no means minimal. ***Therefore for more clarity in the sequel we shall give up all these specific assumptions, merely postulating instead that the main FBSDE of Section 6 has a consistent solution in the sense of Definition 6.4*** (as is for instance the case under the assumptions of Sections 7 to 9).

Part III

Main PDE Results

In this part (see Section 1 for a detailed outline), we derive the companion *variational inequality approach* to the BSDE approach of Part II, working in a suitable space of *viscosity solutions* to the associated *systems of partial integro-differential obstacle problems*.
The results obtained in this part are used in Part I for giving a constructive and computational counterpart to the theoretical BSDE results of Section 2, in the Markovian factor process set-up of Sections 3, 4.1 or 4.2.4.

As announced at the end of Part II, we give up all the specific assumptions made in Sections 7 to 9. We make instead the following standing

Assumption 9.1 The Markovian FBSDE with data $\mathcal{G}$, $\mathcal{C}$, τ has a consistent solution $\mathcal{Z}^t = (\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, χ^t, ν^t), $(\mathcal{X}^t, \mathcal{Y}^t, \bar{\mathcal{Y}}^t)$.

As illustrated in the previous sections, Assumption 9.1 covers various issues such as Lipschitz continuity properties of the forward SDE coefficients b, σ, δ with respect to x, martingale representation properties, some kind of consistency between the drivers B^t, χ^t, ν^t as t varies, and almost sure continuity of the random function τ^t of (t, x, i) on $\mathcal{E}$.

10 Viscosity Solutions of Systems of PIDEs with Obstacles

Our next goal is to establish the connection between $\mathcal{Z}$ and related systems of obstacle problems associated to the data $\mathcal{G}$, $\mathcal{C}$, τ, problems denoted by $(\mathcal{V}1)$ and $(\mathcal{V}2)$ below. In this article we shall consider this issue from the point of view of *viscosity solutions* to the related systems of obstacle problems. We refer the reader to the books by Bensoussan and Lions [11, 12] for alternative results in spaces of weak Sobolev solutions (see also [11, 12, 8, 5, 4]).

We postulate from now on that

Assumption 10.1 **(i)** All the $(t, x, i)-$coefficients of the generator $\mathcal{G}$ are continuous functions;
(ii) The functions δ and f are locally Lipschitz continuous with respect to (t, x), uniformly in y, i;
(iii) τ^t is defined as in our standing Example 8.2 in Part II, so

$$\tau^t = \inf\{s \geq t\,;\, \mathcal{X}^t_s \notin D\} \wedge T \tag{187}$$

where for every $i \in I$:

$$D \cap \left(\mathbb{R}^d \times \{i\}\right) = \{\psi^i > 0\} \text{ for some } \psi^i \in \mathcal{C}^2(\mathbb{R}^d) \text{ with } |\nabla \psi^i| > 0 \text{ on } \{\psi^i = 0\}\ . \tag{188}$$

Let $\mathcal{D} = [0, T] \times \bar{D}$, where $\bar{D}$ denotes the closure[7] of D in $\mathbb{R}^d \times I$. Let also

$$\begin{aligned} \mathrm{Int}_p\mathcal{E} = [0, T) \times \mathbb{R}^d \times I\ ,\ \partial_p\mathcal{E} := \mathcal{E} \setminus \mathrm{Int}_p\mathcal{E} \\ \mathrm{Int}_p\mathcal{D} = [0, T) \times D\ ,\ \partial_p\mathcal{D} := \mathcal{E} \setminus \mathrm{Int}_p\mathcal{D} \end{aligned} \tag{189}$$

[7]In the sense that for every $i \in I$, $\bar{D} \cap \mathbb{R}^d \times \{i\}$ is the closure of $D \cap (\mathbb{R}^d \times \{i\})$, identified to a subset of $\mathbb{R}^d$.

stand for the *parabolic interior* and the *parabolic boundary* of $\mathcal{E}$ and $\mathcal{D}$, respectively.

Remark 10.2 The definition of a 'thick' boundary $\partial_p\mathcal{D} = \mathcal{E} \setminus \mathrm{Int}_p\mathcal{D}$ is made necessary by the presence of the jumps in X.

Given locally bounded test-functions ϕ and φ on $\mathcal{E}$ with φ of class $\mathcal{C}^{1,2}$ around a given point $(t,x,i) \in \mathcal{E}$, we define (cf. (113)–(114)):

$$\widetilde{\mathcal{G}}(\phi,\varphi)^i(t,x) = \partial_t\varphi^i(t,x) + \frac{1}{2}Tr\big[a^i(t,x)\mathcal{H}\varphi^i(t,x)\big] + \partial\varphi^i(t,x)\widetilde{b}^i(t,x) + \mathcal{I}\phi^i(t,x) \quad (190)$$

with

$$\mathcal{I}\phi^i(t,x) := \int_{\mathbb{R}^d}\Big(\phi^i(t,x+\delta^i(t,x,y)) - \phi^i(t,x)\Big)f^i(t,x,y)m(dy)\ . \quad (191)$$

Let also $\widetilde{\mathcal{G}}\varphi$ stand for $\widetilde{\mathcal{G}}(\varphi,\varphi)$. So in particular (cf. (113)):

$$\widetilde{\mathcal{G}}\varphi^i(t,x) + \sum_{j\in I} n^{i,j}(t,x)\Big(\varphi^j(t,x) - \varphi^i(t,x)\Big) = \mathcal{G}\varphi^i(t,x)\ . \quad (192)$$

The problems $(\mathcal{V}2)$ and $(\mathcal{V}1)$ that we now introduce will ultimately constitute a cascade of two PDEs, inasmuch as the boundary (including terminal) condition Ψ in the Cauchy–Dirichlet problem $(\mathcal{V}1)$ will be specified later in the paper as the value function u of Definition 6.4 (cf. Assumption 9.1), characterized as the unique viscosity solution of class $\mathcal{P}$ of $(\mathcal{V}2)$.

We thus denote by $(\mathcal{V}2)$ the following variational inequality with double obstacle problem:

$$\max\Big(\min\Big(-\widetilde{\mathcal{G}}u^i(t,x) - g^i(t,x,u(t,x),(\partial u\sigma)^i(t,x),\mathcal{I}u^i(t,x)),\\ u^i(t,x) - \ell^i(t,x)\Big), u^i(t,x) - h^i(t,x)\Big) = 0$$

on $\mathrm{Int}_p\mathcal{E}$, supplemented by the terminal condition Φ (the terminal cost function in the cost data $\mathcal{C}$) on $\partial_p\mathcal{E}$. We also consider the problem $(\mathcal{V}1)$ obtained by formally replacing h by $+\infty$ in $(\mathcal{V}2)$, that is

$$\min\Big(-\widetilde{\mathcal{G}}u^i(t,x) - g^i(t,x,u(t,x),(\partial u\sigma)^i(t,x),\mathcal{I}u^i(t,x)),\\ u^i(t,x) - \ell^i(t,x)\Big) = 0\ ,$$

on $\mathrm{Int}_p\mathcal{D}$, supplemented by a continuous boundary condition Ψ extending Φ on $\partial_p\mathcal{D}$.

Remark 10.3 We write 'on $\partial_p\mathcal{E}$' rather than 'at T' for $(\mathcal{V}2)$ in order to emphasize the formal analogy between the Cauchy problem $(\mathcal{V}2)$ and the Cauchy–Dirichlet problem $(\mathcal{V}1)$.

The following *continuity property of the integral term* $\mathcal{I}$ of $\widetilde{\mathcal{G}}$ (cf. (191)) is key in the theory of viscosity solutions of nonlinear integro-differential equations (see for instance Alvarez–Tourin [1, page 297]).

Lemma 10.1 *The function* $(t,x,i) \to \mathcal{I}\psi^i(t,x)$ *is continuous on* $\mathcal{E}$*, for any continuous function* ψ *on* $\mathcal{E}$*.*

Proof. One decomposes

$$\mathcal{I}\psi^i(t_n,x_n)-\mathcal{I}\psi^i(t,x)=-\int_{\mathbb{R}^d}\Big(\psi^i(t_n,x_n)f^i(t_n,x_n,y)-\psi^i(t,x)f^i(t,x,y)\Big)m(dy)$$
$$+\int_{\mathbb{R}^d}\Big(\psi^i(t_n,x_n+\delta^i(t_n,x_n,y))f^i(t_n,x_n,y)-\psi^i(t,x+\delta^i(t,x,y))f^i(t,x,y)\Big)m(dy)\ ,$$

where

$$\begin{aligned}&\textstyle\int_{\mathbb{R}^d}\Big(\psi^i(t_n,x_n+\delta^i(t_n,x_n,y))f^i(t_n,x_n,y)-\psi^i(t,x+\delta^i(t,x,y))f^i(t,x,y)\Big)m(dy)\\ &=\textstyle\int_{\mathbb{R}^d}\Big(\psi^i(t_n,x_n+\delta^i(t_n,x_n,y))-\psi^i(t,x+\delta^i(t,x,y))\Big)f^i(t_n,x_n,y)m(dy)\\ &\quad+\textstyle\int_{\mathbb{R}^d}\psi^i(t,x+\delta^i(t,x,y))\Big(f^i(t_n,x_n,y)-f^i(t,x,y)\Big)m(dy)\end{aligned} \tag{193}$$

goes to 0 as $\mathcal{E}\ni(t_n,x_n)\to(t,x)$, by Assumption 10.1(ii), and likewise for

$$\int_{\mathbb{R}^d}\Big(\psi^i(t_n,x_n)f^i(t_n,x_n,y)-\psi^i(t,x)f^i(t,x,y)\Big)m(dy)\ .$$

□

The following definitions are obtained by specifying to problems $(\mathcal{V}1)$ and $(\mathcal{V}2)$ the general definitions of viscosity solutions for nonlinear PDEs (see, for instance, Crandall et al. [39] or Fleming and Soner [54]), adapting further the resulting definitions to finite activity jumps and *systems* of PIDEs as in [1, 87, 6, 30, 65].

Definition 10.4 (a)(i) A locally bounded upper, resp. lower semi-continuous, function u on $\mathcal{E}$, is called a *viscosity subsolution*, resp. *supersolution*, of $(\mathcal{V}2)$ at $(t,x,i)\in\mathrm{Int}_p\mathcal{E}$, if and only if for any $\varphi\in\mathcal{C}^{1,2}(\mathcal{E})$ such that $u^i-\varphi^i$ reaches a global maximum, resp. minimum, at (t,x), one has,

$$\max\Big(\min\Big(-\widetilde{\mathcal{G}}(u,\varphi)^i(t,x)-g^i(t,x,u(t,x),(\partial\varphi\sigma)^i(t,x),\mathcal{I}u^i(t,x)),$$
$$u^i(t,x)-\ell^i(t,x)\Big),u^i(t,x)-h^i(t,x)\Big)\le 0,\ \text{resp.}\ \ge 0\ .$$

Equivalently, u is a viscosity subsolution, resp. supersolution, of $(\mathcal{V}2)$ at (t,x,i), if and only if $u^i(t,x)\le h^i(t,x)$, resp. $u^i(t,x)\ge\ell^i(t,x)$, and if $u^i(t,x)>\ell^i(t,x)$, resp. $u^i(t,x)<h^i(t,x)$, implies that

$$-\widetilde{\mathcal{G}}(u,\varphi)^i(t,x)-g^i(t,x,u(t,x),(\partial\varphi\sigma)^i(t,x),\mathcal{I}u^i(t,x))\le 0,\ \text{resp.}\ \ge 0\ , \tag{194}$$

or inequality (194) with $\widetilde{\mathcal{G}}(u,\varphi)$ and $\mathcal{I}u$ replaced by $\widetilde{\mathcal{G}}\varphi$ and $\mathcal{I}\varphi$, for any $\varphi\in\mathcal{C}^{1,2}(\mathcal{E})$ such that $u^i-\varphi^i$ reaches a global null maximum, resp. minimum, at (t,x), or, in turn, with *global null maximum*, resp. *minimum*, replaced therein by *global null* strict *maximum*, resp. *minimum*.
(ii) A continuous function u on $\mathcal{E}$ is called a *viscosity solution* of $(\mathcal{V}2)$ at $(t,x,i)\in\mathrm{Int}_p\mathcal{E}$, if and only if it is both a viscosity subsolution and a viscosity supersolution of $(\mathcal{V}2)$ at (t,x,i).

(b)(i) By a $\mathcal{P}$ – *viscosity subsolution*, resp. supersolution, u of $(\mathcal{V}2)$ on $\mathcal{E}$ for the boundary condition Φ, we mean an upper, resp. lower semi-continuous function of class $\mathcal{P}$ on $\mathcal{E}$, which is a viscosity subsolution, resp. supersolution of $(\mathcal{V}2)$ on $\mathrm{Int}_p\mathcal{E}$, and such that $u\le\Phi$, resp.

$u \geq \Phi$ pointwise on $\partial_p \mathcal{E}$.
(ii) By a $\mathcal{P}$ – *viscosity solution* u of $(\mathcal{V}2)$ on $\mathcal{E}$, we mean a function that is both a $\mathcal{P}$-subsolution and a $\mathcal{P}$-supersolution of $(\mathcal{V}2)$ on $\mathcal{E}$ — hence $u = \Phi$ on $\partial_p \mathcal{E}$.

(c) The notions of viscosity subsolutions, supersolutions and solutions of $(\mathcal{V}1)$ at $(t, x, i) \in \mathrm{Int}_p \mathcal{D}$, and, given a continuous boundary condition Ψ extending Φ on $\partial_p \mathcal{D}$, $\mathcal{P}$ – viscosity subsolutions, supersolutions and solutions of $(\mathcal{V}1)$ on $\mathcal{E}$, are defined by immediate adaptation of parts (a) and (b) above, substituting $(\mathcal{V}1)$ to $(\mathcal{V}2)$, $+\infty$ to h, $\mathrm{Int}_p \mathcal{D}$ to $\mathrm{Int}_p \mathcal{E}$, $\mathcal{C}^0(\mathcal{E}) \cap \mathcal{C}^{1,2}(\mathcal{D})$ to $\mathcal{C}^{1,2}(\mathcal{E})$, $\partial_p \mathcal{D}$ to $\partial_p \mathcal{E}$ and Ψ to Φ therein.

Comments 10.1 **(i)** We thus consider boundary conditions in the classical sense, rather than in the weak viscosity sense (cf. the proof of Lemma 13.2(ii) for more on this issue, see also Crandall et al. [39]).
(ii) A classical solution (if any) of $(\mathcal{V}1)$, resp. $(\mathcal{V}2)$, is necessarily a viscosity solution of $(\mathcal{V}1)$, resp. $(\mathcal{V}2)$.
(iii) A viscosity solution u of $(\mathcal{V}2)$ necessarily satisfies $\ell \leq u \leq h$. However a viscosity subsolution (resp. supersolution) u of $(\mathcal{V}2)$ does not need to verify $u \geq \ell$ (resp. $u \leq h$). Likewise a viscosity solution v of $(\mathcal{V}1)$ necessarily satisfies $\ell \leq u$, however a viscosity subsolution v of $(\mathcal{V}1)$ does not need to verify $u \geq \ell$.
(iv) The fact that $\widetilde{\mathcal{G}}(u, \varphi)$ and $\mathcal{I}u$ may equivalently be replaced by $\widetilde{\mathcal{G}}\varphi$ and $\mathcal{I}\varphi$ in (194), or in the analogous inequalities regarding $(\mathcal{V}1)$, can be shown by an immediate adaptation to the present context of Barles et al. [6, Lemma 3.3 page 66] (see also 'Definition 2 (Equivalent)' page 300 in Alvarez–Tourin [1]), using the monotonicity assumption (M.1.iii) on g.

Since we only consider solutions in the viscosity sense in this article, (resp. $\mathcal{P}$ –) subsolution, supersolution and solution are to be understood henceforth as (resp. $\mathcal{P}$ –) **viscosity** *subsolution, supersolution and solution.*

11 Existence of a Solution

The value functions u and v appearing in the following results are the ones introduced in Definition 6.4, under Assumption 9.1. This result establishes that u and v are viscosity solutions of the related obstacle problems, with u as boundary Dirichlet condition for v on $\partial_p \mathcal{D}$.

Theorem 11.1 **(i)** *The value function u is a $\mathcal{P}$-solution of $(\mathcal{V}2)$ on $\mathcal{E}$ for the terminal condition Φ on $\partial_p \mathcal{E}$.*
(ii) *The value function v is a $\mathcal{P}$-solution of $(\mathcal{V}1)$ on $\mathcal{E}$ for the boundary condition u on $\partial_p \mathcal{D}$.*

Proof. **(i)** By definition, u is a continuous function of class $\mathcal{P}$ on $\mathcal{E}$. Moreover by definition of u and $\mathcal{Y}$ one has that, the superscript ^{T} referring to an initial condition (T, x, i) for $\mathcal{X}$:

$$u^i(T, x) = Y_T^T = \Phi^i(x)$$
$$\ell^i(t, x) \leq Y_t^t = u^i(t, x) \leq h^i(t, x) \; .$$

So $u = \Phi$ pointwise at T and $\ell \leq u \leq h$ on $\mathcal{E}$. Let us show that u is a subsolution of $(\mathcal{V}2)$ on $\mathrm{Int}_p \mathcal{E}$. We let the reader check likewise that u is a supersolution of $(\mathcal{V}2)$ on $\mathrm{Int}_p \mathcal{E}$. Let thus

$(t,x,i) \in \mathrm{Int}_p\mathcal{E}$ and $\varphi \in \mathcal{C}^{1,2}(\mathcal{E})$ be such that $u^i - \varphi^i$ reaches its maximum at (t,x). Given that $u \le h$, it suffices to prove that

$$-\widetilde{\mathcal{G}}\varphi^i(t,x) - g^i(t,x,u(t,x),(\partial\varphi\sigma)^i(t,x),\mathcal{I}\varphi^i(t,x)) \le 0\ , \tag{195}$$

assuming further that $u^i(t,x) > \ell^i(t,x)$ and $u^i(t,x) = \varphi^i(t,x)$ (cf. Definition 10.4(a)(i)). Suppose by contradiction that (195) does not hold. Then by a continuity argument using in particular Lemma 10.1:

$$\psi(s,y) := \widetilde{\mathcal{G}}\varphi^i(s,y) + g^i(s,y,u(s,y),(\partial\varphi^i\sigma)^i(s,y),\mathcal{I}\varphi^i(s,y)) < 0 \tag{196}$$

for any (s,y) such that $s \in [t,t+\alpha]$ and $|y-x| \le \alpha$, for some small enough $\alpha > 0$ with $t+\alpha < T$. Let

$$\theta = \inf\left\{s \ge t\,;\, |X^t_s - x| \ge \alpha,\ N^t_s \neq i,\ Y^t_s = \ell^i(s,X^t_s)\right\} \wedge (t+\alpha) \tag{197}$$

$$(\widehat{Y}^t,\widehat{Z}^t,\widehat{V}^t,\widehat{\mathcal{V}}^t,\widehat{K}^t) = \Big(\mathbb{1}_{\cdot<\theta}Y^t + \mathbb{1}_{\cdot\ge\theta}u^i(\theta,X^t_\theta)\ ,\ \mathbb{1}_{\cdot\le\theta}Z^t\ ,\ \mathbb{1}_{\cdot\le\theta}V^t\ ,\ \mathbb{1}_{\cdot\le\theta}\mathcal{V}^t\ ,\ K^t_{\cdot\wedge\theta}\Big) \tag{198}$$

$$\begin{aligned}(\widetilde{Y}^t,\widetilde{Z}^t,\widetilde{V}^t) = \Big(&\varphi^i(\cdot X^t_{\cdot\wedge\theta})\ ,\ \mathbb{1}_{\cdot\le\theta}(\partial\varphi\sigma)^i(\cdot,X^t_\cdot)\ ,\\ &\mathbb{1}_{\cdot\le\theta}\big(\big[\varphi^i(\cdot,X^t_{\cdot-} + \delta^i(\cdot,X^t_{\cdot-},y)) - \varphi^i(\cdot,X^t_{\cdot-})\big]\big)_{y\in\mathbb{R}^d}\Big)\ .\end{aligned} \tag{199}$$

Note that $\theta > t$, $\mathbb{P}^t$ – almost surely. Thus, using also the continuity of u^i :

$$\widehat{Y}^t_t = Y^t_t = u^i(t,x) = \varphi^i(t,x) = \widetilde{Y}^t_t\ ,\ \mathbb{P}^t\text{-a.s.} \tag{200}$$

Moreover, by using the R2BSDE equation for $\mathcal{Y}^t$ (in which $K^{t;+} = 0$ on $[t,\theta]$ by the related minimality condition, given that $\ell^i(s,X^t_s) < Y^t_s$ on $[t,\theta)$), one has for $s \le r$ in $[t,\theta)$:

$$\begin{aligned}\widehat{Y}^t_s = u^i(r,X^t_r) &+ \int_s^r g^i(\zeta,X^t_\zeta,u(\zeta,X^t_\zeta),\widehat{Z}^t_\zeta,\widetilde{r}^t_\zeta)d\zeta - (\widehat{K}^{t;-}_r - \widehat{K}^{t;-}_s)\\ &- \int_s^r \widehat{Z}^t_\zeta dB^t_\zeta - \int_s^r\int_{\mathbb{R}^d}\widehat{V}^t_\zeta(y)\widetilde{\chi}^t(dy,d\zeta)\ ,\end{aligned} \tag{201}$$

which also holds true for $s = r = \theta$ by definition of $\widehat{Y}^t$ in (198). Furthermore for $s < r = \theta$:
• either χ^t, whence X^t, do not jump at θ, and (201) holds again at $r = \theta$ by passage to the limit as $r \uparrow \theta$ in (201),
• or (cf. Definition 6.2(i)) N^t does not jump at θ, in which case the R2BSDE equation for $\mathcal{Y}^t$ integrated between s and θ directly gives (201) for $r = \theta$.
In conclusion (201) holds for $s \le r$ in $[t,\theta]$.

Besides, by application of the Itô formula (117) to the function $\widetilde{\varphi}$ defined by $\widetilde{\varphi}^j = \varphi^i$ for all $j \in I$, one gets for any $s \in [t,\theta]$:

$$\begin{aligned}d\varphi^i(s,X^t_s) &= \mathcal{G}\widetilde{\varphi}(s,\mathcal{X}^t_s)ds + (\partial\varphi\sigma)(s,\mathcal{X}^t_s)dB^t_s\\ &\qquad + \int_{\mathbb{R}^d}\big(\varphi^i(s,X^t_{s-} + \delta(s,\mathcal{X}^t_{s-},y)) - \varphi^i(s,\mathcal{X}^t_{s-})\big)\widetilde{\chi}^t(ds,dy)\\ &= \widetilde{\mathcal{G}}\widetilde{\varphi}(s,\mathcal{X}^t_s)ds + (\partial\varphi\sigma)(s,\mathcal{X}^t_s)dB^t_s\\ &\qquad + \int_{\mathbb{R}^d}\big(\varphi^i(s,X^t_{s-} + \delta(s,\mathcal{X}^t_{s-},y)) - \varphi^i(s,\mathcal{X}^t_{s-})\big)\widetilde{\chi}^t(ds,dy)\\ &= \widetilde{\mathcal{G}}\varphi^i(s,X^t_s)ds + (\partial\varphi\sigma)^i(s,X^t_s)dB^t_s\\ &\qquad + \int_{\mathbb{R}^d}\big(\varphi^i(s,X^t_{s-} + \delta^i(s,X^t_{s-},y)) - \varphi^i(s,\mathcal{X}^t_{s-})\big)\widetilde{\chi}^t(ds,dy)\ ,\end{aligned}$$

where the second equality uses (192) applied to $\widetilde{\varphi}$ and the third one exploits the facts that N^t cannot jump before θ and that $\widetilde{\chi}^t$ cannot jump at θ if N^t does. Hence (cf. (199)):

$$\begin{aligned}\widetilde{Y}_s^t &= \varphi^i(\theta, X_\theta^t) - \int_s^\theta \widetilde{\mathcal{G}}\varphi^i(r, X_r^t)dr - \int_s^\theta \widetilde{Z}_r^t dB_r^t - \int_s^\theta \int_{\mathbb{R}^d} \widetilde{V}_r^t(y)\widetilde{\chi}^t(dy, dr) \\ &= \varphi^i(\theta, X_\theta^t) - \int_s^\theta \Big(\psi(r, X_r^t) - g^i(r, X_r^t, u(r, X_r^t), (\partial\varphi)\sigma^i(r, X_r^t), \mathcal{I}\varphi^i(r, X_r^t))\Big)dr \\ &\quad - \int_s^\theta \widetilde{Z}_r^t dB_r^t - \int_s^\theta \int_{\mathbb{R}^d} \widetilde{V}_r^t(y)\widetilde{\chi}^t(dy, dr) \ ,\end{aligned}$$

by definition (196) of ψ. In conclusion one has for $s \in [t, \theta]$:

$$\begin{aligned}\widehat{Y}_s^t &= u^i(\theta, X_\theta^t) + \int_s^\theta g^i(\zeta, X_\zeta^t, u(\zeta, X_\zeta^t), \widehat{Z}_\zeta^t, \widetilde{r}_\zeta^t)d\zeta - (\widehat{K}_\theta^{t;-} - \widehat{K}_s^{t;-}) \\ &\quad - \int_s^\theta \widehat{Z}_\zeta^t dB_\zeta^t - \int_s^\theta \int_{\mathbb{R}^d} \widehat{V}_\zeta^t(y)\widetilde{\chi}^t(dy, d\zeta) \qquad (202) \\ \widetilde{Y}_s^t &= \varphi^i(\theta, X_\theta^t) - \int_s^\theta \Big(\psi(\zeta, X_\zeta^t) - g^i(\zeta, X_\zeta^t, u(\zeta, X_\zeta^t), \widetilde{Z}_\zeta^t, \mathcal{I}\varphi^i(\zeta, X_\zeta^t))\Big)d\zeta \\ &\quad - \int_s^\theta \widetilde{Z}_\zeta^t dB_\zeta^t - \int_s^\theta \int_{\mathbb{R}^d} \widetilde{V}_\zeta^t(y)\widetilde{\chi}^t(dy, d\zeta) \qquad (203)\end{aligned}$$

with by definitions (120) of $\widetilde{r}_\zeta^t = \widetilde{r}_\zeta^t(V_\zeta^t)$, (191) of $\mathcal{I}$ and (199) of $\widetilde{V}$:

$$\begin{aligned}\int_s^\theta g^i\Big(\zeta, X_\zeta^t, u(\zeta, X_\zeta^t), \widehat{Z}_\zeta^t, \widetilde{r}_\zeta^t\Big)d\zeta &= \int_s^\theta g^i\Big(\zeta, X_\zeta^t, u(\zeta, X_\zeta^t), \widehat{Z}_\zeta^t, \int_{\mathbb{R}^d} \widehat{V}_\zeta(y) f^i(\zeta, X_\zeta^t, y)m(dy)\Big)d\zeta \\ \int_s^\theta g^i\Big(\zeta, X_\zeta^t, u(\zeta, X_\zeta^t), \widetilde{Z}_\zeta^t, \mathcal{I}\varphi^i(\zeta, X_\zeta^t\Big)d\zeta &= \int_s^\theta g^i\Big(\zeta, X_\zeta^t, u(\zeta, X_\zeta^t), \widetilde{Z}_\zeta^t, \int_{\mathbb{R}^d} \widetilde{V}_\zeta(y) f^i(\zeta, X_\zeta^t, y)m(dy)\Big)d\zeta \ .\end{aligned}$$

In other words:

• $(\widehat{Y}^t, \widehat{Z}^t, \widehat{V}^t)$ solves the stopped BSDE on $[t, t+\alpha]$ with driver (cf. Definition 5.3(d) and Comments 5.1(i))

$$g^i\Big(s, X_s^t, u(s, X_s^t), z, \int_{\mathbb{R}^d} v(y) f^i(s, X_s^t, y)m(dy)\Big)ds - d\widehat{K}_s^{t;-}$$

and terminal condition $u^i(\theta, X_\theta^t)$ at θ;

• $(\widetilde{Y}^t, \widetilde{Z}^t, \widetilde{V}^t)$ solves the stopped BSDE on $[t, t+\alpha]$ with driver

$$g^i\Big(s, X_s^t, u(s, X_s^t), z, \int_{\mathbb{R}^d} v(y) f^i(s, X_s^t, y)m(dy)\Big)ds - \psi(s, X_s^t)ds$$

and terminal condition $\varphi^i(\theta, X_\theta^t)$ at θ.

