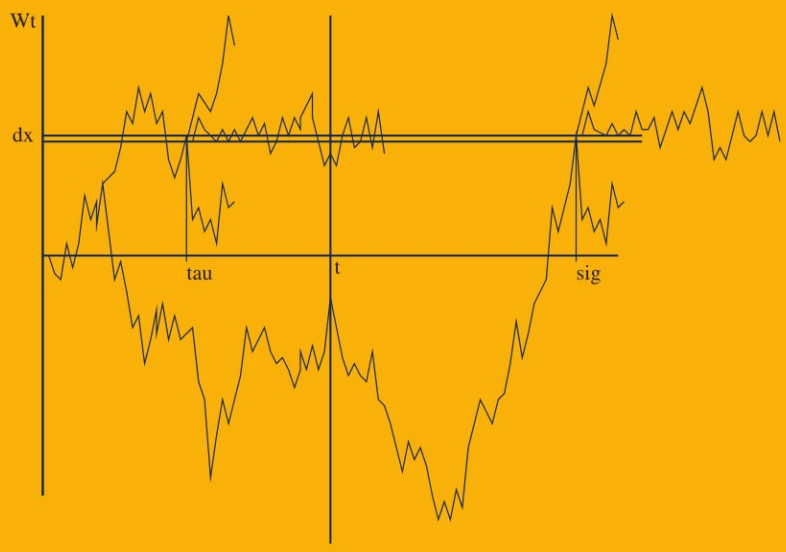
Lecture Notes in Mathematics

2003

Areski Cousin
Stéphane Crépey
Olivier Guéant
David Hobson
Monique Jeanblanc
Jean-Michel Lasry
Jean-Paul Laurent
Pierre-Louis Lions
Peter Tankov

Paris-Princeton Lectures on Mathematical Finance 2010



Springer

Lecture Notes in Mathematics 2003

Editors:
J.-M. Morel, Cachan
F. Takens, Groningen
B. Teissier, Paris

Areski Cousin • Stéphane Crépey
Olivier Guéant • David Hobson
Monique Jeanblanc • Jean-Michel Lasry
Jean-Paul Laurent • Pierre-Louis Lions
Peter Tankov

Paris-Princeton Lectures on Mathematical Finance 2010

Editors:
Rene A. Carmona
Erhan Cinlar
Ivar Ekeland
Elyès Jouini
Jose A. Scheinkman
Nizar Touzi

Springer

Areski Cousin
Université de Lyon
Université Lyon 1
ISFA, 50 Avenue Tony Garnier
69366 Lyon Cedex 07, France
areski.cousin@univ-lyon1.fr

Prof. Stéphane Crépey
Université d'Evry
Département de Mathématiques
Cours Monseigneur Romero
91025 Evry
Bâtiment Maupertuis
France
stephane.crepey@univevry.fr

Dr. Olivier Guéant
34, Rue Mazarine
75006 Paris, France
olivier.gueant@gmail.com

Prof. David Hobson
University of Warwick
Dept of Statistics
Zeeman Building
CV4 7AL Coventry
United Kingdom
D.Hobson@warwick.ac.uk

Prof. Monique Jeanblanc
Université d'Evry
Val d'Essonne
Dépt. Mathématiques
rue du Père Jarlan
91025 Evry Cedex, France
monique.jeanblanc@univ-evry.fr

Jean-Michel Lasry
Université Paris-Dauphine
Institut de Finance
Place du Maréchal de Lattre de
Tassigny
75775 Paris Cedex 16
France
lasry@dauphine.fr

Prof. Jean-Paul Laurent
Université Paris
1 Panthéon-Sorbonne
17 rue de la Sorbonne
7005 Paris
France
laurent.jeanpaul@free.fr

Prof. Pierre-Louis Lions
Collège de France
rue d'Ulm 3
75005 Paris CX 05
France
pierre-louis.lions@college-de-france.fr

Peter Tankov
Ecole Polytechnique
Centre de Mathématiques Appliquées
91128 Palaiseau Cedex
France
peter.tankov@polytechnique.org

[The addresses of the volume editors
appear on page ix]

ISBN 978-3-642-14659-6 e-ISBN 978-3-642-14660-2
DOI 10.1007/978-3-642-14660-2
Springer Heidelberg Dordrecht London New York

Lecture Notes in Mathematics ISSN print edition: 0075-8434
ISSN electronic edition: 1617-9692

Library of Congress Control Number: 2010936197

Mathematics Subject Classification (2010): 91B28, 91B70, 60G49, 49J55, 60H07, 90C46

© Springer-Verlag Berlin Heidelberg 2011, Corrected printing 2011
This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: SPi Publisher Services

Printed on acid-free paper

springer.com

Preface

This is the fourth volume of the Paris-Princeton Lectures in Mathematical Finance. The goal of this series is to publish cutting edge research in self-contained articles prepared by established academics or promising young researchers invited by the editors. Contributions are refereed and particular attention is paid to the quality of the exposition, the goal being to publish articles that can serve as introductory references for research.

The series is a result of frequent exchanges between researchers in finance and financial mathematics in Paris and Princeton. Many of us felt that the field would benefit from timely exposés of topics in which there is important progress. René Carmona, Erhan Cinlar, Ivar Ekeland, Elyes Jouini, José Scheinkman and Nizar Touzi serve in the first editorial board of the Paris-Princeton Lectures in Financial Mathematics. Although many of the chapters involve lectures given in Paris or Princeton, we also invite other contributions. Springer Verlag kindly offered to host the initiative under the umbrella of the Lecture Notes in Mathematics series, and we are thankful to Catriona Byrne for her encouragement and her help.

This fourth volume contains five chapters. In the first chapter, Areski Cousin, Monique Jeanblanc, and Jean-Paul Laurent discuss risk management and hedging of credit derivatives. The latter are over-the-counter (OTC) financial instruments designed to transfer credit risk associated to a reference entity from one counterparty to another. The agreement involves a seller and a buyer of protection, the seller being committed to cover the losses induced by the default. The popularity of these instruments lead a runaway market of complex derivatives whose risk management did not develop as fast. This first chapter fills the gap by providing rigorous tools for quantifying and hedging counterparty risk in some of these markets.

In the second chapter, Stéphane Crépey reviews the general theory of forward backward stochastic differential equations and their associated systems of partial integro-differential obstacle problems and applies it to pricing and hedging financial derivatives. Motivated by the optimal stopping and optimal stopping game formulations of American option and convertible bond pricing, he discusses the well-posedness and sensitivities of reflected and doubly reflected Markovian Backward Stochastic Differential Equations. The third part of the paper is devoted to the variational inequality formulation of these problems and to a detailed discussion of viscosity solutions. Finally he also considers discrete path-dependence issues such as dividend payments.

The third chapter written by Olivier Guéant Jean-Michel Lasry and Pierre-Louis Lions presents an original and unified account of the theory and the applications of the mean field games as introduced and developed by Lasry and Lions in a series of lectures and scattered papers. This chapter provides systematic studies illustrating the application of the theory to domains as diverse as population behavior (the so-called Mexican wave), or economics (management of exhaustible resources). Some of the applications concern optimization of individual behavior when interacting with a large population of individuals with similar and possibly competing objectives. The analysis is also shown to apply to growth models and for example, to their application to salary distributions.

The fourth chapter is contributed by David Hobson. It is concerned with the applications of the famous Skorohod embedding theorem to the proofs of model independent bounds on the prices of options. Beyond the obvious importance of the financial application, the value of this chapter lies in the insightful and extremely pedagogical presentation of the Skorohod embedding problem and its application to the analysis of martingales with given one-dimensional marginals, providing a one-to-one correspondence between candidate price processes which are consistent with observed call option prices and solutions of the Skorokhod embedding problem, extremal solutions leading to robust model independent prices and hedges for exotic options.

The final chapter is concerned with pricing and hedging in exponential Lévy models. Peter Tankov discusses three aspects of exponential Lévy models: absence of arbitrage, including more recent results on the absence of arbitrage in multi-dimensional models, properties of implied volatility, and modern approaches to hedging in these models. It is a self contained introduction surveying all the results and techniques that need to be known to be able to handle exponential Lévy models in finance.

Paris/Princeton The Editors
May 4, 2010

Contents

Hedging CDO Tranches in a Markovian Environment 1
Areski Cousin, Monique Jeanblanc, and Jean-Paul Laurent

About the Pricing Equations in Finance 63
Stéphane Crépey

Mean Field Games and Applications 205
Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions

The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices 267
David Hobson

Pricing and Hedging in Exponential Lévy Models: Review of Recent Results 319
Peter Tankov

Editors

Rene A. Carmona
Princeton University
Dept. Operations Research
& Financial Engineering
08544 Princeton New Jersey
ACE-42, E-Quad, USA
rcarmona@princeton.edu

Erhan Cinlar
Princeton University
Dept. Operations Research
& Financial Engineering
08544 Princeton New Jersey
ACE-42, E-Quad, USA
ecinlar@princeton.edu

Ivar Ekeland
University of British Columbia
Dept. Mathematics
Mathematics Road 1984
V6T 1Z2 Vancouver British
Columbia, Canada
ekeland@math.ubc.ca

Elyès Jouini
University of Paris 9, Dauphine
CEREMADE
Pl. du Maréchal DeLattre De
Tassigny
75775 Paris CX 16, France
jouini@ceremade.dauphine.fr

Jose A. Scheinkman
Princeton University
Dept. Economics
Prospect Ave. 26
08540-5296 Princeton
New Jersey, USA
joses@princeton.edu

Nizar Touzi
École Polytechnique
CNRS UMR 7641
Centre de Mathématiques
Appliquées
91128 Palaiseau CX, France
touzi@cmapx.polytechnique.fr

Hedging CDO Tranches in a Markovian Environment

Areski Cousin, Monique Jeanblanc, and Jean-Paul Laurent

Abstract In this first chapter, we show that a CDO tranche payoff can be perfectly replicated with a self-financed strategy based on the underlying credit default swaps. This extends to any payoff which depends only upon default arrivals, such as basket default swaps. Clearly, the replication result is model dependent and relies on two critical assumptions. First, we preclude the possibility of simultaneous defaults. The other assumption is that credit default swap premiums are adapted to the filtration of default times which therefore can be seen as the relevant information set on economic grounds. Our framework corresponds to a pure contagion model, where the arrivals of defaults lead to jumps in the credit spreads of survived names, the magnitude of which depending upon the names in question, and the whole history of defaults up to the current time. These jumps can be related to the derivatives of the joint survival function of default times. The dynamics of replicating prices of CDO tranches follows the same way. In other words, we only deal with default risks and not with spread risks.

Unsurprisingly, the possibility of perfect hedging is associated with a martingale representation theorem under the filtration of default times. Subsequently, we exhibit a new probability measure under which the short term credit spreads (up to some scaling factor due to positive recovery rates) are the intensities associated with the corresponding default times. For ease of presentation, we introduced first some instantaneous default swaps as a convenient basis of hedging instruments. Eventually, we can exhibit a replicating strategy of a CDO tranche payoff with respect

A. Cousin (✉)
Université de Lyon, Université Lyon 1, ISFA, 50 Avenue Tony Garnier,
69366 Lyon Cedex 07, France
e-mail: areski.cousin@univ-lyon1.fr

M. Jeanblanc
Val d'Essonne, Département de Mathématiques, Université d'Evry, rue du Père Jarlan,
Evry Cedex 91025, France
e-mail: monique.jeanblanc@univ-evry.fr

J.-P. Laurent
Université Paris 1 Panthéon-Sorbonne, 17, rue de la Sorbonne, 7005 Paris, France
e-mail: laurent.jeanpaul@free.fr

A.R. Carmona et al. (eds.), *Paris-Princeton Lectures on Mathematical Finance 2010*,
Lecture Notes in Mathematics 2003, DOI 10.1007/978-3-642-14660-2_1,
© Springer-Verlag Berlin Heidelberg 2011

to actually traded credit default swaps, for instance, with the same maturity as the CDO tranche. Let us note that no Markovian assumption is required for the existence of such a replicating strategy.

However, the practical implementation of actual hedging strategies requires some extra assumptions. We assume that all pre-default intensities are equal and only depend upon the current number of defaults. We also assume that all recovery rates are constant across names and time. In that framework, it can be shown that the aggregate loss process is a homogeneous Markov chain, more precisely a pure death process. Thanks to these restrictions, the model involves as many unknown parameters as the number of underlying names. Such Markovian model is also known as a local intensity model, the simplest form of aggregate loss models. As in local volatility models in the equity derivatives world, there is a perfect match of unknown parameters from a complete set of CDO tranches quotes. Numerical implementation can be achieved through a binomial tree, well-known to finance people, or by means of Markov chain techniques. We provide some examples and show that the market quotes of CDOs are associated with pronounced contagion effects. We can therefore explain the dynamics of the amount of hedging CDS and relate them to deltas computed by market practitioners. The figures are hopefully roughly the same, the discrepancies being mainly explained by contagion effects leading to an increase of dependence between default times after some defaults.

1 Introduction

The risk management and the hedging of credit derivatives and related products are topics of tremendous importance, especially given the recent credit turmoil. The risks at hand are usually split into different categories, which may sometimes overlap, such as credit spread and default risks, correlation and contagion risks.

The credit crisis also drove attention to counterparty risk and related issues such as collateral management, downgrading of guarantors and of course liquidity issues. For simplicity, these will not be dealt within this part.[1]

Credit derivatives are over-the-counter (OTC) financial instruments designed to transfer credit risk of a reference entity between two counterparties by way of a bilateral agreement. The agreement involves a seller of protection and a buyer of protection. The seller of protection is committed to cover the losses induced by the default of a reference entity, typically a corporate. In return, the buyer of protection has to pay at some fixed dates a premium to the seller of protection. By the default, we mean that the entity goes bankrupt or fails to pay a coupon on time, for some of its issued bonds. Even though credit derivatives are traded over-the-counter, credit events are standardized by the International Swap and Derivative Association (ISDA).[2]

[1] See [33] for a discussion of the issues involved.

[2] Although ISDA reports a list of six admissible credit events, most of the contracts only include bankruptcy and failure to pay as credit events. This is the case of contracts referencing companies settled in developed countries. The definitions have been last updated in 2003. An overview of these standardized definitions can be found in [54]. However, these are likely to be updated, for instance due to the ISDA big bang protocol.

Since credit derivatives involve some counterparty risk, the protection seller may be asked to post some collateral. Also, depending on the market value of the contract, the amount of collateral may be dynamically adjusted. However, after the recent credit crisis and subsequent defaults, settlement procedures had to be updated. Various projects including the ISDA, tend to standardize the cash-flows of credit default swaps (CDS), netting and settlement procedures. It is likely that some market features will change. Nevertheless, the main ideas expressed here will still be valid with some minor adaptation.

Financial institutions such as banks, mutual funds, pension funds, insurance and reinsurance companies, monoline insurance companies, corporations or sovereign wealth funds have a natural incentive to use credit derivatives in order to assume, reduce or manage credit exposures.

Surprisingly enough, since pricing at the cost of the hedge is the cornerstone of the derivatives modelling field, models that actually connect pricing and hedging issues for CDOs have been studied after the one factor Gaussian copula model became a pricing standard. This discrepancy with the equity or interest derivatives fields can actually be seen as a weakness and one can reasonably think that further researches in the credit area will aim at closing the gap between pricing and hedging.