Setting $\delta Y^t = \widehat{Y}^t - \widetilde{Y}^t$, we deduce by standard computations (see for instance the proof of the comparison principle in [42]):

$$\Gamma_t^t \delta Y_t^t = \mathbb{E}^t\Big[\Gamma_\theta^t \delta Y_\theta^t + \int_t^\theta \Gamma_s^t dA_s^t\Big] \qquad (204)$$

where:

• $\delta Y_\theta^t = \widehat{Y}_\theta^t - \widetilde{Y}_\theta^t = u^i(\theta, X_\theta^t) - \varphi^i(\theta, X_\theta^t) \leq 0$, by making $s = \theta$ in (202)–(203) and since

$u^i \le \varphi^i$;
• $dA^t_s = \psi(r, X^t_s)ds - d\widehat{K}^{t;-}_s$, so that A^t is (strictly) decreasing on $[t, \theta]$, by (196);
• Γ^t is a (strictly) positive process, the so-called *adjoint* of δY^t (see, for instance, [42]).
Since furthermore $\theta > t$ $\mathbb{P}^t$-a.s., we deduce that $\int_t^\theta \Gamma^t_s dA^t_s < 0$ $\mathbb{P}^t$-a.s., whence $\delta Y^t_t < 0$, by (204). But this contradicts (200).

(ii) v is a continuous function on $\mathcal{E}$, by definition. Moreover by definitions of u, v, $\mathcal{Y}$ and $\bar{\mathcal{Y}}$ (including the definition (187) of τ), we have, for $(t, x, i) \in \partial_p\mathcal{D}$:

$$v^i(t, x) = \bar{Y}^t_t = Y^t_t = u^i(t, x) ,$$

and for any $(t, x, i) \in \mathcal{E}$:

$$\ell^i(t, x) \le \bar{Y}^t_t = v^i(t, x) .$$

So $v = u$ on $\partial_p\mathcal{D}$ and $\ell \le v$ on $\mathcal{E}$. We now show that v is a subsolution of $(\mathcal{V}1)$ on $\mathrm{Int}_p\mathcal{D}$. We let the reader check likewise that v is a supersolution of $(\mathcal{V}1)$ on $\mathrm{Int}_p\mathcal{D}$. Let then $(t, x, i) \in \mathrm{Int}_p\mathcal{D}$ and $\varphi \in \mathcal{C}^0(\mathcal{E}) \cap \mathcal{C}^{1,2}(\mathcal{D})$ be such that $v^i - \varphi^i$ reaches its maximum at (t, x). We need to prove that

$$-\widetilde{\mathcal{G}}\varphi^i(t, x) - g^i(t, x, v(t, x), (\partial\varphi)\sigma^i(t, x), \mathcal{I}\varphi^i(t, x)) \le 0 , \tag{205}$$

assuming further $v^i(t, x) > \ell^i(t, x)$ and $v^i(t, x) = \varphi^i(t, x)$ (cf. Definition 10.4(a)(i)). Suppose by contradiction that (205) does not hold. Then by continuity $(s, y, i) \in \mathrm{Int}_p\mathcal{D}$ and

$$\psi(s, y) := \widetilde{\mathcal{G}}\varphi^i(s, y) + g^i(s, y, v(s, y), (\partial\varphi)\sigma^i(s, y), \mathcal{I}\varphi^i(s, y)) < 0 \tag{206}$$

for any (s, y) such that $s \in [t, t + \alpha]$ and $|y - x| \le \alpha$, for some small enough $\alpha > 0$. Let

$$\theta = \inf\{s \ge t;\, |X^t_s - x| \ge \alpha,\, N^t_s \ne i,\, \bar{Y}^t_s = \ell^i(s, X^t_s)\} \wedge (t + \alpha) \wedge \tau^t \tag{207}$$

$$(\widehat{Y}^t, \widehat{Z}^t, \widehat{V}^t, \widehat{\mathcal{V}}^t, \widehat{K}^t) = \Big(\mathbb{1}_{\cdot<\theta}\bar{Y}^t + \mathbb{1}_{\cdot\ge\theta}v^i(\theta, X^t_\theta)\ ,\ \mathbb{1}_{\cdot\le\theta}\bar{Z}^t\ ,\ \mathbb{1}_{\cdot\le\theta}\bar{V}^t\ ,\ \mathbb{1}_{\cdot\le\theta}\overline{\mathcal{V}}^t\ ,\ \bar{K}^t_{\cdot\wedge\theta}\Big) \tag{208}$$

$$(\widetilde{Y}^t, \widetilde{Z}^t, \widetilde{V}^t) = \Big(\varphi^i(\cdot X^t_{\cdot\wedge\theta})\ ,\ \mathbb{1}_{\cdot\le\theta}(\partial\varphi\sigma)^i(\cdot, X^t_\cdot)\ , \\ \mathbb{1}_{\cdot\le\theta}\big([\varphi^i(\cdot, X^t_{\cdot-} + \delta^i(\cdot, X^t_{\cdot-}, y)) - \varphi^i(\cdot, X^t_{\cdot-})]\big)_{y\in\mathbb{R}^d}\Big) . \tag{209}$$

Using in particular the fact that D is open in (187), one has that $\theta > t$, $\mathbb{P}^t$ – almost surely. Thus, using also the continuity of v^i :

$$\widehat{Y}^t_t = \bar{Y}^t_t = v^i(t, x) = \varphi^i(t, x) = \widetilde{Y}^t_t\ ,\ \mathbb{P}^t\text{-a.s.} \tag{210}$$

Note that by the minimality condition in the stopped RBSDE for $\bar{\mathcal{Y}}^t$, one has that $\bar{K} = 0$ on $[t, \theta]$, since $\ell^i(s, X^t_s) < \bar{Y}^t_s$ on $[t, \theta)$ and $\theta \le \tau^t$. By using the stopped RBSDE equation for $\bar{\mathcal{Y}}^t$, one thus has for $s \le r$ in $[t, \theta)$:

$$\widehat{Y}^t_s = v^i(r, X^t_r) + \int_s^r g^i(\zeta, X^t_\zeta, v(\zeta, X^t_\zeta), \widehat{Z}^t_\zeta, \bar{r}^t_\zeta)d\zeta \\ - \int_s^r \widehat{Z}^t_\zeta dB^t_\zeta - \int_s^r \int_{\mathbb{R}^d} \widehat{V}^t_\zeta(y)\widetilde{\chi}^t(dy, d\zeta) , \tag{211}$$

which also holds true for $s = r = \theta$ by definition of $\widehat{Y}^t$ in (208). Furthermore for $s < r = \theta$ either χ^t, whence X^t, do not jump at θ, and (211) holds again at $r = \theta$ by passage to the limit as $r \uparrow \theta$ in (201), or N^t does not jump at θ, in which case the stopped RBSDE equation for $\bar{\mathcal{Y}}^t$ written between s and θ directly gives (211) for $r = \theta$. In conclusion one has for $s \in [t, \theta]$:

$$\widehat{Y}_s^t = v^i(\theta, X_\theta^t) + \int_s^\theta g^i(\zeta, X_\zeta^t, v(\zeta, X_\zeta^t), \widehat{Z}_\zeta^t, \bar{r}_\zeta^t) d\zeta \\ - \int_s^\theta \widehat{Z}_\zeta^t dB_\zeta^t - \int_s^\theta \int_{\mathbb{R}^d} \widehat{V}_\zeta^t(y) \widetilde{\chi}^t(dy, d\zeta) \tag{212}$$

with (cf. (129)):

$$\int_s^\theta g^i\Big(\zeta, X_\zeta^t, v(\zeta, X_\zeta^t), \widehat{Z}_\zeta^t, \bar{r}_\zeta^t\Big) d\zeta = \int_s^\theta g^i\Big(\zeta, X_\zeta^t, v(\zeta, X_\zeta^t), \widehat{Z}_\zeta^t, \int_{\mathbb{R}^d} \widehat{V}_\zeta(y) f^i(\zeta, X_\zeta^t, y) m(dy)\Big) d\zeta \,.$$

Otherwise said $(\widehat{Y}^t, \widehat{Z}^t, \widehat{V}^t)$ solves the stopped BSDE on $[t, t+\alpha]$ with driver (cf. Comments 5.1(i))

$$g^i\Big(s, X_s^t, v(s, X_s^t), z, \int_{\mathbb{R}^d} v(y) f^i(s, X_s^t, y) m(dy)\Big) ds$$

(where $v(y)$ refers to a generic function $v \in \mathcal{M}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m(dy); \mathbb{R})$, not to be confused with the value function $v = v^i(t,x)$ in $v(\zeta, X_\zeta^t))$ and terminal condition $v^i(\theta, X_\theta^t)$ at θ. Besides one can show as in part (i) above that $(\widetilde{Y}^t, \widetilde{Z}^t, \widetilde{V}^t)$ solves the stopped BSDE on $[t, t+\alpha]$ with driver

$$g^i\Big(s, X_s^t, v(s, X_s^t), z, \int_{\mathbb{R}^d} v(y) f^i(s, X_s^t, y) m(dy)\Big) ds - \psi(s, X_s^t) ds$$

and terminal condition $\varphi^i(\theta, X_\theta^t)$ at θ. We conclude as in part (i). □

12 Uniqueness Issues

In this section we consider the issue of *uniqueness* of a solution to $(\mathcal{V}2)$ and $(\mathcal{V}1)$, respectively. We prove a *semi-continuous solutions comparison principle* for these problems, which implies in particular uniqueness of $\mathcal{P}$-solutions. For related comparison and uniqueness results we refer the reader to Alvarez and Tourin [1], Barles et al. [6, 7], Pardoux et al. [87], Pham [90], Harraj et al. [63], Amadori [2, 3] and Ma and Cvitanic [83], among others.

Assumption 12.1 **(i)** The functions b, σ and δ are locally Lipschitz continuous in (t, x, i), uniformly in y regarding δ;
(ii) There exists, for every $R > 0$, a nonnegative function η_R continuous and null at 0 *(modulus of continuity)* such that

$$|g^i(t, x, u, z, r) - g^i(t, x', u, z, r)| \le \eta_R(|x - x'|(1 + |z|))$$

for any $t \in [0, T]$, $i \in I$, $z \in \mathbb{R}^d$, $r \in \mathbb{R}$ and $x, x' \in \mathbb{R}^d$, $u \in \mathbb{R}^k$ with $|x|$, $|x'|$, $|u| \le R$;
(iii) The function g^i is non-decreasing with respect to u^j, for any $(i, j) \in I^2$ with $i \neq j$.

Comments 12.1 **(i)** By Assumption 12.1(i), one has in particular

$$|b| \vee |\sigma| \vee |\delta| < C(1 + |x|) \tag{213}$$

on $\mathcal{E}$.
(ii) The monotonicity Assumption 12.1(iii) on g means that we deal with a *cooperative* system of PIDEs (see, for instance, Busca and Sirakov [31]).

We are now in position to establish the following

Theorem 12.1 *One has $\mu \le \nu$ on $\mathcal{E}$, for any $\mathcal{P}$-subsolution μ and $\mathcal{P}$-supersolution ν of $(\mathcal{V}2)$ on $\mathcal{E}$ with terminal condition Φ on $\partial_p\mathcal{E}$, respectively of $(\mathcal{V}1)$ on $\mathcal{E}$ with boundary condition u on $\partial_p\mathcal{D}$.*

By a classical trick, one can reduce attention, for the sake of establishing Theorem 12.1, to the special case when g^i is non-decreasing with respect to u^j for any $(i,j) \in I^2$, rather than g^i non-increasing with respect to u^j for any $(i,j) \in I^2$ *with* $i \neq j$ in Assumption 12.1(iii). Note in this regard that g^i being non-decreasing with respect to u^j for any $(i,j) \in I^2$ is in fact equivalent to g being non-increasing with respect to u as a whole, rather than g^i non-increasing with respect to u^j for any $(i,j) \in I^2$ *with* $i \neq j$ in Assumption 12.1(iii).

One thus has,

Lemma 12.2 *If Theorem 12.1 holds in the special case when g^i is non-decreasing with respect to u^j for any $(i,j) \in I^2$, then Theorem 12.1 holds in general.*

Proof. This can be established by application of the special case to the transformed functions $e^{-Rt}\mu^i(t,x)$ and $e^{-Rt}\nu^i(t,x)$ for large enough R. More precisely, under the general assumptions of Theorem 12.1, $e^{-Rt}\mu$ and $e^{-Rt}\nu$ are respectively $\mathcal{P}$-subsolution and $\mathcal{P}$-supersolution of the following transformed problem, for $(\mathcal{V}2)$,

$$\max\Big(\min\Big(-\widetilde{\mathcal{G}}\varphi^i(t,x) - e^{-Rt}g^i(t,x,e^{Rt}\varphi(t,x), e^{Rt}(\partial\varphi)\sigma^i(t,x), e^{Rt}\mathcal{I}\varphi^i(t,x)) - R\varphi^i(t,x),$$
$$\varphi^i(t,x) - e^{-Rt}\ell^i(t,x)\Big), \varphi^i(t,x) - e^{-Rt}h^i(t,x)\Big) = 0$$

on $\mathrm{Int}_p\mathcal{E}$, supplemented by the terminal condition $\varphi = e^{-Rt}\Phi$ on $\partial_p\mathcal{E}$ (and likewise with $h = +\infty$ for $(\mathcal{V}1)$ on $\mathrm{Int}_p\mathcal{D}$, supplemented by the boundary condition $\varphi = e^{-Rt}\Psi$ on $\partial_p\mathcal{D}$). Now, for R large enough, Assumption 12.1(iii) and the Lipschitz continuity property of g with respect to the last variable imply that $g(t,x,e^{Rt}u, e^{Rt}z, e^{Rt}r) + Ru$ is non-decreasing with respect to u. One thus concludes by an application of the assumed restricted version of Theorem 12.1. □

Given Lemma 12.2, one may and do reduce attention, in order to prove Theorem 12.1, to the case where the function g is non-decreasing with respect to u. The statement regarding $(\mathcal{V}2)$ in Theorem 12.1 is then obtained by letting α go to 0 in part (iii) of the next lemma. The proof of the statement regarding $(\mathcal{V}1)$ in Theorem 12.1 would be analogous, substituting $(\mathcal{V}1)$ to $(\mathcal{V}2)$, $+\infty$ to h, $\mathrm{Int}_p\mathcal{D}$ to $\mathrm{Int}_p\mathcal{E}$ and $\mathcal{C}^0(\mathcal{E}) \cap \mathcal{C}^{1,2}(\mathcal{D})$ to $\mathcal{C}^{1,2}(\mathcal{E})$ in Lemma 12.3 below and its proof.

Lemma 12.3 *Let $\Lambda_1 = k\Lambda$ where Λ is the Lipschitz constant of g (cf. Assumption (M.1.ii) in Section 6.4). Given a $\mathcal{P}$-subsolution μ and a $\mathcal{P}$-supersolution ν of $(\mathcal{V}2)$ on $\mathcal{E}$, let q_1 be an integer greater than q_2 such that $\mu, \nu \in \mathcal{P}_{q_2}$, where q_2 is provided by our assumption that $\mu, \nu \in \mathcal{P}$ in Theorem 12.1. Then, assuming g non-decreasing with respect to u :*
(i) *$\omega = \mu - \nu$ is a $\mathcal{P}$-subsolution of*

$$\min\Big(w, -\widetilde{\mathcal{G}}\omega \;-\Lambda_1\big(\max_{j\in I}(\omega^j)^+ + |\partial\omega\sigma| + (\mathcal{I}\omega)^+\big)\Big) = 0$$

on $\mathcal{E}$ with null boundary condition on $\partial_p\mathcal{E}$, in the sense that:
- *$\omega \le 0$ on $\partial_p\mathcal{E}$, and*
- *$\omega^i(t,x) > 0$ implies*

$$-\widetilde{\mathcal{G}}\varphi^i(t,x) \;-\Lambda_1\Big(\max_{j\in I}(\omega^j(t,x))^+ + |\partial\varphi^i(t,x)\sigma^i(t,x)| + (\mathcal{I}\varphi^i(t,x))^+\Big) \le 0 \qquad (214)$$

for any $(t,x,i) \in Int_p\mathcal{E}$ and $\varphi \in \mathcal{C}^{1,2}(\mathcal{E})$ such that $\omega^i - \varphi^i$ reaches a global null maximum at (t,x).
(ii) *There exists $C_1 > 0$ such that the (regular) function*

$$\chi^i(t,x) = (1 + |x|^{q_1})e^{C_1(T-t)}$$

is a strict $\mathcal{P}$-supersolution *of*

$$\min\Big(\chi, -\widetilde{\mathcal{G}}\chi - \Lambda_1\big(\chi + |\partial\chi\sigma| + (\mathcal{I}\chi)^+\big)\Big) = 0$$

on $\mathcal{E}$, in the sense that $\chi > 0$ and

$$-\widetilde{\mathcal{G}}\chi - \Lambda_1\big(\chi + |\partial\chi\sigma| + (\mathcal{I}\chi)^+\big) > 0 \qquad (215)$$

on $\mathcal{E}$.
(iii) $\max_{i\in I}(\omega^i)^+ \le \alpha\chi$ *on $[0,T] \times \mathbb{R}^d$, for any $\alpha > 0$.*

This lemma is an adaptation to our set-up of the analogous result in Barles et al. [6] (see also Pardoux et al. [87] and Harraj et al. [63]). Here are the main differences (our assumptions are fitted to financial applications, see Part I):
(i) We consider a model with jumps in X and regimes represented by N, whereas [6] or [63] only consider jumps in X, and [87] only considers regimes;
(ii) We work with finite jump measures m, jump size δ with linear growth in x, and semi-continuous solutions with polynomial growth in x, whereas [6] or [63] consider general Levy measures, bounded jumps, and continuous solutions with sub-exponential (strictly including polynomial) growth in x;
(iii) [6] deals with classical BSDEs (without barriers);
(iv) We consider time-dependent coefficients b, σ, δ whereas [6] considers homogeneous dynamics.
Because of these differences we provide a detailed proof in Appendix B.1.

To conclude this section we can state the following proposition, which sums-up the results of Theorems 11.1 and 12.1.

Proposition 12.4 **(i)** *The value function u is the* unique *$\mathcal{P}$-solution, the* maximal $\mathcal{P}$-subsolution *and the* minimal $\mathcal{P}$-supersolution *of $(\mathcal{V}2)$ on $\mathcal{E}$ with terminal condition Φ on $\partial_p\mathcal{E}$;*
(ii) *The value function v is the* unique *$\mathcal{P}$-solution, the* maximal $\mathcal{P}$-subsolution, *and the* minimal $\mathcal{P}$-supersolution *of $(\mathcal{V}1)$ on $\mathcal{E}$ with boundary condition u on $\partial_p\mathcal{D}$.* □

13 Approximation

An important feature of semi-continuous viscosity solutions comparison principles like Theorem 12.1 above is that they ensure the *stability* of the related PIDE problem, providing in particular generic conditions ensuring the convergence of a wide family of deterministic approximation schemes. These are the so called *stability, monotonicity and consistency conditions* originally introduced for PDEs by Barles and Souganidis [9]. See also Briani, La Chioma and Natalini [30], Cont and Voltchkova [38] or Jakobsen et al. [69] for various extensions of these results to PIDEs.

The following results thus extend to models with regimes, thus *systems* of PIDEs, the results of [9, 30], among others.

The following lemma is standard and elementary, and thus stated without proof.

Lemma 13.1 *Let $(\mathcal{E}_h)_{h>0}$ denote a family of subsets of $\mathcal{E}$, such that for any $(t,x,i) \in \mathcal{E}$, there exists sequences h_n, t_n, x_n verifying*

$$h_n \to 0 \text{ and } \mathcal{E}^{h_n} \ni (t_n, x_n, i) \to (t,x,i) \text{ as } n \to \infty.$$

Let $(u_h)_{h>0}$ be a family of uniformly locally bounded real functions with u_h defined on the set $\mathcal{E}_h$, for any $h>0$.
(i) *For any $(t,x,i) \in \mathcal{E}$, the set of limits of the following kind:*

$$\lim_{n\to+\infty} u^i_{h_n}(t_n,x_n) \text{ with } h_n \to 0 \text{ and } \mathcal{E}^{h_n} \ni (t_n,x_n,i) \to (t,x,i) \text{ as } n \to \infty \;, \tag{216}$$

is non empty and compact in $\mathbb{R}$. It admits as such a smallest and a greatest element: $\underline{u}^i(t,x) \le \overline{u}^i(t,x)$ in $\mathbb{R}$.
(ii) *The function $\underline{u}$, respectively $\overline{u}$, defined in this way, is locally bounded and lower semi-continuous on $\mathcal{E}$, respectively locally bounded and upper semi-continuous on $\mathcal{E}$. We call it the lower limit, respectively upper limit, of $(u_h)_{h>0}$ at (t,x,i) as $h \to 0^+$. We say that u_h converges to l at $(t,x,i) \in \mathcal{E}$ as $h \to 0$, and we denote :*

$$\lim_{\substack{h\to 0^+ \\ \mathcal{E}_h \ni (t_h,x_h,i)\to(t,x,i)}} u^i_h(t_h,x_h) \;=\; l \;,$$

if and only if $\underline{u}^i(t,x) = \overline{u}^i(t,x) = l$, or, equivalently:

$$\lim_{n\to+\infty} u^i_{h_n}(t_n,x_n) \;=\; l$$

for any $h_n \to 0$ et $\mathcal{E}^{h_n} \ni (t_n,x_n,i) \to (t,x,i)$.
(iii) *If u_h converges pointwise everywhere to a continuous function u on $\mathcal{E}$, then this convergence is locally uniform:*

$$\max_{\mathcal{E}_h \cap \mathcal{C}} |u_h - u| \;\to 0$$

as $h \to 0^+$, for any compact subset $\mathcal{C}$ of $\mathcal{E}$. □

Definition 13.1 Let us be given families of operators

$$\widetilde{\mathcal{G}}_h = \widetilde{\mathcal{G}}_h u^i(t_h,x_h) \;,\; \delta_h = \delta_h u^i(t_h,x_h) \;,\; \mathcal{I}_h = \mathcal{I}_h u^i(t_h,x_h)$$

devoted to approximate $\widetilde{\mathcal{G}}u^i(t_h, x_h)$, $\partial u^i(t_h, x_h)$ and $\mathcal{I}u^i(t_h, x_h)$ on $\mathcal{E}_h$ for real-valued functions u on $\mathcal{E}$, respectively. For $\mathcal{L} = \partial$, $\mathcal{I}$ or $\widetilde{\mathcal{G}}$:
(i) we say that the related discretisation $\mathcal{L}_h = \delta_h$, $\mathcal{I}_h$ or $\widetilde{\mathcal{G}}_h$ is *monotone*, if

$$\mathcal{L}_h u_1^i(t_h, x_h) \leq \mathcal{L}_h u_2^i(t_h, x_h) \tag{217}$$

for any functions $u_1 \leq u_2$ on $\mathcal{E}_h$ with $u_1^i(t_h, x_h) = u_2^i(t_h, x_h)$;
(ii) we say that the related discretisation scheme $(\mathcal{L}_h)_{h>0}$ is *consistent* with $\mathcal{L}$ if and only if for any continuous function φ on $\mathcal{E}$ of class $\mathcal{C}^{1,2}$ around (t, x, i), we have:

$$\mathcal{L}_h(\varphi + \xi_h)^i(t_h, x_h) \to \mathcal{L}\varphi^i(t, x) \tag{218}$$

whenever $h \to 0^+$, $\mathcal{E}_h \ni (t_h, x_h, i) \to (t, x, i) \in \mathcal{E}$ and $\mathbb{R} \ni \xi_h \to 0$.

Assumption 13.2 The function

$$\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \ni (u, p, r) \mapsto g^{\cdot}(\cdot, \cdot, u, p\sigma, r) \in \mathbb{R}^{\mathcal{E}} \tag{219}$$

is non-decreasing, in the sense that for any $(u, p, r) \leq (u', p', r')$ coordinate by coordinate in $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}$, one has $g^i(t, x, u, p\sigma^i(t, x), r) \leq g^i(t, x, u', p'\sigma^i(t, x), r')$ for any $(t, x, i) \in \mathcal{E}$.

Comments 13.1 **(i)** The previous monotonicity and consistency assumptions are abstract conditions which need to be verified carefully on a case-by-case basis for any concrete approximation scheme under consideration (like specific finite difference schemes). We refer the reader to Cont and Voltchkova [38] (see also Jakobsen et al. [69]) for the complete analysis of specific schemes under various sets of assumptions. In our case the most stringent condition seems to be the one regarding the monotonicity of g with respect to p in (219). Note however that this condition is obviously satisfied in every of the following three cases:
• the function $g = g^i(t, x, u, z, r)$ does not depend on the argument z, which is typically the case with risk-neutral pricing problems in finance (see Section 6.6);
• σ is equal to zero, which corresponds to the situation of pure jump models; note however that our continuity Assumption 8.1 on τ^t fails to be satisfied in this case for domains as simple as $D = \{|x| < R\} \times I$, τ being defined as in Assumption 10.1(iii);
• the dimension q of the jump-diffusion component X of $\mathcal{X}$ is equal to one and ∂ is discretized by *decentered forward finite differences*, yielding an *upwind discretization scheme* for $\partial\varphi\sigma$, by non-negativity of σ in the scalar case (see, for instance, Kushner and Dupuis [77]).
(ii) Under the weaker assumption that $g^i(t, x, u, p\sigma^i(t, x), r)$ is non-decreasing with respect to (p, r) and non-decreasing with respect to u^j for $j \neq i$, then the mapping $u^i(t, x) \mapsto \widetilde{u}^i(t, x) := e^{-Rt}u^i(t, x)$ for R large enough transforms the problem into one in which Assumption 13.2 holds (see the proof of Lemma 12.2). Suitable approximation schemes may then be applied to the transformed problem, and a convergent approximation to the solution of the original problem is recovered by setting $u_h^i(t, x) := e^{Rt}\widetilde{u}_h^i(t, x)$.

By $(u_h)_{h>0}$ uniformly polynomially bounded in (a) below we mean that u_h is bounded by $C(1 + |x|^q)$ for some C and q independent of h.