Before proceeding further, let us recall the main features in a hedging and risk management problem, which come to light whatever the underlying risks:

- A first issue is related to the choice and the liquidity of the hedging instruments: typically, one could think of credit index default swaps, CDS on names with possibly different maturities, standardized synthetic single tranche CDOs and even other products such as equity put options, though this will not be detailed in this part. We reckon that the use of equity products to mitigate risks can be useful in the high yield market, but this is seemingly not the case for CDO tranches related to investment grade portfolios.
- A second issue is related to the products to be hedged. In the remainder, we will focus either on single name CDS or basket credit derivatives, such as First to Default Swaps, CDO tranches, bespoke CDOs or tranchelets. We will leave aside interest rate or foreign exchange hybrid products, credit spread options and exotic basket derivatives such as leveraged tranches, forward starting CDOs or tranche options.
- A third issue relies on the choice of the hedging method. The mainstream theoretical approach in mathematical finance favors the notion of replication of complex products through dynamic hedging strategies based on plain underlying instruments. However, it is clear that in many cases, risk can be mitigated by offsetting long and short positions, providing either a complete clearing or more usually leaving the dealer exposed to some basis albeit small risk. Moreover, such an approach is obviously quite robust to model risk. Unfortunately, there are some imbalances in customer demand and investment banks can be left with rather large outstanding positions on parts of the capital structure that must be managed up to maturity.

2 Hedging Instruments

This section is a primer about hedging of defaultable securities. It aims at presenting a general model of prices and hedging of defaultable claims, in a pure jump setting (there is no Brownian motion involved in our presentation). It also introduces the main hedging instruments we will consider throughout this part. We will particularly describe the cash-flows of CDS and derive the dynamics of their price. We also stress the impact of a credit event on the price dynamics of the surviving names.

2.1 Credit Default Swap

A CDS is a bilateral over-the-counter agreement which transfers the credit risk of a defined reference entity from a buyer of protection to a seller of protection up to a fixed maturity time T. The reference entity denoted **C** is typically a corporate or a sovereign obligor.

We assume that **C** may default at a particular time τ which is a non negative random variable constructed on a probability space $(\Omega, \mathcal{G}, \mathbb{P})$. The default time τ corresponds to a credit event leading to payment to the protection buyer. Moreover, if **C** defaults, only a fraction R (the recovery rate) of the initial investment is recovered. Figure 1 illustrates the structure of a CDS.

2.1.1 Cash-Flow Description

Let us consider a CDS initiated at time $t = 0$ with maturity T and nominal value E. The cash-flows of a CDS can be divided in two parts (or legs): the default leg which corresponds to the cash-flows generated by the seller of protection and the premium leg which is the set of cash-flows generated by the protection buyer. For simplicity, we will assume that nor the protection seller, neither the protection buyer can default.

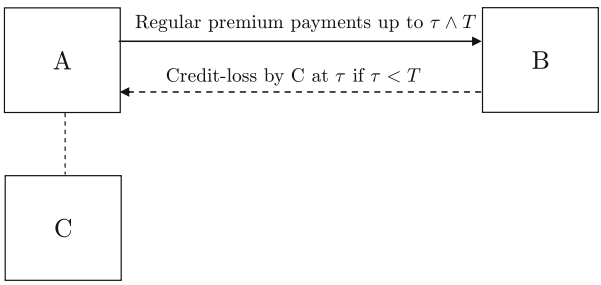


Fig. 1 Structure of a credit default swap

Default Leg

The seller of credit protection (denoted **B** in Fig. 1) agrees to cover losses induced by the default of the obligor **C** at time τ if the latter occurs before maturity ($\tau < T$). In that case, the payment is exactly equal to the fraction of the loss that is not recovered, i.e., the loss given default $E(1 - R)$. The settlement procedures in order to determine the recovery rate are not detailed here. The contract is worthless after the default of **C**.

Premium Leg or Fee Leg

In return, the buyer of protection (denoted **A** in Fig. 1) pays a periodic fee to **B** up to default time τ or until maturity T, whichever comes first. Each premium payment is proportional to a contractual credit spread[3] κ and to the nominal value E. More precisely, the protection buyer pays $\kappa \cdot \Delta_i \cdot E$ to the protection seller **B**, at every premium payment date $0 < T_1 < \cdots < T_p = T$ or until $\tau < T$, where $\Delta_i = T_i - T_{i-1}$, $i = 1, \ldots, p$ are the time intervals between two premium payment dates.[4] Let us remark that premium payments are made in arrears and begin at the end of the fist period (at T_1). If default happens between two premium payment dates, say $\tau \in]T_{i-1}, T_i[$, the protection fee has not been paid yet for the period $]T_{i-1}, \tau]$. In that case **A** will pay **B** an accrued premium equal to $\kappa \cdot (\tau - T_{i-1}) \cdot E$. The accrued premium payment is usually made at time τ. After default of C ($t > \tau$), there are no more cash-flows on the premium leg which is worthless.

It is noteworthy that the contractual spread κ is fixed at inception (at $t = 0$) and remains the same until maturity. It is determined so that the expected discounted cash-flows (under a proper pricing measure to be detailed below) between **A** and **B** are the same when the CDS contract is settled.

Due to the credit turmoil, some major market participants encourage a change in market convention for single name CDS quotes. In the proposal, the contractual spread will be fixed at $\kappa = 100$ bps or $\kappa = 500$ bps depending on the quality of the credit. The buyer of protection will have to make an immediate premium payment (upfront payment) to enter the contract (see [5] for more details).

2.2 Theoretical Framework

2.2.1 Default Times

In what follows, we consider n default times τ_i, $i = 1, \ldots, n$, that is, non-negative and finite random variables constructed on the same probability space $(\Omega, \mathcal{G}, \mathbb{P})$. For

[3] The contractual spread is quoted in basis points per annum.

[4] With the convention that $T_0 = 0$.

any $i = 1, \ldots, n$, we denote by $(N_t^i = \mathbb{1}_{\tau_i \leq t},\ t \geq 0)$ the ith default process, and by $\mathcal{H}_t^i = \sigma(N_s^i,\ s \leq t)$ the natural filtration of N^i (after completion and regularization on right). We introduce $\mathbb{H}$, the filtration generated by the processes $N^i, i = 1, \ldots, n$, defined as $\mathbb{H} = \mathbb{H}^1 \vee \cdots \vee \mathbb{H}^n$, i.e., $\mathcal{H}_t = \vee_{i=1}^n \mathcal{H}_t^i$.

We denote by $\tau_{(1)}, \ldots, \tau_{(n)}$ the ordered default times.

Hypothesis 1. We assume that no simultaneous defaults can occur, i.e., $\mathbb{P}(\tau_i = \tau_j) = 0,\ \forall i \neq j$. This assumption is important with respect to the completeness of the market. As shown below, it allows to dynamically hedge credit derivatives referencing a pool of defaultable entities with n credit default swaps.[5]

Hypothesis 2. We assume that, for any $i = 1, \ldots, n$, there exists a non-negative $\mathbb{H}$-adapted process $(\alpha_t^{i,\mathbb{P}},\ t \geq 0)$ such that the process

$$M_t^{i,\mathbb{P}} := N_t^i - \int_0^t \alpha_s^{i,\mathbb{P}} ds \tag{1}$$

is a $(\mathbb{P}, \mathbb{H})$-martingale. The process $\alpha^{i,\mathbb{P}}$ is called the $(\mathbb{P}, \mathbb{H})$-intensity of τ_i (Note that the value of the intensity depends strongly of the underlying probability). This process vanishes after τ_i (otherwise, after τ_i, the martingale $M^{i,\mathbb{P}}$ would be continuous and strictly decreasing, which is impossible) and can be written $\alpha_t^{i,\mathbb{P}} = (1 - N_t^i)\widetilde{\alpha}_t^{i,\mathbb{P}}$ for some $\mathbb{H}^1 \vee \cdots \vee \mathbb{H}^{i-1} \vee \mathbb{H}^{i+1} \vee \cdots \vee \mathbb{H}^n$-adapted process $\widetilde{\alpha}^{i,\mathbb{P}}$ (see [6] for more details). In particular, for $n = 1$, the process $\widetilde{\alpha}^{1,\mathbb{P}}$ is deterministic. In terms of the process $\widetilde{\alpha}^{i,\mathbb{P}}$, one has

$$M_t^{i,\mathbb{P}} = N_t^i - \int_0^{t \wedge \tau_i} \alpha_s^{i,\mathbb{P}} ds = N_t^i - \int_0^t (1 - N_s^i)\widetilde{\alpha}_s^{i,\mathbb{P}} ds.$$

Comments. (a) Let us remark that the latter hypothesis is not as strong as it seems to be. Indeed, the process N^i is an increasing $\mathbb{H}$-adapted process, hence an $\mathbb{H}$-submartingale. The Doob–Meyer decomposition implies that there exists a unique increasing $\mathbb{H}$-predictable process Λ^i such that $(N_t^i - \Lambda_t^i,\ t \geq 0)$ is an $\mathbb{H}$-martingale. We do not enter into details here,[6] it's enough to know that a left-continuous adapted process is predictable. It is also well known that the process Λ^i is continuous if and only if τ_i is totally inaccessible.[7] Here, we restrict our attention to processes Λ^i which are absolutely continuous with respect to Lebesgue measure. (b) It will be important to keep in mind that the martingale $M^{i,\mathbb{P}}$ has only one jump of size 1 at time τ_i.

[5] In the general case where multiple defaults could occur, we have to consider possibly 2^n states, and we would require non standard credit default swaps with default payments conditionally on all sets of multiple defaults to hedge multiname credit derivatives.

[6] The reader is referred to [56] for the definition of a predictable process. A stopping time ϑ is predictable if there exists a sequence of stopping times ϑ_n such that $\vartheta_n < \vartheta$ and ϑ_n converges to ϑ as n goes to infinity.

[7] A stopping time τ is totally inaccessible if $\mathbb{P}(\tau = \vartheta) = 0$ for any predictable stopping time ϑ.

2.2.2 Market Assumptions

For the sake of simplicity, let us assume that instantaneous digital default swaps are traded on the names. An instantaneous digital credit default swap on name i traded at time t is a stylized bilateral agreement between a buyer and a seller of protection. More precisely, the protection buyer receives one monetary unit at time $t + dt$ if name i defaults between t and $t + dt$. If α_t^i denotes the contractual spread of this stylized CDS, the seller of protection receives in return a fee equal to $\alpha_t^i dt$ which is paid at time $t + dt$ by the buyer of protection. The *cash-flows* associated with a buy protection position on an instantaneous digital default swaps on name i traded at time t are summarized in Fig. 2.

Let us also remark that there is no charge at inception (at time t) to enter an instantaneous digital credit default swap trade. Then, its payoff is equal to $dN_t^i - \alpha_t^i dt$ at $t + dt$ where dN_t^i is the payment on the default leg and $\alpha_t^i dt$ is the (short term) premium on the default swap.

Hypothesis 3. We assume that contractual spreads $\alpha^1, \ldots, \alpha^n$ are adapted to the filtration $\mathbb{H}$ of default times. The natural filtration of default times can thus be seen as the relevant information on economic grounds.

Moreover, since the instantaneous digital credit default swap is worthless after default of name i, credit spreads must vanish after τ_i, i.e., $\alpha_t^i = 0$ on the set $\{t > \tau_i\}$.

Note that considering such instantaneous digital default swaps rather than actually traded credit default swaps is not a limitation of our purpose. This can rather be seen as a convenient choice of basis from a theoretical point of view.

For simplicity, we further assume that (continuously compounded) default-free interest rates are constant and equal to r. Given some initial investment V_0 and some $\mathbb{H}$-predictable bounded processes $\delta^1, \ldots, \delta^n$ associated with some self-financed trading strategy in instantaneous digital credit default swaps, we attain at time T the payoff:

$$V_0 e^{rT} + \sum_{i=1}^{n} \int_0^T \delta_s^i e^{r(T-s)} \left(dN_s^i - \alpha_s^i ds \right).$$

By definition, δ_s^i is the nominal amount of instantaneous digital credit default swap on name i held at time s. This induces a net cash-flow of $\delta_s^i \cdot \left(dN_s^i - \alpha_s^i ds \right)$ at time $s + ds$, which has to be invested in the default-free savings account up to time T.

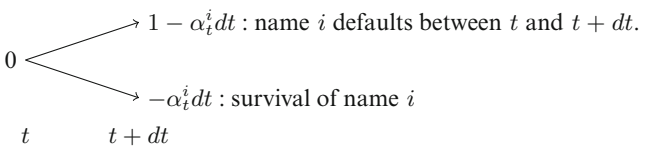


Fig. 2 *Cash-flows* of an instantaneous digital credit default swap (buy protection position)

2.2.3 Hedging and Martingale Representation Theorem

In our framework (we do not have any extra noise in our model, and the intensities do no depend on an exogenous factor), individual default intensities are not driven by a specific spread risk but by the arrival of new defaults: default intensities $\alpha^{i,\mathbb{P}}$, $i = 1, \ldots, n$ are deterministic functions of the past default times between two default dates. More precisely, as we shall prove later on, the intensity of τ_i on the set $\{t; \tau_{(j)} \leq t < \tau_{(j+1)}\}$ is a deterministic function of $\tau_{(1)}, \ldots, \tau_{(j)}$.

The main mathematical result of the study derives from the predictable representation theorem (see Theorem 9 in [10], Chap. III or [42]).

Theorem 1. *Let $A \in \mathcal{H}_T$ be a $\mathbb{P}$-integrable random variable. Then, there exists $\mathbb{H}$-predictable processes $\theta^i, i = 1, \ldots, n$ such that*

$$A = \mathbb{E}_{\mathbb{P}}[A] + \sum_{i=1}^{n} \int_0^T \theta_s^i \left(dN_s^i - \alpha_s^{i,\mathbb{P}} ds \right) = \mathbb{E}_{\mathbb{P}}[A] + \sum_{i=1}^{n} \int_0^T \theta_s^i dM_s^{i,\mathbb{P}}, \tag{2}$$

and $\mathbb{E}_{\mathbb{P}} \left(\int_0^T |\theta_s^i| \alpha_s^{i,\mathbb{P}} ds \right) < \infty$.

Proof. We do not enter into details. The idea is to prove that the set of random variables

$$Y = \exp \left(\sum_{i=1}^{n} \int_0^T \varphi_s^i dM_s^i - \int_0^T (e^{\varphi_s^i} - 1) \alpha_s^{i,\mathbb{P}} ds \right)$$

where φ^i are deterministic functions, is total in $L^2(\mathcal{H}_T)$ and to note that Y satisfies (2): indeed,

$$Y = 1 + \sum_{i=1}^{n} \int_0^T \varphi_s^i Y_{s-}^i dM_s^i \,.$$

Due to the integrability assumption on the r.v. A, and the predictable property of the θ's, the processes $\int_0^t \theta_s^i dM_s^i$, $i = 1, \ldots, n$ are $(\mathbb{P}, \mathbb{H})$-martingales. □

Let us remark that relation (2) implies that the predictable representation theorem (PRT) holds: any $(\mathbb{P}, \mathbb{H})$-martingale can be written in terms of the fundamental martingales $M^{i,\mathbb{P}}$. Indeed, if $M^{\mathbb{P}}$ is a $(\mathbb{P}, \mathbb{H})$-martingale, applying (2) to $A = M_T^{\mathbb{P}}$ and using the fact that $\int_0^t \theta_s^i dM_s^{i,\mathbb{P}}$ are martingales,

$$M_t^{\mathbb{P}} = \mathbb{E}_{\mathbb{P}} \left[M_T^{\mathbb{P}} \mid \mathcal{H}_t \right] = \mathbb{E}_{\mathbb{P}} \left[M_T^{\mathbb{P}} \right] + \sum_{i=1}^{n} \int_0^t \theta_s^i dM_s^{i,\mathbb{P}}. \tag{3}$$

From the PRT, any strictly positive $(\mathbb{P}, \mathbb{H})$-martingale ζ with expectation equal to 1 (as any Radon–Nikodym density) can be written as

$$d\zeta_t = \zeta_{t-} \sum_{i=1}^{n} \theta_t^i dM_t^{i,\mathbb{P}}, \ \zeta_0 = 1. \tag{4}$$

Indeed, as any martingale, ζ admits a representation as

$$d\zeta_t = \sum_{i=1}^{n} \hat{\theta}_t^i dM_t^{i,\mathbb{P}},\ \zeta_0 = 1$$

Since ζ is assumed to be strictly positive, introducing the predictable processes θ^i as $\theta_s^i = \frac{1}{\zeta_{s-}^i}\hat{\theta}_s^i$ allows to obtain the equality (4). We emphasize that the predictable property of θ is essential to guarantee that the processes $\int \theta_s dM_s^i$ are (local) martingales.

Conversely, the Doléans–Dade exponential, (unique) solution of

$$d\zeta_t = \zeta_{t-} \sum_{i=1}^{n} \theta_t^i dM_t^{i,\mathbb{P}},\ \zeta_0 = 1$$

is a (local) martingale. Note that, in order that ζ is indeed a non-negative local martingale, one needs that $\theta_t^i > -1$. Indeed, the solution of (4) is

$$\zeta_t = \exp\left(-\int_0^t \sum_{i=1}^{n} \theta_s^i \alpha_s^{i,\mathbb{P}} ds\right) \prod_{i=1}^{n} (1+\theta_{\tau_i}^i)^{N_t^i}.$$

The process ζ is a true martingale under some integrability conditions on θ (e.g., θ bounded) or if $\mathbb{E}^{\mathbb{P}}[\zeta_t] = 1$ for any t. Note that the jump of ζ at time $t = \tau_i$ is $\Delta\zeta_t = \zeta_t - \zeta_{t-} = \zeta_{t-}\theta_t^i$ (so that $\zeta_t = \zeta_{t-}(1+\theta_t^i)$ at time τ_i, hence the condition on θ to preserve non-negativity of ζ).

Theorem 2. *Let ζ satisfying (4) with $\theta_t^i > -1$ and $\mathbb{E}^{\mathbb{P}}[\zeta_t] = 1$, and define the probability measure $\mathbb{Q}$ as*

$$d\mathbb{Q}|_{\mathcal{H}_t} = \zeta_t d\mathbb{P}|_{\mathcal{H}_t}.$$

Then, the process

$$M_t^i := M_t^{i,\mathbb{P}} - \int_0^t \theta_s^i \alpha_s^{i,\mathbb{P}} ds = N_t^i - \int_0^t (1+\theta_s^i)\alpha_s^{i,\mathbb{P}} ds$$

is a $\mathbb{Q}$-martingale. In particular, the $(\mathbb{Q},\mathbb{H})$-intensity of τ_i is $\alpha_t^i = (1+\theta_t^i)\alpha_t^{i,\mathbb{P}}$.

Proof. The process M^i is an $(\mathbb{Q},\mathbb{H})$-martingale if and only if the process $M^i\zeta$ is a $(\mathbb{P},\mathbb{H})$-martingale. Using integration by parts formula

$$\begin{aligned}
d(M_t^i\zeta_t) &= M_{t-}^i d\zeta_t + \zeta_{t-} dM_t^i + \Delta M_t^i \Delta\zeta_t \\
&= M_{t-}^i d\zeta_t + \zeta_{t-} dM_t^{i,\mathbb{P}} - \zeta_{t-}\theta_t^i \alpha_t^{i,\mathbb{P}} dt + \zeta_{t-}\theta_t^i dN_t^i \\
&= M_{t-}^i d\zeta_t + \zeta_{t-} dM_t^{i,\mathbb{P}} - \zeta_{t-}\theta_t^i \alpha_t^{i,\mathbb{P}} dt + \zeta_{t-}\theta_t^i (dM_t^{i,\mathbb{P}} + \alpha^{i,\mathbb{P}} dt) \\
&= M_{t-}^i d\zeta_t + \zeta_{t-} dM_t^{i,\mathbb{P}} - \zeta_{t-}\theta_t^i \alpha_t^{i,\mathbb{P}} dt + \zeta_{t-}\theta_t^i dM_t^{i,\mathbb{P}} + \zeta_{t-}\theta_t^i \alpha_t^{i,\mathbb{P}} dt \\
&= M_{t-}^i d\zeta_t + \zeta_{t-}\left(1+\theta_t^i\right) dM_t^{i,\mathbb{P}}.
\end{aligned}$$

□

Note that the predictable processes θ^i's used to define the Radon–Nikodym density in (4) can be chosen such that the instantaneous credit default swap spreads $\alpha^1, \ldots, \alpha^n$ (defined in Sect. 2.2.2) are exactly the $(\mathbb{Q}, \mathbb{H})$-intensities associated with the default times. Let us recall that contractual spreads $\alpha^1, \ldots, \alpha^n$ are assumed to be adapted to the natural filtration $\mathbb{H}$ of default times. Moreover, from the absence of arbitrage opportunities, the cost of protection is positive if and only if a default risk exists. The latter argument implies that $\alpha^1, \ldots, \alpha^n$ are non negative $\mathbb{H}$-adapted processes and $\left\{\alpha_t^i > 0\right\} \overset{\mathbb{P}-a.s.}{=} \left\{\alpha_t^{i,\mathbb{P}} > 0\right\}$ for all time t and all name $i = 1, \ldots, n$. The processes θ^i, $i = 1, \ldots, n$ defined by

$$\theta_t^i = \left(\frac{\alpha_t^i}{\alpha_t^{i,\mathbb{P}}} - 1\right)(1 - N_{t-}^i),\ t \geq 0,\ i = 1, \ldots, n, \tag{5}$$

are therefore positive $\mathbb{H}$-predictable processes strictly greater than -1. They are admissible processes to define an equivalent change of probability measure. In the rest of the study, we will work under the probability $\mathbb{Q}$ obtained from $\mathbb{P}$ through the change of probability measure defined by (4) and (5).

It can be proved, using standard arguments that any $(\mathbb{Q}, \mathbb{H})$-martingale can be written as a sum of integrals with respect to M^i. Let M be a $(\mathbb{Q}, \mathbb{H})$-martingale. Then, there exists $\mathbb{H}$-predictable processes θ^i such that:

$$M_t = \mathbb{E}\left[M_T \mid \mathcal{H}_t\right] = \mathbb{E}\left[M_T\right] + \sum_{i=1}^{n} \int_0^t \theta_s^i dM_s^i, \tag{6}$$

where $\mathbb{E}$ is the expectation under $\mathbb{Q}$. Indeed, the process $(\widehat{M}_t := M_t\zeta_t, t \geq 0)$ being a $(\mathbb{P}, \mathbb{H})$-martingale admits a representation as $\widehat{M}_t = M_0 + \sum_{i=1}^n \int_0^t \widehat{\theta}_s^i dM_s^{i,\mathbb{P}}$. It suffices to apply integration by parts formula to $M_t = \widehat{M}_t(\zeta_t)^{-1}$ to obtain the result.

In particular, for $A \in \mathcal{H}_T$, one has

$$A = \mathbb{E}\left[A \mid \mathcal{H}_t\right] + \sum_{i=1}^{n} \int_t^T \theta_s^i dM_s^i. \tag{7}$$

Starting from time t, we can thus replicate the claim A with the initial investment $V_t = \mathbb{E}\left[Ae^{-r(T-t)} \mid \mathcal{H}_t\right]$ (in the savings account) and the trading strategy based on instantaneous digital credit default swaps defined by $\delta_s^i = \theta_s^i e^{-r(T-s)}$ for $t \leq s \leq T$ and $i = 1, \ldots, n$. As there is no initial charge to enter an instantaneous digital credit default swap, $V_t = \mathbb{E}\left[Ae^{-r(T-t)} \mid \mathcal{H}_t\right]$ corresponds to the time-t replication price of A. Since A depends upon the default indicators of the names up to time T, this encompasses the cases of multiname credit derivatives such as CDO tranches and basket default swaps, provided that recovery rates are deterministic.

We can also remark that for a small time interval dt,

$$V_{t+dt} \approx V_t(1 + rdt) + \sum_{i=1}^{n} \delta_t^i \left(dN_t^i - \alpha_t^i dt\right) \tag{8}$$

which is consistent with market practice and regular rebalancing of the replicating portfolio. An investor who wants to be compensated at time t against the price fluctuations of A during a small period dt has to invest V_t in the risk-free asset and take positions $\delta^1, \ldots, \delta^n$ in the n instantaneous digital credit default swaps.

Thanks to the PRT, it is also possible to describe the dynamics of a traditional credit default swap in terms of the dynamics of instantaneous credit default swaps. In the rest of this section we propose to build a general model of default times from the risk-neutral probability $\mathbb{Q}$ under which any defaultable claim can be replicated using instantaneous credit default swaps. There are various ways to construct such models. One of them, may be the most general in the case of non common defaults, is to start with the joint law of default times, and to make some regularity assumptions on that law (more precisely, that $G(t_1, \ldots, t_n) := \mathbb{Q}(\tau_1 > t_1, \ldots, \tau_i > t_i, \ldots, \tau_n > t_n)$ is n-time differentiable with respect to $(t_1, \ldots, t_n)$ and such that G and its derivatives do not vanish). We shall present this approach below, which is closely related to the well-known copula approach. Another way, more tractable but less general, is to specify the form of the intensities (in a Markov setting), and to construct the default times from these intensities. This approach will be presented in the following section and may be connected in some cases to the Markov chain used as a first step in a class of top-down models. A third method is to construct the random times as the first passage times at a random level for an increasing process. This last method is interesting for simulation, and allows correlation between the default times, via correlation of the random levels (see [59]). These three approaches both allow to derive the individual CDS spread dynamics as well as the dynamics of the portfolio loss, which will be needed for the pricing and hedging of CDO tranches.

In a first part, we shall present computations in the case $n = 1$. Then, we shall study the case $n = 2$.

2.3 The Single Default Case

We study the case $n = 1$. Here, τ is a non-negative random variable on the probability space $(\Omega, \mathcal{G}, \mathbb{Q})$ with risk-neutral survival function

$$G(t) := \mathbb{Q}(\tau > t) = 1 - \mathbb{Q}(\tau \leq t) = 1 - F(t)$$

where F is the cumulative distribution function of τ, under $\mathbb{Q}$. We assume that $G(t) > 0, \forall t$, and that G is continuous. Here $\mathbb{H} = \mathbb{H}^1$.

2.3.1 Some Important Martingales

Lemma 1. *For any (integrable) random variable X*

$$\mathbb{E}(X|\mathcal{H}_t)\mathbb{1}_{t<\tau} = \mathbb{1}_{t<\tau}\frac{1}{G(t)}\mathbb{E}(X\mathbb{1}_{t<\tau}) \tag{9}$$

and for any Borelian (bounded) function h

$$\mathbb{E}(h(\tau)|\mathcal{H}_t) = \mathbb{1}_{\tau \le t} h(\tau) - \mathbb{1}_{t<\tau} \frac{1}{G(t)} \int_t^\infty h(u) dG(u).$$

Proof. This well known result is established in a more general setting in [21]. We give here a proof for completeness. For fixed t, the σ-algebra $\mathcal{H}_t$ being generated by the random variable $\tau \wedge t$, any $\mathcal{H}_t$-measurable random variable can be written as $h(\tau \wedge t)$ where h is a bounded Borel function. It is then obvious that, on the set $\{t < \tau\}$, any $\mathcal{H}_t$-measurable random variable is deterministic. Hence, there exists a constant k such that $\mathbb{E}(X|\mathcal{H}_t)\mathbb{1}_{t<\tau} = k\mathbb{1}_{t<\tau}$. Taking expectation of both members leads to $k = \frac{1}{G(t)}\mathbb{E}(X\mathbb{1}_{t<\tau})$. The second formula follows from

$$\mathbb{E}(h(\tau)|\mathcal{H}_t) = h(\tau)\mathbb{1}_{\tau \le t} + \mathbb{1}_{t<\tau} \frac{\mathbb{E}(h(\tau)\mathbb{1}_{t<\tau})}{G(t)}$$

where we have used (9). The result is obtained with a computation of the last expectation. Note that the minus sign in front of the integral w.r.t. dG is due to the fact that G is decreasing. □

We now assume that G is differentiable (i.e., that τ admits a density f, so that $G'(t) = -f(t)$) (see [6] for the general case).

Proposition 1. *The process* $(M_t, t \ge 0)$ *defined as*

$$M_t = N_t - \int_0^{\tau \wedge t} \frac{f(s)}{G(s)} ds = N_t - \int_0^t (1 - N_s) \frac{f(s)}{G(s)} ds$$

is an $\mathbb{H}$*-martingale. In other terms, the intensity of* τ *is* $(1 - N_t)\widetilde{\alpha}(t)$ *where* $\widetilde{\alpha}$ *is the deterministic function* $\widetilde{\alpha}(t) = \frac{f(t)}{G(t)}$.

Proof. Let $s < t$. Then, from (9),

$$\mathbb{E}(N_t - N_s|\mathcal{H}_s) = \mathbb{1}_{\{s<\tau\}}\mathbb{E}(\mathbb{1}_{\{s<\tau\le t\}}|\mathcal{H}_s) = \mathbb{1}_{\{s<\tau\}} \frac{F(t) - F(s)}{G(s)}\,. \tag{10}$$

On the other hand, the quantity

$$C := \mathbb{E}\left[\int_s^t (1 - N_u) \frac{f(u)}{G(u)} du \,\middle|\, \mathcal{H}_s \right]$$

is equal to

$$\begin{aligned} C &= \int_s^t \frac{f(u)}{G(u)} \mathbb{E}\left[\mathbb{1}_{\{\tau>u\}}\middle|\mathcal{H}_s\right] du = \mathbb{1}_{\{\tau>s\}} \int_s^t \frac{f(u)}{G(u)} \frac{G(u)}{G(s)} du \\ &= \mathbb{1}_{\{\tau>s\}} \frac{F(t) - F(s)}{G(s)} \end{aligned}$$

which, from (10), proves $\mathbb{E}(M_t - M_s|\mathcal{H}_s) = 0$, hence the desired result. □

One should not confuse the intensity α and $\widetilde{\alpha}$, called the predefault-intensity. The intensity α is stochastic, and vanishes after τ, the predefault intensity is deterministic. The survival function G can be expressed in terms of the predefault intensity $\widetilde{\alpha}$. Indeed, we have proved in Proposition 1 that

$$\widetilde{\alpha}(t) = \frac{f(t)}{G(t)} = -\frac{G'(t)}{G(t)}.$$

Solving this ODE with initial condition $G(0) = 1$ leads to

$$G(t) = \mathbb{Q}(\tau > t) = \exp\left(-\int_0^t \widetilde{\alpha}(u)\, du\right).$$

Note that $\widetilde{\alpha}(t)dt = \mathbb{Q}(\tau \in dt | \tau > t)$: this is the probability that τ occurs in the interval $[t, t + dt]$ knowing that τ has not yet occurred.

2.3.2 CDS Market Value

For the sake of notational simplicity, we assume in this section that the interest rate r is null, so that the price of a savings account is $B_t = 1$ for every t. We moreover assume that the contractual spread κ is paid in continuous time (i.e., during the time interval $[t, t + dt]$ the amount κdt is paid by the protection buyer to the protection seller). We also consider that the payment at default time is a deterministic function of the default time, i.e., $\xi(\tau)$, which allows to deal with time dependent recovery rates.[8] Let us remark that the results described below can be easily extended to the case of a constant interest rate r or if cash-flows on the premium leg are more realistic.

We saw in Sect. 2.1 that the cash-flows of a CDS could be divided in two legs: the default leg and the premium leg. The time-t market value of a buy protection position on a CDS is equal to:

$$V_t(\kappa) = D_t - \kappa \cdot P_t, \tag{11}$$

where D_t is the time-t present value of the default leg and P_t is the time-t present value of the premium leg per unit of κ. This corresponds to the amount a buyer of protection is willing to pay (or gain) in order to close his position at time t. Let us recall that the contractual spread κ is such that the CDS market value is equal to zero at inception ($V_0(\kappa) = 0$).

[8] $\xi(\tau)$ is equal to the loss given default associated with the reference entity times the notional of the CDS.