Lemma 13.2 *Let us be given monotone and consistent approximation schemes*

$$(\widetilde{\mathcal{G}}_h)_{h>0},\ (\delta_h)_{h>0} \text{ and } (\mathcal{I}_h)_{h>0}$$

for $\widetilde{\mathcal{G}}$, ∂ *and* $\mathcal{I}$ *respectively,* g *satisfying the monotonicity Assumption 13.2.*
(a) *Let* $(u_h)_{h>0}$ *be uniformly polynomially bounded and satisfy*

$$\max\Big(\min\Big(-\widetilde{\mathcal{G}}_h {u_h}^i(t_h,x_h) - g^i(t_h,x_h,u_h(t_h,x_h),(\delta_h u_h\sigma)^i(t_h,x_h),\mathcal{I}_h u_h^i(t_h,x_h)), \quad (220)$$

$$u_h^i(t_h,x_h)-\ell^i(t_h,x_h)\Big), u_h^i(t_h,x_h)-h^i(t_h,x_h)\Big)=0 \quad (221)$$

on $Int_p\mathcal{E}\cap\mathcal{E}_h$ *and* $u_h=\Phi$ *on* $\partial_p\mathcal{E}\cap\mathcal{E}_h$ *for any* $h>0$*. Then:*
(i) *The upper and lower limits* $\overline{u}$ *and* $\underline{u}$ *of* u_h *as* $h\to 0$*, are respectively viscosity subsolutions and supersolutions of* $(\mathcal{V}2)$ *on* $Int_p\mathcal{E}$*;*
(ii) *One has* $\overline{u}\le\Phi\le\underline{u}$ *pointwise at* T*.*

(b) *Let* $(v_h)_{h>0}$ *be uniformly polynomially bounded and satisfy*

$$\min\Big(-\widetilde{\mathcal{G}}_h v_h^i(t_h,x_h) - g^i(t_h,x_h,v_h(t_h,x_h),(\delta_h v_h\sigma)^i(t_h,x_h),\mathcal{I}_h v_h^i(t_h,x_h)), \quad (222)$$

$$v_h^i(t_h,x_h)-\ell^i(t_h,x_h)\Big)=0 \quad (223)$$

on $Int_p\mathcal{D}\cap\mathcal{E}_h$ *and* $v_h=u$ *on* $\partial_p\mathcal{D}\cap\mathcal{E}_h$ *for any* $h>0$*. Then:*
(i) *The upper and lower limits* $\overline{v}$ *and* $\underline{v}$ *of* v_h *as* $h\to 0$*, are respectively viscosity subsolutions and supersolutions of* $(\mathcal{V}1)$ *on* $Int_p\mathcal{D}$*;*
(ii) *One has* $\overline{v}\le u(=\Phi)\le\underline{v}$ *pointwise at* T*.*

Proof. We only prove (a), since the proof of (b) is similar (cf. the comments preceding Lemma 12.3). *Note that one only has* $\overline{v}\le u\le\underline{v}$ at T *in (b), and not necessarily* $\overline{v}\le u\le\underline{v}$ *on* $\partial_p\mathcal{D}$*; see comments in part (ii) below.*
(i) We prove that $\overline{u}$ is a viscosity subsolution of $(\mathcal{V}2)$ on $\mathrm{Int}_p\mathcal{E}$. The fact that $\underline{u}$ is a viscosity supersolution of $(\mathcal{V}2)$ on $\mathrm{Int}_p\mathcal{E}$ can be shown likewise. First note that $\overline{u}\le h$, by (220) on $\mathrm{Int}_p\mathcal{E}\cap\mathcal{E}_h$, inequality $\Phi\le h$ on $\partial_p\mathcal{E}$ (cf. (M.2.ii)) and continuity of h and Φ. Let then $(t^\star,x^\star,i)\in\mathrm{Int}_p\mathcal{E}$ be such that $\overline{u}^i(t^\star,x^\star)>\ell^i(t^\star,x^\star)$ and $(t^\star,x^\star)$ maximizes strictly $\overline{u}^i-\varphi^i$ at zero for some function $\varphi\in\mathcal{C}^{1,2}(\mathcal{E})$. We need to show that (cf. (194)):

$$-\widetilde{\mathcal{G}}\varphi^i(t^\star,x^\star)-g^i\left(t^\star,x^\star,\overline{u}(t^\star,x^\star),(\partial\varphi\sigma)^i(t^\star,x^\star),\mathcal{I}\varphi^i(t^\star,x^\star)\right)\le 0\,. \quad (224)$$

By a classical argument in the theory of viscosity solutions (see, e.g., Barles and Souganidis [9]), there exists, for any $h>0$, a point (t,x) in $[0,T]\times\bar{B}_R$, where $\bar{B}_R$ is a ball with large radius R around $x^\star$, such that (we omit the dependence of t,x in h for notational simplicity):

$$u_h^i\le\varphi^i+(u_h-\varphi)^i(t,x) \quad (225)$$

with equality at (t,x), and $\xi_h:=(u_h-\varphi)^i(t,x)$ goes to $0=(\overline{u}-\varphi)^i(t^\star,x^\star)$, whence $u_h^i(t,x)$ goes to $\overline{u}^i(t^\star,x^\star)$, as $h\to 0$ (cf. an analogous statement and its justification in the second part of the proof of part (ii) below). Therefore $\overline{u}^i(t^\star,x^\star)>\ell^i(t^\star,x^\star)$ implies that $u_h^i(t,x)>\ell^i(t,x)$ for h small enough, whence by (220):

$$-\widetilde{\mathcal{G}}_h u_h^i(t,x)-g^i(t,x,u_h(t,x),(\delta_h u_h\sigma)^i(t,x),\mathcal{I}_h u_h^i(t,x))\le 0\,. \quad (226)$$

Given (225), one thus has by monotonicity of the scheme and of g (Assumption 13.2):

$$\begin{aligned}
-\widetilde{\mathcal{G}}_h(\varphi+\xi_h)^i(t,x) &\le g^i\left(t,x,u_h(t,x),(\delta_h(\varphi+\xi_h)\sigma)^i(t,x),\mathcal{I}_h(\varphi+\xi_h)^i(t,x)\right)\\
&\le g^i\left(t^\star,x^\star,\overline{u}(t^\star,x^\star),(\partial\varphi\sigma)^i(t^\star,x^\star),\mathcal{I}\varphi^i(t^\star,x^\star)\right)\\
&\quad+\eta(|t-t^\star|)+\eta_R(|x-x^\star|(1+|(\partial\varphi\sigma)^i(t^\star,x^\star)|))+k\Lambda\max_{j\in I}(u_h^j(t,x)-\overline{u}^j(t^\star,x^\star))^+\\
&\quad+\Lambda|(\delta_h(\varphi+\xi_h)\sigma)^i(t,x)-(\partial\varphi\sigma)^i(t^\star,x^\star)|+\Lambda(\mathcal{I}_h(\varphi+\xi_h)^i(t,x)-\mathcal{I}\varphi^i(t^\star,x^\star))^+\,,
\end{aligned}$$

where in the last inequality (cf. proof of Lemma 12.3(i) in Appendix B.1):
• η is a modulus of continuity of g^i on a 'large' compact set around

$$(t^\star, x^\star, \overline{u}(t^\star, x^\star), (\partial\varphi\sigma)^i(t^\star, x^\star), \mathcal{I}\varphi^i(t^\star, x^\star)) \ ;$$

• η_R is the modulus of continuity standing in Assumption 12.1(ii);
• the three last terms come from the Lipschitz continuity and monotonicity properties of g.
Inequality (224) follows by sending h to zero in the previous inequality, using the consistency (218) of the scheme.
(ii) Let us show further that $\overline{u}$ and $\underline{u}$ satisfy the boundary condition in the so-called *weak viscosity sense* on $\partial_p\mathcal{E}$, namely in the case of $\overline{u}$ (the related statement and proof are similar in the case of $\underline{u}$): Inequality (224) holds for any $(t^\star, x^\star, i) \in \partial_p\mathcal{E}$ and $\varphi \in \mathcal{C}^{1,2}(\mathcal{E})$ such that

$$\overline{u}^i(t^\star, x^\star) > \ell^i(t^\star, x^\star) \vee \Phi^i(t^*, x^*) \tag{227}$$

and $(t^\star, x^\star)$ maximizes globally and strictly $\overline{u}^i - \varphi^i$ at zero. As in part (i), there exists, for any $h > 0$, a point (t, x) in $[0, T] \times \bar{B}_R$ (we omit the dependence of t, x in h for notational simplicity), where $\bar{B}_R$ is a ball with large radius R around $x^\star$, such that inequality (225) holds with equality at (t, x), and $\xi_h = (u_h - \varphi)^i(t, x)$, whence $u_h^i(t, x) - \overline{u}^i(t^\star, x^\star)$, goes to zero as $h \to 0$. Therefore inequality (227) implies that

$$u_h^i(t, x) > \ell^i(t, x) \vee \Phi^i(t, x)$$

for h small enough, whence $(t, x, i) \in \mathrm{Int}_p\mathcal{E}$ and by (220):

$$-\widetilde{\mathcal{G}}_h u_h^i(t, x) - g^i(t, x, u_h(t, x), (\delta_h u_h \sigma)^i(t, x), \mathcal{I}_h u_h^i(t, x)) \leq 0 \ . \tag{228}$$

Inequality (224) follows like in part (i) above.

Now *(note that the following argument only works at T and cannot be adapted to the case of problem $(\mathcal{V}1)$ on the whole of $\partial_p\mathcal{D}$, cf. comment at the beginning of the proof)*, by a classical argument in the theory of viscosity solutions (see Alvarez and Tourin [1, bottom of page 303] or Amadori [2, 3]), any viscosity subsolution or supersolution of $(\mathcal{V}2)$ on $\mathrm{Int}_p\mathcal{E}$ satisfying the boundary condition in the weak viscosity sense on $\partial_p\mathcal{E}$, satisfies it pointwise at T. So, in our case, suppose for instance by contradiction that

$$\overline{u}^i(T, x^\star) > \Phi^i(T, x^*) \tag{229}$$

for some $x^\star$ with $(T, x^\star) \in \partial_p\mathcal{E}$. Let us then introduce the function

$$\varphi_\varepsilon^i(t, x) = \overline{u}^i(t, x) - \frac{|x^* - x|^2}{\varepsilon} - C_\varepsilon(T - t) \tag{230}$$

in which

$$C_\varepsilon > \sup_{(t,x)\in[t-\eta,T]\times\bar{B}_1(x^*)} \tag{231}$$
$$\widetilde{\mathcal{G}}\left(\frac{|y - x^*|^2}{\varepsilon}\right)^i(t, x) + g^i\left(t, x, \overline{u}(t, x), \left(\frac{2(y - x^*)\sigma}{\varepsilon}\right)^i(t, x), \mathcal{I}\left(\frac{|y - x^*|^2}{\varepsilon}\right)^i(t, x)\right)$$

goes to ∞ as $\varepsilon \to 0$, where $\bar{B}_1(x^*)$ denotes the closed unit ball centered at x^* in $\mathbb{R}^d$. There exists, for any $\varepsilon > 0$, a point (t, x) in $[0, T] \times \bar{B}_R$ (we omit the dependence of (t, x) in ε for

notational simplicity), where $\bar{B}_R$ is a ball with large radius R around $x^\star$, such that:
• for any $\varepsilon > 0$ the related point (t,x) maximizes φ^i_ε over $[0,T] \times \bar{B}_R$,
• $(t,x) \to (T,x^\star)$ and $\overline{u}^i(t,x) \to \overline{u}^i(T,x^\star)$ as $\varepsilon \to 0$.
To justify the last point, note that by the maximizing property of (t,x) one has that

$$\varphi^i_\varepsilon(T,x^\star) \le \varphi^i_\varepsilon(t,x)$$

whence in particular (cf. (230))

$$0 \le \frac{|x^* - x|^2}{\varepsilon} + C_\varepsilon(T-t) \le \overline{u}^i(t,x) - \overline{u}^i(T,x^\star) \tag{232}$$

so

$$\overline{u}^i(T,x^\star) \le \overline{u}^i(t,x) . \tag{233}$$

Since $\overline{u}$ is locally bounded, (232) implies that $(t,x) \to (T,x^\star)$ as $\varepsilon \to 0$, which, joint to the upper semi-continuity of $\overline{u}$ and to (233), implies that $\overline{u}^i(t,x) \to \overline{u}^i(T,x^\star)$ as $\varepsilon \to 0$.

Now one has $\ell \le \Phi$ pointwise at T, therefore (229) joint to the fact that $\lim_{\varepsilon\to 0} \overline{u}^i(t,x) = \overline{u}^i(T,x^\star)$ imply that $\overline{u}^i(t,x) > \ell^i(t,x)$, for ε small enough. In virtue of the results already established at this point of the proof, the function $(s,y) \mapsto \frac{|x^*-y|^2}{\varepsilon} + C_\varepsilon(T-s)$ thus satisfies the related viscosity subsolution inequality at (t,x,i), so

$$C_\varepsilon - \widetilde{\mathcal{G}}\left(\frac{|y-x^*|^2}{\varepsilon}\right)^i(t,x) - g^i\left(t,x,\overline{u}(t,x),\left(\frac{2(y-x^*)\sigma}{\varepsilon}\right)^i(t,x),\mathcal{I}\left(\frac{|y-x^*|^2}{\varepsilon}\right)^i(t,x)\right) \le 0 ,$$

which for ε small enough contradicts (231). □

Proposition 13.3 *Let $(u_h)_{h>0}$, resp. $(v)_{h>0}$, denote a stable, monotone and consistent approximation scheme, in the sense that all conditions in Lemma 13.2(a), resp. (b) are satisfied for the value function u, resp. v. Then:*
(a) *$u_h \to u$ locally uniformly on $\mathcal{E}$ as $h \to 0$.*
(b) *$v_h \to v$ locally uniformly on $\mathcal{E}$ as $h \to 0$, provided $v_h \to v(= u)$ on $\partial_p\mathcal{D} \cap \{t < T\}$.*

Proof. **(a)** By Lemma 13.2(a), the upper and lower limits $\overline{u}$ and $\underline{u}$ are $\mathcal{P}$-subsolutions and $\mathcal{P}$-supersolutions of $(\mathcal{V}2)$ on $\mathcal{E}$. So $\overline{u} \le \underline{u}$, by Theorem 12.1. Moreover $\underline{u} \le \overline{u}$ by Lemma 13.1(i). Thus finally $\underline{u} = \overline{u}$, which implies that $u_h \to u$ locally uniformly on $\mathcal{E}$ as $h \to 0$, by Lemma 13.1(iii).
(b) By Lemma 13.2(b)(i), $\overline{v}$ and $\underline{v}$ are respectively viscosity subsolutions and supersolutions of $(\mathcal{V}1)$ on $\mathrm{Int}_p\mathcal{D}$. Moreover, they satisfy $\overline{v} \le u \le \underline{v}$ at T, by Lemma 13.2(b)(ii). If, in addition, $v_h \to v(= u)$ on $\partial_p\mathcal{D} \cap \{s < T\}$, then $\overline{v} \le u \le \underline{v}$ on $\partial_p\mathcal{D}$, and $\overline{v}$ and $\underline{v}$ are $\mathcal{P}$-subsolutions and $\mathcal{P}$-supersolutions of $(\mathcal{V}1)$ on $\mathcal{E}$. We conclude like in part (a). □

Remark 13.3 The convergence result regarding v in Proposition 13.3(b) can only be considered as a partial result, since one only gets the convergence on $\mathcal{E}$ conditionally on the convergence on $\partial_p\mathcal{D} \cap \{t < T\}$, for which no explicit criterion is given. Moreover the related approximation scheme v_h is written under the working assumption that the true value for u is plugged on $\partial_p\mathcal{D}$ in the approximation scheme for v (cf. the boundary condition '$v_h = u$ on $\partial_p\mathcal{D} \cap \mathcal{D}_h$' in Lemma 13.2(b)).

Part IV
Further Applications

In this part we provide various extensions to the BSDE and PDE results of Parts II and III which are needed for dealing with important practical issues like *discrete dividends* or *discrete path-dependence* in the context of pricing problems in finance.

Let us thus be given a set $\mathcal{T} = \{T_0, T_1 \ldots, T_m\}$ of fixed times with $0 = T_0 < T_1 < \cdots < T_{m-1} < T_m = T$, representing in the financial interpretation *discrete dividends dates*, or *monitoring dates* in the case of discretely path-dependent payoffs. Let us set, for $l = 1, \ldots, m$ (recall (189)),

$$\mathcal{E}_l = \mathcal{E} \cap \{T_{l-1} \le t \le T_l\} \ , \ \mathrm{Int}_p\mathcal{E}_l = \mathrm{Int}_p\mathcal{E} \cap \{T_{l-1} \le t < T_l\} \ , \ \partial_p\mathcal{E}_l = \mathcal{E}_l \setminus \mathrm{Int}_p\mathcal{E}_l$$
$$\mathcal{D}_l = \mathcal{D} \cap \{T_{l-1} \le t \le T_l\} \ , \ \mathrm{Int}_p\mathcal{D}_l = \mathrm{Int}_p\mathcal{D} \cap \{T_{l-1} \le t < T_l\} \ , \ \partial_p\mathcal{D}_l = \mathcal{E}_l \setminus \mathrm{Int}_p\mathcal{D}_l$$
$$\mathcal{E}_l^* = \mathcal{E} \cap \{T_{l-1} \le t < T_l\} \ , \ \mathcal{D}_l^* = \mathcal{D} \cap \{T_{l-1} \le t < T_l\} \ .$$

Note that the sets $\mathrm{Int}_p\mathcal{E}_l$s, resp. $\mathrm{Int}_p\mathcal{D}_l$s, partition $\mathrm{Int}_p\mathcal{E}$, resp. $\mathrm{Int}_p\mathcal{D}$, and that the sets $\mathcal{E}_l^*$s and $\mathcal{E}_{(T)}$, resp. $\mathcal{D}_l^*$s and $\mathcal{D} \cap \mathcal{E}_{(T)}$, with $\mathcal{E}_{(T)} = \{T\} \times \mathbb{R}^d \times I$, partition $\mathcal{E}$, resp. $\mathcal{D}$.

Discrete dividends on a financial derivative or on an underlying asset (component of the factor process $\mathcal{X}$) motivate separate developments presented in Sections 14 and 15, respectively. Section 16 deals with the issue of discretely monitored call protection (*intermittent* call protection, as opposed to call protection *before a stopping time* earlier in this article).

14 Time-Discontinuous Running Cost Function

Many derivative payoffs, like for instance convertible bonds (see Section 4.2.1.1), entail discrete coupon tenors, that is, coupons paid at specific coupon dates T_ls, rather than theoretical coupon streams that would be paid in continuous-time. But discrete coupons imply predictable jumps, by the coupon amounts, of the related financial derivatives arbitrage price processes at the T_ls. At first sight the resulting pricing problems are not amenable to the methods of this paper anymore. Note in particular that all the BSDEs introduced in this paper have time-differentiable driver coefficients (the place for dividends in the case of pricing equations, see Part I), and that the state-process Y of the solution to a BSDE, which is intended to represent the price process of a financial derivative, can only jump at totally unpredictable stopping times.

However, as demonstrated in [15, 16, 17, 18], this apparent difficulty can be handled by working with a suitable notion of *clean* (instead of *ex-dividend*) price process for a financial derivative (price less accrued interest at time t, a notion of price commonly used by market practitioners). This simple transformation allows one to restore the continuity in time (but for totally unpredictable jumps) of the price processes, and to fit the set-up and assumptions of the present paper.

Yet an aside of this transformation is that the resulting running cost function g is not continuous anymore, but presents left-discontinuities in time at the T_ls. This motivates an extension of the results of this paper to the case of a running cost function g defined by concatenation on the $\mathcal{E}_l^*$s of functions g_ls satisfying our usual assumptions relative to the

$\mathcal{E}_l$s. Definition 10.4 for viscosity solutions of $(\mathcal{V}2)$ and $(\mathcal{V}1)$ then needs to be amended as follows.

Definition 14.1 **(i)** A locally bounded upper semi-continuous, resp. lower semi-continuous, resp. resp. continuous, function u on $\mathcal{E}$, is called a *viscosity subsolution*, resp. *supersolution*, resp. resp. *solution*, of $(\mathcal{V}2)$ at $(t,x,i) \in \mathrm{Int}_p\mathcal{E}$, if and only if the restriction of u to $\mathcal{E}_l$ with $(t,x,i) \in \mathrm{Int}_p\mathcal{E}_l$ is a viscosity subsolution, resp. supersolution, resp. resp. solution, of $(\mathcal{V}2)$ at (t,x,i), relative to $\mathcal{E}_l$ (cf. Definition 10.4(a)).
(ii) A $\mathcal{P}$ – *viscosity subsolution*, resp. *supersolution*, resp. resp. solution u to $(\mathcal{V}2)$ on $\mathcal{E}$ for the boundary condition Φ at T is then formally defined as in Definition 10.4(b), with the embedded notions of viscosity subsolution, resp. supersolution, resp. resp. solution, of $(\mathcal{V}2)$ at any (t,x,i) in $\mathrm{Int}_p\mathcal{E}$ defined as in (i) above.
(iii) The notions of viscosity subsolutions, supersolutions and solutions of $(\mathcal{V}1)$ at $(t,x,i) \in \mathrm{Int}_p\mathcal{D}$, and, given a further continuous boundary condition Ψ on $\partial_p\mathcal{D}$ such that $\Psi = \Phi$ at T, of $\mathcal{P}$ – viscosity subsolutions, supersolutions and solutions of $(\mathcal{V}1)$ on $\mathcal{E}$, are defined similarly (cf. Definition 10.4(c)).

Proposition 14.1 *Using Definition 14.1 for the involved notions of viscosity solutions, all the results of this paper still hold true under the currently relaxed assumption on g.*

Proof. In Part II, the continuity of g was used first, to ensure well-definedness of the process $\widetilde{g}(s, \mathcal{X}_s^t, Y_s^t, Z_s^t, \mathcal{V}_s^t)$ (cf. (119)) for any $(Y^t, Z^t, \mathcal{V}^t) \in \mathcal{S}^2 \times \mathcal{H}_d^2 \times \mathcal{H}_{\mu^t}^2$, and second, for the stability results of Propositions 8.2(ii) and 8.5(ii). But it can be checked by inspection of the related proofs that these stability results are still true under the current assumption on g. Moreover the process $\widetilde{g}(s, \mathcal{X}_s^t, Y_s^t, Z_s^t, \mathcal{V}_s^t)$ is obviously still well-defined under the currently relaxed assumption on g, for any $(Y^t, Z^t, \mathcal{V}^t) \in \mathcal{S}^2 \times \mathcal{H}_d^2 \times \mathcal{H}_{\mu^t}^2$.
In Part III, Theorem 11.1 still holds true, by immediate inspection of its proof. Moreover, under the l by l version of Assumption 12.1(ii) on the g_ls, Lemma 12.3 and Theorem 12.1 (whence Proposition 12.4) can be proven together iteratively on l as follows. Let μ and ν denote a $\mathcal{P}$-subsolution and a $\mathcal{P}$-supersolution v of $(\mathcal{V}2)$ on $\mathcal{E}$ (the proof would be analogous for $(\mathcal{V}1)$). Lemma 12.3 relative to $\mathcal{E}_m$ is proven in exactly the same way as before. We thus have (cf. Theorem 12.1) $\mu \le \nu$ on $\mathcal{E}_m$. We can then establish likewise the version of Lemma 12.3 relative to $\mathcal{E}_{m-1}$ (note that $\mu - \nu \le 0$ on $\partial_p\mathcal{E}_{m-1}$, by the first step of the proof). So $\mu \le \nu$ on $\mathcal{E}_{m-1}$, and so on until $l = 1$. Lemma 13.1 is of course not affected by the relaxation of the assumption on g. Finally, given Definition 14.1, Lemma 13.2(a)(i) can be proven exactly as before, on each $\mathrm{Int}_p\mathcal{E}_l$, and the proof of Lemma 13.2(a)(ii) does not change. Lemma 13.2(a) is thus still true, and so is likewise Lemma 13.2(b), hence Proposition 13.3 follows as before. □

15 Deterministic Jumps in $\mathcal{X}$

15.1 Deterministic Jumps in X

After having considered dividends on a financial derivative with factor process $\mathcal{X}$ in Section 14, we now want to deal with pricing problems involving discrete dividends at times T_ls on a primary asset, specifically given as a component of X in our generic factor process $\mathcal{X} = (X, N)$, underlying a financial derivative.

Note that our basic model $\mathcal{X}$ cannot jump at the T_ls, since the jump times of the driving random measures χ and ν are totally inaccessible. We thus enrich our model $\mathcal{X}$ by the introduction of deterministic jumps in X at the T_ls (instead of discontinuities in the running cost function g in Section 14), specifically,

$$X_{T_l} = \theta_l(\mathcal{X}_{T_l-}) ,$$

where the jump function θ is given as a system of Lipschitz functions $y \to \theta_l^j(y)$ from $\mathbb{R}^d$ into itself, for every $i \in I$ and $l = 1, \ldots, m$.

Definition 15.1 **(i)** A *Cauchy cascade* Φ, ν on $\mathcal{E}$ is a pair made of a terminal condition Φ of class $\mathcal{P}$ at T, along with a sequence $\nu = (u_l)_{1 \le l \le m}$ of functions u_ls of class $\mathcal{P}$ on the $\mathcal{E}_l$s, satisfying the following jump condition on $\mathbb{R}^d \times I$, for every $l = 1, \ldots, m$:

$$u_l^i(T_l, x) = u_{l+1}^i(T_l, \theta_l^i(x)) \tag{234}$$

where, in case $l = m$, u_{l+1}^i is to be understood as Φ in the right-hand-side of (234). A *continuous* Cauchy cascade is a Cauchy cascade with continuous ingredients Φ, u_ls;
(ii) The function defined by a Cauchy cascade Φ, ν is the function u on $\mathcal{E}$ given as the concatenation on the $\mathcal{E}_l^*$s of the u_ls, along with the terminal condition Φ at T.

The formal analogue of Definition 6.4 for a consistent solution to the Markovian decoupled Forward Backward stochastic differential equation with data $\mathcal{G}$ (including here the jumps defined by θ in X), $\mathcal{C}$ and τ may thus be formulated, where :
• A 'model $\mathcal{X}$ with generator $\mathcal{G}$' in Definition 6.4(a) is to be understood in the sense that for every $l = 1, \ldots, m$ with $t \le T_l$,
— $\mathcal{X}^t$ obeys the dynamics (152) on the time interval $[T_{l-1} \vee t, T_l)$,
— $X_{T_l}^t = \theta_l(\mathcal{X}_{T_l-}^t)$ and $N_{T_l}^t = N_{T_l-}^t$,
where the superscript ^{t} refers as usual to a constant initial condition (t, x, i) for $\mathcal{X}$, so $\mathcal{X}_t^t = (x, i)$;
• In Definition 6.4(b):
— The deterministic value function u in Definition 6.4(b)(i) is no longer continuous on $\mathcal{E}$, but defined by a continuous Cauchy cascade $\Phi, (u_l)_{1 \le l \le m}$;
— The deterministic value function v in Definition 6.4(b)(ii) is defined likewise by a continuous Cauchy cascade $\Phi, (v_l)_{1 \le l \le m}$.

One assumes in this section that the lower and upper cost functions ℓ and h are not continuous on $\mathcal{E}$, but are defined by continuous Cauchy cascades $\Lambda, (\ell_l)_{1 \le l \le m}$ and $\Upsilon, (h_l)_{1 \le l \le m}$ such that $\ell_l \le h_l$ for every $l = 1, \ldots, m$, and $\Lambda \le \Phi \le \Upsilon$, whence in particular

$$\ell_m^i(T, x) = \Lambda^i(T, \theta_m^i(x)) \le \Phi^i(T, \theta_m^i(x)) \le \Upsilon^i(T, \theta_m^i(x)) = h_m^i(T, x) . \tag{235}$$

Note that $\ell(s, \mathcal{X}_s^t)$ and $h(s, \mathcal{X}_s^t)$ are then quasi-left continuous processes satisfying our standing assumption (H.2) in Section 5.1, as should be in view of application of general reflected BSDE results.

Suitable semi-group properties analogous to Propositions 8.3 and 8.6 in Part II, and existence of a consistent solution in the above sense to the Markovian decoupled Forward Backward SDE with data $\mathcal{G}$, $\mathcal{C}$ and τ (cf. Theorems 9.1, 9.3 and Proposition 9.4 in Part II), can then be established like in Part II (see also Theorem 16.12 in Part IV below).

Remark 15.2 The fact that the value functions u and v are defined by continuous Cauchy cascades can be established much like Theorem 16.12 below. However the proof is much simpler here. For this reason we provide no proof here, referring the reader to the proof of Theorem 16.12 for similar arguments in a more complex situation.

The next step consists in deriving analytic characterizations of the value functions u and v in terms of viscosity solutions to related doubly reflected and reflected partial integro-differential problems.

Reasoning as in Part III (cf. the proof of Proposition 14.1 for a review of the main arguments), one can thus show,

Proposition 15.1 *Under the currently extended model dynamics for $\mathcal{X}$ (with deterministic jumps in X as specified by θ):*
(i) *All the results of Part II still hold true, using the previously amended notions of solutions to the related FBSDEs;*
(ii) *For every $l = 1, \ldots, m$,*
• u_l *is the* unique $\mathcal{P}$*-solution, the* maximal $\mathcal{P}$-subsolution *and the* minimal $\mathcal{P}$-supersolution *of $(\mathcal{V}2)$ on $\mathcal{E}_l$ with terminal condition $u^i_{l+1}(T_l, \theta^i_l(x))$ — with u_{l+1} in the sense of Φ, in case $l = m$ — on $\partial_p\mathcal{E}_l$,*
• v_l *is the* unique $\mathcal{P}$*-solution, the* maximal $\mathcal{P}$-subsolution *and the* minimal $\mathcal{P}$-supersolution *of $(\mathcal{V}1)$ on $\mathcal{E}_l$ with boundary condition u_l on $\partial_p\mathcal{D}_l$.*

Part (ii) of this Proposition is thus the generalization to the present set-up of Proposition 12.4 in Part III. As for the approximation arguments of Section 13, they can only be used in the present set-up for establishing that, for l decreasing from m to 1:
• $u_{l,h} \to u_l$ locally uniformly on $\mathcal{E}_l$ as $h \to 0$, under the working assumption that the true value for $u^i_l(T_l, x) = u^i_{l+1}(T_l, \theta^i_l(x))$ is plugged at T_l in the approximation scheme for u_l;
• $v_{l,h} \to v_l$ locally uniformly on $\mathcal{E}_l$ as $h \to 0$, under the working assumption that the true value for u_l is plugged on $\partial_p\mathcal{D}_l$ in the approximation scheme for v_l, and provided $v_{l,h} \to v_l(= u_l)$ on $\partial_p\mathcal{D}_l \cap \{t < T_l\}$.
Of course, in practice (cf. also Remark 13.3):
• u_l is only approximately known at T_l (except for $l = m$) when it comes to approximating u_l on $\mathcal{E}_l$, using the already computed function $u_{l+1,h}$ at T_l as input data;
• v_l is only approximately known on $\partial_p\mathcal{D}^l$ when it comes to approximating v_l on $\mathcal{E}_l$, using the already computed function $u_{l,h}$ on $\partial_p\mathcal{D}_l$ as input data.
There is thus need for improvement in these approximation results.

15.2 Case of a Marker Process N

We motivated the introduction of deterministic jumps in the factor process X in Section 15.1 by its use in modeling discrete dividends on a primary asset underlying a financial derivative, the primary asset being given as one of the components of X in our generic factor process $\mathcal{X} = (X, N)$.

Still in the context of pricing problems in finance, there is another important motivation for introducing deterministic jumps in the factor process X, related to the issue of extension of the state space when dealing with *discretely path-dependent* financial derivatives.