We first focus on price dynamics of a CDS with spread κ initiated at time 0. The time-t market price of a CDS maturing at T with contractual spread κ is then given by the formula

$$V_t(\kappa) = \mathbb{E}\Big(\xi(\tau)\mathbb{1}_{\{t<\tau\leq T\}} - \mathbb{1}_{\{t<\tau\}}\kappa\big((\tau\wedge T)-t\big)\,\Big|\,\mathcal{H}_t\Big). \tag{12}$$

Proposition 2. *The price at time $t\in[0,T]$ of a credit default swap with spread κ is*

$$V_t(\kappa) = \mathbb{1}_{\{t<\tau\}}\widetilde{V}_t(\kappa), \quad \forall\, t\in[0,T],$$

where $\widetilde{V}_t(\kappa)$, a deterministic function, stands for the pre-default value of a CDS and equals

$$\widetilde{V}_t(\kappa) = \frac{1}{G(t)}\left(-\int_t^T \xi(u)\,dG(u) - \kappa\int_t^T G(u)\,du\right).$$

Proof. From Lemma 1, we have, on the set $\{t<\tau\}$,

$$\begin{aligned} V_t(\kappa) &= -\frac{\int_t^T \xi(u)\,dG(u)}{G(t)} - \kappa\left(\frac{-\int_t^T u\,dG(u) + TG(T)}{G(t)} - t\right) \\ &= \frac{1}{G(t)}\left(-\int_t^T \xi(u)\,dG(u) - \kappa\Big(TG(T) - tG(t) - \int_t^T u\,dG(u)\Big)\right). \end{aligned}$$

where, in the last equality, we have used an integration by parts to obtain

$$\int_t^T G(u)\,du = TG(T) - tG(t) - \int_t^T u\,dG(u)$$

□

2.3.3 CDS Market Spreads

Like traditional interest-rate swaps, CDS quotations are based on spreads, though this is likely to be modified after the ISDA big bang protocol. Quoted spreads will be after that only a way to express upfront premiums. Let us consider a CDS initiated at time 0 with maturity T and contractual spread κ. The time-t market spread is defined as the contractual spread of the contract if it would have been initiated at time t. In other words, this is the level of the spread $\kappa = \kappa(t,T)$ that makes a T-maturity CDS worthless at time t. A CDS market spread at time t is thus determined by the equation $V_t(\kappa(t,T)) = 0$ where V_t is defined in Proposition 2.

The T-maturity market spread $\kappa(t,T)$ is therefore a solution to the equation

$$\int_t^T \xi(u)\,dG(u) + \kappa(t,T)\int_t^T G(u)\,du = 0,$$

and thus for every $t \in [0,T]$,

$$\kappa(t,T) = -\frac{\int_t^T \xi(u)\, dG(u)}{\int_t^T G(u)\, du}. \tag{13}$$

There exists liquidly quoted CDS spreads on most big companies, and standard maturities are $T = 3, 5, 7, 10$ years. Given (13), it is possible to extract a market-implied survival distribution $G(t) = \mathbb{Q}(\tau > t)$, $t \geq 0$ from the term structure of CDS market spreads. See Chap. 3 of [59] or Chaps. 2 and 3 of [19] for more details.

In what follows, we fix the maturity date T, and we write briefly $\kappa(t)$ instead of $\kappa(t,T)$. There is a simple relationship between credit spreads and market values. The market price of a CDS with payment ξ at default, maturity T and contractual spread κ equals, for every $t \in [0,T]$,

$$V_t(\kappa) = \mathbb{1}_{\{t<\tau\}}\, (\kappa(t) - \kappa)\, \frac{\int_t^T G(u)\, du}{G(t)},$$

or more explicitly,

$$V_t(\kappa) = \mathbb{1}_{\{t<\tau\}}\, \frac{\int_t^T G(u)\, du}{G(t)} \left(\frac{\int_0^T \xi(u)\, dG(u)}{\int_0^T G(u)\, du} - \frac{\int_t^T \xi(u)\, dG(u)}{\int_t^T G(u)\, du} \right).$$

The latter expression simply means that the value of a CDS contract for a buyer of protection is positive when the current market spread $\kappa(t)$ is greater than the contractual spread κ.

2.3.4 Dynamics of CDS Prices in a Single Default Setting

Proposition 3. *The dynamics of the (ex-dividend) price $V_t(\kappa)$ on $[0,T]$ are*

$$dV_t(\kappa) = -V_{t-}(\kappa)\, dM_t + (1 - N_t)(\kappa - \xi(t)\widetilde{\alpha}(t))\, dt,$$

where the $(\mathbb{Q}, \mathbb{H})$-martingale M is given in Proposition 1.

Proof. It suffices to recall that

$$V_t(\kappa) = (1 - N_t)\widetilde{V}_t(\kappa)$$

with $\widetilde{V}$ given in Proposition 2, so that, using integration by parts formula,

$$dV_t(\kappa) = (1 - N_t)\, d\widetilde{V}_t(\kappa) - \widetilde{V}_{t-}(\kappa)\, dN_t.$$

Using the explicit expression of $\widetilde{V}_t(\kappa)$, we find easily that we have

$$d\widetilde{V}_t(\kappa) = \widetilde{\alpha}(t)\widetilde{V}_t(\kappa)\,dt + (\kappa - \xi(t)\widetilde{\alpha}(t))\,dt.$$

The SDE for $V(\kappa)$ follows. □

It is worthwhile to note that the price dynamics is not a martingale under the risk-neutral probability, despite the fact that the interest rate is null. This is because we are dealing with the ex-dividend price. The premium κ is similar to a dividend to be paid, hence the quantity $\kappa(1 - N_t)dt$ appears. The quantity $\xi(t)$ can be interpreted as a dividend to be received, at time t, with probability $\widetilde{\alpha}(t)dt$. At default time, the price jumps from $V_{\tau-}(\kappa)$ to 0, as can be seen in the right-hand side of the dynamics.

2.4 Two Default Times

Let us now study the case with two random times τ_1, τ_2. We denote by $(N_t^i,\ t \geq 0)$ the default process associated with τ_i, $i = 1, 2$. The filtration generated by the process N^i is denoted $\mathbb{H}^i$ and the filtration generated by the two processes N^1, N^2 is $\mathbb{H} = \mathbb{H}^1 \vee \mathbb{H}^2$.

Note that an $\mathcal{H}_t^1 \vee \mathcal{H}_t^2$-measurable random variable is

- A constant on the set $t < \tau_1 \wedge \tau_2$,
- A $\sigma(\tau_1 \wedge \tau_2)$-measurable random variable on the set $\tau_1 \wedge \tau_2 \leq t < \tau_1 \vee \tau_2$, i.e., a $\sigma(\tau_1)$-measurable random variable on the set $\tau_1 \leq t < \tau_2$, and a $\sigma(\tau_2)$-measurable random variable on the set $\tau_2 \leq t < \tau_1$. We recall that a $\sigma(\tau_1)$-measurable random variable is a Borel function of τ_1.
- A $\sigma(\tau_1, \tau_2)$-measurable random variable (i.e., a Borel function $h(\tau_1, \tau_2)$) on the set $\tau_1 \vee \tau_2 \leq t$.

To summarize, for fixed t, any $\mathcal{H}_t^1 \vee \mathcal{H}_t^2$-measurable random variable Z admits a representation as

$$Z = h\mathbb{1}_{t<\tau_1\wedge\tau_2} + h_1(\tau_1)\mathbb{1}_{\tau_1\leq t<\tau_2} + h_2(\tau_2)\mathbb{1}_{\tau_2\leq t<\tau_1} + h(\tau_1,\tau_2)\mathbb{1}_{\tau_1\vee\tau_2\leq t}\,.$$

We denote by $G(t, s) = \mathbb{Q}(\tau_1 > t, \tau_2 > s)$ the survival probability of the pair (τ_1, τ_2) and we assume that this function is twice differentiable. We denote by $\partial_i G$, the partial derivative of G with respect to the ith variable, $i = 1, 2$. The density of the pair (τ_1, τ_2) is denoted by f. Simultaneous defaults are precluded in this framework, i.e., $\mathbb{Q}(\tau_1 = \tau_2) = 0$.

Even if the case of two default times is more involved, closed form expressions for the intensities are available. It is important to take into account that the choice of the filtration is very important. Indeed, in general, an $\mathbb{H}^1$-martingale is not an $\mathbb{H}^1 \vee \mathbb{H}^2$-martingale. We shall illustrate this important fact below.

2.4.1 Intensities

We present the computation of martingales associated with default times τ_i, $i = 1, 2$, in different filtrations. In particular, we shall obtain the computation of the intensities in various filtrations.

- **Filtration** $\mathbb{H}^i$: We study, for any fixed i, the Doob–Meyer decomposition of the submartingale N^i in the filtration $\mathbb{H}^i$. From Proposition 1, the process

$$N_t^i - \int_0^{t\wedge\tau_i} \frac{f_i(s)}{G_i(s)} ds \tag{14}$$

  is an $\mathbb{H}^i$-martingale. Here, $1 - G_i(s) = F_i(s) = \mathbb{Q}(\tau_i \leq s) = \int_0^s f_i(u)du$. In other terms, the process $(1 - N_t^i)\frac{f_i(t)}{G_i(t)}$ is the $\mathbb{H}^i$-intensity of τ^i.
- **Filtration** $\mathbb{H}$: We recall a general result which allows to compute the intensities of a default time (see [27]).

Lemma 2. *Let $\mathbb{G} = \mathbb{F}\vee\mathbb{H}$, where $\mathbb{F}$ is a reference filtration and $\mathcal{H}_t = \sigma(\tau\wedge t)$ where τ is a random time. Assume that the supermartingale $G_t := \mathbb{P}(\tau > t|\mathcal{F}_t)$ admits the Doob–Meyer decomposition $G_t = Z_t - A_t$ where Z is a martingale and A is a predictable increasing process absolutely continuous with respect to Lebesgue measure. Then*

$$M_t = N_t - \int_0^{t\wedge\tau} \frac{dA_s}{G_s}$$

is a $\mathbb{G}$-martingale.

Proof. The proof relies on the computation of $\mathbb{E}(M_t - M_s|\mathcal{H}_s)$ for $t > s$. See [27] for details. □

In order to find the intensity of τ_1 in a general two defaults setting, we apply the previous lemma to the case $\mathbb{F} = \mathbb{H}^2$ and $\mathbb{H} = \mathbb{H}^1$. The first step is to compute the associated supermartingale (under the risk-neutral probability $\mathbb{Q}$).

Lemma 3. *The $\mathbb{H}^2$- supermartingale $\mathbb{Q}(\tau_1 > t|\mathcal{H}_t^2)$ equals*

$$G_t^{1|2} := \mathbb{Q}(\tau_1 > t|\mathcal{H}_t^2) = N_t^2 h(t, \tau_2) + (1 - N_t^2)\psi(t) \tag{15}$$

where $\psi(t) = G(t, t)/G(0, t)$, and $h(t, v) = \frac{\partial_2 G(t,v)}{\partial_2 G(0,v)}$.

Proof. From Proposition 1,

$$\mathbb{Q}(\tau_1 > t|\mathcal{H}_t^2) = \mathbb{1}_{t<\tau_2} \frac{\mathbb{Q}(\tau_1 > t, \tau_2 > t)}{\mathbb{Q}(\tau_2 > t)} + \mathbb{1}_{\tau_2\leq t}\mathbb{Q}(\tau_1 > t|\tau_2).$$

It is easy to check that

$$\mathbb{Q}(\tau_1 > t|\tau_2 = u) = \frac{\mathbb{Q}(\tau_1 > t, \tau_2 \in du)}{\mathbb{Q}(\tau_2 \in du)} = h(t,u)$$

and the result follows. □

Proposition 4. *Let*

$$a_t = N_t^2 \partial_1 h(t,\tau_2) + (1-N_t^2)\frac{\partial_1 G(t,t)}{G(0,t)}\,.$$

The process M^1 defined as

$$\begin{aligned} M_t^1 &:= N_t^1 + \int_0^{t\wedge\tau_1} \frac{a_s}{G_s^{1|2}} ds \\ &= N_t^1 + \int_0^{t\wedge\tau_1\wedge\tau_2} \frac{\partial_1 G(s,s)}{G(s,s)} ds + \int_{t\wedge\tau_1\wedge\tau_2}^{t\wedge\tau_1} \frac{\partial_{1,2} G(s,\tau_2)}{\partial_2 G(s,\tau_2)} ds \end{aligned}$$

is an $\mathbb{H}$-martingale.

Proof. The proof relies on some Itô's calculus to obtain the Doob–Meyer decomposition of $\mathbb{Q}(\tau_1 > t|\mathcal{H}_t^2)$ and to prove that $dA_t = a_t dt$. We refer the reader to [8] for details. □

This means that the $\mathbb{H}$-intensity of τ_1 takes into account the knowledge of τ_2 and is equal to the deterministic function $-\frac{\partial_1 G(t,t)}{G(t,t)}$ on the set $t < \tau_2$ and to the random quantity $\varphi(t,\tau_2)$ where $\varphi(t,s) = -\frac{\partial_{1,2} G(t,s)}{\partial_2 G(t,s)}$ on the set $t \geq \tau_2$. In a closed form, the processes $N_t^i - \int_0^t \alpha_s^i ds$, $i = 1,2$, are martingales in the same filtration $\mathbb{H}$, where

$$\begin{aligned} \alpha_t^1 &= (1-N_t^1)\left((1-N_t^2)\frac{-\partial_1 G(t,t)}{G(t,t)} - N_t^2 \frac{\partial_{1,2} G(t,\tau_2)}{\partial_2 G(t,\tau_2)}\right) \\ &= (1-N_t^1)(1-N_t^2)\widetilde{\alpha}^1(t) + (1-N_t^1)N_t^2 \widetilde{\alpha}^{1|2}(t,\tau_2) \\ \alpha_t^2 &= (1-N_t^2)\left((1-N_t^1)\frac{-\partial_2 G(t,t)}{G(t,t)} - N_t^1 \frac{\partial_{1,2} G(\tau_1,t)}{\partial_1 G(\tau_1,t)}\right) \\ &= (1-N_t^1)(1-N_t^2)\widetilde{\alpha}^2(t) + N_t^1(1-N_t^2)\widetilde{\alpha}^{2|1}(\tau_1,t) \end{aligned}$$

where

$$\widetilde{\alpha}^i(t) = -\frac{\partial_i G(t,t)}{G(t,t)} \tag{16}$$

$$\widetilde{\alpha}^{1|2}(t,s) = -\frac{\partial_{1,2} G(t,s)}{\partial_2 G(t,s)}, \quad \widetilde{\alpha}^{2|1}(s,t) = -\frac{\partial_{1,2} G(s,t)}{\partial_1 G(s,t)}. \tag{17}$$

Note that the minus signs in the value of the intensity are due to the fact that G is decreasing with respect to its component, hence the first order derivatives are non-positive and the second order derivative $\partial_1\partial_2 G$ – equal to the density of the pair (τ_1, τ_2) – is non-negative. The quantity $\widetilde{\alpha}^1(t)dt$ is equal to $\mathbb{Q}(\tau_1 \in dt | \tau_1 \wedge \tau_2 > t)$. The quantity $\widetilde{\alpha}^{1|2}(t,s) = -\frac{f(t,s)}{\partial_2 G(t,s)}$ evaluated at $s = \tau_2$, represents the value of the predefault intensity process of τ_1 with respect to the filtration $\mathbb{H}$ on the event $\{\tau_2 < t\}$.

Let us remark that, in the particular case where τ_1 and τ_2 are independent (or if $\tau_1 < \tau_2$), the $\mathbb{H}$ intensity of τ_1 equals its $\mathbb{H}^1$ intensity.

This model is very general. Let us note that it is not a Markov model, except if $h(t,s)$ does not depend on s (see [6] for a formal proof). Moreover, it can be extended at the price of notational complexity to n names but computations are not so easy, since they involve partial derivatives of the joint survival function and do not usually lead to tractable Markov processes.

Since we are working in the same filtration[9] the compensated martingale of the counting process $N_t = N_t^1 + N_t^2 = \sum_{i=1}^2 \mathbb{1}_{\tau_i \le t}$ is $M_t := N_t - \int_0^t \alpha_s ds$ where

$$\begin{aligned}\alpha_t &= \alpha_t^1 + \alpha_t^2 \\ &= (1 - N_t^1)(1 - N_t^2)\left(-\frac{\partial_1 G(t,t) + \partial_2 G(t,t)}{G(t,t)}\right) \\ &\quad -(1 - N_t^1)N_t^2 \frac{\partial_{1,2} G(t,\tau_2)}{\partial_2 G(t,\tau_2)} - (1 - N_t^2)N_t^1 \frac{\partial_{1,2} G(\tau_1,t)}{\partial_2 G(\tau_1,t)}.\end{aligned}$$

It is proved in Bielecki et al. [7] that the process N is Markov if and only if the quantities $\frac{\partial_{1,2} G(t,\tau_2)}{\partial_2 G(t,\tau_2)}$ and $\frac{\partial_{1,2} G(\tau_1,t)}{\partial_2 G(\tau_1,t)}$ are deterministic.

2.4.2 Dynamics of CDS Prices in a Two Defaults Setting

Let us now examine the valuation of a single-name CDS written on the default τ_1. Our aim is to show that the dynamics of this CDS will be affected by the information on τ_2: when τ_2 occurs, the intensity of τ_1 changes, and this will change the parameters of the price dynamics.

We consider a CDS

- With a continuously paid constant premium κ,
- Which delivers $\xi(\tau_1)$ at time τ_1 if $\tau_1 < T$, where ξ is a deterministic function. In the simplest case ξ is constant, corresponding to constant recovery rates. We recall that ξ corresponds to the loss given default times the nominal of the CDS.

[9] The sum of two martingales in the same filtration is a martingale.

The value of the CDS takes the form

$$V_t(\kappa) = \widetilde{V}_t(\kappa)\mathbb{1}_{t<\tau_2\wedge\tau_1} + \widehat{V}_t(\kappa)\mathbb{1}_{\tau_1\wedge\tau_2\leq t<\tau_1}.$$

First, we restrict our attention to the case $t < \tau_2 \wedge \tau_1$.

Proposition 5. *On the set $t < \tau_2 \wedge \tau_1$, the value of the CDS is*

$$\widetilde{V}_t(\kappa) = \frac{1}{G(t,t)}\left(-\int_t^T \xi(u)\partial_1 G(u,t)\,du - \kappa\int_t^T G(u,t)\,du\right).$$

Proof. The value $V(\kappa)$ of this CDS, computed in the filtration $\mathbb{H}$ including the information related to the second default, is

$$V_t(\kappa) = \mathbb{1}_{t<\tau_1}\mathbb{E}\left(\xi(\tau_1)\mathbb{1}_{\tau_1\leq T} - \kappa((T\wedge\tau_1)-t)|\mathcal{H}_t\right).$$

Let us denote by $\tau = \tau_1 \wedge \tau_2$ the first default time. Then, $\mathbb{1}_{t<\tau}V_t(\kappa) = \mathbb{1}_{t<\tau}\widetilde{V}_t(\kappa)$, where

$$\begin{aligned}
\widetilde{V}_t(\kappa) &= \frac{1}{\mathbb{Q}(\tau>t)}\mathbb{E}\left(\xi(\tau_1)\mathbb{1}_{\tau_1\leq T}\mathbb{1}_{t<\tau} - \kappa((T\wedge\tau_1)-t)\mathbb{1}_{t<\tau}\right)\\
&= \frac{1}{G(t,t)}\mathbb{E}\left(\xi(\tau_1)\mathbb{1}_{\tau_1\leq T}\mathbb{1}_{t<\tau} - \kappa((T\wedge\tau_1)-t)\mathbb{1}_{t<\tau}\right)\\
&= \frac{1}{G(t,t)}\left(\int_t^T \xi(u)\mathbb{Q}(\tau_1\in du,\tau_2>t)\right.\\
&\quad\left.-\kappa\int_t^T (u-t)\mathbb{Q}(\tau_1\in du,\tau_2>t) - (T-t)\kappa\int_T^\infty \mathbb{Q}(\tau_1\in du,\tau_2>t)\right).
\end{aligned}$$

In other terms, using integration by parts formula

$$\widetilde{V}_t(\kappa) = \frac{1}{G(t,t)}\left(-\int_t^T \xi(u)\partial_1 G(u,t)\,du - \kappa\int_t^T G(u,t)\,du\right).$$

□

On the event $\{\tau_2 \leq t < \tau_1\}$, the CDS price equals

$$\begin{aligned}
V_t(\kappa) = \hat{V}_t &= \mathbb{1}_{t<\tau_1}\mathbb{E}\left(\xi(\tau_1)\mathbb{1}_{\tau_1\leq T} - \kappa((T\wedge\tau_1)-t)|\sigma(\tau_2)\right)\\
&= \frac{1}{\partial_2 G(t,\tau_2)}\left(-\int_t^T \xi(u)f(u,\tau_2)\,du - \kappa\int_t^T \partial_2 G(u,\tau_2)\,du\right) := V_t^{1|2}(\tau_2)
\end{aligned}$$

where

$$V_t^{1|2}(s) = \frac{1}{\partial_2 G(t,s)} \left(- \int_t^T \xi(u) f(u,s)\, du - \kappa \int_t^T \partial_2 G(u,s)\, du \right).$$

In the financial interpretation, $V_t^{1|2}(s)$ is the market price at time t of a CDS on the first credit name, under the assumption that the default τ_2 occurred at time s and the first name has not yet defaulted (recall that simultaneous defaults are excluded, since we have assumed that G is differentiable).

Differentiating the deterministic function which gives the value of the CDS leads to the following result:

Proposition 6. *The price of a CDS is* $V_t(\kappa) = \widetilde{V}_t(\kappa)\mathbb{1}_{t<\tau_2\wedge\tau_1} + \hat{V}_t(\kappa)\mathbb{1}_{\tau_2\wedge\tau_1\le t<\tau_1}$. *The dynamics of* $\widetilde{V}(\kappa)$ *are*

$$d\widetilde{V}_t(\kappa) = \Big(\big(\widetilde{\alpha}_1(t) + \widetilde{\alpha}_2(t)\big)\widetilde{V}_t(\kappa) + \kappa - \widetilde{\alpha}_1(t)\xi(t) - \widetilde{\alpha}_2(t) V_t^{1|2}(t) \Big)\, dt,$$

where for $i = 1, 2$ *the function* $\widetilde{\alpha}_i(t)$ *is the (deterministic) pre-default intensity of* τ_i *given in (16). The dynamics of* $\widehat{V}(\kappa)$ *are*

$$d\widehat{V}_t(\kappa) = \Big(\widetilde{\alpha}^{1|2}(t,\tau_2) \Big(\widehat{V}_t(\kappa) - \xi(t) \Big) + \kappa \Big)\, dt$$

where $\widetilde{\alpha}^{1|2}(t,s)$ *is given in (17).*

Hence, differentiating $V_t = \widetilde{V}_t(1-N_t^1)(1-N_t^2) + \hat{V}_t(1-N_t^1)N_t^2$ one obtains

$$\begin{aligned} dV_t = {} & (1-N_t^1)(1-N_t^2)d\widetilde{V}_t + (1-N_t^1)N_t^2 d\widehat{V}_t - V_{t-}dN_t^1 \\ & + (1-N_t^1)(V_t^{1|2}(t) - \widetilde{V}_t)dN_t^2 \end{aligned}$$

which leads after light computations[10] to

$$\begin{aligned} dV_t &= (1-N_t^1)(1-N_t^2)(\kappa - \xi(t)\widetilde{\alpha}^1(t))dt + (1-N_t^1)N_t^2(\kappa - \xi(t)\widetilde{\alpha}^{1|2}(t,\tau_2))dt \\ &\quad - V_{t-}dM_t^1 + (1-N_t^1)(V_t^{1|2}(t) - \widetilde{V}_t)dM_t^2 \\ &= \text{dividend part} - V_{t-}dM_t^1 + (1-N_t^1)(V_t^{1|2}(t) - \widetilde{V}_t)dM_t^2\,. \end{aligned} \tag{18}$$

[10] From the definition, one has $dV_t = (1-N_t^1)(1-N_t^2)\cdots + (1-N_t^1)(\widehat{V}_t(\tau_2) - \widetilde{V}_t)dN_t^2$. It is important to note that $\widehat{V}_t(\tau_2)dN_t^2 = V_t^{1|2}(t)dN_t^2$: a computation using $\widehat{V}_t(\tau_2)dN_t^2 = \widehat{V}_t(\tau_2)(dM_t^2 + \ldots dt)$ would lead to a quantity $\widehat{V}_t(\tau_2)dM_t^2$ which has a meaning, but which is NOT a martingale, due to the lack of adapteness of the coefficient $\widehat{V}_t(\tau_2)$.

Assume now that a CDS written on τ_2 is also traded in the market. We denote by $V^i, i = 1, 2$ the prices of the two CDS. Since the CDS are paying premiums, a self financing strategy consisting in ϑ^i units of CDS's has value $X_t = \vartheta_t^1 V_t^1 + \vartheta_t^2 V_t^2$ and dynamics

$$\begin{aligned} dX_t &= \vartheta_t^1 \left(-V_{t-}^1 dM_t^1 + (1 - N_t^1)(V_t^{1|2}(t) - \widetilde{V}_t^1) dM_t^2 \right) \\ &\quad + \vartheta_t^2 \left(-V_{t-}^2 dM_t^2 + (1 - N_t^2)(V_t^{2|1}(t) - \widetilde{V}_t^2) dM_t^1 \right) \\ &= \left(-\vartheta_t^1 V_{t-}^1 + \vartheta_t^2 (1 - N_t^2)(V_t^{2|1}(t) - \widetilde{V}_t^2 \right) dM_t^1 \\ &\quad + \left(\vartheta_t^1 (1 - N_t^1)(V_t^{1|2}(t) - \widetilde{V}_t^1) - \vartheta_t^2 V_{t-}^2 \right) dM_t^2 \,. \end{aligned}$$

In order to duplicate a claim with value

$$A_t = \mathbb{E}(A) + \int_0^t \delta_t^1 dM_t^1 + \int_0^t \delta_t^2 dM_t^2$$

it remains to solve the linear system

$$\begin{cases} -\vartheta_t^1 V_{t-}^1 + \vartheta_t^2 (1 - N_t^2)(V_t^{2|1}(t) - \widetilde{V}_t^2) = \delta_t^1, \\ \vartheta_t^1 (1 - N_t^1)(V_t^{1|2}(t) - \widetilde{V}_t^1) - \vartheta_t^2 V_{t-}^2 \quad = \delta_t^2 \,. \end{cases}$$

Thus, under standard invertibility conditions, one can easily use actually traded CDS instead of instantaneous digital CDS when replicating the claim A.

3 Hedging Default Risks of CDOs in Markovian Contagion Models

When dealing with CDO tranches, the market approach to the derivation of credit default swap deltas consists in bumping the credit curves of the names and computing the ratios of changes in present value of the CDO tranches and the hedging credit default swaps. This involves a pricing engine for CDO tranches, usually some mixture of copula and base correlation approaches, leading to some "market delta."

The only rationale of this modus operandi is local hedging with respect to credit spread risks, provided that the trading books are marked-to-market with the same pricing engine. Even when dealing with small changes in credit spreads, there is no guarantee that this would lead to appropriate hedging strategies, especially to cover large spread widenings and possibly defaults. Also, one could think of changes in base correlation correlated with changes in credit spreads.

A number of CDO hedging anomalies in the base correlation approach are reported in [52]. Moreover, the standard approach is not associated with a replicating theory, thus inducing the possibility of unexplained drifts and time decay effects in the present value of hedged portfolios (see [55]).

Unfortunately, the trading desks cannot rely on a sound theory to determine replicating prices of CDO tranches. This is partly due to the dimensionality issue, partly to the stacking of credit spread and default risks. Laurent [45] considers the case of multivariate intensities in a conditionally independent framework and shows that for large portfolios where default risks are well diversified, one can concentrate on the hedging of credit spread risks and control the hedging errors. In this approach, the key assumption is the absence of contagion effects which implies that credit spreads of survival names do not jump at default times, or equivalently that defaults are not informative. Whether one should rely on this assumption is to be considered with caution, as discussed in the empirical studies [4] and [16].[11] Moreover, anecdotal evidence such as the failures of Delphi, Enron, Parmalat and WorldCom shows mixed results.

In this section, we adopt the framework of Laurent et al. [47], concentrating on default risks, credit spreads and dependence dynamics being driven by the arrival of defaults. We will calculate so-called "credit deltas," that are the present value impacts of some default event on a given CDO tranche, divided by the present value impact of the hedging instrument (here the underlying index) under the same scenario. Contagion models were introduced to the credit field by Davis and Lo [17], Jarrow and Yu [38] and further studied by Yu [65]. Schönbucher and Schubert [61] show that copula models exhibit some contagion effects and relate jumps of credit spreads at default times to the partial derivatives of the copula. This is also the framework used by Bielecki et al. [8] to address the hedging issue. We refer to Sect. 2.4 of this contribution for a detailed discussion of this topic. A similar but somehow more tractable approach has been considered by Frey and Backhaus [31], since the latter paper considers some Markovian models of contagion. In a copula model, the contagion effects are computed from the dependence structure of default times, while in contagion models the intensity dynamics are the inputs from which the dependence structure of default times is derived. In both approaches, credit spreads shifts occur only at default times. Thanks to this quite simplistic assumption, and provided that no simultaneous defaults occurs, it can be shown that the CDO market is complete, i.e., CDO tranche cash-flows can be fully replicated by dynamically trading individual credit default swaps or, in some cases, by trading the credit default swap index (see Sect. 2.2.3 of this contribution for a presentation of the theoretical ideas).

In this section we focus on the hedging of synthetic CDO tranches. For the section to be self-contained, we briefly describe in Sect. 3.1 the cash-flows of a

[11] The conclusions of this paper have been disputed by [44] in which the conditional independence assumption has not been rejected when tested on the same default database. These discrepancies are explained by an alternative specification of individual default intensities.

synthetic CDO tranche. While the use of the representation Theorem 1 guarantees that, in our framework, any basket default swap can be perfectly hedged with respect to default risks, it does not provide a practical way of constructing hedging strategies. In Sect. 3.2, we restrict ourselves to the case of homogeneous portfolios with Markovian intensities which results in a dramatic dimensionality reduction for the (risk-neutral) valuation of CDO tranches and the hedging of such tranches as well. We find out that the aggregate loss is associated with a pure birth process, which is now well documented in the credit literature. Section 3.3 provides an overview of the calibration methods proposed in the literature on contagion credit risk models. We investigate in particular a calibration method based on the marginal distributions of the number of defaults. Section 3.4 details the computation of replicating strategies of CDO tranches with respect to the credit default swap index, through a recombining tree on the aggregate loss. We discuss how hedging strategies are related to dependence assumptions in Gaussian copula and base correlation frameworks. We also compare the replicating strategies obtained in the contagion model with the hedging ratios (spread sensitivity ratios) provided by the Gaussian copula approach or computed in alternative credit risk models.

3.1 Synthetic CDO Tranches

Synthetic CDOs are structured products based on an underlying portfolio of reference entities subject to credit risk. It allows investors to sell protection on specific risky portion or tranche of the underlying credit portfolio depending on their desired risk-profile. A synthetic CDO structure is initially arranged by a financial institution (typically an investment bank) which holds a credit portfolio composed of CDS (see Fig. 3). This CDS portfolio is then transferred to a subsidiary company commonly called a special purpose vehicle (SPV). The SPV redistributes the credit risk of the underlying portfolio by raising specific credit-protection products corresponding to different levels of risk. The SPV liability side is defined by the different tranches that have been sold and the asset side corresponds to the portfolio of CDS. The incomes generated by the pool of CDS (premium payments) are re-allocated to the different tranches using a precise prioritization scheme. An investor (seller of protection) on a CDO tranche receives a higher premium if the tranche has a lower level of subordination. For example, the equity tranche which covers the fist losses on the underlying portfolio receives the highest income.