To make it as simple as possible, let us thus consider an European option with payoff $\Phi(S_{T_0}, S_{T_1}, \ldots, S_{T_m})$ at maturity time $T_m = T$, where S represents an underlying stock price process. Such payoffs are for instance to be found in *cliquet options*, *volatility and variance swaps*, or *discretely monitored Asian options*. As is well-known, these can often be priced efficiently by PDE methods after an appropriate extension of the state space. We refer the reader to Windcliff et al. [95, 94] for illustrations in the cases of cliquet options and volatility and variance swaps, respectively.

Now, provided one works with a suitably extended state space, the methods and results of the present paper are indeed applicable to such forms of path-dependence, with all the consequences in terms of pricing and hedging developed in Part I.

Let us thus assume S to be given as a standard jump-diffusion, to fix ideas. A first possibility would be to introduce the extended factor process $X_t = (S_t, S^0_t, \ldots, S^{m-1}_t)$, where the auxiliary factor processes S^ls are equal to 0 before T_l and to S_{T_l} on $[T_l, T]$. Since this extended factor process X exhibits deterministic jumps at times T_ls, we are in the set-up of Section 15.1 (case of a degenerate model $\mathcal{X} = (X, N) = X$ therein), which provides a second and important motivation for the developments and results of Section 15.1.
But this state space extension is not the only possible one. Exploiding the specific nature of the payoff function Φ, more parsimonious alternatives in state spaces like $\mathbb{R}^d$ for some $d < m$ rather then $\mathbb{R}^m$ above can often be found (see, e.g., Windcliff et al. [94, 95]).
An extreme situation in this regard is the one where it is enough to know whether the values of S at the T_ls are above or below some trigger levels, so that it is enough to extend the factor process into $\mathcal{X}_t = (X_t, N_t)$, where $X_t = S_t$ and where the *marker process* N_t represents a vector of indicator processes with deterministic jumps at the T_ls. By deterministic jumps here we mean jumps given by deterministic functions of the S_{T_l-}s, or equivalently, since S can only jump at totally unpredictable stopping times, of the S_{T_l}s.

One would thus like to be able to address the issue of *discretely monitored call protection* τ, like for instance,

Example 15.3 Given a constant trigger level $\bar{S}$ and an integer $\imath$,
(i) Call possible from the first time τ that S has been $\geq \bar{S}$ at the last $\imath$ monitoring times, Call protection before τ,
Or more generally, given a further integer $\jmath \geq \imath$,
(ii) Call possible from the first time τ that S has been $\geq \bar{S}$ on at least $\imath$ of the last $\jmath$ monitoring times, Call protection before τ.

As we shall see as an aside of the results of Section 16 (cf. Section 16.3.5), it is actually possible to deal with such (and even more general, see Example 4.4-16.6) forms of path-dependence, resorting to a 'degenerate variant' $\mathcal{X} = (X, N)$ of the general jump-diffusion setting with regimes of this paper, in which X is a Markovian jump-diffusion not depending on N, and the I-valued pure jump marker process N is constant except for deterministic jumps at the T_ls, from $N^t_{T_l-}$ to

$$N^t_{T_l} = \theta_l(\mathcal{X}^t_{T_l-}) , \tag{236}$$

for a suitable *jump function* θ.

Comments 15.1 In this set-up,
(i) In the notation of Section 7.1, $\mathbb{F}_{\nu^t}$ is embedded into $\mathbb{F}_{X^t}$ which is itself embedded into $\mathbb{F}_{B^t} \vee \mathbb{F}_{\chi^t}$. Therefore $\mathbb{F}^t = \mathbb{F}_{B^t} \vee \mathbb{F}_{\chi^t} \vee \mathbb{F}_{\nu^t} = \mathbb{F}_{B^t} \vee \mathbb{F}_{\chi^t}$, where $(\mathbb{F}_{B^t} \vee \mathbb{F}_{\chi^t}, \mathbb{P}^t; B^t, \chi^t)$ has the local martingale predictable representation property (same proof as Proposition 7.2(ii)). As a consequence, there are no ν^t – martingale components in any of the related forward or backward SDEs.
(ii) Since X does not depend on N, the error estimate (137) on X and the estimates on $\widetilde{\mathcal{Y}}$ in Proposition 8.2 are still valid, independently of the error estimate (136) on N. Incidentally note that the latter estimate does not hold anymore, since N now depends on X via (236), even under the original measure $\mathbb{P}$ (before the change of measure to $\mathbb{P}^t$).

15.3 General Case

The situations of Sections 15.1 and 15.2 can both be regarded as special cases, covering many practical pricing applications, of deterministic jumps of the factor process $\mathcal{X}$ at fixed times T_ls. However from the mathematical point of view it is interesting to note that the general case of deterministic jumps of $\mathcal{X}$ from $\mathcal{X}_{T_l-}$ to $\mathcal{X}_{T_l} = \theta_l(\mathcal{X}_{T_l-})$ at the T_ls, for a suitable function θ, seems difficult to deal with. Indeed, as soon as N depends on X via its jumps at the T_ls (even under the original measure $\mathbb{P}$):
• First, the error estimate (136) on N is not valid anymore. The error estimate (137) on X and the continuity results on $\widetilde{\mathcal{Y}}$ and $\overline{\mathcal{Y}}$ in Propositions 8.2(ii) and 8.5(ii), which all relied on (136), are therefore not available either (at least, not by the same arguments as before), unless we are in the special case of Section 15.3 where X does not depend on N;
• Second, the martingale representation property of Proposition 7.2(ii) under the original measure $\mathbb{P}$, which was used to derive the martingale representation property under the equivalent measure $\mathbb{P}^t$ at Proposition 7.5(ii), becomes subject to caution, inasmuch as N and B are not independent anymore (not even under the original measure $\mathbb{P}$), unless we are in the special case of Section 15.2 where $\mathbb{F}^t = \mathbb{F}_{B^t} \vee \mathbb{F}_{\chi^t}$.

16 Intermittent Upper Barrier

16.1 Financial Motivation

A more general form of call protection than those considered earlier in Parts II and III consists in 'intermittent' (or 'Bermudan') call protection. In the financial set-up of Part I, this involves considering generalized upper payoff processes of the form

$$\bar{U}_t = \Omega_t^c \infty + \Omega_t U_t \tag{237}$$

for given càdlàg event indicator processes $\Omega_t, \Omega_t^c = 1 - \Omega_t$, rather than more specifically (cf. (107))

$$\bar{U}_t = \mathbb{1}_{\{t<\tau\}} \infty + \mathbb{1}_{\{t \geq \tau\}} U_t \tag{238}$$

for a stopping time τ.

Let a non-decreasing sequence of $[0,T]$-valued stopping times τ_ls be given, with $\tau_0 = 0$ and $\tau_l = T$ for l large enough, almost surely. We assume that a call protection is active at time

0, and that every subsequent time τ_l is a time of *switching* between call protection and no protection. Thus, for $t \in [0,T]$,

$$\Omega_t = \mathbb{1}_{\{l_t \text{ odd}\}}\ , \tag{239}$$

where l_t is the index l of the random time interval $[T_l, T_{l+1})$ containing t.

Remark 16.1 Considering sequences τ such that $\tau_0 = \tau_1 = 0$ and $\tau_2 > 0$ almost surely, one recovers in this formalism the case where the protection is inactive on the first non-empty time interval.

In the special case of a doubly reflected BSDE of the form (15) with a generalized effective call payoff process $\bar{U}$ as of (237), (239) therein, the identification between the *arbitrage* or *infimal super-hedging price process* of the related financial derivative and the state-process $Y = \Pi$ of a solution, *assumed to exist*, to (15), can be established by a straightforward adaptation of the arguments developed in Part I (see Section 16.2.1).

Remark 16.2 We shall see shortly that in the present set-up the possibility of jumps from finite to infinite values in $\bar{U}$ leads to relax the continuity condition on the process K in the Definition 2.7 of a solution (see Definition 16.3 below). This is why one is led to a notion of infimal (rather than minimal) super-hedging price in the financial interpretation. See Bielecki et al. [16, Long Preprint Version] or Chassagneux et al. or [32] for more about this.

However doubly reflected BSDEs with generalized upper barriers as of (237), (239) are not handled in the literature. This section aims at filling this gap by showing that such BSDEs are well-posed under suitable assumptions, and by establishing the related analytic approach in the Markovian case.

To start with, the results of Section 16.2 extend to more general RIBSDEs (see Definition 16.3 and Remark 5.5) the abstract RDBSDE results of Crépey and Matoussi [42]: general well-posedness (in the sense of existence, uniqueness and a priori estimates) and comparison results. In order to recover the results of [42], simply consider in Section 16.2 the special case of a non-decreasing sequence of stopping time $\tau = (\tau_l)_{l \geq 0}$ such that $\tau_2 = T$ almost surely, so $\tau_l = \tau_2 = T$ for $l \geq 2$ — with the only difference that the component K of a solution to an RDBSDE is by definition given as a continuous process, whereas this continuity condition on K has to be relaxed in the case of a more general RIBSDE.

We then deal with the Markovian case in Section 16.3.

16.2 General Set-Up

In this section one works in the general set-up and under the assumptions of Section 5. Let us further be given a non-decreasing sequence $\tau = (\tau_l)_{l \geq 0}$ of $[0,T]$-valued predictable stopping times τ_ls, with $\tau_0 = 0$ and $\tau_l = T$ for l large enough, almost surely. The *RIBSDE with data* (g, ξ, L, U, τ), where the 'I' in RIBSDE stands for 'intermittent', is the generalization of a R2BSDE in which the upper barrier U is only active on the 'odd' random time intervals $[\tau_{2l+1}, \tau_{2l+2})$. Essentially, we replace U by $\bar{U}$ in Definition 5.3(a)(iii), with for $t \in [0,T]$,

$$\bar{U}_t = \mathbb{1}_{\{l_t \text{ even}\}}\infty + \mathbb{1}_{\{l_t \text{ odd}\}}U_t \tag{240}$$

with l_t defined by $\tau_{l_t} \le t < \tau_{l_t+1}$. However this generalization leads to relax the continuity assumption on K in the solution. Let thus A^2 stand for the space of finite variation but not necessarily continuous processes K vanishing at time 0, with (non-decreasing, null at time 0) Jordan components denoted as usual by $K^{\pm}$.

Definition 16.3 An $(\Omega, \mathbb{F}, \mathbb{P}), (B, \mu)$-*solution* $\mathcal{Y}$ *to the RIBSDE with data* (g, ξ, L, U, τ) is a quadruple $\mathcal{Y} = (Y, Z, V, K)$, such that:

(i) $Y \in \mathcal{S}^2, Z \in \mathcal{H}^2_d, V \in \mathcal{H}^2_\mu, K \in A^2$,

(ii) $$Y_t = \xi + \int_t^T g_s(Y_s, Z_s, V_s)ds + K_T - K_t - \int_t^T Z_s dB_s - \int_t^T \int_E V_s(e)\widetilde{\mu}(ds, de)\ ,\ t \in [0, T]\ ,$$

(iii) $L \le Y$ on $[0, T]$, $Y \le \bar{U}$ on $[0, T]$ and $$\int_0^T (Y_{t-} - L_{t-})dK^+_t = \int_0^T (\bar{U}_{t-} - Y_{t-})dK^-_t = 0\ ,$$

where $\bar{U}$ is defined by (240), and with the convention that $0 \times \pm\infty = 0$ in (iii).

Remark 16.4 In the special case when $\tau_2 = T$ a.s. (so $\tau_l = \tau_2 = T$ for $l \ge 2$), the RIBSDE with data (g, ξ, L, U, τ) reduces to the RDBSDE with data (g, ξ, L, U, τ_1) (see Definition 5.4(ii)). If moreover $\tau_1 = 0$, then one deals with an R2BSDE.

16.2.1 Verification Principle

Given $t \in [0, T]$, let $\mathcal{T}_t$ denote the set of $[t, T]$-valued stopping times. The following *Verification Principle*, stated without proof, is an easy generalization of Proposition 5.2 in Part II. From the point of view of the financial application, this result can be used to establish the abovementioned connection between the arbitrage price process of a game option with call protection τ and the state-process Y of a solution, *assumed to exist*, to the related RIBSDE (see Remark 16.2).

Proposition 16.1 (Verification Principle) *If* $\mathcal{Y} = (Y, Z, V, K)$ *solves the RIBSDE with data* (g, ξ, L, U, τ), *then the state process* Y *is the conditional value process of the Dynkin game with payoff functional given by, for any* $t \in [0, T]$ *and* $\rho, \theta \in \mathcal{T}_t$,

$$J(t; \rho, \theta) = \int_t^{\rho\wedge\theta} g_s(Y_s, Z_s, V_s)ds + L_\theta \mathbb{1}_{\{\rho\wedge\theta=\theta<T\}} + \bar{U}_\rho \mathbb{1}_{\{\rho<\theta\}} + \xi \mathbb{1}_{[\rho\wedge\theta=T]}\ .$$

More precisely, for every $\varepsilon > 0$, *an* ε *– saddle-point of the game at time t is given by:*

$$\rho^\varepsilon_t = \inf\Big\{ s \in [t, T]\,;\, Y_s \ge \bar{U}_u - \varepsilon \Big\} \wedge T\ ,\ \theta^\varepsilon_t = \inf\Big\{ s \in [t, T]\,;\, Y_s \le L_u + \varepsilon \Big\} \wedge T\ .$$

So, for any $\rho, \theta \in \times\mathcal{T}_t$,

$$\mathbb{E}\big[J(t; \tau, \theta)\big|\mathcal{F}_t\big] - \varepsilon \le Y_t \le \mathbb{E}\big[J(t; \rho, \theta)\big|\mathcal{F}_t\big] + \varepsilon\ . \tag{241}$$

Of course, given the definition of $\bar{U}$ in (240), this Dynkin game effectively reduces to a 'constrained Dynkin game' with upper payoff process U (instead of $\bar{U}$ in Proposition 16.1), which would be posed over the constrained set of stopping policies $(\rho,\theta)\in\bar{\mathcal{T}}_t\times\mathcal{T}_t$, where $\bar{\mathcal{T}}_t$ denotes the set of the $\cup_{l\geq 0}[\tau_{2l+1}\vee t,\tau_{2l+2}\vee t)\cup\{T\}$ – valued stopping times. In particular, one has,

$$\rho_t^\varepsilon=\inf\Big\{\, s\in\cup_{l\geq 0}[\tau_{2l+1}\vee t,\tau_{2l+2}\vee t)\,;\,Y_s\geq U_u-\varepsilon\Big\}\wedge T\;.$$

16.2.2 A Priori Estimates and Uniqueness

Recall that a *quasimartingale* L is a difference of two non-negative supermartingales. The following classical results about quasimartingales can be found, for instance, in Dellacherie and Meyer [46] (see also Protter [91]).

Lemma 16.2 **(i)** (See Section VI.40 of [46]) *Among the various decompositions of a quasimartingale X as a difference of two non-negative supermartingales X^1 and X^2, there exists a unique decomposition $X=\bar{X}^1-\bar{X}^2$, called* the Rao decomposition *of X, which is* minimal *in the sense that $X^1\geq\bar{X}^1$, $X^2\geq\bar{X}^2$, for any such decomposition $X=X^1-X^2$.*
(ii) (See Appendix 2.4 of [46]) *Any quasimartingale* X belonging to $\mathcal{S}^2$ *is a* special semimartingale *with canonical decomposition*

$$X_t=X_0+M_t+A_t\;,\;t\in[0,T] \tag{242}$$

for a uniformly integrable martingale M and a predictable process of integrable variation A.

The following estimates are immediate extensions to RIBSDEs of the analogous results which were established for R2BSDEs and RDBSDEs in [42].

Theorem 16.3 *We consider a sequence of RIBSDEs with data and solutions indexed by n, but for a* common *sequence τ of stopping times, and with lower barrier L_ns given as quasimartingales in $\mathcal{S}^2$, with predictable finite variation components denoted by A_ns (cf. (242)). The data are assumed to be bounded in the sense that the driver coefficients $g^n=g^n_t(y,z,v)$'s are uniformly Λ – Lipschitz continuous in (y,z,v), and one has for some constant c_1 :*

$$\|\xi^n\|_2^2+\|g^n_\cdot(0,0,0)\|^2_{\mathcal{H}^2}+\|L^n\|^2_{\mathcal{S}^2}+\|U^n\|^2_{\mathcal{S}^2}+\|A^{n,-}\|^2_{\mathcal{S}^2}\;\leq\;c_1\;. \tag{243}$$

Then one has for some constant $c(\Lambda)$:

$$\|Y^n\|^2_{\mathcal{S}^2}+\|Z^n\|^2_{\mathcal{H}^2_d}+\|V^n\|^2_{\mathcal{H}^2_\mu}+\|K^{n,+}\|^2_{\mathcal{S}^2}+\|K^{n,-}\|^2_{\mathcal{S}^2}\leq c(\Lambda)c_1\;. \tag{244}$$

Indexing by ^{n,p} the differences $\cdot^n-\cdot^p$, one also has:

$$\begin{aligned}&\|Y^{n,p}\|^2_{\mathcal{S}^2}+\|Z^{n,p}\|^2_{\mathcal{H}^2_d}+\|V^{n,p}\|^2_{\mathcal{H}^2_\mu}+\|K^{n,p}\|^2_{\mathcal{S}^2}\leq\\ &\quad c(\Lambda)c_1\Big(\|\xi^{n,p}\|_2^2+\|g^{n,p}_\cdot(Y^n_\cdot,Z^n_\cdot,V^n_\cdot)\|^2_{\mathcal{H}^2}+\|L^{n,p}\|_{\mathcal{S}^2}+\|U^{n,p}\|_{\mathcal{S}^2}\Big)\;.\end{aligned} \tag{245}$$

Assume further $dA^{n,-}\leq\alpha^n_t dt$ for some progressively measurable processes α^n with $\|\alpha^n\|_{\mathcal{H}^2}$ finite for every $n\in\mathbb{N}$. Then one may replace $\|L^n\|^2_{\mathcal{S}^2}$ and $\|L^{n,p}\|_{\mathcal{S}^2}$ by $\|L^n\|^2_{\mathcal{H}^2}$ and $\|L^{n,p}\|_{\mathcal{H}^2}$ in (243) and (245).

Suppose additionally that $\|\alpha^n\|_{\mathcal{H}^2}$ is bounded over $\mathbb{N}$ and that when $n \to \infty$:
• *$g^n_\cdot(Y_\cdot, Z_\cdot, V_\cdot)$ $\mathcal{H}^2$-converges to $g_\cdot(Y_\cdot, Z_\cdot, V_\cdot)$ locally uniformly w.r.t. $(Y, Z, V) \in \mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu$, and*
• *(ξ^n, L^n, U^n) $\mathcal{L}^2 \times \mathcal{H}^2 \times \mathcal{S}^2$-converges to (ξ, L, U).*
Then (Y^n, Z^n, V^n, K^n) $\mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu \times \mathcal{S}^2$-converges to a solution (Y, Z, V, K) of the limiting RIBSDE with data (g, ξ, L, U, τ). Moreover, (Y, Z, V, K) also satisfies (244)–(245) 'with $n = \infty$' therein.

Moreover, in the special case $L^{n,p} = U^{n,p} = 0$, one has like for R2BSDEs (cf. Appendix A of [42]) that estimate (245) holds, with $L^{n,p} = U^{n,p} = 0$ therein, irrespectively of the specific assumptions on the L_ns in Theorem 16.3. In particular,

Proposition 16.4 *Uniqueness holds for an RIBSDE satisfying the standing assumptions (H.0)–(H.1)–(H.2).*

16.2.3 Comparison

In this section we specialize the general assumption (H.1) in Section 5.1 to the case where (cf. section 4 of [42])

$$g_t(y, z, v) = \widetilde{g}_t\Big(y, z, \int_E v(e)\eta_t(e)\zeta_t(e)\,\rho(de)\Big)\,, \tag{246}$$

for a $\widetilde{\mathcal{P}}$-measurable non-negative function $\eta_t(e)$ with $|\eta_t|_t$ uniformly bounded, and a $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^{1\otimes d}) \otimes \mathcal{B}(\mathbb{R})$-measurable function $\widetilde{g} : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^{1\otimes d} \times \mathbb{R} \to \mathbb{R}$ such that:
(H.1.i)' $\widetilde{g}_\cdot(y, z, r)$ is a progressively measurable process, for any $y \in \mathbb{R}$, $z \in \mathbb{R}^{1\otimes d}$, $r \in \mathbb{R}$;
(H.1.ii)' $\|\widetilde{g}_\cdot(0, 0, 0)\|_{\mathcal{H}^2} < +\infty$;
(H.1.iii)' $|\widetilde{g}_t(y, z, r) - \widetilde{g}_t(y', z', r')| \le \Lambda\big(|y - y'| + |z - z'| + |r - r'|\big)$, for any $t \in [0, T]$, $y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^{1\otimes d}$ and $r, r' \in \mathbb{R}$;
(H.1.iv)' $r \mapsto \widetilde{g}_t(y, z, r)$ is non-decreasing, for any $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^{1\otimes d}$.

Using in particular the fact that

$$\left|\int_E (v(e) - v'(e))\eta_t(e)\zeta_t(e)\,\rho(de)\right| \le |v - v'|_t |\eta_t|$$

with $|\eta_t|_t$ uniformly bounded, so g defined by (246) satisfies (H.1).

The following RIBSDE comparison result is then an easy generalization of the R2BSDE comparison result of Crépey and Matoussi [42].

Theorem 16.5 *Let (Y, Z, V, K) and (Y', Z', V', K') be solutions to the RIBSDEs with data (g, ξ, L, U, τ) and $(g', \xi', L', U', \tau')$ satisfying assumptions (H.0)–(H.1)–(H.2). We assume further that g satisfies (H.1)'. Then $Y \le Y'$, $d\mathbb{P} \otimes dt$ – almost everywhere, whenever:*
(i) $\xi \le \xi'$, $\mathbb{P}$ – almost surely,
(ii) $g_\cdot(Y'_\cdot, Z'_\cdot, V'_\cdot) \le g'_\cdot(Y'_\cdot, Z'_\cdot, V'_\cdot)$, $d\mathbb{P} \otimes dt$ – almost everywhere,
(iii) $L \le L'$ and $\bar{U} \le \bar{U}'$, $d\mathbb{P} \otimes dt$ – almost everywhere, where $\bar{U}$ is defined by (240) and $\bar{U}'$ is the analogous process relative to τ'.

16.2.4 Existence

We work here under the following *square integrable martingale predictable representation* assumption:
(**H**) Every square integrable martingale M admits a representation

$$M_t = M_0 + \int_0^t Z_s\,dB_s + \int_0^t \int_E V_s(e)\widetilde{\mu}(ds,de)\,,\quad t\in[0,T] \tag{247}$$

for some $Z\in\mathcal{H}^2_d$ and $V\in\mathcal{H}^2_\mu$.

We also strengthen Assumption (H.2.i) into:
$(\mathbf{H.2.i})'$ L and U are càdlàg *quasi-left continuous* processes in $\mathcal{S}^2$.
Recall that for a càdlàg process X, quasi-left continuity is equivalent to the existence of sequence of totally inaccessible stopping times which exhausts the jumps of X, whence ${}^pX = X_{\cdot-}$ (see, e.g., Jacod–Shiryaev [67, Propositions I.2.26 page 22 and I.2.35 page 25]). We thus work in this section under assumptions (H)–(H.0)–(H.1)–(H.2)', where (H.2)' denotes (H.2) with (H.2.i) replaced by (H.2.i)' therein.

Finally we postulate the so-called *the Mokobodski condition* (see [42]), namely the existence of a quasimartingale X with Rao components in $\mathcal{S}^2$ and such that $L\le X\le U$ over $[0,T]$. In view of Lemma 16.2, This is tantamount to the existence of non-negative supermartingales X^1,X^2 belonging to $\mathcal{S}^2$ and such that $L\le X^1-X^2\le U$ over $[0,T]$. X is then obviously a quasimartingale in $\mathcal{S}^2$. The Mokobodski condition is of course satisfied when L is a quasimartingale with Rao components in $\mathcal{S}^2$, as for instance under the general assumptions of Theorem 16.3.

The following two lemmas establish existence of a solution in the special cases of RIBSDEs that are effectively reducible to problems with only one call protection switching time involved.

The first case of this kind is that of a RDBSDE (or RIBSDE with $\tau_2=T$, see Remark 16.4).

Lemma 16.6 *Assuming (H)–(H.0)–(H.1)–(H.2)' and the Mokobodski condition, then, in the special case when $\tau_2 = T$ almost surely, the RIBSDE with data (g,ξ,L,U,τ) has a (unique) solution (Y,Z,V,K). Moreover K is continuous.*

Proof. Under the present assumptions, existence of a solution to a RDBSDE was established in Crépey and Matoussi [42] (in which continuity of the related process K is part of the definition of a solution), by 'pasting' in a suitable way the solution of a related R2BSDE over $[\tau_2,T]$ with that of a related RBSDE over $[0,\tau_2]$. □

We now consider the case where $\tau_1=0$ and $\tau_3=T$ almost surely, so that the upper barrier U is effectively active on $[0,\tau_2)$, and inactive on $[\tau_2,T)$ (cf. Remark 16.1).

Let $[\![\theta]\!]$ denotes the graph of a stopping time θ.

Lemma 16.7 *Assuming (H)–(H.0)–(H.1)–(H.2)' and the Mokobodski condition, then, in the special case when $0=\tau_1\le\tau_2\le\tau_3=T$ almost surely, the RIBSDE with data (g,ξ,L,U,τ) has a solution (Y,Z,V,K). Moreover, K^+ is a continuous process, and*

$$\{(\omega,t)\,;\,\Delta K^-_t\neq 0\}\subseteq[\![\tau_2]\!]\,,\ \Delta Y_{\tau_2}=\Delta K^-_{\tau_2}\,.$$

Proof. The solution (Y,Z,V,K) can be obtained by an elementary two-stages construction analogous to that used for establishing existence of a solution to a RDBSDE in [42], by 'pasting' appropriately the solution $(\widehat{Y},\widehat{Z},\widehat{V},\widehat{K})$ of a related RBSDE over the random time interval $[\tau_2,T]$, with the solution $(\bar{Y},\bar{Z},\bar{V},\bar{K})$ of a related R2BSDE with terminal condition $\bar{Y}_{\tau_2}=\min(Y_{\tau_2},U_{\tau_2})$ over the random time interval $[0,\tau_2]$. The detail of this construction appears in the statement of Theorem 16.8(i) below. In particular $\Delta K^-_{\tau_2}$ is set in order to satisfy the constraint that $Y_{\tau_2}\le \bar{U}_{\tau_2}=U_{\tau_2}$. Note in this respect that the process U cannot jump at τ_2, by Assumption (H.2.i)' and the fact that the τ_ls are predictable stopping times. The random measure μ cannot jump at the predictable stopping time τ_2 either. Therefore the process (Y,Z,V,K) resulting from this construction is such that

$$Y_{\tau_2-}=U_{\tau_2}=U_{\tau_2-}=\bar{U}_{\tau_2-},$$

so that the jump of K^- at τ_2 respects the minimality condition corresponding to the last identity in Definition 16.3(iii). □

Iterated and alternate applications of Lemmas 16.6 and 16.7 yield the following existence result for an RIBSDE,

Theorem 16.8 *Let us be given a RIBSDE with data (g,ξ,L,U,τ). We assume (H)–(H.0)–(H.1)–(H.2)', the Mokobodski condition, and $\tau_{m+1}=T$ almost surely for some fixed index m.*
(i) *The following iterative construction is well-defined, for l decreasing from m to 0: $\mathcal{Y}^l=(Y^l,Z^l,V^l,K^l)$ is the $(\Omega,\mathbb{F},\mathbb{P}),(B,\mu)$ – solution, with K^l continuous, to the stopped RBSDE (for l even) or R2BSDE (for l odd) on $[0,T]$ with data*

$$\left\{\begin{array}{ll} g\ ,\ Y^{l+1}_{\tau_{l+1}}\ ,\ L\ ,\ \tau_{l+1} & (l\ \mathit{even})\\ g\ ,\ \min(Y^{l+1}_{\tau_{l+1}},U_{\tau_{l+1}})\ ,\ L\ ,\ U\ ,\ \tau_{l+1} & (l\ \mathit{odd})\end{array}\right. \tag{248}$$

where, in case $l=m$, $Y^{l+1}_{\tau_{l+1}}$ is to be understood as ξ (so $\min(Y^{l+1}_{\tau_{l+1}},U_{\tau_{l+1}})=\min(\xi,U_T)=\xi$).
(ii) *Let us define $\mathcal{Y}=(Y,Z,V,K)$ on $[0,T]$ by, for every $l=0,\dots,m$:*
- *$(Y,Z,V)=(Y^l,Z^l,V^l)$ on $[\tau_l,\tau_{l+1})$, and also at $\tau_{m+1}=T$ in case $l=m$,*
- *$dK=dK^l$ on (τ_l,τ_{l+1}),*

$$\Delta K_{\tau_l}=Y^l_{\tau_l}-\min(Y^l_{\tau_l},U_{\tau_l})=\Delta Y_{\tau_l}(=0\ \mathit{for}\ l\ \mathit{odd})$$

and $\Delta K_T=\Delta Y_T=0$.