3.1.1 Credit Default Swap Indices

A credit default swap Index (CDS Index) is a multi-name credit derivative which allows market participants to buy and sell protection directly on a pool of CDS.

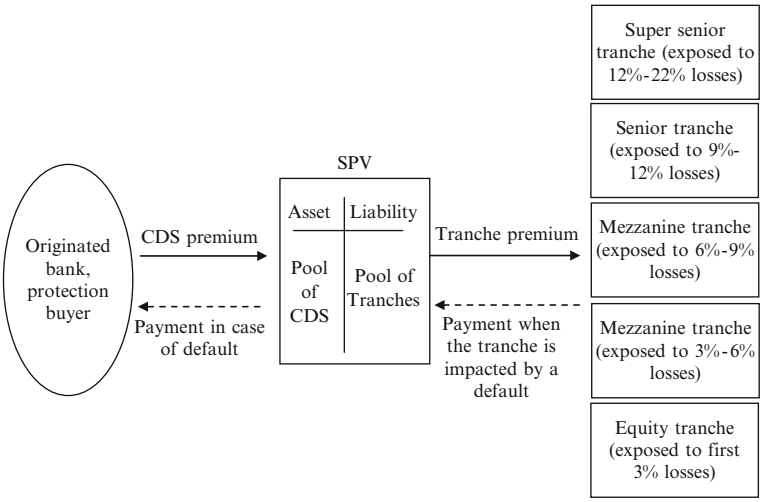


Fig. 3 Structure of a synthetic CDO

CDS indices are actively traded. This means that it can be easier to hedge a credit derivatives referencing a portfolio of CDS with an index than it would be to buy many CDS to achieve a similar effect. This is the reason why a popular use of CDS indices is to hedge multi-name credit positions.

There are currently two main families of CDS indices: CDX and iTraxx. CDX indices contain North American and Emerging Market companies and iTraxx contain companies from the rest of the world (mainly Europe and Asia). The iTraxx Europe Main and the CDX North America Main are the most liquid CDS indices. Each Main index includes 125 equally weighted CDS issuers from their respective region.[12] These issuers are investment grade at the time an index series is launched, with a new series launched every 6 months. In practice, "on the run" Main indices are mostly composed of A-rated and BBB-rated issuers.

3.1.2 Standardized CDO Tranches

Market-makers of these indices have also agreed to quote standard tranches on these portfolios from equity or first loss tranches to the most senior tranches.

Each tranche is defined by its attachment point which defines the level of subordination and its detachment point which defines the maximum loss of the underlying portfolio that would result in a full loss of tranche notional. The first-loss 0–3% equity tranche is exposed to the first several defaults in the underlying portfolio. This tranche is the riskiest as there is no benefit of subordination but it also offers

[12] The proportion of each issuer in the pool is equal to $1/125 = 0.08\%$.

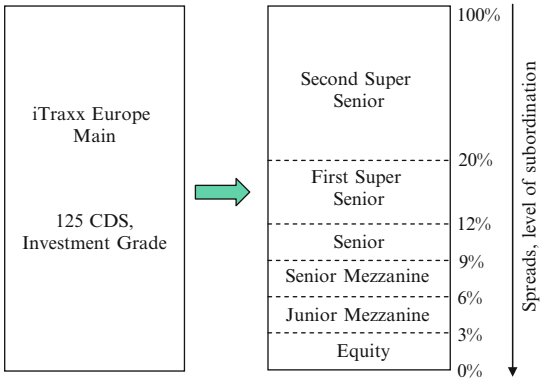


Fig. 4 Standardized CDO tranches on iTraxx Europe Main

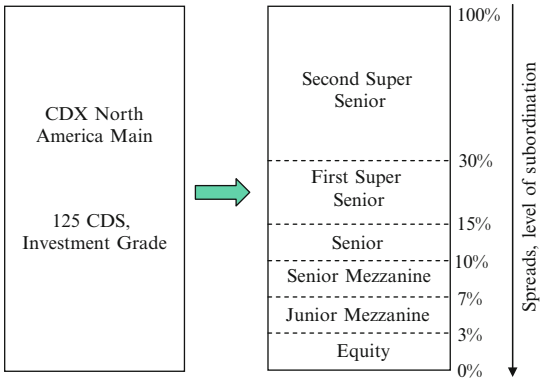


Fig. 5 Standardized CDO tranches on CDX North America Main

high returns if no defaults occur. The junior mezzanine 3–6% and the senior mezzanine 6–9% tranches are less immediately exposed to the portfolio defaults but the premium received by the protection seller is smaller than for the equity tranche. The 9–12% tranche is the senior tranche, while the 12–20% tranche is the low-risk super senior piece. As illustrated in Figs. 4 and 5, the tranching of the indices in Europe and North America is different. In North America, the CDX index is tranched into standard classes representing equity 0–3%, junior mezzanine 3–7%, senior mezzanine 7–10%, senior 10–15% and super senior 15–30% tranche.

For a detailed description of the credit derivatives market, the reader is referred to the textbooks [15, 19, 39, 49, 59]. Before addressing the hedging issue of CDO tranches, let us describe the cash-flows associated with these products.

3.1.3 Cash-Flows Description

We adopt the same notation as in Sect. 2 and we work under the risk-neutral probability $\mathbb{Q}$ defined in Sect. 2.2.3. We consider a portfolio of n credit references and we denote by $(\tau_1, \ldots, \tau_n)$ the random vector of default times defined on the probability space $(\Omega, \mathcal{G}, \mathbb{Q})$. If name i defaults, it drives a loss of $\xi_i = E_i\,(1 - R_i)$ where E_i denotes the nominal amount and R_i the recovery rate. The loss given default ξ_i is assumed here to be constant over time. The key quantity for the pricing of synthetic CDO tranches is the cumulative loss

$$L_t = \sum_{i=1}^{n} \xi_i N_t^i,$$

where $N_t^i = 1_{\{\tau_i \leq t\}}$, $i = 1, \ldots, n$ are the default indicator processes associated with default time τ_i, $i = 1, \ldots, n$. Let us recall that the processes N^i, $i = 1, \ldots, n$ are adapted to the global filtration $\mathbb{H} = (\mathcal{H}_t,\, t \geq 0)$ where $\mathcal{H}_t = \vee_{i=1}^n \mathcal{H}_t^i$ and $\mathcal{H}_t^i = \sigma(N_s^i,\, s \leq t)$. Let us remark at this stage that the loss process L is an increasing right-continuous pure jump process.

The cash-flows associated with a synthetic CDO tranche only depend upon the realized path of the cumulative loss process L. Default losses on the reference portfolio are split along some thresholds (attachment and detachment points) and allocated to the various tranches. A synthetic CDO tranche can be viewed as a bilateral contract between a protection seller and a protection buyer. In what follows, we consider a synthetic CDO tranche with attachment point a, detachment point b and maturity T and we describe the cash-flows associated with the default payment leg (payments received by the protection buyer) and the cash-flows associated with the premium payment leg (payments received by the protection seller).

Default Payments Leg

The protection seller agrees to pay the protection buyer default losses each time they impact the tranche (a, b) of the reference portfolio. More precisely, the cumulative default payment $L_t^{(a,b)}$ on the tranche $[a, b]$ is equal to zero if $L_t \leq a$, to $L_t - a$ if $a \leq L_t \leq b$ and to $b - a$ if $L_t \geq b$. Let us remark that $L_t^{(a,b)}$ has a call spread payoff with respect to L_t and can be expressed as $L_t^{(a,b)} = (L_t - a)^+ - (L_t - b)^+$. Default payments are simply equal to the increment of $L_t^{(a,b)}$, i.e., there is a payment of $L_t^{(a,b)} - L_{t-}^{(a,b)}$ from the protection seller at every time a jump of $L_t^{(a,b)}$ occurs before contract maturity T. Figure 6 shows a realized path of the loss process $(L_t, t \geq 0)$ and the corresponding path of losses affecting CDO tranche $[a, b]$.

For simplicity we assume that the continuously compounded default free interest rate r_t is deterministic and we denote $B_t(t') = \exp\left(-\int_t^{t'} r_s ds\right)$ the time-t

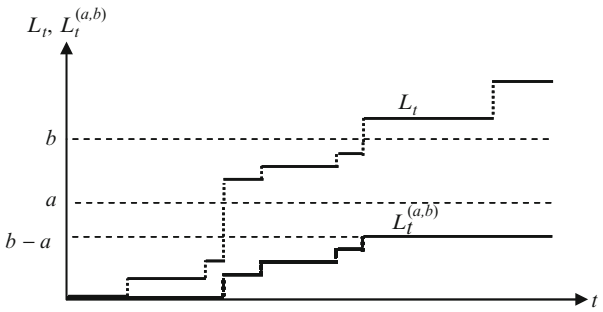


Fig. 6 A realized path of the reference portfolio losses and the corresponding path of losses affecting CDO tranche $[a, b]$. Jumps occur at default times

discount factor up to time t' ($t \leq t'$). At time t, the discounted payoff corresponding to default payments after time t can written as:

$$\int_t^T B_t(s) dL_s^{(a,b)} := \sum_{i=1}^{n} B_t(\tau_i) \left(L_{\tau_i}^{(a,b)} - L_{\tau_i-}^{(a,b)} \right) 1_{\{t<\tau_i \leq T\}}. \tag{19}$$

Thanks to Stieltjes integration by parts formula and Fubini theorem, the time-t price of the default payment leg under the risk-neutral measure can be expressed as:

$$\begin{aligned} D_t &= \mathbb{E}\left[\int_t^T B_t(s) dL_s^{(a,b)} \mid \mathcal{H}_t \right] \\ &= B_t(T)\mathbb{E}\left[L_T^{(a,b)} \mid \mathcal{H}_t \right] - L_t^{(a,b)} + \int_t^T r_s B_t(s) \mathbb{E}\left[L_s^{(a,b)} \mid \mathcal{H}_t \right] ds. \end{aligned}$$

Premium Payments Leg

The protection buyer has to pay the protection seller a periodic premium payment (quarterly for standardized indexes) based on a fixed contractual spread κ and proportional to the current outstanding nominal of the tranche $b - a - L_t^{(a,b)}$. Let us denote by $T_1 < \ldots < T_p$, the premium payment dates with $T_p = T$ and by Δ_i the length of the i-th period $[T_{i-1}, T_i]$ (in fractions of a year and with convention $T_0 = 0$). The CDO premium payments are equal to $\kappa \Delta_i \left(b - a - L_{T_i}^{(a,b)} \right)$ at regular payment dates T_i, $i = 1, \ldots, p$. Moreover, when a default occurs between two premium payment dates and when it affects the tranche, an additional payment (the accrued coupon) must be made at default time to compensate the change in value of the tranche outstanding nominal. For example, if name j defaults between T_{i-1}

and T_i, the associated accrued coupon is equal to $\kappa\left(\tau_j - T_{i-1}\right)\left(L_{\tau_j}^{(a,b)} - L_{\tau_j-}^{(a,b)}\right)$. Eventually, at time t, the discounted payoff corresponding to premium payments can be expressed as:

$$\sum_{i=p_t}^{p}\left(B_t(T_i)\kappa\Delta_i\left(b-a-L_{T_i}^{(a,b)}\right)+\int_{T_{i-1}}^{T_i}B_t(s)\kappa\left(s-T_{i-1}\right)dL_s^{(a,b)}\right), \quad (20)$$

where $p_t = \inf\{i = 1, \ldots, p \mid T_i > t\}$ is the index of the first premium payment date after time t and $T_{p_t-1} = t$ by convention.

Using the same computational method as for the default leg, the time-t present value of the premium leg under the risk-neutral measure, denoted P_t, can be expressed as:

$$P_t = \kappa \cdot P_t^u, \quad (21)$$

with

$$P_t^u = \sum_{i=p_t}^{p}\left(B_t(T_i)\Delta_i\left(b-a-\mathbb{E}\left[L_{T_i}^{(a,b)} \mid \mathcal{H}_t\right]\right)+AC_{i,t}\right), \quad (22)$$

and where

$$AC_{i,t} = B_t(T_i)\Delta_i\mathbb{E}\left[L_{T_i}^{(a,b)} \mid \mathcal{H}_t\right] - \int_{T_{i-1}}^{T_i}B_t(s)\left(1-r_s\left(s-T_{i-1}\right)\right) \times \mathbb{E}\left[L_s^{(a,b)} \mid \mathcal{H}_t\right]ds. \quad (23)$$

The quantity P_t^u corresponds to the time-t present value of the unitary premium leg (contractual spread κ equal to 1 bp).

The CDO tranche (contractual) spread κ is chosen such that the contract is fair at inception, i.e., such that the default payment leg is equal to the premium payment leg:

$$\kappa = \frac{D_0}{P_0^u}.$$

The spread κ is quoted in basis point per annum.[13] Let us remark that the computation of κ only involves the expected losses on the tranche, $\mathbb{E}\left[L_t^{(a,b)}\right]$ at different time horizons. These can readily be derived from the marginal distributions of the aggregate loss on the reference portfolio.

[13] Let us remark that market conventions are different for the pricing of equity tranches (CDO tranches $(0, b)$ with $0 < b \leq 1$). Due to the high level of risk embedded in these "first losses tranches," the premium κ is fixed beforehand at 500 bps per annum and the protection seller receives an additional payment at inception based on an "upfront premium" and proportional to the size of the tranche. This "upfront premium" is quoted in percentage of the nominal value.

3.1.4 CDO Tranche Price and Market Spread

The time-t price (buy protection position) of a CDO tranche (a, b) is such that $V_t(\kappa) = D_t - \kappa \cdot P_t^u$. This corresponds to the amount a buyer of protection is willing to pay (or gain) in order to close his position at time t. Let us note that this is consistent with the definition of the contractual spread κ for which the market value must be equal to zero at inception ($V_0(\kappa) = 0$).

Like CDS, most CDO tranche quotations are based on spreads. The time-t market spread is defined as the contractual spread of a tranche with the same characteristic but initiated at time t:

$$\kappa_t = \frac{D_t}{P_t^u}.$$

Let us also note that there is a simple relationship between credit spreads and market values:

$$V_t(\kappa) = P_t^u \, (\kappa_t - \kappa).$$

The latter expression simply means that the value of a CDO tranche contract for a buyer of protection is positive when the current market spread is greater than the contractual spread.

As illustrated in Sect. 3.1.2, there exists liquidly quoted CDO tranches on most CDS indices. Figure 7 shows the dynamics of credit spreads on the 5 year iTraxx index between May and November 2007.[14] It is interesting to observe a sharp bump corresponding to the summer 2007 credit crisis.

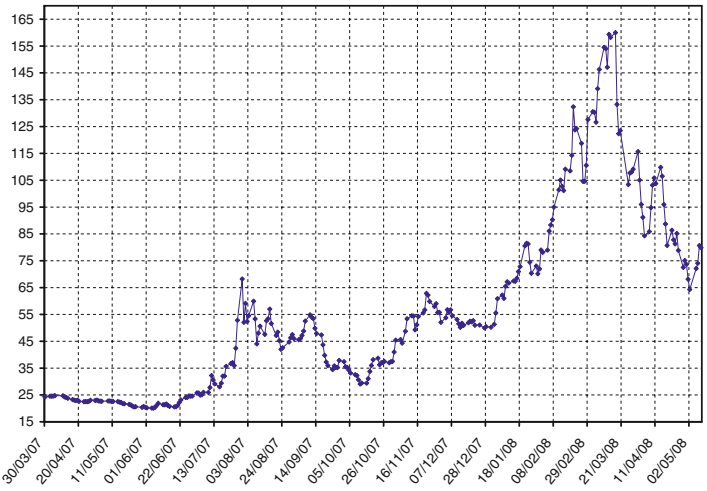


Fig. 7 Credit spreads on the 5 years iTraxx index (Series 7, 8 and 9) in bps, source Markit

[14] Apart from details regarding the premium leg, cash-flows generated by a CDS index can be considered to be the same as the ones of a $[0, 100\%]$ CDO tranche.

3.2 Homogeneous Markovian Contagion Models

While the use of the representation Theorem 1 guarantees that, in our framework, any basket default swap can be perfectly hedged with respect to default risks, it does not provide a practical way of constructing hedging strategies. As is the case with interest rate or equity derivatives, exhibiting hedging strategies involves some Markovian assumptions.

3.2.1 Intensity Specification

In the contagion approach, one starts from a specification of the risk-neutral predefault intensities[15] $\tilde{\alpha}^1, \ldots, \tilde{\alpha}^n$. In Sect. 2, risk-neutral predefault intensities depend upon the complete history of defaults. More simplistically, it is often assumed that they depend only upon the current credit status, i.e., the default indicators; thus $\tilde{\alpha}_t^i$, $i = 1, \ldots, n$ are deterministic functions of $N_t^1, \ldots, N_t^n$. In this paper, we will further remain in this Markovian framework, i.e., the default intensities will take the form

$$\tilde{\alpha}_t^i = \tilde{\alpha}^i \left(t, N_t^1, \ldots, N_t^n\right), \ i = 1, \ldots, n. \tag{24}$$

This Markovian assumption may be questionable, since the contagion effect of a default event may vanish as time goes by. The Hawkes process, that was used in the credit field by Giesecke and Goldberg (see [29] or [34]), provides such an example of a more complex time dependence.

Other specifications with the same aim are discussed in [48]. Popular examples are the models of [38,42,65], where the intensities are affine functions of the default indicators.

The connection between contagion models and Markov chains is described in the book of Lando [43]. More recently, Herbertsson and Rootzén [37] proved that default times with default intensities defined by (24) could be represented as the times until absorption in a finite state absorbing Markov chain. According to Assaf et al. [3] terminology, default times follow a multivariate phase-type distribution in this framework.

Another practical issue is related to name heterogeneity. Modelling all possible interactions amongst names leads to a huge number of contagion parameters and high-dimensional problems, thus to numerical issues. For this practical purpose, we will further restrict to models where all the names share the same risk-neutral intensity.[16] This can be viewed as a reasonable assumption for CDO tranches on large indices, although this is an issue with equity tranches for which idiosyncratic risk is an important feature. Since risk-neutral predefault intensities, $\tilde{\alpha}^1, \ldots, \tilde{\alpha}^n$ are equal, we will further denote these individual predefault intensities by $\tilde{\alpha}$.

For further tractability, we will further rely on a strong name homogeneity assumption, that individual predefault intensities only depend upon the number of

[15] Let us recall that the default intensity of name i vanishes after τ_i, i.e., $\alpha_t^i = 0$ on the set $\{t > \tau_i\}$.

[16] This means that the predefault intensities have the same functional dependence to the default indicators.

defaults. Let us denote by $N_t = \sum_{i=1}^{n} N_t^i$ the number of defaults at time t within the pool of assets. Predefault intensities thus take the form[17] $\tilde{\alpha}_t^i = \tilde{\alpha}(t, N_t)$. This is related to mean-field approaches (see [31]). As for parametric specifications, we can think of some additive effects, i.e., the predefault name intensities take the form $\tilde{\alpha}(t, N_t) = a + b \times N_t$ for some constants a, b as mentioned in [31], corresponding to the "linear counterparty risk model," or multiplicative effects in the spirit of Davis and Lo [17], i.e., the predefault intensities take the form $\tilde{\alpha}(t, N_t) = a \times b^{N_t}$. Of course, we could think of a non-parametric model. We provide a calibration procedure of such unconstrained intensities onto market inputs in Sect. 3.3.

For simplicity, we will further assume a constant recovery rate equal to R and a constant exposure among the underlying names. The aggregate fractional loss at time t is given by: $L_t = (1 - R)\frac{N_t}{n}$. As a consequence of the no simultaneous defaults assumption, the intensity of L_t or of N_t is simply the sum of the individual default intensities and is itself only a function of the number of defaults process. Let us denote by $\lambda(t, N_t)$ the risk-neutral loss intensity. It is related to the individual predefault intensities by:

$$\lambda(t, N_t) = (n - N_t) \times \tilde{\alpha}(t, N_t).$$

We are thus typically in a bottom-up approach, where one starts with the specification of name intensities and thus derives the dynamics of the aggregate loss.

3.2.2 Risk-Neutral Pricing

Let us remark that in a Markovian homogeneous contagion model, the process N_t is a continuous time Markov chain (under the risk-neutral probability $\mathbb{Q}$), and more precisely a pure birth process, according to Karlin and Taylor [40] terminology,[18] since only single defaults can occur.[19] The generator of the chain, $\Lambda(t)$ is quite simple:

$$\Lambda(t) = \begin{pmatrix} -\lambda(t,0) & \lambda(t,0) & 0 & \cdots & 0 \\ 0 & -\lambda(t,1) & \lambda(t,1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & & & -\lambda(t,n-1) & \lambda(t,n-1) \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Such a simple model of the number of defaults dynamics was considered in [60] where it is called the "one-step representation of the loss distribution." The approach

[17] Let us remark that on $\{\tau_i > t\}$, $N_t = \sum_{j \neq i} N_t^j$, so that the predefault intensity of name i, actually only depends on the credit status of the other names.

[18] According to Feller's terminology, we should speak of a pure death process. Since, we later refer to [40], we prefer their terminology.

[19] Regarding the assumption of no simultaneous defaults, we also refer to [11, 57, 64]. Allowing for multiple defaults could actually ease the calibration to senior CDO tranche quotes.

described in this section can be seen as a bottom-up view of the previous model, where the risk-neutral prices can actually be viewed as replicating prices. As an example of this approach, let us consider the replication price of a European payoff with payment date T, such as a "zero-coupon tranchelet," paying $1_{\{N_T=k\}}$ at time T for some $k \in \{0, 1, \ldots, n\}$. Let us denote by $V(t, N_t) = e^{-r(T-t)}\mathbb{Q}(N_T = k \,|N_t)$ the time-t replication price and by $V(t, .)$ the price vector whose components are $V(t, 0), V(t, 1), \ldots, V(t, n)$ for $0 \leq t \leq T$. We can thus relate the price vector $V(t, .)$ to the terminal payoff, using the transition matrix $\mathbf{Q}(t, T)$ between dates t and T:

$$V(t, .) = e^{-r(T-t)}\mathbf{Q}(t, T)V(T, .), \tag{25}$$

where $V(T, j) = 1_{\{j=k\}}$, $j = 0, 1, \ldots, n$. The transition matrix solves for the Kolmogorov backward and forward equations:

$$\frac{\partial \mathbf{Q}(t, T)}{\partial t} = -\Lambda(t)\mathbf{Q}(t, T), \quad \frac{\partial \mathbf{Q}(t, T)}{\partial T} = \mathbf{Q}(t, T)\Lambda(T). \tag{26}$$

In the time homogeneous case, i.e., when the generator is a constant $\Lambda(t) = \Lambda$, the transition matrix can be written in exponential form:

$$\mathbf{Q}(t, T) = \exp\left((T - t)\Lambda\right). \tag{27}$$

These ideas have been put in practice by [1, 18, 28, 36, 37, 48, 63]. These papers focus on the pricing of credit derivatives, while our concern here is the feasibility and implementation of replicating strategies.

3.2.3 Computation of Credit Deltas

We recall that the credit delta with respect to name i is the amount of hedging instruments (the index here, but possibly a ith credit default swap) that should be bought to be protected against a sudden default of name i. A nice feature of homogeneous contagion models is that the credit deltas are the same for all (the non-defaulted) names, which results in a dramatic dimensionality reduction. In that case, it is enough to consider the index portfolio as a single hedging instrument, which is consistent with some market practices.

Let us consider a European type payoff[20] and denote its replication price at time t by $V(t, .)$. In order to compute the credit deltas, let us remark that, by Itô's lemma,

$$\begin{aligned} dV(t, N_t) &= \frac{\partial V(t, N_{t-})}{\partial t}dt + \left(V(t, N_t) - V(t, N_{t-})\right)dN_t \\ &= \frac{\partial V(t, N_{t-})}{\partial t}dt + \left(V(t, N_{t-} + 1) - V(t, N_{t-})\right)dN_t. \end{aligned}$$

[20] For notational simplicity, we assume that there are no intermediate payments. This corresponds for instance to the case of zero-coupon CDO tranches with up-front premiums. The more general case is considered in Sect. 3.4.

The second term in the right hand side of the latter expression, $V(t, N_{t-}+1) - V(t, N_{t-})$ is associated with the jump in the price process when a default occurs in the credit portfolio, i.e., $dN_t = 1$. Thanks to the fact that $dN_t = \sum_{i=1}^{n} dN_t^i$ and, since $e^{-rt}V(t, N_t)$ is a $(\mathbb{Q}, \mathbb{H})$–martingale, it can be seen using Itô's lemma that V solves for the backward Kolmogorov equations:

$$\frac{\partial V(t,k)}{\partial t} + \lambda(t,k) \times (V(t,k+1) - V(t,k)) = rV(t,k), \ k = 0, \ldots, n-1,$$

where the terminal conditions are given by the payoff function at time T and with $V(t,n) = 0$, for all $t \geq 0$. We end up with:

$$\begin{aligned} dV(t,N_t) &= rV(t,N_{t-})\,dt + \sum_{i=1}^{n} (V(t,N_{t-}+1) \\ &\quad - V(t,N_{t-}))\left(dN_t^i - \tilde{\alpha}(t,N_{t-})(1-N_t^i)dt\right). \end{aligned} \tag{28}$$

As a consequence the credit deltas with respect to the individual instantaneous default swaps are equal to:

$$\delta_t^i = \delta^i(t,N_{t-}) = V(t,N_{t-}+1) - V(t,N_{t-}),$$

for $0 \leq t \leq T$ and $i = 1, \ldots, n$.

Let us denote by $V^I(t,k) = e^{-r(T-t)}E\left[1 - \frac{N_T}{n} \,|N_t = k\right]$ the time-t price of the equally weighted portfolio involving defaultable bonds and set

$$\delta_t^I = \delta^I(t,N_{t-}) = \frac{V(t,N_{t-}+1) - V(t,N_{t-})}{V^I(t,N_{t-}+1) - V^I(t,N_{t-})}. \tag{29}$$

As the dynamics of V^I also satisfies SDE (28) and using (29), we can deduce that:

$$dV(t,N_t) = r \times \left(V(t,N_{t-}) - \delta^I(t,N_{t-})V^I(t,N_{t-})\right)dt + \delta^I(t,N_{t-})dV^I(t,N_t).$$

As a consequence, we can perfectly hedge a European type payoff, say a zero-coupon CDO tranche, using only the index portfolio and the risk-free asset. The hedge ratio, with respect to the index portfolio is actually equal to (29). The previous hedging strategy is feasible provided that $V^I(t,N_{t-}+1) \neq V^I(t,N_{t-})$. The usual case corresponds to some positive dependence, thus $\alpha(t,0) \leq \alpha(t,1) \leq \cdots \leq \alpha(t,n-1)$. Therefore $V^I(t,N_{t-}+1) < V^I(t,N_{t-})$.[21] The decrease in

[21] In the case where $\alpha(t,0) = \alpha(t,1) = \ldots = \alpha(t,n)$, there are no contagion effects and default dates are independent. We still have $V^I(t,N_{t-}+1) < V^I(t,N_{t-})$ since $V^I(t,N_{t-})$ is linear in the number of surviving names.

the index portfolio value is the consequence of a direct default effect (one name defaults) and an indirect effect related to a positive shift in the credit spreads associated with the non-defaulted names.

The idea of building a hedging strategy based on the change in value at default times was introduced in [2]. The rigorous construction of a dynamic hedging strategy in a univariate case can be found in [9]. Our result can be seen as a natural extension to the multivariate case, provided that we deal with Markovian homogeneous models: we simply need to deal with the number of defaults N_t and the index portfolio $V^I(t, N_t)$ instead of a single default indicator N_t^i and the corresponding defaultable discount bond price.

3.3 Calibration of Loss Intensities

Another nice feature of the homogeneous Markovian contagion model is that the loss dynamics or equivalently the default intensities can be determined from market inputs such as CDO tranche premiums. Since the risk neutral dynamics are unambiguously derived from market inputs, so will be for dynamic hedging strategies of CDO tranches. This greatly facilitates empirical studies, since the replicating figures do not depend upon unobserved and difficult to calibrate parameters.

The construction of the implied Markov chain for the aggregate loss parallels the one made by Dupire [25] to construct a local volatility model from call option prices. Similar ideas are used in [23, 58] to build up implied trees. Laurent and Leisen [46] have shown how an implied Markov chain can be derived from a discrete set of option prices. In these approaches, the calibration of the implied dynamics on market inputs involves forward Kolmogorov equations. Starting from a complete set of CDO tranche premiums or equivalently from a complete set of number of default distributions, [60] provided the construction of the loss intensities. Similarities between the Dupire's approach and the building of the one step Markov chain of [60] have also been reported in [13, 20, 48]. We propose now to detail and comment the latter calibration approach of loss intensities.

3.3.1 Calibration of Loss Intensities on a Complete Set of Number of Defaults Probabilities

While the pricing and thus the hedging involves a backward procedure, calibration is associated with forward Kolmogorov differential equations. We show here a non-parametric fitting procedure of a possibly non time homogeneous pure birth process onto a complete set of marginal distributions of number of defaults. This is quite similar to the one described in [60], though the purpose is somehow different since the aim of [60] was to construct arbitrage-free, consistent with some complete loss surface, Markovian models of aggregate losses, possibly in incomplete markets, without detailing the feasibility and implementation of replicating strategies.

We will further denote the marginal default probabilities by $p(t,k) = \mathbb{Q}(N_t = k)$ for $0 \leq t \leq T$, $k = 0, 1, \ldots, n$. Clearly, this involves more information that one could directly access through the quotes of liquid CDO tranches, especially with respect to small and large number of defaults. As for the computation of the number of default probabilities from quoted CDO tranche premiums, we refer to [32,41,50,53,62,64]. Practical issues related to the calibration inputs are also discussed in [63].

In the case of a pure birth process, the forward Kolmogorov equations can be written as:

$$\begin{cases} \frac{dp(t,0)}{dt} = -\lambda(t,0)p(t,0), & k = 0, \\ \frac{dp(t,k)}{dt} = \lambda(t,k-1)p(t,k-1) - \lambda(t,k)p(t,k), & k = 1, \ldots, n. \end{cases} \tag{30}$$

Since the space state is finite, there are no regularity issues and these equations admit a unique solution[22] (see below for practical implementation). These forward equations can be used to compute the loss intensity dynamics $t \in [0,T] \to \lambda(t, N_t)$, thanks to:

$$\begin{cases} \lambda(t,0) = -\frac{1}{p(t,0)} \frac{dp(t,0)}{dt}, & k = 0, \\ \lambda(t,k) = \frac{1}{p(t,k)} \left[\lambda(t,k-1)p(t,k-1) - \frac{dp(t,k)}{dt} \right], & k = 1, \ldots, n, \end{cases} \tag{31}$$

for $0 \leq t \leq T$. Let us remark that we can also write:

$$\lambda(t,k) = -\frac{1}{p(t,k)} \frac{d \sum_{m=0}^{k} p(t,m)}{dt} = -\frac{1}{\mathbb{Q}(N_t = k)} \frac{d\mathbb{Q}(N_t \leq k)}{dt}, \quad k = 0, \ldots, n. \tag{32}$$

Eventually, the name intensities are provided by: $\tilde{\alpha}(t, N_t) = \frac{\lambda(t,N_t)}{n - N_t}$. This shows that we can fully recover the loss intensities from the marginal distributions of the number of defaults, if the latters do not occur simultaneously. This has to be related to the result of Cont and Minca [13] which states that, under the assumption of no simultaneous defaults, the flow of marginal loss distributions associated with a general point process can be matched with the one of a Markovian jump process.

On practical grounds, the computation of the $p(t,k)$ usually involves some arbitrary smoothing procedure and hazardous extrapolations for small time horizons. For these reasons, we think that it is more appropriate and reasonable to calibrate the Markov chain of aggregate losses on a discrete set of meaningful market inputs corresponding to liquid maturities.