Then $\mathcal{Y}=(Y,Z,V,K)$ is the $(\Omega,\mathbb{F},\mathbb{P}),(B,\mu)$ – solution to the RIBSDE with data (g,ξ,L,U,τ). Moreover, K^+ is a continuous process, and

$$\{(\omega,t)\,;\,\Delta K^-_t\neq 0\}\subseteq\bigcup_{\{l\ \mathit{even}\}}[\![\tau_l]\!]\ ,\ \Delta Y=\Delta K^-\ \mathit{on}\ \bigcup_{\{l\ \mathit{even}\}}[\![\tau_l]\!]\ .$$

Remark 16.5 A natural conjecture is that under (H)–(H.0)–(H.1)–(H.2)' and the Mokobodski condition, the RIBSDE with data (g,ξ,L,U,τ) has a solution (Y,Z,V,K). In other words, our guess is that it is possible to get rid of the condition that $\tau_{m+1}=T$ almost surely for some fixed index m in Theorem 16.8. In the case of a Brownian filtration (so $\mathbb{F}=\mathbb{F}_B$ and there is no random measure μ involved), this actually follows by application of the results of

Peng and Xu [89]. More precisely, this follows from an immediate extension of these results to the case of an $\mathbb{R}\cup\{+\infty\}$ – valued upper barrier $\bar{U}$, noting that the results of Peng and Xu [89], which are based on Peng [88], even if stated for real-valued barriers, only use the fact that $\bar{U}^- = U^-$ lies in $\mathcal{S}^2$. This is of course verified under the standing assumption (H.2.i) of this paper (see section 5.1). Moreover it is apparent that the penalization approach and the related results of Peng [88] and Peng and Xu [89] can be extended in a rather straightforward way to the more general case of a filtration $\mathbb{F} = \mathbb{F}_B \vee \mathbb{F}_\mu$, which would then establish the above conjecture. Since Theorem 16.8 is enough for our purposes in this article, we shall not push this further however.

16.3 Markovian Set-Up

16.3.1 Jump-Diffusion Set-Up with Marker Process

We now specify the previous set-up to a Markovian jump-diffusion model with marker $\mathcal{X} = (X, N)$ as of Section 15.2, in which X is a Markovian jump-diffusion not depending on N, and the I-valued pure jump marker process N is constant except for deterministic jumps at the times T_ls, from $N^t_{T_l-}$ to

$$N^t_{T_l} = \theta_l(\mathcal{X}^t_{T_l-}) \; , \tag{249}$$

for a suitable jump function θ.

Again (see Remark 15.1), in this set-up:
- $(\mathbb{F}^t = \mathbb{F}_{B^t} \vee \mathbb{F}_{\chi^t}, \mathbb{P}^t; B^t, \chi^t)$ has the local martingale predictable representation property,
- The error estimate (137) on X is still valid.

Let us set, for a regular function u over $[0, T] \times \mathbb{R}^d$ (cf. (113) and the related comments):

$$\begin{aligned}\mathcal{G}u(t,x) = \partial_t u(t,x) + \frac{1}{2}\mathrm{Tr}[a(t,x)\mathcal{H}v(t,x)] + \partial u(t,x)\widetilde{b}(t,x) \\ + \int_{\mathbb{R}^d}\Big(u(t,x+\delta(t,x,y)) - u(t,x)\Big)f(t,x,y)m(dy)\end{aligned} \tag{250}$$

with

$$\widetilde{b}(t,x) = b(t,x) - \int_{\mathbb{R}^d}\delta(t,x,y)f(t,x,y)m(dy) \; . \tag{251}$$

In the present set-up $\mathcal{G}$ defined by (250) is thus the generator of the Markov process X.

We now consider a Markovian RIBSDE with underlying factor process $\mathcal{X} = (X, N)$. More precisely, let us be given a family of RIBSDEs parameterized by the initial condition (t, x, i) of $\mathcal{X}^t$ (where the superscript ^{t} stands as usual in this article in reference to (t, x, i)), with the following data:
- The generator $\mathcal{G}$ of X defined by (250), and the specification of the jump size function θ of N in (249),
- Cost data $\mathcal{C}$ as of Section 6.4, assumed here not to depend on $i \in I$,
- τ defined by $\tau^t_0 = t$ and, for every $l \geq 0$ (to be compared with Hypothesis 10.1(iii) in Part III):

$$\tau^t_{2l+1} = \inf\{s > \tau^t_{2l}\,;\, N^t_s \notin \Delta\} \wedge T \; , \; \tau^t_{2l+2} = \inf\{s > \tau^t_{2l+1}\,;\, N^t_s \in \Delta\} \wedge T \; , \tag{252}$$

for a given subset Δ of I, resulting in an effective upper payoff process $\bar{U}$ of the Markovian form of (237) corresponding to the event-process

$$\Omega_s^t = \mathbb{1}_{N_s^t \notin \Delta} \; . \qquad (253)$$

Observe that since the cost data do not depend on i, the only impact of the marker process N is via its influence on τ. Also note that the τ_ls effectively reduce to $\mathcal{T}$-valued stopping times, and that one almost surely has $\tau_{m+1} = T$.

This Markovian set-up allows one to account for various forms of *intermittent path-dependent call protection*. Denoting by S_s^t the first component of the $\mathbb{R}^d$-valued process X_s^t and by S the first component of the mute vector-variable $x \in \mathbb{R}^d$, one may thus consider the following clauses of call protection, which correspond to Example 4.4 in Part I.

Example 16.6 Given a constant *trigger level* $\bar{S}$ and an integer $\imath \leq m$, τ^t of the form (252) above, with:
(i) $I = \{0, \ldots, \imath\}$, $\Delta = \{0, \ldots, \imath - 1\}$ and θ defined by

$$\theta_l^i(x) = \begin{cases} (i+1) \wedge \imath, & S \geq \bar{S} \\ 0, & S < \bar{S} \end{cases}$$

(which in this case does not depend on l). With the initial condition $N_t^t = 0$, N_s^t then represents the number of consecutive monitoring dates T_ls with $S_{T_l}^t \geq \bar{S}$ from time s backwards since the initial time t, capped at $\imath$. Call is possible whenever $N_s^t \geq \imath$, which means that S_s^t has been $\geq \bar{S}$ at the last $\imath$ monitoring times since the initial time t; Otherwise call protection is in force;
Or more generally, given a further integer $\jmath \geq \imath$,
(ii) $I = \{0,1\}^\jmath$ for some given integer $\jmath \in \{\imath, \ldots, m\}$, $\Delta = \{i \in I\,; |i| < \imath\}$ with $|i| = \sum_{1 \leq j \leq \jmath} i_j$, and θ defined by

$$\theta_l^i(x) = (\mathbb{1}_{S \geq \bar{X}}, i_1, \ldots, i_{d-1})$$

(independently of l). With the initial condition $N_t^t = 0_\jmath$, N_s^t then represents the vector of the indicator functions of the events $S_{T_l}^t \geq \bar{X}$ at the last $\jmath$ monitoring dates preceding time s since the initial time t. Call is possible whenever $|N_s^t| \geq \imath$, which means that S_s^t has been $\geq \bar{S}$ on at least $\imath$ of the last $\jmath$ monitoring times since the initial time t; Otherwise call protection is in force.

16.3.2 Well-Posedness of the Markovian RIBSDE

In the present set-up where $\mathbb{F}^t = \mathbb{F}_{B^t} \vee \mathbb{F}_{\chi^t}$, there are no ν^t – martingale components in any of the related forward or backward SDEs, and the definitions of $\widetilde{g}$ and $\widehat{g}$ (cf. (119), (160)) reduce to the following expressions, where in particular v denotes a generic element $v \in \mathcal{M}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m(dy); \mathbb{R})$:

$$\begin{aligned} \widetilde{g}(s, \mathcal{X}_s^t, y, z, v) &= g(s, \mathcal{X}_s^t, y, z, \widetilde{r}_s^t) \text{ with } \widetilde{r}_s^t = \widetilde{r}_s^t(v) = \int_{\mathbb{R}^d} v(y) f(s, \mathcal{X}_s^t, y) m(dy) \\ \widehat{g}(s, \mathcal{X}_s^t, y, z, \widehat{v}) &= g(s, \mathcal{X}_s^t, y, z, \widetilde{r}_s^t) + (\widetilde{r}_s^t - \widehat{r}) \text{ with } \widehat{r} = \widehat{r}(v) = \int_{\mathbb{R}^d} v(y) m(dy) \; . \end{aligned} \qquad (254)$$

Accordingly, the V^t-component of a solution to any Markovian BSDE (cf. Theorem 16.9) lives in $\mathcal{H}_{\mu^t}^2 = \mathcal{H}_{\chi^t}^2$.

Proposition 16.9 **(i)** *The following iterative construction is well-defined, for l decreasing from m to 0: $\mathcal{Y}^{l,t} = (Y^{l,t}, Z^{l,t}, V^{l,t}, K^{l,t})$ is the $(\Omega, \mathbb{F}^t, \mathbb{P}^t), (B^t, \mu^t)$ – solution, with $K^{l,t}$ continuous, to the stopped RBSDE (for l even) or R2BSDE (for l odd) on $[t, T]$ with data*

$$\begin{cases} \mathbb{1}_{\{s>t\}}\widetilde{g}(s, \mathcal{X}_s^t, y, z, v) \ , \ Y^{l+1,t}_{\tau^t_{l+1}} \ , \ \ell(s \vee t, \mathcal{X}^t_{s\vee t}) \ , \ \tau^t_{l+1} & (l \text{ even}) \\ \mathbb{1}_{\{s>t\}}\widetilde{g}(s, \mathcal{X}_s^t, y, z, v) \ , \ \min(Y^{l+1,t}_{\tau^t_{l+1}}, h(\tau^t_{l+1}, \mathcal{X}^t_{\tau^t_{l+1}})) \ , \ \ell(s \vee t, \mathcal{X}^t_{s\vee t}) \ , \ h(s \vee t, \mathcal{X}^t_{s\vee t}) \ , \ \tau^t_{l+1} & (l \text{ odd}) \end{cases} \tag{255}$$

where, in case $l = m$, $Y^{l+1,t}_{\tau^t_{l+1}}$ is to be understood as $\Phi(\mathcal{X}^t_T)$.

Let $\mathcal{Y}^t = (Y^t, Z^t, V^t, K^t)$ be defined in terms of the $\mathcal{Y}^{l,t}$s as $\mathcal{Y}$ in terms of the $\mathcal{Y}^l$s in Theorem 16.8(ii). So in particular $Y^t = Y^{l,t}$ on $[\tau^t_l, \tau^t_{l+1})$, for every $l = 0, \ldots, m$, and

$$Y^t_t = \begin{cases} Y^{0,t}_t, & i \in \Delta \\ Y^{1,t}_t, & i \notin \Delta \end{cases} \tag{256}$$

Then $\mathcal{Y}^t$ is the $(\Omega, \mathbb{F}^t, \mathbb{P}^t), (B^t, \mu^t)$ – solution to the RIBSDE on $[t, T]$ with data

$$\widetilde{g}(s, \mathcal{X}^t_s, y, z, v) \ , \ \Phi(\mathcal{X}^t_T) \ , \ \ell(s, \mathcal{X}^t_s) \ , \ h(s, \mathcal{X}^t_s) \ , \ \tau^t \ . \tag{257}$$

(ii) *For every $l = 0, \ldots, m$, we extend $Y^{l,t}$ by $Y^{l,t}_t$, and $K^{l,t}_t$, $Z^{l,t}$ and $V^{l,t}$ by 0 on $[0, t]$. Then, for every $l = m, \ldots, 0$: $\mathcal{Y}^{l,t} = (Y^{l,t}, Z^{l,t}, V^{l,t}, K^{l,t})$ is the $(\Omega, \mathbb{F}, \mathbb{P}), (B, \mu)$ – solution, with $K^{l,t}$ continuous, to the stopped RBSDE (for l even) or R2BSDE (for l odd) on $[0, T]$ with data as of (255), with $\widehat{g}$ instead of $\widetilde{g}$ therein.*

Proof. Part (i) follows by application of Proposition 16.8. Identity (256) simply results from the fact that, since $Y^t = Y^{l,t}$ on $[\tau^t_l, \tau^t_{l+1})$,

$$\begin{cases} Y^t_t = Y^{0,t}_t, & N^t_t \in \Delta \\ Y^t_t = Y^{1,t}_t, & N^t_t \notin \Delta \end{cases} \tag{258}$$

with $N^t_t = i$. Part (ii) then follows from part (i) as in the proof of Theorem 8.4. □

Our next goal is to derive stability results on $\mathcal{Y}^t$, or, more precisely, on the $\mathcal{Y}^{l,t}$s. Toward this end a suitable stability assumption on τ^t is needed. Note that in the present set-up assuming the τ^t_ls continuous, which would be the 'naive analog' of Assumption 8.1, would be too strong for practical purposes. This is for instance typically not satisfied in the situations of Example 16.6. One is thus led to introduce the following weaker

Assumption 16.7 Viewed as a random function of the initial condition (t, x, i) of $\mathcal{X}$, then, at every (t, x, i) in $\mathcal{E}$, τ is, almost surely:
(i) continuous at (t, x, i) if $t \notin \mathcal{T}$, and right-continuous at (t, x, i) if $t \in \mathcal{T}$,
(ii) left-limited at (t, x, i) if $t = T_l \in \mathcal{T}$ and θ_l is continuous at (x, i).

By this, we mean that:
- $\tau^{t_n} \to \tau^t$ if $(t_n, x_n, i) \to (t, x, i)$ with $t \notin \mathcal{T}$, or, for $t = T_l \in \mathcal{T}$, if $\mathcal{E}_{l+1} \ni (t_n, x_n, i) \to (T_l, x, i)$;
- if $\mathcal{E}^*_l \ni (t_n, x_n, i) \to (t = T_l, x, i)$ and that θ_l is continuous at (x, i), then τ^{t_n} converges to some non-decreasing sequence, denoted by $\widetilde{\tau}^t$, of predictable stopping times, such that in particular $\widetilde{\tau}^t_l = T$ for $l \geq m + 1$.

Definition 16.8 One denotes by $\widetilde{\mathcal{Y}}^t = (\widetilde{\mathcal{Y}}^{l,t})_{0\le l\le m}$, with $\widetilde{\mathcal{Y}}^{l,t} = (\widetilde{Y}^{l,t}, \widetilde{Z}^{l,t}, \widetilde{V}^{l,t}, \widetilde{K}^{l,t})$ and $\widetilde{K}^{l,t}$ continuous for every $l = 0, \dots, m$, the sequence of solutions of stopped RBSDEs (for l even) or R2BSDEs (for l odd) which is obtained by substituting $\widetilde{\tau}^t$ to τ^t in the construction of $\mathcal{Y}^t$ in Theorem 16.9(i).

Observe that since the τ_ls are in fact $\mathcal{T}$-valued stopping times:
• The continuity assumption effectively means that $\tau_l^{t_n} = \tau_l^t$ for n large enough, almost surely, for every $l = 1, \dots, m+1$ and $\mathcal{E} \ni (t_n, x_n, i) \to (t, x, i) \in \mathcal{E}$ with $t \notin \mathcal{T}$;
• The right-continuity, resp. left-limit assumption, effectively means that for n large enough $\tau_l^{t_n} = \tau_l^t$, resp. $\widetilde{\tau}_l^t$, almost surely, for every $l = 1, \dots, m+1$ and $\mathcal{E}_{l+1} \ni$, resp. $\mathcal{E}_l^* \ni$ $(t_n, x_n, i) \to (T_l, x, i) \in \mathcal{E}$.

Remark 16.9 It is intuitively clear, though we shall not try to formally establish at the level of this article, that Assumption 16.7 is satisfied in the situations of Example 16.6, in case the jump-diffusion X is uniformly elliptic in the direction of its first component S (cf. Example 8.2). We refer the reader to [32] for a precise statement and proof in a diffusion set-up.

Moreover we make the following additional hypothesis on the upper payoff function h, whereas the lower payoff function ℓ is still supposed to satisfy assumption (M.3). Also recall that in this section the cost data $\mathcal{C}$, including the function h, do not depend on $i \in I$.

Assumption 16.10 h is a Lipschitz function of (t, x).

Then,

Theorem 16.10 *Let $\mathcal{Y}^t = (\mathcal{Y}^{l,t})_{0\le l\le m}$ and $\widetilde{\mathcal{Y}}^t = (\widetilde{\mathcal{Y}}^{l,t})_{0\le l\le m}$ be defined as in Theorem 16.9(i) and Definition 16.8, respectively. Then, for every $l = m, \dots, 0$:*
(i) *One has the following estimate on $\mathcal{Y}^{l,t}$,*

$$\|Y^{l,t}\|^2_{\mathcal{S}^2} + \|Z^{l,t}\|^2_{\mathcal{H}^2_d} + \|V^{l,t}\|^2_{\mathcal{H}^2_\mu} + \|K^{l,t}\|^2_{\mathcal{S}^2} \le C(1 + |x|^{2q}) . \tag{259}$$

Moreover, an analogous bound estimate is satisfied by $\widetilde{\mathcal{Y}}^{l,t}$;
(ii) *^{t_n} referring to a perturbed initial condition (t_n, x_n, i) of $\mathcal{X}$, then:*
• *in case $t \notin \mathcal{T}$, $\mathcal{Y}^{l,t_n}$ $\mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu \times \mathcal{S}^2$ – converges to $\mathcal{Y}^{l,t}$ as $\mathcal{E} \ni (t_n, x_n, i) \to (t, x, i)$;*
• *in case $t = T_{l'} \in \mathcal{T}$:*
— *$\mathcal{Y}^{l,t_n}$ $\mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu \times \mathcal{S}^2$ – converges to $\mathcal{Y}^{l,t}$ as $\mathcal{E}_{l'+1} \ni (t_n, x_n, i) \to (t, x, i)$;*
— *if $\theta_{l'}$ is continuous at (x, i), then $\mathcal{Y}^{l,t_n}$ $\mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu \times \mathcal{S}^2$ – converges to $\widetilde{\mathcal{Y}}^{l,t}$ as $\mathcal{E}^*_{l'} \ni$ $(t_n, x_n, i) \to (t, x, i)$.*

Proof. Under Assumption 16.7, these results can be established, recursively on l decreasing from m to 0, by easy amendments to the proof of Proposition 8.5 in Appendix A.3, using Assumption 16.10 for controlling new terms in $\|h(t\vee\cdot\wedge\tau^t_{l+1}, \mathcal{X}^t_{t\cdot\wedge\tau^t_{l+1}}) - h(t_n\cdot\wedge\tau^{t_n}_{l+1}, \mathcal{X}^{t_n}_{t_n\cdot\wedge\tau^{t_n}_{l+1}})\|_{\mathcal{S}^2}$ and $\|h(t\vee\cdot\wedge\widetilde{\tau}^t_{l+1}, \mathcal{X}^t_{t\cdot\wedge\widetilde{\tau}^t_{l+1}}) - h(t_n\cdot\wedge\widetilde{\tau}^{t_n}_{l+1}, \mathcal{X}^{t_n}_{t_n\cdot\wedge\widetilde{\tau}^{t_n}_{l+1}})\|_{\mathcal{S}^2}$ that arise (for l odd) in the right-hand-side of the analogs of inequality (280). □

16.3.3 Semi-Group and Markov Properties

Let ^{t} refer to the constant initial condition (t, x, i) as usual. Let $\mathcal{X}^t = (X^t, N^t)$ and $\mathcal{Y}^t = (Y^t, Z^t, V^t, K^t)$ be defined as in Section 16.3.1 and Theorem 16.9, respectively. Given $t' \geq t$, let $\widetilde{\mathbb{F}}^{t'}$ stand for $(\widetilde{\mathcal{F}}_r^{t'})_{r \geq t'}$ with for $r \geq t'$

$$\widetilde{\mathcal{F}}_r^{t'} = \sigma(\mathcal{X}_{t'}^t) \bigvee \mathcal{F}_r^{t'} \;;$$

Let $\tau' = t' \vee \tau^t$, in the sense that $\tau'_l = t' \vee \tau_l^t$, for $l = 1, \dots, m+1$. As for $\mathbb{F}^{t'} = (\mathcal{F}_r^{t'})_{r \geq t'}$, $\mathbb{P}^{t'}$, $B^{t'}$ and $\mu^{t'}$, they are defined as usual as in Sections 7.1-7.2, with t' instead of t therein. Note in particular that $\widetilde{\mathbb{F}}^{t'}$ is smaller than or equal to the restriction $\mathbb{F}^t_{|[t',T]}$ of $\mathbb{F}^t$ to $[t', T]$.

We then have the following semi-group properties which are the analogs in the present set-up of Propositions 8.3, 8.6 in Part II.

Proposition 16.11 **(i)** *The Jump-Diffusion model with Marker Process on $[t', T]$ with initial condition $\mathcal{X}_{t'}^t$ at t' admits a unique $(\Omega, \widetilde{\mathbb{F}}^{t'}, \mathbb{P})$ – solution $\mathcal{X}^{t'} = (X^{t'}, N^{t'})$, which coincides with the restriction of $\mathcal{X}^t$ to $[t', T]$, so:*

$$\mathcal{X}^{t'} = (X_r^{t'}, N_r^{t'})_{t' \leq r \leq T} = (\mathcal{X}_r^t)_{t' \leq r \leq T} \; .$$

(ii) *For t and t' in the same monitoring time strip so that $T_{l-1} \leq t < t' < T_l$ for some $l \in \{1, \dots, m\}$, then $\tau' = t' \vee \tau^t$ is an $\widetilde{\mathbb{F}}^{t'}$ – stopping time, and the RIBSDE on $[t', T]$ with data*

$$\widetilde{g}(s, \mathcal{X}_s^{t'}, y, z, \widetilde{v}) \; , \; \Phi(\mathcal{X}_T^{t'}) \; , \; \ell(s, \mathcal{X}_s^{t'}) \; , \; h(s, \mathcal{X}_s^{t'}), \tau' \tag{260}$$

has a unique $(\Omega, \widetilde{\mathbb{F}}^{t'}, \mathbb{P}^{t'})$, $(B^{t'}, \mu^{t'})$ – solution $\mathcal{Y}^{t'} = (Y^{t'}, Z^{t'}, V^{t'}, K^{t'})$ such that, with $\mathcal{Y}^t = (Y_r^t, Z_r^t, V_r^t, K_r^t)_{t \leq r \leq T}$ defined as in Theorem 16.9:

$$\mathcal{Y}^{t'} = (Y_r^{t'}, Z_r^{t'}, V_r^{t'}, K_r^{t'})_{t' \leq r \leq T} = (Y_r^t, Z_r^t, V_r^t, K_r^t - K_{t'}^t)_{t' \leq r \leq T} \; . \tag{261}$$

Proof. Part (i) can be shown much like Proposition 8.3(i). It implies in particular that whenever $T_{l-1} \leq t < t' < T_l$ for some $l \in \{1, \dots, m\}$, then $N_r^t = N_r^{t'} = i$ for $r \in [t', T_l)$. In view of (252) one thus has $\tau'_0 = t'$ and, for every $l \geq 0$:

$$\tau'_{2l+1} = \inf\{s > \tau'_{2l} \; ; \; N_s^{t'} \notin \Delta\} \wedge T \; , \; \tau'_{2l+2} = \inf\{s > \tau'_{2l+1} \; ; \; N_s^{t'} \in \Delta\} \wedge T \; . \tag{262}$$

This shows that τ' is an $\widetilde{\mathbb{F}}^{t'}$ – stopping time, namely the analog of τ^t relative to $N^{t'}$. Knowing this, part (ii) can then be established much like Proposition 8.3(ii) or 8.6(ii) in Part II. □

In the present set-up the suitable notion of a Cauchy cascade (cf. Definition 15.1) takes the following form.

Definition 16.11 **(i)** A *Cauchy cascade* Φ, ν on $\mathcal{E}$ is pair made of a terminal condition Φ of class $\mathcal{P}$ at T, along with a sequence $\nu = (v_l)_{1 \leq l \leq m}$ of functions v_ls of class $\mathcal{P}$ on the $\mathcal{E}_l$s, satisfying the following jump condition, at every point of continuity of θ_l^i in x:

$$v_l^i(T_l, x) = \begin{cases} \min(v_{l+1}(T_l, x, \theta_l^i(x)), h(x)) & \text{if } i \notin \Delta \text{ and } \theta_l^i(x) \in \Delta, \\ v_{l+1}(T_l, x, \theta_l^i(x)) & \text{else} \end{cases} \tag{263}$$

where, in case $l = m$, v_{l+1} is to be understood as Φ.
A *continuous Cauchy cascade* is a Cauchy cascade with continuous ingredients Φ at T and v_ls on the $\mathcal{E}_l$s, except maybe for discontinuities of the v_ls at the points (T_l, x, i) of discontinuity of θ_l^i in x;
(ii) The function defined by a Cauchy cascade is the function on $\mathcal{E}$ given as the concatenation on the $\mathcal{E}_l^*$s of the v_ls, and by the terminal condition Φ at T.

Remark 16.12 So, at points (T_l, x, i) of discontinuity of θ_l^i in x, $v_l^i(t_n, x_n)$ may fail to converge to $v_l^i(T_l, x)$ as $\mathcal{E}_l \ni (t_n, x_n, i) \to (T_l, x, i)$. Note that in the specific situations of Examples 15.3 or 16.6–4.4, the set of discontinuity points x of θ_l^i is given by the hyperplane $\{x_1 = \bar{S}\}$ of $\mathbb{R}^d$, for every l, i.

We are now in a position to state the Markov properties of $\mathcal{Y}$. The notion of ε – saddle-point in part (iii) was introduced in the general RIBSDEs verification principle of Proposition 16.1.

Theorem 16.12 **(i)** *Given $(t, x, i) \in \mathcal{E}$, let $\mathcal{Y}^t = (Y^t, Z^t, V^t, K^t)$ be defined as in Theorem 16.9. As (t, x, i) varies in $\mathcal{E}$, Y_t^t is a deterministic function v defined by a continuous Cauchy cascade $\Phi, (v_l)_{1 \le l \le m}$ on $\mathcal{E}$.*
(ii) *One has, $\mathbb{P}^t$-a.s.,*

$$Y_s^t = v(s, \mathcal{X}_s^t) \ , \ s \in [t, T] \ . \tag{264}$$

(iii) *For every $\varepsilon > 0$, an ε – saddle-point of the related Dynkin game at time t is given by,*

$$\rho_t^\varepsilon = \inf \left\{ s \in \cup_{l \ge 0} [\tau_{2l+1}^t, \tau_{2l+2}^t) \, ; \, (s, \mathcal{X}_s^t) \in \mathcal{E}_\varepsilon^- \right\} \wedge T \ , \ \theta_t^\varepsilon = \inf \left\{ s \in [t, T] \, ; \, (s, \mathcal{X}_s^t) \in \mathcal{E}_\varepsilon^+ \right\} \wedge T$$

with

$$\mathcal{E}_\varepsilon^- = \{(t, x, i) \in \mathcal{E} \, ; \, v^i(t, x) \ge h^i(t, x) - \varepsilon\} \ , \ \mathcal{E}_\varepsilon^+ = \{(t, x, i) \in \mathcal{E} \, ; \, v^i(t, x) \ge \ell^i(t, x) + \varepsilon\} \ .$$

Proof. Let us prove parts (i) and (ii), which immediately imply (iii) by an application of Proposition 16.1. Let us prove part (i).[8] By taking $r = t'$ in the semi-group property (261) of $\mathcal{Y}$, one gets, for every $l = 1, \ldots, m$ and $T_{l-1} \le t \le r < T_l$,

$$Y_r^t = v_l(r, \mathcal{X}_r^t) \ , \ \mathbb{P}^t\text{–}a.s. \tag{265}$$

for a deterministic function v_l on $\mathcal{E}_l^*$. In particular,

$$Y_t^t = v^i(t, x) \ , \text{ for any } (t, x, i) \in \mathcal{E} \ , \tag{266}$$

where v is the function defined on $\mathcal{E}$ by the concatenation of the v_ls and the terminal condition Φ at T. In view of (256), the fact that v is of class $\mathcal{P}$ then directly follows from the bound estimates (259) on $\mathcal{Y}^{0,t}$ and $\mathcal{Y}^{1,t}$.

Let us show that the v_ls are continuous over the $\mathcal{E}_l^*$s. Given $\mathcal{E} \ni (t_n, x_n, i) \to (t, x, i)$ with $t \notin \mathcal{T}$ or $t_n \ge T_l = t$, one decomposes by (256):

$$|v^i(t, x) - v^i(t_n, x_n)| = |Y_t^t - Y_{t_n}^{t_n}| \le$$
$$\begin{cases} |\mathbb{E}(Y_t^{0,t} - Y_{t_n}^{0,t})| + \mathbb{E}|Y_{t_n}^{0,t} - Y_{t_n}^{0,t_n}|, & i \in \Delta \\ |\mathbb{E}(Y_t^{1,t} - Y_{t_n}^{1,t})| + \mathbb{E}|Y_{t_n}^{1,t} - Y_{t_n}^{1,t_n}|, & i \notin \Delta \ . \end{cases}$$

[8] This proof is an immediate extension to jump-diffusions X of the same result which is shown in case of a diffusion X in [32]. We also give the proof here for the reader's convenience.

In either case we conclude as in the proof of Theorem 9.3(i), using Proposition 16.10 as a main tool, that $|v^i(t,x) - v^i(t_n,x_n)|$ goes to zero as $n \to \infty$.

It remains to show that the v_ls can be extended by continuity over the $\mathcal{E}_l$s, except maybe at the boundary points (T_l, x, i) such that θ_l^i is discontinuous at x. Given $\mathcal{E}_l^* \ni (t_n, x_n, i) \to (T_l, x, i)$ with θ_l continuous at (x,i), one needs to show that $v_l^i(t_n,x_n) = v^i(t_n,x_n) \to v_l^i(T_l,x)$, where $v_l^i(T_l,x)$ is given by (263)). We distinguish four cases.

• In case $i \notin \Delta$ and $\theta_l^i(x) \in \Delta$, one has, denoting $\widetilde{v}^j(s,y) = \min(v(s,y,\theta_l^j(y)), h(y))$,

$$
\begin{aligned}
|\widetilde{v}^i(T_l,x) - v^i(t_n,x_n)|^2 &= |\widetilde{v}^i(T_l,x) - Y_{t_n}^{1,t_n}|^2 \leq \\
&2\mathbb{E}|\widetilde{v}^i(T_l,x) - \widetilde{v}(T_l,\mathcal{X}_{T_l}^{t_n})|^2 + 2|\mathbb{E}\big(\widetilde{v}(T_l,\mathcal{X}_{T_l}^{t_n}) - Y_{t_n}^{1,t_n}\big)|^2 \,. \qquad (267)
\end{aligned}
$$

By continuity of θ_l at (x,i), one has $\theta_l(\mathcal{X}_{T_l}^{t_n}) = \theta_l^i(x) \in \Delta$ for $\mathcal{X}_{T_l}^{t_n}$ close enough to x, say $||\mathcal{X}_{T_l}^{t_n} - x|| \leq c$. In this case $T_l = \tau_2^{t_n}$, therefore (cf. (255)) $Y_{T_l}^{1,t_n} = \widetilde{v}(T_l,\mathcal{X}_{T_l}^{t_n})$. So

$$
\mathbb{E}|\mathbb{1}_{||\mathcal{X}_{T_l}^{t_n} - x|| \leq c}\big(\widetilde{v}(T_l,\mathcal{X}_{T_l}^{t_n}) - Y_{t_n}^{1,t_n}\big)|^2 \leq \mathbb{E}|Y_{T_l}^{1,t_n} - Y_{t_n}^{1,t_n}|^2 \,,
$$

which can be shown to converge to zero as $n \to \infty$ by using the R2BSDE satisfied by Y^{1,t_n} and the convergence of $\mathcal{Y}^{1,t_n}$ to $\widetilde{\mathcal{Y}}^{1,t}$. Moreover $\mathbb{E}|\mathbb{1}_{||\mathcal{X}_{T_l}^{t_n} - x|| > c}\big(\widetilde{v}(T_l,\mathcal{X}_{T_l}^{t_n}) - Y_{t_n}^{1,t_n}\big)|^2$ goes to zero as $n \to \infty$ by the a priori estimates on X and Y^{1,t_n} and the continuity of $\widetilde{v}$ already established over $\mathcal{E}_{l+1}^*$. Moreover by this continuity and the a priori estimates on X the first term in (267) also goes to zero as $n \to \infty$. So, as $n \to \infty$,

$$
v^i(t_n,x_n) \to \widetilde{v}^i(T_l,x) = \min(v(T_l,x,\theta_l^i(x)), h(x)) = v_l^i(T_l,x) \,.
$$

• In case $i \in \Delta$ and $\theta_l^i(x) \notin \Delta$, one can show likewise, using $\widehat{v}^j(s,y) := v(s,y,\theta_l^j(y))$ instead of $\widetilde{v}^j(s,y)$ and Y^0 instead of Y^1 above, that

$$
v^i(t_n,x_n) \to v(T_l,x,\theta_l^i(x)) = v_l^i(T_l,x) \qquad (268)
$$

as $n \to \infty$.

• If $i, \theta_l^i(x) \notin \Delta$, it comes,

$$
\begin{aligned}
|\widehat{v}^i(T_l,x) - v^i(t_n,x_n)|^2 &= |\widehat{v}^i(T_l,x) - Y_{t_n}^{1,t_n}|^2 \\
&\leq 2\mathbb{E}|\widehat{v}^i(T_l,x) - \widehat{v}(T_l,\mathcal{X}_{T_l}^{t_n})|^2 + 2|\mathbb{E}\big(\widehat{v}(T_l,\mathcal{X}_{T_l}^{t_n}) - Y_{t_n}^{1,t_n}\big)|^2 \\
&\leq 2\mathbb{E}|\widehat{v}^i(T_l,x) - \widehat{v}(T_l,\mathcal{X}_{T_l}^{t_n})|^2 + 2|\mathbb{E}\big(Y_{T_l}^{1,t_n} - Y_{t_n}^{1,t_n}\big)|^2 \,.
\end{aligned}
$$

which goes to zero as $\to \infty$ by an analysis similar to (but simpler than) that of the first bullet point. Hence (268) follows.

• If $i, \theta_l^i(x) \in \Delta$, (268) can be shown as in the above bullet point. □

16.3.4 Viscosity Solutions Approach

The next step consists in deriving an analytic characterization of the value function v, or, more precisely, of $\nu = (v_l)_{1 \leq l \leq m}$, in terms of viscosity solutions to a related partial integro-differential problem. In the present case this problem assumes the form of the following cascade of variational inequalities:

For l decreasing from m to 1,
• At $t = T_l$, for every $i \in I$ and $x \in \mathbb{R}^d$,

$$v_l^i(T_l, x) = \begin{cases} \min(v_{l+1}(T_l, x, \theta_l^i(x)), h(x)), & i \notin \Delta \text{ and } \theta_l^i(x) \in \Delta \\ v_{l+1}(T_l, x, \theta_l^i(x)), & \text{else} \end{cases} \quad (269)$$

with v_{l+1} in the sense of Φ in case $l = m$;
• On the time interval $[T_{l-1}, T_l)$, for every $i \in I$,

$$\begin{cases} \min\Big(-\mathcal{G}v_l^i - g^{v_l^i}, v_l^i - \ell \Big) = 0 \ , \ i \in \Delta \\ \max\Big(\min\Big(-\mathcal{G}v_l^i - g^{v_l^i}, v_l^i - \ell \Big), v_l^i - h \Big) = 0 \ , \ i \notin \Delta \end{cases} \quad (270)$$

where $\mathcal{G}$ is given by (250) and where we set, for any function $u = u(t, x)$,

$$g^u = g^u(t, x) = g(t, x, u(t, x), (\partial u \sigma)(t, x), \mathcal{I}u(t, x)) \ . \quad (271)$$

In the special case of a jump size function θ independent of x, so $\theta_l^i(x) = \theta_l^i$, then the v_ls are in fact continuous functions over the $\mathcal{E}^l$s. This can be shown by a simplified version of the proof of Theorem 16.12. Using the notions of viscosity solutions introduced in Definition 14.1, one then has in virtue of arguments already used in Part III (cf. also Proposition 15.1(ii)) that for every $l = 1, \ldots, m$ and $i \in I$, the function v_l^i is the unique $\mathcal{P}$-solution, the maximal $\mathcal{P}$-subsolution and the minimal $\mathcal{P}$-supersolution of the related problem $(\mathcal{V}1)$ or $(\mathcal{V}2)$ on $\mathcal{E}_l$ which is visible in (269)–(270), with terminal condition at T_{l+1} dictated by v_{l+1}, h and/or Φ. Moreover, under the working assumption that the true value for v_{l+1} is plugged at T_{l+1} in an approximation scheme for v_l, then $v_{l,h} \to v_l$ locally uniformly on $\mathcal{E}_l$ as $h \to 0$.

But, thinking for instance of the situations of Example 16.6, the case of θ not depending on x is of course too specific. Now, as soon as θ depends x, θ presents discontinuities in x, and, under Assumption 16.7, the functions v_ls typically present discontinuities at the points (T_l, x, i) of discontinuity of the θ_l^is. There is then no chance to characterize the v_ls in terms of continuous viscosity solutions to (269)–(270) anymore. It would be possible however, though we shall not develop this further in this article, to characterize ν in terms of a suitable notion of *discontinuous viscosity solution* to (269)–(270).

16.3.5 Protection Before a Stopping Time Again

We finally consider the special case where the marker process N is stopped at its first exit time of Δ, which corresponds to jump functions $\theta_l^i(x)$ such that $\theta_l^i(x) = i$ for $i \notin \Delta$. The sequence $\tau^t = (\tau_l^t)_{l \geq 0}$ is then stopped at rank $l = 2$, so $\tau_l^t = T$ for $l \geq 2$. In this case the Markovian form of (237) reduces to (cf. (253)),

$$\Omega_s^t = \mathbb{1}_{N_s^t \notin \Delta} = \mathbb{1}_{s \geq \tau_1^t} \ . \quad (272)$$

From the point of view of financial interpretation we recover a case of call protection *before a stopping time* as of Parts II and III. If $N_t^t = i \notin \Delta$, one has $\tau_1^t = t$, and call protection on $[0, \tau_1^t)$ actually reduces to no protection. For less trivial examples (provided $N_t^t = i \in \Delta$) we refer the reader to Example 15.3, which corresponds to the 'stopped' version of Example 16.6–4.4.

From a mathematical point of view one is back to an RDBSDE as of Definition 5.4(ii) (cf. (107), (16)). But this is for a stopping time, τ_1^t, which falls outside the scope of Assumption 9.1 in Part III, so that the PDE results of Part III cannot be applied as such. However, assuming (272), one can check by inspection in the arguments of sections 16.3.2 to 16.3.4, that:
• For $i \notin \Delta$, the $\mathcal{Y}^{l,t}$s do not depend on i, and $\mathcal{Y}^t$ in Theorem 16.9 coincides with $\mathcal{Y}^t$ in Theorem 8.1(i) (special case of $\mathcal{X}^t$ therein given as X^t here);
• The $\mathcal{Y}^{l,t}$s have continuous $K^{l,t}$s components (since the discontinuities of the $K^{l,t}$s occurred because of the switchings from no call protection to call protection, and that such switchings are not possible for τ^t stopped at rank 2),
• Theorem 16.10 is true independently of Assumption 16.10 (since again this Assumption was only used for taking care of the case where a call protection period follows a no call protection period), so that Assumption 16.10 is in fact not required in this section.

16.3.5.1 No-Protection Price

Regarding the *no-protection* period $[\tau_1^t, T]$ one thus has the following result, either by application of the results of Parts II and III, or by inspection of the proofs in Sections 16.3.2 to 16.3.4,

Proposition 16.13 (i) *For $i \notin \Delta$, $Y_t^{1,t} =: u(t,x)$ defines a continuous function u on $[0,T] \times \mathbb{R}^d$.*
(ii) *This function u corresponds to a* no call protection pricing function *in the sense that one has, starting from every initial condition $(t,x,i) \in \mathcal{E}$,*

$$Y_s^t = u(s, X_s^t) \text{ on } [\tau_1^t, T] \,,$$

with $\tau_1^t = \inf\{s > t\,;\, N_s^t \notin \Delta\}$;
(iii) *The no protection value function u thus defined is the unique $\mathcal{P}$-solution, the maximal $\mathcal{P}$-subsolution, and the minimal $\mathcal{P}$-supersolution of*

$$\max\Big(\min\Big(-\mathcal{G}u - g^u, u - \ell\Big), u - h\Big) = 0 \tag{273}$$

on $\mathcal{E}$ with boundary condition Φ at T, where $\mathcal{G}$ is given by (250) and where g^u is defined by (271).
(iv) *Stable, monotone and consistent approximation schemes u_h for u converge to u locally uniformly on $\mathcal{E}$ as $h \to 0$.*

Note that the no-protection pricing function u is but the function v^i for $i \notin \Delta$ of Theorem 16.12, which for $i \notin \Delta$ does not depend on i (v^i is constant in i outside Δ, under (272)).

16.3.5.2 Protection Price

As for the *protection period* $[0, \tau_1^t)$, since the v_l^is for $i \notin \Delta$ all reduce to u, the Cauchy cascade (269)–(270) in $\nu = (v_l)_{1 \le l \le m} = (v_l^i)_{1 \le l \le m}^{i \in I}$ effectively reduces to the following *Cauchy–Dirichlet cascade* in $(v_l^i)_{1 \le l \le m}^{i \in \Delta}$, with the function u as boundary condition, and where in view of identity (264) in Theorem 16.12, $(v_l^i)_{1 \le l \le m}^{i \in \Delta}$ can be interpreted as the *protection pricing function*:

For l decreasing from m to 1,
• At $t = T_l$, for every $i \in \Delta$ and $x \in \mathbb{R}^d$,

$$v_l^i(T_l, x) = \begin{cases} u(T_l, x), & l = m \text{ or } \theta_l^i(x) \notin \Delta \\ v_{l+1}(T_l, x, \theta_l^i(x)), & \text{else }, \end{cases} \tag{274}$$

• On the time interval $[T_{l-1}, T_l)$, for every $i \in \Delta$,

$$\min\Big(- \mathcal{G}v_l^i - g^{v_l^i}, v_l^i - \ell \Big) = 0 \,. \tag{275}$$

Given a pertaining notion of discontinuous viscosity solution of (274)-(275), $(v_l^i)_{1\le l\le m}^{i\in\Delta}$ could then be characterized as the unique solution in this sense to (274)–(275).

Remark 16.13 The Cauchy-Dirichlet cascade (273)-(274)-(275) involves less equations than the Cauchy cascade (269)-(270). However 'less' here is still often far too much (see for instance Example 15.3(ii)) from the point of view of a practical resolution by deterministic numerical schemes. For 'very large' sets Δ simulation schemes are then the only viable alternative.

A Proofs of Auxiliary BSDE Results

A.1 Proof of Lemma 7.4

Recall that a càdlàg process Z is a $\mathbb{P}^t$ – local martingale if and only if $\Gamma^t Z$ is a $\mathbb{P}$ – local martingale (see, e.g., Proposition III.3.8 in Jacod–Shiryaev [67]). Now for

$$Z = B^t \,, \text{ resp. } \int_t^{\cdot} \int_{\mathbb{R}^d} V_s(y)\, \widetilde{\chi}^t(ds, dy) \,, \text{ resp. resp. } \sum_{j\in I} \int_t^{\cdot} W_s(j)\, d\widetilde{\nu}_s^t(j)$$

with V, W in the related spaces of predictable integrands, we have, "$\triangleq$" standing for "equality up to an $(\mathbb{F}^t, \mathbb{P})$ – local martingale term":

$$d(\Gamma^t Z)_s \triangleq \Gamma^t_{s-} dZ_s + \Delta\Gamma^t_s \Delta Z_s$$

where

$$\Delta Z_s = 0 \,, \text{ resp. } \textstyle\int_{\mathbb{R}^d} V_s(y)\, \chi(ds, dy) \,, \text{ resp. resp. } \sum_{j\in I} W_s(j)\, d\nu_s^t(j) \,.$$

In case $Z = B^t$, $\Gamma^t Z$ is obviously a $\mathbb{P}$ – local martingale. B^t is thus a continuous $\mathbb{P}^t$ – local martingale null at time t with $\langle B^t, B^t\rangle_s = (s-t)\mathrm{Id}_{d\otimes d}$. Therefore B^t is a $\mathbb{P}^t$ – Brownian motion starting at time t on $[t, T]$.

In case $Z = \int_t^{\cdot} \int_{\mathbb{R}^d} V_s(y)\, \widetilde{\chi}^t(ds, dy)$, since χ and ν cannot jump together (see Remark 7.2) , one has by (148):

$$\Delta\Gamma^t_s \Delta Z_s = \Delta Z_s \Gamma^t_{s-} \int_{\mathbb{R}^d} \left(\frac{f(s, \mathcal{X}^t_{s-}, y)}{\widehat{f}(s, \mathcal{X}^t_{s-}, y)} - 1 \right) \chi(ds, dy) \,.$$

So

$$
\begin{aligned}
d(\Gamma^t Z)_s &\triangleq \Gamma^t_{s-}\int_{\mathbb{R}^d} V_s(y)\,\widetilde{\chi}^t(ds,dy) + \Gamma^t_{s-}\int_{\mathbb{R}^d} V_s(y)\left(\frac{f(s,\mathcal{X}^t_{s-},y)}{\widehat{f}(s,\mathcal{X}^t_{s-},y)} - 1\right)\chi(ds,dy)\\
&= -\Gamma^t_{s-}\int_{\mathbb{R}^d} V_s(y)\,f(s,\mathcal{X}^t_s,y)m(dy)ds + \Gamma^t_{s-}\int_{\mathbb{R}^d} V_s(y)\,\frac{f(s,\mathcal{X}^t_{s-},y)}{\widehat{f}(s,\mathcal{X}^t_{s-},y)}\chi(ds,dy)\\
&= \Gamma^t_{s-}\int_{\mathbb{R}^d} V_s(y)\frac{f(s,\mathcal{X}^t_{s-},y)}{\widehat{f}(s,\mathcal{X}^t_{s-},y)}\widetilde{\chi}(ds,dy)
\end{aligned}
$$

and $\Gamma^t Z$ is also a $\mathbb{P}$ – local martingale in this case.
In case $Z = \sum_{j\in I}\int_t^{\cdot} W_s(j)\,d\widetilde{\nu}^t_s(j)$ one gets likewise

$$
d(\Gamma^t Z)_s \triangleq \Gamma^t_{s-}\sum_{j\in I} W_s(j)\,\frac{n(s,\mathcal{X}^t_{s-},j)}{\widehat{n}(N^t_{s-},j)}d\widetilde{\nu}_s(j)
$$

and $\Gamma^t Z$ is again a $\mathbb{P}$ – local martingale.

A.2 Proof of Proposition 8.2

First we have, using the facts that f (cf. Lemma 7.3(i)) and $\widehat{n}$ are bounded, with f positively bounded for (H.1.ii)":
(H.1.i)" $\mathbb{1}_{\{\cdot>t\}}\widehat{g}(\cdot,\mathcal{X}^t_\cdot,y,z,\widehat{v})$ is a progressively measurable process with

$$
\|\mathbb{1}_{\{\cdot>t\}}\widehat{g}(\cdot,\mathcal{X}^t_\cdot,y,z,\widehat{v})\|_{\mathcal{H}^2} < \infty \text{ for any } y\in\mathbb{R},\ z\in\mathbb{R}^{1\otimes d},\ \widehat{v}\in\mathcal{M}_\pi\ ;
$$

(H.1.ii)" $\mathbb{1}_{\{\cdot>t\}}\widehat{g}(\cdot,\mathcal{X}^t_\cdot,y,z,\widehat{v})$ is uniformly Λ – Lipschitz continuous with respect to $(y,z,\widehat{v})$, in the sense that for every $s\in[0,T], y,y'\in\mathbb{R},\ z,z'\in\mathbb{R}^{1\otimes d},\ \widehat{v},\widehat{v}'\in\mathcal{M}_\pi$:

$$
|\widehat{g}(s,\mathcal{X}^t_s,y,z,\widehat{v}) - \widehat{g}(s,\mathcal{X}^t_s,y',z',\widehat{v}')| \le \Lambda\big(|y-y'| + |z-z'| + |\widehat{v}-\widehat{v}'|\big)
$$

(cf. (156) for the definition of $|\widehat{v}-\widehat{v}'|$).
So the driver $\mathbb{1}_{\{\cdot>t\}}\widehat{g}$ satisfies the general assumptions (H.1), hence the data (159) satisfy the general assumptions (H.0)–(H.1)–(H.2), relative to $(\Omega,\mathbb{F},\mathbb{P}),(B,\mu)$.

(i) By the general results of [42], one thus has the following bound estimate on $\widetilde{\mathcal{Y}}^t$:

$$
\|Y^t\|^2_{\mathcal{S}^2} + \|Z^t\|^2_{\mathcal{H}^2_d} + \|\widetilde{\mathcal{V}}^t\|^2_{\mathcal{H}^2_\mu} + \|K^{t,+}\|^2_{\mathcal{S}^2} + \|K^{t,-}\|^2_{\mathcal{S}^2} \le c(\Lambda)c_1
$$

with

$$
\begin{aligned}
c_1 := \|\Phi(\mathcal{X}^t_T)\|^2_2 + \|\mathbb{1}_{\{\cdot>t\}}\widehat{g}(\cdot,\mathcal{X}^t_\cdot,0,0,0)\|^2_{\mathcal{H}^2} +&\\
\|\ell(\cdot\vee t,\mathcal{X}^t_{\cdot\vee t})\|^2_{\mathcal{S}^2} + \|h(\cdot\vee t,\mathcal{X}^t_{\cdot\vee t})\|^2_{\mathcal{S}^2} + \|\int_{\cdot\wedge t}^{\cdot}\mathcal{G}\phi(r,\mathcal{X}^t_r)dr\|^2_{\mathcal{S}^2}&\ ,
\end{aligned}
$$

where ϕ is the function introduced at Assumption (M.3). Estimate (161) then follows by standard computations, given the Lipschitz continuous and growth assumptions on the data and the bound estimate (134) on X^t.

(ii) By the general results of [42], we also have the following error estimate in which c_1 is as above:

$$\|Y^t - Y^{t_n}\|^2_{\mathcal{S}^2} + \|Z^t - Z^{t_n}\|^2_{\mathcal{H}^2_d} + \|\widetilde{\mathcal{V}}^t - \widetilde{\mathcal{V}}^{t_n}\|^2_{\mathcal{H}^2_\mu} + \|K^t - K^{t_n}\|^2_{\mathcal{S}^2} \leq c(\Lambda)c_1 \times$$
$$\Big(\|\Phi(\mathcal{X}^t_T) - \Phi(\mathcal{X}^{t_n}_T)\|^2_2 + \|\mathbb{1}_{\{\cdot>t\}}\widehat{g}(\cdot, \mathcal{X}^t_\cdot, Y^t_\cdot, Z^t_\cdot, \widetilde{\mathcal{V}}^t_\cdot) - \mathbb{1}_{\{\cdot>t_n\}}\widehat{g}(\cdot, \mathcal{X}^{t_n}_\cdot, Y^t_\cdot, Z^t_\cdot, \widetilde{\mathcal{V}}^t_\cdot)\|^2_{\mathcal{H}^2}$$
$$+\|\ell(\cdot \vee t, \mathcal{X}^t_{\cdot\vee t}) - \ell(\cdot \vee t_n, \mathcal{X}^{t_n}_{\cdot\vee t_n})\|_{\mathcal{S}^2} + \|h(\cdot \vee t, \mathcal{X}^t_{\cdot\vee t}) - h(\cdot \vee t_n, \mathcal{X}^{t_n}_{\cdot\vee t_n})\|_{\mathcal{S}^2} \Big) . \qquad (276)$$

First note that $c(\Lambda)c_1 \leq C(1 + |x|^{2q})$, by part (i). It thus simply remains to show that each term of the sum goes to 0 as $n \to \infty$ in the right hand side of (276). We provide a detailed proof for the term

$$\|\mathbb{1}_{\{\cdot>t\}}\widehat{g}(\cdot, \mathcal{X}^t_\cdot, Y^t_\cdot, Z^t_\cdot, \widetilde{\mathcal{V}}^t_\cdot) - \mathbb{1}_{\{\cdot>t_n\}}\widehat{g}(\cdot, \mathcal{X}^{t_n}_\cdot, Y^t_\cdot, Z^t_\cdot, \widetilde{\mathcal{V}}^t_\cdot)\|^2_{\mathcal{H}^2} .$$

The other terms can be treated along the same lines. Introducing a sequence (R_m) of positive numbers going to infinity as $m \to \infty$, let thus

$$\Omega^{m,n}_s := \{s \geq t \vee t_n\} \cap \{N^t_s = N^{t_n}_s\} \cap \{|X^t_s| \vee |X^{t_n}_s| \vee |Y^t_s| \vee |Z^t_s| \vee r^t_s \leq R_m\} ,$$

with $r^t_s := |\widehat{r}^t_s| \vee |\widetilde{r}^t_s| \vee |\widetilde{r}^{t_n}_s|$, where

$$\widehat{r}^t_s = \int_{\mathbb{R}^d} V^t_s(y)m(dy),\ \widetilde{r}^t_s = \int_{\mathbb{R}^d} V^t_s(y)f(s, X^t_s, N^t_s, y)m(dy),\ \widetilde{r}^{t_n}_s = \int_{\mathbb{R}^d} V^t_s(y)f(s, X^{t_n}_s, N^t_s, y)m(dy) \qquad (277)$$

and let $\bar{\Omega}^{m,n}_s$ denote the complement of the set $\Omega^{m,n}_s$. One has for any m, n :

$$\begin{aligned}
&\|\mathbb{1}_{\{\cdot>t\}}\widehat{g}(\cdot, \mathcal{X}^t_\cdot, Y^t_\cdot, Z^t_\cdot, \widetilde{\mathcal{V}}^t_\cdot) - \mathbb{1}_{\{\cdot>t_n\}}\widehat{g}(\cdot, \mathcal{X}^{t_n}_\cdot, Y^t_\cdot, Z^t_\cdot, \widetilde{\mathcal{V}}^t_\cdot)\|^2_{\mathcal{H}^2} \\
&= \mathbb{E}\int_{t\wedge t_n}^T \Big[\mathbb{1}_{\{s>t\}}\widehat{g}(s, \mathcal{X}^t_s, Y^t_s, Z^t_s, \widetilde{\mathcal{V}}^t_s) - \mathbb{1}_{\{s>t_n\}}\widehat{g}(s, \mathcal{X}^{t_n}_s, Y^t_s, Z^t_s, \widetilde{\mathcal{V}}^t_s)\Big]^2 ds \\
&= \mathbb{E}\int_{t\wedge t_n}^T \Big[\mathbb{1}_{\{s>t\}}\widehat{g}(s, \mathcal{X}^t_s, Y^t_s, Z^t_s, \widetilde{\mathcal{V}}^t_s) - \mathbb{1}_{\{s>t_n\}}\widehat{g}(s, \mathcal{X}^{t_n}_s, Y^t_s, Z^t_s, \widetilde{\mathcal{V}}^t_s)\Big]^2 \mathbb{1}_{\bar{\Omega}^{m,n}_s} ds + \\
&\qquad \mathbb{E}\int_{t\wedge t_n}^T \Big[\mathbb{1}_{\{s>t\}}\widehat{g}(s, \mathcal{X}^t_s, Y^t_s, Z^t_s, \widetilde{\mathcal{V}}^t_s) - \mathbb{1}_{\{s>t_n\}}\widehat{g}(s, \mathcal{X}^{t_n}_s, Y^t_s, Z^t_s, \widetilde{\mathcal{V}}^t_s)\Big]^2 \mathbb{1}_{\Omega^{m,n}_s} ds \\
&\leq 2\mathbb{E}\int_{t\wedge t_n}^T \Big[\widehat{g}(s, \mathcal{X}^t_s, Y^t_s, Z^t_s, \widetilde{\mathcal{V}}^t_s)^2 + \widehat{g}(s, \mathcal{X}^{t_n}_s, Y^t_s, Z^t_s, \widetilde{\mathcal{V}}^t_s)^2\Big] \mathbb{1}_{\bar{\Omega}^{m,n}_s} ds + \\
&\qquad \mathbb{E}\int_0^T \Big[\widehat{g}(s, \mathcal{X}^t_s, Y^t_s, Z^t_s, \widetilde{\mathcal{V}}^t_s) - \widehat{g}(s, \mathcal{X}^{t_n}_s, Y^t_s, Z^t_s, \widetilde{\mathcal{V}}^t_s)\Big]^2 \mathbb{1}_{\Omega^{m,n}_s} ds =: I_{m,n} + II_{m,n} .
\end{aligned}$$

Now,

$$\widehat{g}(s, \mathcal{X}^t_s, Y^t_s, Z^t_s, \widetilde{\mathcal{V}}^t_s)^2 + \widehat{g}(s, \mathcal{X}^{t_n}_s, Y^t_s, Z^t_s, \widetilde{\mathcal{V}}^t_s)^2 \leq C\Big(1 + |X^t_s|^{2q} + |X^{t_n}_s|^{2q} + |Y^t_s|^2 + |Z^t_s|^2 + |\widetilde{\mathcal{V}}^t_s|^2\Big) . \qquad (278)$$

Note that $|X^{t_n}_s|^{2q}$ is equi $-\, d\mathbb{P} \times dt$-integrable, by estimate (134) on X applied for $p > 2q$. So are therefore the right hand side, and in turn the left hand side, in (278), since $\widetilde{\mathcal{Y}}^t \in \mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu \times \mathcal{A}^2$. Besides, one has that

$$\mathbb{E}\int_{t\wedge t_n}^T \mathbb{1}_{\bar{\Omega}^{m,n}_s} ds \leq T|t - t_n| + \mathbb{E}\int_{t\vee t_n}^T \mathbb{1}_{\bar{\Omega}^{m,n}_s} ds , \qquad (279)$$

where for $s \geq t \vee t_n$:

$$\bar{\Omega}^{m,n}_s \subseteq \{N^t_s \neq N^{t_n}_s\} \cup \{|X^t_s| \vee |X^{t_n}_s| \vee |Y^t_s| \vee |Z^t_s| \vee |r^t_s| \geq R_m\} .$$

Note that $\|r^t\|_{\mathcal{H}^2} < \infty$. Using also estimates (136) on N, (134) on X and (161) on $\widetilde{\mathcal{Y}}$, we thus get by Markov's inequality:

$$\begin{aligned}
&\mathbb{E}\int_{t\vee t_n}^T \mathbb{1}_{\bar{\Omega}_s^{m,n}} ds \\
&\leq C|t-t_n| + \mathbb{E}\int_{t\vee t_n}^T \left(\mathbb{1}_{\{|X_s^t|\geq R_m\}} + \mathbb{1}_{\{|X_s^{t_n}|\geq R_m\}} + \mathbb{1}_{\{|Y_s^t|\geq R_m\}} + \mathbb{1}_{\{|Z_s^t|\geq R_m\}} + \mathbb{1}_{\{|r_s^t|\geq R_m\}}\right) ds \\
&\leq C\big(|t-t_n| + \tfrac{1}{R_m^2}\big) .
\end{aligned}$$

Therefore, given (279), $\mathbb{E}\int_{t\wedge t_n}^T \mathbb{1}_{\bar{\Omega}_s^{m,n}} ds$ goes to 0 as $m, n \to \infty$.

Note that $\mathbb{E}\int_{t\wedge t_n}^T \mathbb{1}_{\bar{\Omega}_s^{m,n}} ds = \mathbb{E}\int_{t\wedge t_n}^T \mathbb{1}_{\widetilde{\Omega}_s^{m,n}} ds$, with $\widetilde{\Omega}_s^{m,n} = \bar{\Omega}_s^{m,n} \cap \{s > t\wedge t_n\}$. By standard results, the fact that $\mathbb{E}\int_0^T \mathbb{1}_{\widetilde{\Omega}_s^{m,n}} ds \to 0$ as $m, n \to \infty$ implies that $\mathbb{E}\int_0^T f_s^{n,m} \mathbb{1}_{\widetilde{\Omega}_s^{m,n}} ds \to 0$ as $m, n \to \infty$, for any equi $-$ $d\mathbb{P} \times dt$-integrable family of non-negative processes $f = (f_s^{n,m})_{m,n}$. Applying this to

$$f^{n,m} = \widehat{g}(s, \mathcal{X}_s^t, Y_s^t, Z_s^t, \widetilde{\mathcal{V}}_s^t)^2 + \widehat{g}(s, \mathcal{X}_s^{t_n}, Y_s^t, Z_s^t, \widetilde{\mathcal{V}}_s^t)^2 ,$$

we conclude that $I_{m,n} \to 0$ as $m, n \to \infty$.

On the other hand, since $N_s^t = N_s^{t_n}$ on $\Omega_s^{m,n}$, and using the form (160) of $\widehat{g}$ in which g satisfies (M.1), we have:

$$\begin{aligned}
II_{m,n} &= \mathbb{E}\int_0^T \left[\widehat{g}(s, X_s^t, N_s^t, Y_s^t, Z_s^t, \widetilde{\mathcal{V}}_s^t) - \widehat{g}(s, X_s^{t_n}, N_s^t, Y_s^t, Z_s^t, \widetilde{\mathcal{V}}_s^t)\right]^2 \mathbb{1}_{\Omega_s^{m,n}} ds \\
&\leq \mathbb{E}\int_0^T \eta_m \left(|X_s^t - X_s^{t_n}| + |\widetilde{r}_s^t - \widetilde{r}_s^{t_n}|\right) ds
\end{aligned}$$

where η_m is a non-negative bounded function continuous and null at 0. Given $\varepsilon > 0$, let m_ε, n_ε be such that $I_{m_\varepsilon,n} \leq \frac{\varepsilon}{2}$ for $n \geq n_\varepsilon$. Let further μ_ε be such $\eta_{m_\varepsilon}(\rho) \leq \varepsilon$ for $\rho \leq \mu_\varepsilon$. C_ε denoting an upper bound on η_{m_ε}, it comes, for every n:

$$\begin{aligned}
II_{m_\varepsilon,n} &\leq \mathbb{E}\int_0^T \eta_{m_\varepsilon} \left(|X_s^t - X_s^{t_n}| + |\widetilde{r}_s^t - \widetilde{r}_s^{t_n}|\right) ds \\
&\leq \mathbb{E}\int_0^T \left(\varepsilon + C_\varepsilon \mathbb{1}_{\{|X_s^t - X_s^{t_n}|\geq \mu_\varepsilon\}} + C_\varepsilon \mathbb{1}_{\{|\widetilde{r}_s^t - \widetilde{r}_s^{t_n}|\geq \mu_\varepsilon\}}\right) ds \\
&\leq T\Big(\varepsilon + C_\varepsilon \mathbb{P}[\sup_{[0,T]} |X_\cdot^t - X_\cdot^{t_n}| \geq \mu_\varepsilon]\Big) + C_\varepsilon \mathbb{E}\int_0^T \mathbb{1}_{\{|\widetilde{r}_s^t - \widetilde{r}_s^{t_n}| ds \geq \mu_\varepsilon\}} .
\end{aligned}$$

Now, given estimate (137), one has that $\mathbb{P}[\sup_{[0,T]} |X_\cdot^t - X_\cdot^{t_n}| \geq \mu_\varepsilon] \to 0$ as $n \to \infty$, by Markov's inequality. Moreover (cf. (277))

$$|\widetilde{r}_s^t - \widetilde{r}_s^{t_n}| \leq \int_{\mathbb{R}^d} |V_s^t(y)||f(s, X_s^t, N_s^t, y) - f(s, X_s^{t_n}, N_s^t, y)| m(dy) ,$$

so $\|\widetilde{r}^t - \widetilde{r}^{t_n}\|_{\mathcal{H}^2} \to 0$ as $n \to \infty$, by dominated convergence using the Lipschitz continuity property of f in Lemma 7.3(i). Thus by Markov's inequality:

$$\mathbb{E}\int_0^T \mathbb{1}_{\{|\widetilde{r}_s^t - \widetilde{r}_s^{t_n}| ds \geq \mu_\varepsilon\}} \leq \frac{\|\widetilde{r}^t - \widetilde{r}^{t_n}\|_{\mathcal{H}^2}^2}{\mu_\varepsilon^2}$$

converges to 0 as $n \to \infty$.

In conclusion $I_{m_\varepsilon,n} + II_{m_\varepsilon,n} \leq \varepsilon$ for $n \geq n_\varepsilon \vee n'_\varepsilon$, for any $\varepsilon > 0$, which proves that

$$\|\mathbb{1}_{\{\cdot > t\}} \widehat{g}(\cdot, \mathcal{X}_\cdot^t, Y_\cdot^t, Z_\cdot^t, \widetilde{\mathcal{V}}_\cdot^t) - \mathbb{1}_{\{\cdot > t_n\}} \widehat{g}(\cdot, \mathcal{X}_\cdot^{t_n}, Y_\cdot^t, Z_\cdot^t, \widetilde{\mathcal{V}}_\cdot^t)\|_{\mathcal{H}^2}^2 \to 0 \text{ as } n \to \infty .$$

A.3 Proof of Proposition 8.5

By the bound estimate (161) on $\widetilde{\mathcal{Y}}^t$, $Y^t_{\tau^t} \in \mathcal{L}^2$. Moreover, one checks as in the proof of Proposition 8.2 that the driver $\mathbb{1}_{\{t<\cdot<\tau^t\}}\widehat{g}(\cdot, \mathcal{X}^t_\cdot, y, z, \widehat{v})$ satisfies the general assumptions (H.1). Hence the data

$$\mathbb{1}_{\{t<s<\tau^t\}}\widehat{g}(s, \mathcal{X}^t_s, y, z, \widehat{v}) \ , \ Y^t_{\tau^t} \ , \ \ell(t \vee s \wedge \tau^t, \mathcal{X}^t_{s\vee t\wedge\tau^t})$$

satisfy the general assumptions (H.0)–(H.1)–[assumptions regarding L in](H.2) relative to $(\Omega, \mathbb{F}, \mathbb{P})$, (B, μ).

(i) By the general results of [42], one thus has the following bound estimate on $\overline{\mathcal{Y}}^t$:

$$\|\bar{Y}^t\|^2_{\mathcal{S}^2} + \|\bar{Z}^t\|^2_{\mathcal{H}^2_d} + \|\overline{\mathcal{V}}^t\|^2_{\mathcal{H}^2_\mu} + \|\bar{K}^t\|^2_{\mathcal{S}^2} \le c(\Lambda)c_1$$

with

$$c_1 := \|Y^t_{\tau^t}\|^2_2 + \|\widehat{g}(\cdot, \mathcal{X}^t_\cdot, 0, 0, 0)\|^2_{\mathcal{H}^2} + \|\ell(t \vee \cdot \wedge \tau^t, \mathcal{X}^t_{t\vee\cdot\wedge\tau^t})\|^2_{\mathcal{S}^2} \ .$$

Estimate (169) then follows by standard computations, given the Lipschitz continuous and growth assumptions on the data and estimate (134) on X^t.
(ii) Given the assumptions made on ℓ, one has the following error estimate in which c_1 is as above, by the general results of [42]:

$$\begin{aligned}
&\|\bar{Y}^t - \bar{Y}^{t_n}\|^2_{\mathcal{S}^2} + \|\bar{Z}^t - \bar{Z}^{t_n}\|^2_{\mathcal{H}^2_d} + \|\overline{\mathcal{V}}^t - \overline{\mathcal{V}}^{t_n}\|^2_{\mathcal{H}^2_\mu} + \|\bar{K}^t - \bar{K}^{t_n}\|^2_{\mathcal{S}^2} \le c(\Lambda)c_1 \times \\
&\quad \Big(\|Y^t_{\tau^t} - Y^{t_n}_{\tau^{t_n}})\|^2_2 \\
&\quad + \|\mathbb{1}_{\{t<\cdot<\tau^t\}}\widehat{g}(\cdot, \mathcal{X}^t_\cdot, \bar{Y}^t_\cdot, \bar{Z}^t_\cdot, \overline{\mathcal{V}}^t_\cdot) - \mathbb{1}_{\{t_n<\cdot<\tau^{t_n}\}}\widehat{g}(\cdot, \mathcal{X}^{t_n}_\cdot, \bar{Y}^t_\cdot, \bar{Z}^t_\cdot, \overline{\mathcal{V}}^t_\cdot)\|^2_{\mathcal{H}^2} \\
&\quad + \|\ell(t \vee \cdot \wedge \tau^t, \mathcal{X}^t_{t\vee\cdot\wedge\tau^t}) - \ell(t_n \vee \cdot \wedge \tau^{t_n}, \mathcal{X}^{t_n}_{t_n\vee\cdot\wedge\tau^{t_n}})\|_{\mathcal{H}^2} \Big)
\end{aligned} \tag{280}$$

(with in particular $\|\cdot\|_{\mathcal{H}^2}$, better than $\|\cdot\|_{\mathcal{S}^2}$, in the last term, thanks to the regularity assumption (M.3) on ℓ, cf. [42]). Since $c(\Lambda)c_1 \le C(1 + |x|^{2q})$ by (i), it simply remains to show that each term of the sum goes to 0 as $n \to \infty$ in the right hand side of (280). We provide a detailed proof for the term

$$\|\mathbb{1}_{\{t<\cdot<\tau^t\}}\widehat{g}(\cdot, \mathcal{X}^t_\cdot, \bar{Y}^t_\cdot, \bar{Z}^t_\cdot, \overline{\mathcal{V}}^t_\cdot) - \mathbb{1}_{\{t_n<\cdot<\tau^{t_n}\}}\widehat{g}(\cdot, \mathcal{X}^{t_n}_\cdot, \bar{Y}^t_\cdot, \bar{Z}^t_\cdot, \overline{\mathcal{V}}^t_\cdot)\|^2_{\mathcal{H}^2}$$

(the other terms can be treated along the same lines). Introducing a sequence (R_m) of positive numbers going to infinity as $m \to \infty$, let $\Omega^{m,n}_s$ and $\bar{\Omega}^{m,n}_s$ be defined as in the proof of Proposition 8.2(ii), with $(\bar{Y}^t, \bar{Z}^t, \overline{\mathcal{V}}^t)$ instead of $(Y^t, Z^t, \mathcal{V}^t)$ therein. One has for any m, n :

$$\begin{aligned}
&\|\mathbb{1}_{\{t<s\cdot<\tau^t\}}\widehat{g}(\cdot, \mathcal{X}^t_\cdot, \bar{Y}^t_\cdot, \bar{Z}^t_\cdot, \overline{\mathcal{V}}^t_\cdot) - \mathbb{1}_{\{t_n<\cdot<\tau^{t_n}\}}\widehat{g}(\cdot, \mathcal{X}^{t_n}_\cdot, \bar{Y}^t_\cdot, \bar{Z}^t_\cdot, \overline{\mathcal{V}}^t_\cdot)\|^2_{\mathcal{H}^2} \\
&= \mathbb{E}\int_0^T \Big[\mathbb{1}_{\{t<s<\tau^t\}}\widehat{g}(s, \mathcal{X}^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s) - \mathbb{1}_{\{t_n<s<\tau^{t_n}\}}\widehat{g}(s, \mathcal{X}^{t_n}_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s)\Big]^2 ds \\
&\le 2\mathbb{E}\int_0^T \Big[\widehat{g}(s, \mathcal{X}^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s)^2 + \widehat{g}(s, \mathcal{X}^{t_n}_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s)^2\Big] \mathbb{1}_{\bar{\Omega}^{m,n}_s} ds + \\
&\qquad \mathbb{E}\int_0^T \Big[\mathbb{1}_{\{t<s<\tau^t\}}\widehat{g}(s, \mathcal{X}^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s) - \mathbb{1}_{\{t_n<s<\tau^{t_n}\}}\widehat{g}(s, \mathcal{X}^{t_n}_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s)\Big]^2 \mathbb{1}_{\Omega^{m,n}_s} ds \\
&=: I_{m,n} + II_{m,n} \ .
\end{aligned}$$

As in the proof Proposition 8.2(ii) (using the fact that $\overline{\mathcal{Y}}^t \in \mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu \times \mathcal{A}^2$ instead of $\widetilde{\mathcal{Y}}^t$ therein), $I_{m,n} \to 0$ as $m, n \to \infty$. Moreover since $N^t_s = N^{t_n}_s$ on $\Omega^{m,n}_s$ one has that

$$
\begin{aligned}
II_{m,n} &= \mathbb{E}\int_0^T \Big[\mathbb{1}_{\{t<s<\tau^t\}}\widehat{g}(s, X^t_s, N^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s) - \mathbb{1}_{\{t_n<s<\tau^{t_n}\}}\widehat{g}(s, X^{t_n}_s, N^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s)\Big]^2 \mathbb{1}_{\Omega^{m,n}_s} ds \\
&\leq 2\mathbb{E}\int_0^T \Big[\widehat{g}(s, X^t_s, N^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s) - \widehat{g}(s, X^{t_n}_s, N^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s)\Big]^2 \mathbb{1}_{\Omega^{m,n}_s} ds \\
&\quad + 2\mathbb{E}\int_0^T |\mathbb{1}_{\{t<s<\tau^t\}} - \mathbb{1}_{\{t_n<s<\tau^{t_n}\}}| \widehat{g}(s, X^t_s, N^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s)^2 ds \;,
\end{aligned}
$$

where in the last inequality:

• $\mathbb{E}\int_0^T \Big[\widehat{g}(s, X^t_s, N^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s) - \widehat{g}(s, X^{t_n}_s, N^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s)\Big]^2 \mathbb{1}_{\Omega^{m,n}_s} ds \leq \mathbb{E}\int_0^T \eta_m\left(|X^t_s - X^{t_n}_s|\right) ds$ for a non-negative bounded function η_m continuous and null at 0 (cf. the proof of Proposition 8.2(ii));

• $\mathbb{E}\int_0^T |\mathbb{1}_{\{t<s<\tau^t\}} - \mathbb{1}_{\{t_n<s<\tau^{t_n}\}}| (\widehat{g}(s, X^t_s, N^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s)^2 ds$ goes to 0 as $n, m \to \infty$, by $d\mathbb{P}\times dt$-integrability of $\widehat{g}(s, X^t_s, N^t_s, \bar{Y}^t_s, \bar{Z}^t_s, \overline{\mathcal{V}}^t_s)^2$ joint to the fact that

$$
\mathbb{E}\int_0^T |\mathbb{1}_{\{t<s<\tau^t\}} - \mathbb{1}_{\{t_n<s<\tau^{t_n}\}}| ds = \mathbb{E}|\tau^t - \tau^{t_n}| + |t - t_n| \to 0 \text{ as } n, m \to \infty \;,
$$

by dominated convergence (under Assumption 8.1).
We conclude the proof as for Proposition 8.2(ii).

B Proofs of Auxiliary PDE Results

B.1 Proof of Lemma 12.3

(i) Let $(t^\star, x^\star, i) \in (0, T) \times \mathbb{R}^d \times I$ be such that $\omega^i(t^\star, x^\star) > 0$ and $(t^\star, x^\star)$ maximizes $\omega^i - \varphi^i$ for some function $\varphi \in \mathcal{C}^{1,2}(\mathcal{E})$. We need to show that (214) holds at $(t^\star, x^\star, i)$. We first assume $t^\star > 0$. By a classical argument, we may and do reduce attention to the case where $(t^\star, x^\star)$ maximizes *strictly* $\omega^i - \varphi^i$. Let us then introduce the function

$$
\varphi^i_{\varepsilon,\alpha}(t, x, s, y) = \mu^i(t, x) - \nu^i(s, y) - \frac{|x-y|^2}{\varepsilon^2} - \frac{|t-s|^2}{\alpha^2} - \varphi^i(t, x) \tag{281}
$$

on $[0, T] \times \mathbb{R}^d$, in which ε, α are positive parameters devoted to tend to zero in some way later in the proof. By a classical argument in the theory of viscosity solutions known as *the Jensen–Ishii Lemma* (see, e.g., Crandall et al. [39] or Fleming and Soner [54]), there exists, for any positive ε, α, points (t, x), (s, y) in $[0, T] \times \bar{B}_R$ (we omit the dependence of t, x, s, y in ε, α, for notational simplicity), where $\bar{B}_R$ is a ball around $x^\star$ with a large radius R which will be fixed throughout in a way made precise later, such that:

• for any positive ε, α, the related quadruple (t, x, s, y) maximizes $\varphi^i_{\varepsilon,\alpha}$ over $([0, T] \times \bar{B}_R)^2$. In particular,

$$
\begin{aligned}
&\mu^i(t^\star, x^\star) - \nu^i(t^\star, x^\star) - \varphi^i(t^\star, x^\star) = \varphi^i_{\varepsilon,\alpha}(t^\star, x^\star, t^\star, x^\star) \\
&\quad \leq \varphi^i_{\varepsilon,\alpha}(t, x, s, y) = \mu^i(t, x) - \nu^i(s, y) - \frac{|x-y|^2}{\varepsilon^2} - \frac{|t-s|^2}{\alpha^2} - \varphi^i(t, x) \;; \qquad (282)
\end{aligned}
$$

- (t, x), $(s, y) \to (t^\star, x^\star)$ as $\varepsilon, \alpha \to 0$;
- $\frac{|x-y|^2}{\varepsilon^2}$, $\frac{|t-s|^2}{\alpha^2}$ are bounded and tend to zero as $\varepsilon, \alpha \to 0$.

It follows from [39, Theorem 8.3] that there exists symmetric matrices $X, Y \in \mathbb{R}^{d \otimes d}$ such that

$$
\begin{gathered}
(a + \partial_t \varphi(t, x), p + \partial \varphi^i(t, x), X) \in \bar{\mathcal{P}}^{2,+} \mu^i(t, x) \\
(a, p, Y) \in \bar{\mathcal{P}}^{2,-} \nu^i(s, y) \\
\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq \tfrac{4}{\varepsilon^2} \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + \begin{pmatrix} \mathcal{H}\varphi(t, x) & 0 \\ 0 & 0 \end{pmatrix}
\end{gathered} \tag{283}
$$

where $\bar{\mathcal{P}}^{2,+} \mu^i(t, x)$, resp. $\bar{\mathcal{P}}^{2,-} \nu^i(s, y)$, denotes the *closure of the parabolic superjet* of μ^i at (t, x), resp. *subjet* of ν^i at (s, y) (see [39, 54]), and

$$
a = \frac{2(t-s)}{\alpha^2} \ , \ p = \frac{2(x-y)^\mathsf{T}}{\varepsilon^2} \, . \tag{284}
$$

Modifying if necessary $\varphi^i_{\varepsilon,\alpha} = \varphi^i_{\varepsilon,\alpha}(t', x', s', y')$ by adding terms of the form $\xi(x')$ and $\xi(y')$ with supports in the complement $\bar{B}^c_{R/2}$ of $\bar{B}_{R/2}$, we may assume that (t, x, s, y) is a global maximum point of $\varphi^i_{\varepsilon,\alpha}$ over $([0, T] \times \mathbb{R}^d)^2$. Since $\omega^i(t^\star, x^\star) > 0$, then by (282) there exists $\rho > 0$ such that $\mu^i(t, x) - \nu^i(s, y) \geq \rho > 0$ for (ε, α) small enough. Combining this inequality with the fact that $\ell \leq \nu$ and $\mu \leq h$, we deduce by continuity of the obstacles ℓ and h that for (ε, α) small enough:

$$
\ell^i(t, x) < \mu^i(t, x) \ , \ \nu^i(s, y) < \ell^i(s, y)
$$

so that the related sub- and super-solution inequalities are satisfied by μ at (t, x, i) and ν at (s, y, i). Thus

$$
\begin{aligned}
&- a - \partial_t \varphi^i(t, x) - \frac{1}{2} \mathrm{Tr}(a^i(t, x) X) - p b^i(t, x) - \partial \varphi^i(t, x) \Big(b^i(t, x) - \int_{\mathbb{R}^d} \delta^i(t, x, z) f^i(t, x, z) m(dz) \Big) \\
&\quad - \int_{\mathbb{R}^d} \Big(\mu^i(t, x + \delta^i(t, x, z)) - \mu^i(t, x) - p \delta^i(t, x, z) \Big) f^i(t, x, z) m(dz) \\
&\quad - g^i(t, x, \mu(t, x), (p + \partial \varphi^i(t, x)) \sigma^i(t, x), \mathcal{I} \mu^i(t, x)) \leq 0 \\
&- a - \frac{1}{2} \mathrm{Tr}(a^i(s, y)) - p b^i(s, y) \\
&\quad - \int_{\mathbb{R}^d} \Big(\nu^i(s, y + \delta^i(s, y, z)) - \nu^i(s, y) - p \delta^i(s, y, z) \Big) f^i(s, y, z) m(dz) \\
&\quad - g^i(s, y, \nu(s, y), p \sigma^i(s, y), \mathcal{I} \nu^i(s, y)) \geq 0
\end{aligned}
$$

Comments B.1 **(i)** The ξ terms that one has added to $\varphi_{\varepsilon,\alpha}$ to have a global maximum point do not appear in these inequalities because δ has linear growth in x and is thus locally bounded, whereas these terms have a support which is included in $\bar{B}^c_{R/2}$ with R large.
(ii) Since we restrict ourselves to finite jump measures $m(dz)$, the Jensen–Ishii Lemma is indeed applicable in its ‘differential’ form (such as it is stated in [39]) as done here. In the case of unbounded Levy measures however, Barles and Imbert [7] (see also Jakobsen and Karlsen [68]) recently established that this Lemma (and thus the related uniqueness proofs in Barles et al. [6], and then in turn in Harraj et al. [63]) has to be amended in a rather involved way.

By substracting the previous inequalities, there comes:

$$
\begin{aligned}
&-\partial_t\varphi^i(t,x) - \frac{1}{2}\Big(\mathrm{Tr}(a^i(t,x)X) - \mathrm{Tr}(a^i(s,y))\Big) \\
&\quad - p\Big(b^i(t,x) - b^i(s,y))\Big) - \partial\varphi^i(t,x)\Big(b^i(t,x) - \int_{\mathbb{R}^d}\delta^i(t,x,z)f^i(t,x,z)m(dz)\Big) \\
&\quad - \int_{\mathbb{R}^d}\Big[(\mu^i(t,x+\delta^i(t,x,z)) - \mu^i(t,x)) - (\nu^i(s,y+\delta^i(s,y,z)) - \nu^i(s,y)) \\
&\qquad\quad - p(\delta^i(t,x,z) - \delta^i(s,y,z))\Big]f^i(t,x,z)m(dz) \\
&\quad + \int_{\mathbb{R}^d}\Big[\nu^i(s,y+\delta^i(s,y,z)) - \nu^i(s,y) - p\delta^i(s,y,z))\Big]\Big[f^i(t,x,z) - f^i(s,y,z)\Big]m(dz) \\
&\quad - \Big(g^i(t,x,\mu(t,x),(p+\partial\varphi^i(t,x))\sigma^i(t,x),\mathcal{I}\mu^i(t,x)) - g^i(s,y,\nu(s,y),p\sigma^i(s,y),\mathcal{I}\nu^i(s,y))\Big) \le 0
\end{aligned}
$$

Now, by straightforward computations analogous to those in [6, page 76-77] (see also [87]) using the maximization property of (t,x,s,y), the definition of p (cf. (284)), the matrix inequality (283) and the Lipschitz continuity properties of the data (and accounting for the fact that we deal here with inhomogeneous coefficients $b^i(t,x)$, $\sigma^i(t,x)$, and $\delta^i(t,x,z)$ here, instead of $b(x)$, $\sigma(x)$, and $c(x,z)$ in [6, 87]), we have:

$$
\begin{aligned}
&|p|(|t-s|+|x-y|) \le C\frac{|t-s|^2+|x-y|^2}{\varepsilon^2} \\
&\mathrm{Tr}(a^i(t,x)X) - \mathrm{Tr}(a^i(s,y)Y) \le C\frac{|t-s|^2+|x-y|^2}{\varepsilon^2} + \mathrm{Tr}(a^i(t,x)\mathcal{H}\varphi^i(t,x)) \\
&|p\Big(b^i(t,x) - b^i(s,y))\Big)| \le C\frac{|t-s|^2+|x-y|^2}{\varepsilon^2} \\
&(\mu^i(t,x+\delta^i(t,x,z)) - \mu^i(t,x)) - (\nu^i(s,y+\delta^i(s,y,z)) - \nu^i(s,y)) \\
&\qquad \le (\varphi^i(t,x+\delta^i(t,x,z)) - \varphi^i(t,x)) + \Big(\frac{|x+\delta^i(t,x,z)-y-\delta^i(s,y,z)|^2}{\varepsilon^2} - \frac{|x-y|^2}{\varepsilon^2}\Big)
\end{aligned}
$$

where in the last inequality

$$
\begin{aligned}
&\frac{|x+\delta^i(t,x,z)-y-\delta^i(s,y,z)|^2}{\varepsilon^2} - \frac{|x-y|^2}{\varepsilon^2} \\
&\quad = \frac{1}{\varepsilon^2}\Big[2(x-y)^{\mathsf{T}}(\delta^i(t,x,z)-\delta^i(s,y,z)) + |\delta^i(t,x,z)-\delta^i(s,y,z)|^2\Big] \\
&\quad = p(\delta^i(t,x,z)-\delta^i(s,y,z)) + \frac{1}{\varepsilon^2}|\delta^i(t,x,z)-\delta^i(s,y,z)|^2 \\
&\quad \le C\frac{|t-s|^2+|x-y|^2}{\varepsilon^2}
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&-\partial_t\varphi^i(t,x)-\frac{1}{2}\mathrm{Tr}(a^i(t,x)\mathcal{H}\varphi^i(t,x))-\partial\varphi^i(t,x)\Big(b^i(t,x)-\int_{\mathbb{R}^d}\delta^i(t,x,z)f^i(t,x,z)m(dz)\Big)\\
&\quad-\int_{\mathbb{R}^d}\Big(\varphi^i(t,x+\delta^i(t,x,z))-\varphi^i(t,x)\Big)f^i(t,x,z)m(dz)\\
&\quad-\Big(g^i(t,x,\mu(t,x),(p+\partial\varphi^i(t,x))\sigma^i(t,x),\mathcal{I}\mu^i(t,x))-g^i(s,y,\nu(s,y),p\sigma^i(s,y),\mathcal{I}\nu^i(s,y))\Big)\\
&\qquad\le C\Big(|t-s|+|x-y|+\frac{|t-s|^2+|x-y|^2}{\varepsilon^2}\Big)
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{I}\mu^i(t,x)-\mathcal{I}\nu^i(s,y))&=\int_{\mathbb{R}^d}\Big[\nu^i(s,y+\delta^i(s,y,z))-\nu^i(s,y)\Big]\Big[f^i(t,x,z)-f^i(s,y,z)\Big]m(dz)\\
&\quad+\int_{\mathbb{R}^d}\Big[(\mu^i(t,x+\delta^i(t,x,z))-\mu^i(t,x))-(\nu^i(s,y+\delta^i(s,y,z))-\nu^i(s,y))\Big]f^i(t,x,z)m(dz)\\
&\le\int_{\mathbb{R}^d}\Big[(\varphi^i(t,x+\delta^i(t,x,z))-\varphi^i(t,x))\Big]f^i(t,x,z)m(dz)\\
&\qquad+C\Big(|t-s|+|x-y|+\frac{|t-s|^2+|x-y|^2}{\varepsilon^2}\Big)\\
&=\mathcal{I}\varphi^i(t,x)+C\Big(|t-s|+|x-y|+\frac{|t-s|^2+|x-y|^2}{\varepsilon^2}\Big)
\end{aligned}
$$

$$
\begin{aligned}
&g^i(t,x,\mu(t,x),(p+\partial\varphi^i(t,x))\sigma^i(t,x),\mathcal{I}\mu^i(t,x))-g^i(s,y,\nu(s,y),p\sigma^i(s,y),\mathcal{I}\nu^i(s,y))\\
&\quad\le\eta_\varepsilon(|t-s|)+\eta_R(|x-y|(1+|p\sigma^i(s,y)|))+k\Lambda\max_{j\in I}(\mu^j(t,x)-\nu^j(s,y))^+\\
&\qquad+\Lambda|p(\sigma^i(t,x)-\sigma^i(s,y))+(\partial\varphi\sigma)^i(t,x)|+\Lambda(\mathcal{I}\mu^i(t,x)-\mathcal{I}\nu^i(s,y))^+
\end{aligned}
$$

where in the last inequality:
- η_ε is a is a modulus of continuity of g^i on a compact set parameterized by ε, obtained by using the fact that p in (284) is bounded independently of α, for given ε;
- η_R is the modulus of continuity standing in Assumption 12.1(ii);
- the three last terms come from the Lipschitz continuity and/or monotonicity properties of g with respect to its three last variables. Therefore

$$
\begin{aligned}
-\widetilde{\mathcal{G}}\varphi^i(t,x)&=-\partial_t\varphi^i(t,x)-\frac{1}{2}\mathrm{Tr}(a^i(t,x)\mathcal{H}\varphi^i(t,x))-\partial\varphi^i(t,x)\Big(b^i(t,x)-\int_{\mathbb{R}^d}\delta^i(t,x,z)f^i(t,x,z)m(dz)\Big)\\
&\qquad-\int_{\mathbb{R}^d}\Big(\varphi^i(t,x+\delta^i(t,x,z))-\varphi^i(t,x)\Big)f^i(t,x,z)m(dz)\\
&\le\Lambda_1\Big(\max_{j\in I}(\mu^j(t,x)-\nu^j(s,y))^++|(\partial\varphi\sigma)^i(t,x)|+\mathcal{I}\varphi^i(t,x)^+\Big)\\
&\qquad+\eta_\varepsilon(|t-s|)+\eta_R(|x-y|(1+|p\sigma^i(s,y)|))\\
&\qquad+C\Big(|t-s|+|x-y|+\frac{|t-s|^2+|x-y|^2}{\varepsilon^2}\Big)
\end{aligned}
$$

Given $\rho>0$ one thus has for $\varepsilon\le\varepsilon_\rho$ and $\alpha\le\varepsilon$, using the properties of (t,x,s,y) in the Jensen–Ishii Lemma and the regularity of φ^i :

$$
\begin{aligned}
&-\widetilde{\mathcal{G}}\varphi^i(t^*,x^*)-\Lambda_1\Big(\max_{j\in I}(\mu^j(t,x)-\nu^j(s,y))^++|(\partial\varphi\sigma)^i(t^*,x^*)|+\mathcal{I}\varphi^i(t^*,x^*)^+\Big)\\
&\quad\le\rho+\eta_\varepsilon(|t-s|)\ .
\end{aligned}
$$

Note that $t-s \to 0$ for fixed ε as $\alpha \to 0$, by boundness of $\frac{|t-s|^2}{\alpha^2}$ in the Jensen–Ishii Lemma. Whence for $\alpha \le \alpha_\varepsilon (\le \varepsilon)$:

$$-\widetilde{\mathcal{G}}\varphi^i(t^*,x^*) - \Lambda_1\Big(\max_{j\in I}(\mu^j(t,x)-\nu^j(s,y))^+ + |(\partial\varphi\sigma)^i(t^*,x^*)| + \mathcal{I}\varphi^i(t^*,x^*)^+\Big) \le 2\rho$$

Sending $\rho, \varepsilon, \alpha$ to zero with $\varepsilon \le \varepsilon_\rho$ and $\alpha \le \alpha_\varepsilon$, inequality (214) at $(t^\star, x^\star, i)$ follows by upper semi continuity of the function $(t',x',s',y') \mapsto \max_{j\in I}(\mu^j(t',x') - \nu^j(s',y'))^+$. This finishes to prove that (214) holds at $(t^\star, x^\star, i)$ in case $t^\star > 0$.

Now in case $t^\star = 0$ let us introduce the function

$$\varphi^i_\varepsilon(t,x) = \omega^i(t,x) - (\varphi^i(t,x) + \frac{\varepsilon}{t}) \tag{285}$$

on $[0,T]\times \bar{B}_R$, in which ε is a positive parameter devoted to tend to zero. Assuming again w.l.o.g. that $(t^\star = 0, x^\star)$ maximizes *strictly* $\omega^i - \varphi^i$, there exists, for any $\varepsilon > 0$, a point (t,x) in $[0,T]\times\bar{B}_R$ (we omit the dependence of (t,x) in ε, for notational simplicity), where $\bar{B}_R$ is a ball with large radius R around $x^\star$, such that:

- for any $\varepsilon > 0$ the related point (t,x) maximizes φ^i_ε over $[0,T]\times\bar{B}_R$, *and one has* $t>0$, *for* ε *small enough;*
- $(t,x) \to (t^\star, x^\star)$ as $\varepsilon \to 0$.

In virtue of the part of the result already established in $t^* > 0$, we may thus apply (214) to the function $(s,y) \mapsto \varphi^i(s,y) + \frac{\varepsilon}{s}$ at (t,x,i), whence:

$$-\widetilde{\mathcal{G}}\varphi^i(t,x) \ - \Lambda_1\Big(\max_{j\in I}(\omega^j(t,x))^+ + |\partial\varphi^i(t,x)\sigma^i(t,x)| + (\mathcal{I}\varphi^i(t,x))^+\Big) \le -\frac{\varepsilon}{t^2} \le 0 \ .$$

Sending ε to 0 in the left hand side we conclude by upper semi-continuity of $\max_{j\in I}(\omega^j)^+$ that (214) holds at $(t^\star = 0, x^\star, i)$.

(ii) Straightforward computations give:

$$-\partial_t\chi(t,x) = C_1\chi(t,x)$$
$$(1+|x|)|\partial\chi(t,x)| \ \vee \ (1+|x|^2)|\mathcal{H}\chi(t,x)| \ \vee \ \chi(t, x+\delta^i(t,x,z))| \le C|\chi(t,x)|$$

on $\mathcal{E}$, for a constant C independent of C_1. Therefore for $C_1 > 0$ large enough

$$-\widetilde{\mathcal{G}}\chi - \Lambda_1\big(\chi + |\partial\chi\sigma| + (\mathcal{I}\chi)^+\big) > 0$$

on $\mathcal{E}$.

(iii) First note that $\frac{|\omega|}{\chi}$ goes to 0 uniformly in t,i as $|x|\to\infty$, since $q_1 > q_2$. Given $\alpha > 0$, let us prove that

$$\sup_{(t,x,i)\in\mathcal{E}} \Big((\omega^i(t,x))^+ - \alpha\chi(t,x)\Big)e^{-\Lambda_1(T-t)} \le 0 \ . \tag{286}$$

Assume by contradiction that one has $>$ instead of $\le$ in (286). Then by upper semi-continuity of ω^+ the supremum is reached at some point $(t^*, x^*, i) \in \mathrm{Int}_p\mathcal{E}$ in the left hand side of (286), and

$$(\omega^i(t^*,x^*))^+ \ge \omega^i(t^*,x^*))^+ - \alpha\chi(t^*,x^*) > 0 \ . \tag{287}$$

Therefore one has on $[0,T]\times\mathbb{R}^d$:

$$\Big(\omega^i(t,x)-\alpha\chi(t,x)\Big)e^{-\Lambda_1(T-t)}\leq\Big((\omega^i(t,x))^+-\alpha\chi(t,x)\Big)e^{-\Lambda_1(T-t)}$$
$$\leq\Big((\omega^i(t^*,x^*))^+-\alpha\chi(t^*,x^*)\Big)e^{-\Lambda_1(T-t^*)}=\Big(\omega^i(t^*,x^*)-\alpha\chi(t^*,x^*)\Big)e^{-\Lambda_1(T-t^*)}$$

thus

$$\omega^i(t,x)-\alpha\chi(t,x)\leq\Big(\omega^i(t^*,x^*)-\alpha\chi(t^*,x^*)\Big)e^{-\Lambda_1(t-t^*)}\ .$$

In other words, (t^*,x^*) maximizes globally at zero $\omega^i-\varphi^i$ over $[0,T]\times\mathbb{R}^d$, with

$$\varphi^i(t,x)=\alpha\chi(t,x)+\Big(\omega^i(t^*,x^*)-\alpha\chi(t^*,x^*)\Big)e^{-\Lambda_1(t-t^*)}\ .$$

Whence by part (i) (given that $\omega^i(t^*,x^*)>0$, by (287)):

$$-\widetilde{\mathcal{G}}\varphi^i(t^*,x^*)\ -\Lambda_1\Big(\max_{j\in I}(\omega^j(t^*,x^*))^++|\partial\varphi^i(t^*,x^*)\sigma^i(t^*,x^*)|+(\mathcal{I}\varphi^i(t^*,x^*))^+\Big)\leq 0\ . \tag{288}$$

But the left hand side in this inequality is nothing but

$$-\alpha\widetilde{\mathcal{G}}\chi(t^*,x^*)+\Lambda_1\Big(\omega^i(t^*,x^*)-\alpha\chi(t^*,x^*)\Big)$$
$$-\Lambda_1\Big(\omega^i(t^*,x^*)+\alpha|\partial\chi(t^*,x^*)\sigma^i(t^*,x^*)|+\alpha(\mathcal{I}\chi^i(t^*,x^*))^+\Big)$$
$$=-\alpha\widetilde{\mathcal{G}}\chi(t^*,x^*)-\Lambda_1\Big(\alpha\chi(t^*,x^*)+\alpha|\partial\chi(t^*,x^*)\sigma^i(t^*,x^*)|+\alpha(\mathcal{I}\chi^i(t^*,x^*))^+\Big)$$

which should be positive by (215) in (ii), in contradiction with (288).

References

[1] Alvarez O., Tourin, A.: Viscosity solutions of nonlinear integro-differential equations. *Annales de l'institut Henri Poincaré (C) Analyse non linéaire*, 13 no. 3 (1996), p. 293-317.

[2] Amadori, A.L.: Nonlinear integro-differential evolution problems arising in option pricing: a viscosity solutions approach. *J. Differential and Integral Equations* 16(7), 787–811 (2003).

[3] Amadori, A.L.: The obstacle problem for nonlinear integro-differential operators arising in option pricing. *Ricerche di matematica*, Volume 56, Number 1, June 2007.

[4] Bally V., Caballero E., Fernandez, B. and El-Karoui N.: Reflected BSDE's PDE's and Variational Inequalities, To appear in *Bernoulli*.

[5] Bally, V. and Matoussi, A.: Weak solutions for SPDEs and Backward doubly stochastic differential equations, *Journal of Theoretical Probability,* Vol. 14, No.1, 125-164 (2001).

[6] Barles, G., Buckdahn, R. and Pardoux, E.: Backward Stochastic Differential Equations and Integral-Partial Differential Equations. *Stochastics and Stochastics Reports*, Vol. 60, pp. 57-83 (1997).

[7] Barles, G., Imbert, C.: Second-Order Elliptic Integro-Differential Equations: Viscosity Solutions' Theory Revisited. *Annales de l'IHP*, Volume 25, No 3, Pages 567-585 (2008).

[8] Barles, G. and L. Lesigne: SDE, BSDE and PDE. *Backward Stochastic differential Equations. Pitman Research Notes in Mathematics Series* 364, Eds El Karoui, N. and Mazliak, L., pp. 47-80 (1997).

[9] Barles, G. and Souganidis, P.E.: Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Analysis* 4 (1991), 271–283.

[10] Becherer, D. and Schweizer, M.: Classical Solutions to Reaction-Diffusion Systems for Hedging with Interacting Itô and Point Processes. *Annals of Applied Probability*, 15, 1111-1144, 2005.

[11] Bensoussan, A. and Lions, J.L.: *Applications of Variational Inequalities in Stochastic Control*, North-Holland, Amsterdam, 1982.

[12] Bensoussan, A. and Lions, J.L.: *Impulse Control and Quasi-variational Inequalities*, Gauthier-Villars, Paris, 1984.

[13] Bichteler, K.: *Stochastic Integration with Jumps,* Cambridge University Press, 2002.

[14] Bielecki, T.R., Crépey, S. and Jeanblanc, M.: Up and Down Credit Risk. Forthcoming in *Quantitative Finance.*

[15] Bielecki, T.R., Crépey, S., Jeanblanc, M. and Rutkowski, M.: Arbitrage pricing of defaultable game options with applications to convertible bonds. *Quantitative Finance*, Vol. 8, Issue 8 (December 2008), pp. 795–810 (Preprint version available online at `www.defaultrisk.com`).

[16] Bielecki, T.R., Crépey, S., Jeanblanc, M. and Rutkowski, M.: Valuation and hedging of defaultable game options in a hazard process model. *Journal of Applied Mathematics and Stochastic Analysis*, Article ID 695798, 2009 (Long Preprint version available online at `www.defaultrisk.com`).

[17] Bielecki, T.R., Crépey, S., Jeanblanc, M. and Rutkowski, M.: Defaultable options in a Markovian intensity model of credit risk. *Mathematical Finance*, vol. 18, pp. 493-518, 2008 (Updated version available online at `www.defaultrisk.com`).

[18] Bielecki, T.R., Crépey, S., Jeanblanc, M. and Rutkowski, M.: Convertible bonds in a defaultable diffusion model. *Working Paper* (available online at `www.defaultrisk.com`).

[19] Bielecki, T.R., Crépey, S., Jeanblanc, M. and Rutkowski, M.: Valuation of basket credit derivatives in the credit migrations environment. *Handbook of Financial Engineering*, J. Birge and V. Linetsky eds., Elsevier, 2007.

[20] Bielecki, T.R. and Rutkowski, M.: *Credit Risk: Modeling, Valuation and Hedging.* Springer, 2002.

[21] Bielecki, T.R., Vidozzi, A. and Vidozzi, L.: An efficient approach to valuation of credit basket products and ratings triggered step-up bonds. *Working Paper*, 2006.

[22] Bielecki, T.R., Vidozzi, A. and Vidozzi, L.: A Markov Copulae Approach to Pricing and Hedging of Credit Index Derivatives and Ratings Triggered Step–Up Bonds, *J. of Credit Risk*, 2008.

[23] Black, F. and Scholes, M.: The pricing of options and corporate liabilities, *J. Pol. Econ.*, 81 (1973), pp. 637–659.

[24] Boel, R., Varaiya, P., Wong, E.: Martingales on jump processes: I Representation results, *SIAM JC,* Vol. 13, No. 5, 1975.

[25] Boel, R., Varaiya, P., Wong, E.: Martingales on jump processes: II Applications, *SIAM JC,* Vol. 13, No. 5, 1975.

[26] Bouchard, B. and Chassagneux, J.-F.: Discrete time approximation for continuously and discretely reflected BSDE's. Forthcoming in *Stochastic Processes and their Applications.*

[27] Bouchard B. and Elie R: Discrete time approximation of decoupled Forward-Backward SDE with jumps, *Stoch. Process. Appl.*, 118, 53–75, 2008.

[28] Bouchard, B. and Touzi, N.: Discrete-time approximation and Monte Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their Applications* 111 (2004), 175–206.

[29] Brémaud, P.: *Point processes and queues, martingale dynamics*, Springer, 1981.

[30] Briani, M., La Chioma, C. and Natalini, R.: Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in financial theory, *Numer. Math.*, 2004, vol. 98, no4, pp. 607-646.

[31] Busca, J. and Sirakov, B.: Harnack type estimates for nonlinear elliptic systems and applications, *Annales de l'institut Henri Poincaré (C) Analyse non linéaire*, 21 no. 5 (2004), p. 543-590.

[32] Chassagneux, J.-F. and Crépey, S.: Doubly reflected BSDEs with Call Protection and their Approximation. *In Preparation.*

[33] Chassagneux, J.-F., Crépey, S. and Rahal, A.: Pricing Game Options with Call Protection. *In Preparation.*

[34] Cherny, A. and Shiryaev, A. Vector stochastic integrals and the fundamental theorems of asset pricing, *Proceedings of the Steklov Institute of mathematics*, 237 (2002), pp. 6–49.

[35] Coculescu, D. and Nikeghbali, A.: Hazard processes and martingale hazard processes. *Working Paper*, 2008.

[36] Cont, R. and Minca, A.: Recovering Portfolio Default Intensities Implied by CDO Quotes, *Working Paper*, 2008.

[37] Cont, R. and Tankov, P.: *Financial Modelling with Jump Processes*, Chapman & Hall, 2003.

[38] Cont, R., Voltchkova, K.: Finite difference methods for option pricing methods in exponential Levy models, *SIAM Journal on Numerical Analysis*, Vol 43, 2005.

[39] Crandall, M., Ishii, H. and Lions, P.-L.: User's guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc.*, 1992.

[40] Crandall, M.G., Kocan, M. and Swiech, A.: L^p-theory for fully nonlinear parabolic equations, *CPDE*, 25 (2000), 11 and 12, pp. 1997–2053.

[41] Carmona, R. and Crépey, S.: Importance Sampling and Interacting Particle Systems for the Estimation of Markovian Credit Portfolios Loss Distribution, Forthcoming in *International Journal of Theoretical and Applied Finance.*

[42] Crépey, S. and Matoussi, A.: Reflected and doubly reflected BSDEs with jumps: A priori estimates and comparison principle. *Annals of Applied Probability*, Vol 18, Issue 5 (October 2008), pp. 2041–69.

[43] Cvitanic, J. and Karatzas, I.: Backward Stochastic Differential Equations with reflection and Dynkin Games, *Annals of Probability* 24 (4), 2024-2056 (1996).

[44] Darling, R. and Pardoux, E.: Backward SDE with random terminal time and applications to semilinear elliptic PDE. *Annals of Probability* 25 (1997), 1135–1159.

[45] Delbaen, F. and Schachermayer, W.: The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische Annalen* 312 (1997), 215–250.

[46] Dellacherie, C. and Meyer, P.-A.: *Probabilities and Potential*, North-Holland, 1982.

[47] Dynkin, E.B.: Game variant of a problem on optimal stopping. *Soviet Math. Dokl.* 10 (1969), 270–274.

[48] Dynkin, E.B.: *Markov processes*, Springer-Verlag, 1965.

[49] El Karoui, N., Hamadène, S. and Matoussi, A.: Backward Stochastic Differential Equations and Applications, in *Paris-Princeton Lecture Notes on Mathematical Finance*, Edts. R. Carmona, Springer-Verlag, Forthcoming.

[50] El Karoui, N., Kapoudjian, E., Pardoux, C., Peng, S., and Quenez, M.-C.: Reflected solutions of backward SDEs, and related obstacle problems for PDEs. *Annals of Probability* 25 (1997), 702–737.

[51] El Karoui, N., Peng, S., and Quenez, M.-C.: Backward stochastic differential equations in finance. *Mathematical Finance* 7 (1997), 1–71.

[52] El Karoui, N. and Quenez, M.-C.: Nonlinear pricing theory and backward stochastic differential equations. In: *Financial Mathematics, Bressanone, 1996,* W. Runggaldier, ed. Lecture Notes in Math. 1656, Springer, Berlin Heidelberg New York, 1997, pp. 191–246.

[53] Ethier, H.J. and Kurtz, T.G.: *Markov Processes. Characterization and Convergence.* Wiley, 1986.

[54] Fleming, W. and Soner, H.: *Controlled Markov processes and viscosity solutions*, Second edition, Springer, 2006.

[55] Föllmer, H., and Sondermann, D.: Hedging of non-redundant contingent claims, *Contributions to Mathematical Economics*, A. Mas-Colell and W. Hildenbrand ed., North-Holland, Amsterdam, 205–223, 1986.

[56] Freidlin, M.: *Markov Processes and Differential Equations: asymptotic problems*, Birkhaüser, 1996.

[57] Frey, R., Backhaus, J.: Dynamic hedging of synthetic CDO-tranches with spread- and contagion risk, Forthcoming in in *Journal of Economic Dynamics and Control.*

[58] Frey, R., Backhaus, J.: Pricing and hedging of portfolio credit derivatives with interacting default intensities. *International Journal of Theoretical and Applied Finance*, 11 (6), 611-634, 2008.

[59] Fujiwara, T. and Kunita, H.: Stochastic Differential Equations of Jump Type and Lévy Processes in Diffeomorphisms Group, *J.Math. Kyoto Univ*, 25, 71-106 (1985)

[60] Hamadène, S.: Mixed zero-sum differential game and American game options. *SIAM J. Control Optim.* 45 (2006), 496–518.

[61] Hamadène, S. and Hassani, M.: BSDEs with two reflecting barriers driven by a Brownian motion and an independent Poisson noise and related Dynkin game. *Electronic Journal of Probability*, vol. 11 (2006), paper no. 5, pp. 121-145.

[62] Hamadène, S. and Ouknine, Y.: Reflected backward stochastic differential equation with jumps and random obstacle. *Electronic Journal of Probability*, Vol. 8 (2003), p. 1–20.

[63] Harraj., N., Ouknine, Y. and Turpin, I.: Double-barriers-reflected BSDEs with jumps and viscosity solutions of parabolic integrodifferential PDEs, *Journal of Applied Mathematics and Stochastic Analysis* (2005), 1, 37-53. *The Journal of Computational Finance*, (1-30) Volume 12/Number 1, Fall 2008.

[64] Herbertsson, A.: Pricing portfolio credit derivatives. *PhD Thesis*, Göteborg University, 2007.

[65] Ishii, H. and Koike, S.: Viscosity solutions for monotone systems of second-order elliptic PDEs, *Comm. Partial Differential Equations*, 16(6-7):1095-1128, 1991.

[66] Jacod, J.: *Calcul Stochastique et Problèmes de Martingales.* Springer, Berlin Heidelberg New York, 2003.

[67] Jacod, J. and Shiryaev, A. N.: *Limit Theorems for Stochastic Processes.* Springer, Berlin Heidelberg New York, 2003.

[68] Jakobsen, E. R. and Karlsen, K. H.: A "maximum principle for semi-continuous functions" applicable to integro-partial differential equations. *Nonlinear Differential Equations Appl.*, 13:137-165, 2006.

[69] Jakobsen, E. R., Karlsen, K. H. and La Chioma, C.: Error estimates for approximate solutions to Bellman equations associated with controlled jump-diffusions. *Numer. Math.*, Volume 110, Number 2 / August 2008.

[70] Jarrow, R.A. and Turnbull, S.: Pricing Derivatives on Financial Securities Subject to Credit Risk. *Journal of Finance* vol. 50, March, 1995.

[71] Jeanblanc, M., Yor, M. and Chesney, M.: *Mathematical methods for Financial Markets.* Springer, 2009.

[72] Kallsen, J. and Kühn, C.: Convertible bonds: financial derivatives of game type. In: *Exotic Option Pricing and Advanced Lévy Models*, Edited by Kyprianou, A., Schoutens, W. and Wilmott, P., Wiley, 2005, pp. 277–288.

[73] Karatzas, I.: On the pricing of American options. *Appl. Math. Optim.* 17 (1988), 37–60.

[74] Kifer, Y.: Game options. *Finance and Stochastics* 4 (2000), 443–463.

[75] Kunita, H.: Stochastic differential equations and stochastic flows of diffeomorphisms. Ecole dété de Probabilité de Saint-Flour, Lect. Notes Math. 1097, 143-303 (1982).

[76] Kunita, H. and Watanabe, T: Notes on transformations of Markov processes connected with multiplicative functionals, *Mem. Fac. Sci. Kyushu Univ*, A17, 181-191 (1963).

[77] Kushner, H. and Dupuis, B.: *Numerical methods for stochastic control problems in continuous time*, Springer, 1992.

[78] Kusuoka, S.: A remark on default risk models. *Advances in Mathematical Economics* 1 (1999), 69–82.

[79] Lando, D.: Three Essays on Contingent Claims Pricing. *Ph.D. Thesis*, Cornell University, 1994.

[80] Last, G., Brandt, A.: *Marked Point Processes on the Real Line: The Dynamical Approach*, Springer, 1995.

[81] Laurent, J.-P., Cousin, A. and Fermanian, J-D.: Hedging default risks of CDOs in Markovian contagion models. Forthcoming in *Quantitative Finance*.

[82] Lepeltier, J.-P. and Maingueneau, M.: Le jeu de Dynkin en théorie générale sans l'hypothèse de Mokobodski. *Stochastics* 13 (1984), 25–44.

[83] Ma, J. and Cvitanić, J.: Reflected Forward-Backward SDEs and obstacle problems with boundary conditions. *Journal of Applied Mathematics and Stochastic Analysis* 14(2) (2001), 113–138.

[84] Merton, R.C.: The theory of rational option pricing, *Bell J. Econ. Man. Sc.*, 4 (1973), pp. 141–183.

[85] Nikeghbali, A. and Yor, M.: A definition and some characteristic properties of pseudo-stopping times. *Annals of Probability* 33 (2005), 1804–1824.

[86] Palmowski, Z. and Rolski, T.: A technique for exponential change of measure for Markov processes. *Bernoulli*, 8 (6), 767–785, 2002.

[87] Pardoux, E., Pradeilles, F. and Rao, Z.: Probabilistic interpretation of systems of semilinear PDEs. *Annales de l'Institut Henri Poincaré, série Probabilités-Statistiques*, 33, 467–490, 1997.

[88] Peng, S.: Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob–Meyer's type, *Probab. Theory Related Fields*, 113 (1999), no. 4, 473–499.

[89] Peng, S. and Xu, M.: The smallest g-supermartingale and reflected BSDE with single and double L_2 obstacles, *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, Volume 41, Issue 3, May-June 2005, Pages 605-630.

[90] Pham, H.: Optimal stopping of controlled jump diffusion processes: A viscosity solution approach, *Journal of Mathematical Systems, Estimation and Control*, 8, 1–27, 1998.

[91] Protter, P.E.: *Stochastic Integration and Differential Equations, Second Edition.* Springer, Berlin Heidelberg New York, 2004.

[92] Protter, P. E.: A partial introduction to financial asset pricing theory, *Stochastic Processes and their Applications*, 91, 2, pp. 169-203(35), 2001.

[93] Schweizer, M.: Option hedging for semimartingales. *Stochastic Processes and their Applications* 37 (1991), 339–363.

[94] Windcliff, H.A., Forsyth, P.A. and Vetzal, K.R.: Pricing methods and hedging strategies for volatility derivatives, *Journal of Banking and Finance*, 30 (2006) 409–431.

[95] Windcliff, H.A., Forsyth, P.A. and Vetzal, K.R.: Numerical Methods for Valuing Cliquet Options. *Applied Mathematical Finance*, 13 (2006), 353–386.