[22] We refer to [40] for more details about the forward equations in the case of a pure birth process.

3.3.2 Calibration of Time Homogeneous Loss Intensities

In practical applications, we can only rely on a discrete set of loss distributions corresponding to liquid CDO tranche maturities. In the examples below, we will calibrate the loss intensities given a single calibration date T. For simplicity, we will be given the default probabilities $p(T,k), k = 0,1,\ldots,n$. Now and in the sequel, we assume that the loss intensities are time homogeneous: the intensities do not depend on time but only on the number of realized defaults. We further denote by $\lambda_k = \lambda(t,k)$ for $0 \leq t \leq T$, the loss intensity for $k = 0,1,\ldots,n-1$. Let us note that [28] have also developed a similar computation of the loss intensities λ_k from the values of default probabilities.

Solving the forward equations (30) provides

$$\begin{cases} p(T,0) = e^{-\lambda_0 T}, & k = 0, \\ p(T,k) = \lambda_{k-1} \int_0^T e^{-\lambda_k (T-s)} p(s,k-1) ds, & k = 1,\ldots,n-1. \end{cases} \tag{33}$$

The previous equations can be used to determine $\lambda_0,\ldots,\lambda_{n-1}$ iteratively, even if our calibration inputs are the defaults probabilities at the single date T.

Assume for the moment that the intensities $\lambda_0,\ldots,\lambda_{n-1}$ are known, positive and distinct.[23] To solve the forward equations, we assume that the default probabilities can be written as

$$p(t,k) = \sum_{i=0}^{k} a_{k,i} e^{-\lambda_i t}, \quad 0 \leq t \leq T, \ k = 0,\ldots,n-1, \tag{34}$$

where the parameters $a_{k,i}$, $i = 0,1,\ldots,k$, $k = 0,\ldots,n-1$ satisfy the following recurrence equations:

$$\begin{cases} a_{0,0} = 1, \\ a_{k,i} = \frac{\lambda_{k-1}}{\lambda_k - \lambda_i} a_{k-1,i}, \quad i = 0,1,\ldots,k-1, \ k = 1,\ldots,n-1, \\ a_{k,k} = -\sum_{i=0}^{k-1} a_{k,i} \end{cases} \tag{35}$$

Then, we check easily that, the functions $t \mapsto p(t,k)$, $k = 0,\ldots,n-1$ described by (34) and (35) provide some solutions of the forward PDE. Since it is well-known that these solutions are unique, it means we have obtained explicitly the solutions of the forward PDE, knowing the intensities $(\lambda_k)_{k=0,\ldots,n-1}$. Therefore, using $p(0,k) = 0$ and $\lambda_0 = -\frac{\ln(p(T,0))}{T}$, we can compute iteratively $\lambda_1,\ldots,\lambda_{n-1}$ by solving the univariate non-linear implicit equations $p(T,k) = \sum_{i=0}^{k} a_{k,i} e^{-\lambda_i T}$, or equivalently:

$$\sum_{i=0}^{k-1} a_{k-1,i} e^{-\lambda_i T} \left(\frac{1 - e^{-(\lambda_k - \lambda_i)T}}{\lambda_k - \lambda_i} \right) = \frac{p(T,k)}{\lambda_{k-1}}, \quad k = 1,\ldots,n-1. \tag{36}$$

[23] Due to the last assumption, the described calibration approach is not highly regarded by numerical analysts (see [51] for a discussion). However, it is well suited in our case studies.

It can be seen easily that for any $k \in \{0, \ldots, n-1\}$, $p(T,k)$ is a decreasing function of λ_k, taking value $\lambda_{k-1} \int_0^T p(s, k-1)ds$ for $\lambda_k = 0$ and with the limit equal to zero as λ_k tends to infinity. In other words, the system of equations (36) has a unique solution provided that:

$$p(T,k) < \lambda_{k-1} \left(\sum_{i=0}^{k-1} a_{k-1,i} \left(\frac{1 - e^{-\lambda_i T}}{\lambda_i} \right) \right), \quad k = 1, \ldots, n-1. \tag{37}$$

Note that, in practice, all the intensities λ_k will be different. Thus, starting from the T–default probabilities only, we have found the explicit solutions of the forward equations and the intensities $(\lambda_k)_{k=0,\ldots,n-1}$ that would be consistent with these probabilities.

3.3.3 Other Calibrating Approaches

Alternative calibrating approaches based on minimization algorithms have been proposed by several authors.

In Herbertsson [36], the name intensities $\alpha(t, N_t)$ are time homogeneous, piecewise constant in the number of defaults (the node points are given by standard detachment points) and they are fitted to spread quotes by a least square numerical procedure. Arnsdorf and Halperin [1] propose a piecewise constant parameterization of name intensities (which are referred to as "contagion factors") in time. When intensities are piecewise linear in the number of defaults too, they use a "multi-dimensional solver" to calibrate onto the observed tranche prices. In the same vein, Frey and Backhaus [30,31] introduce a parametric form for the function $\lambda(t,k)$, a variant of the "convex counterparty risk model," and fit the parameters to some tranche spreads. Lopatin and Misirpashaev [48] express the loss intensity $\lambda(t,k)$ as a polynomial function of an auxiliary variable involving the number of defaults.

Cont and Minca [13] propose an alternative method based on the principle of minimum relative entropy. The name intensities has to be chosen in such a way that the loss process is close enough to a simple prior process in the sens of an entropy distance. In the same time, the usual calibration constraints have to be satisfied. However, the main drawback of this approach is the fact that the fitted intensities strongly rely on the choice of the prior.

In the spirit of Dupire [25], Cont et al. [14] show that loss intensities $\lambda(t,k)$, $0 \leq t \leq T$, $k = 0, \ldots, n$ can be formulated using prices of put options on the aggregate loss, i.e., $\mathbb{E}[\min(N_t, k)]$, $0 \leq t \leq T$, $k = 0, \ldots, n$. It allows to transform the calibration of the loss intensities into the calibration of the put option values. Given the small number of available quotes, they remark that there can be several sets of put values that are consistent with the market CDO data. Therefore, a calibration algorithm based on quadratic programming is proposed in order to pinpoint a unique set of put values if it exists. They also compare their method

with the calibration approaches introduced by Herbertsson [36] and Cont and Minca [13] and show that calibrated intensity surfaces can be significantly different across algorithms.

3.4 Computation of Credit Deltas Through a Recombining Tree

We now address the computation of CDO tranche deltas with respect to the credit default swap index of the same maturity. As for the hedging instrument, the premium is set at the inception of the deal and remains fixed which corresponds to market conventions. We do not take into account roll dates every 6 months and trade the same index series up to maturity. Switching from one hedging instrument to another could be dealt with very easily in our framework and closer to market practice but we thought that using the same underlying across the tree would simplify the exposition.[24]

3.4.1 Building Up a Tree

Let us recall that the (fractional) loss at time t is given by $L_t = (1-R)\frac{N_t}{n}$. In what follows, we consider a tranche with attachment point a and detachment point b, $0 \leq a \leq b \leq 1$. Up to some minor adjustment for the premium leg (see below), the credit default swap index is assimilated to a $[0, 100\%]$ tranche. We denote by $O(N_t)$ the outstanding nominal on a tranche. It is equal to $b-a$ if $L(t) < a$, to $b - L(t)$ if $a \leq L(t) < b$ and to 0 if $L(t) \geq b$.

Let us recall that, for a European type payoff the price vector fulfils $V(t,.) = e^{-r(t'-t)}\mathbf{Q}(t,t')V(t',)$ for $0 \leq t \leq t' \leq T$. The transition matrix can be expressed as $\mathbf{Q}(t,t') = \exp(\Lambda(t'-t))$ where Λ is the generator matrix associated with the number of defaults process. Note that, in the time homogeneous framework discussed in the previous section, the generator matrix does not depend on time.

For practical implementation, we will be given a set of node dates $t_0 = 0$, $\ldots, t_i, \ldots, t_{n_s} = T$. For simplicity, we will further consider a constant time step $\Delta = t_1 - t_0 = \cdots = t_i - t_{i-1} = \cdots$; this assumption can easily be relaxed. The most simple discrete time approximation one can think of is $\mathbf{Q}(t_i, t_{i+1}) \simeq Id + \Lambda(t_i) \times (t_{i+1} - t_i)$, which leads to $\mathbb{Q}\left(N_{t_{i+1}} = k+1 \,|N_{t_i} = k\right) \simeq \lambda_k \Delta$ and $\mathbb{Q}\left(N_{t_{i+1}} = k \,|N_{t_i} = k\right) \simeq 1 - \lambda_k \Delta$. For large λ_k, the transition probabilities can become negative. Thus, we will rather use the following approximations:

$$\begin{cases} \mathbb{Q}\left(N_{t_{i+1}} = k+1 \,|N_{t_i} = k\right) \simeq 1 - e^{-\lambda_k \Delta}, \\ \mathbb{Q}\left(N_{t_{i+1}} = k \,|N_{t_i} = k\right) \simeq e^{-\lambda_k \Delta}. \end{cases} \tag{38}$$

Given the latter approximations and as illustrated in Fig. 8, the dynamics of the number of defaults process can be described through a recombining tree.

[24] Actually, the credit deltas at inception are the same whatever the choice.

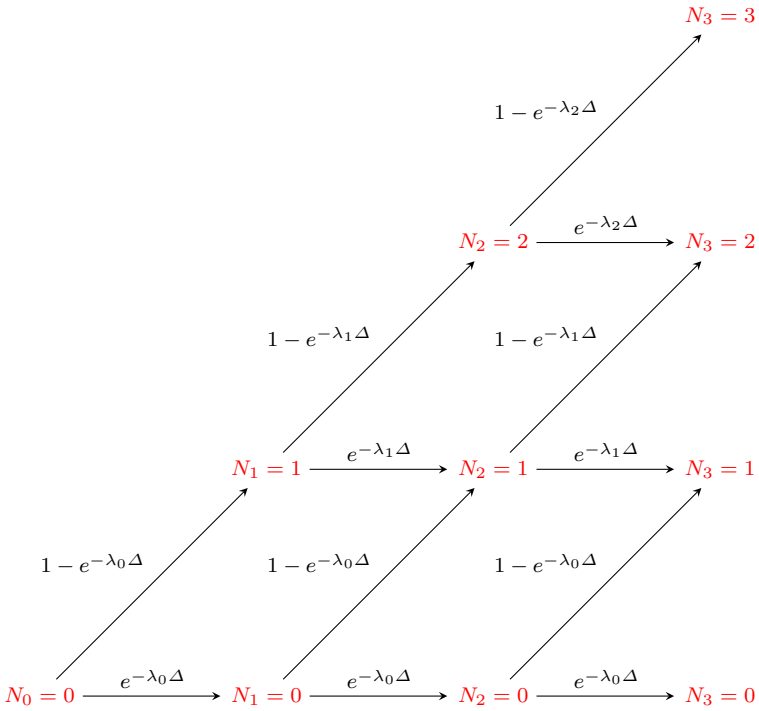


Fig. 8 Number of defaults tree

This idea has also been exploited by van der Voort [63]. One could clearly think of using continuous Markov chain techniques to compute present values of derivative products at hand, but the tree implementation is quite intuitive from a financial point of view as it corresponds to the implied binomial tree of Derman and Kani [23]. Convergence of the discrete time Markov chain to its continuous limit is a rather standard issue and will not be detailed here.

3.5 Computation of Hedge Ratios for CDO Tranches

3.5.1 Present Values of a CDO Tranche in the Tree Nodes

Let us denote by $D(i,k)$ the value at time t_i when $N_{t_i} = k$ of the default payment leg of the CDO tranche.[25] The default payment at time t_{i+1} is equal to $O\left(N_{t_i}\right) -$

[25] We consider the value of the default leg immediately after t_i. Thus, we do not consider a possible default payment at t_i in the calculation of $D(i,k)$.

$O\left(N_{t_{i+1}}\right)$. Thus, $D(i,k)$ is given by the following recurrence equation[26]:

$$D(i,k) = e^{-r\Delta} \cdot \left(\left(1-e^{-\lambda_k \Delta}\right)\{D(i+1,k+1)+O(k)-O(k+1)\} + e^{-\lambda_k \Delta} D(i+1,k)\right). \quad (39)$$

Let us now deal with a (unitary) premium leg. We denote the regular premium payment dates by $T_1,\ldots,T_p$ and for simplicity we assume that: $\{T_1,\ldots,T_p\} \subset \{t_0,\ldots,t_{n_s}\}$. Let us consider some date t_{i+1} and set l such that $T_l < t_{i+1} \leq T_{l+1}$. Whatever t_{i+1}, there is an accrued premium payment of $\left(O\left(N_{t_i}\right) - O\left(N_{t_{i+1}}\right)\right) \times (t_{i+1}-T_l)$. If $t_{i+1} = T_{l+1}$, i.e., t_{i+1} is a regular premium payment date, there is an extra premium cash-flow at time t_{i+1} of $O\left(N(T_{l+1})\right) \times (T_{l+1}-T_l)$. Thus, if t_{i+1} is a regular premium payment date, the total premium payment is equal to $O\left(N_{t_i}\right) \times (T_{l+1}-T_l)$. Let us denote by $P(i,k)$ the value at time t_i when $N_{t_i} = k$ of the unitary premium leg.[27] If $t_{i+1} \in \{T_1,\ldots,T_p\}$, $P(i,k)$ is provided by:

$$P(i,k) = e^{-r\Delta} \cdot \Big(O(k)\,(T_{l+1}-T_l) + \left(1-e^{-\lambda_k \Delta}\right) P(i+1,k+1) + e^{-\lambda_k \Delta} P(i+1,k)\Big). \quad (40)$$

If $t_{i+1} \notin \{T_1,\ldots,T_p\}$, then[28]:

$$P(i,k) = e^{-r\Delta} \cdot \left(\left(1-e^{-\lambda_k \Delta}\right)\{P(i+1,k+1) + (O(k)-O(k+1))\,(t_{i+1}-T_l)\} + e^{-\lambda_k \Delta} P(i+1,k)\right). \quad (41)$$

The CDO tranche premium is equal to $\kappa = \frac{D(0,0)}{P(0,0)}$. The value of the CDO tranche (buy protection case) at time t_i when $N_{t_i} = k$ is given by $V(i,k) = D(i,k) - \kappa \cdot P(i,k)$. The equity tranche needs to be dealt with slightly differently since its spread is set to $\kappa = 500$ bp. However, the value of the CDO equity tranche is still given by $D(i,k) - \kappa \cdot P(i,k)$.

3.5.2 Present Values of a CDS Index in the Tree Nodes

As for the credit default swap index, we will denote by $P^I(i,k)$ and $D^I(i,k)$ the values of the premium and default legs. We define the credit default swap index spread at time t_i when $N_{t_i} = k$ by $\kappa^I(i,k) \cdot P^I(i,k) = D^I(i,k)$. The value of the credit default swap index, bought at inception, at node (i,k) is given by

[26] This relation holds for $i = 0,\ldots,n_s-1$, $k = 0,\ldots,\min(i,n-1)$ and with $D(n_s,k) = 0$ when $k = 0,\ldots,n$ and $D(i,n) = 0$ when $i = n,\ldots,n_s-1$.

[27] As for the default leg, we consider the value of the premium leg immediately after t_i. Thus, we do not take into account a possible premium payment at t_i in the calculation of $P(i,k)$ either.

[28] Relations (40) and (41) hold for $i = 0,\ldots,n_s-1$, $k = 0,\ldots,\min(i,n-1)$ and with $P(n_s,k) = 0$ when $k = 0,\ldots,n$ and $P(i,n) = 0$ when $i = n,\ldots,n_s-1$.

$V^I(i,k) = D^I(i,k) - \kappa^I(0,0) \cdot P^I(i,k)$.[29] The default leg of the credit default swap index is computed as a standard default leg of a $(0, 100\%)$ CDO tranche. Thus, in the recursion equation (39) giving $D^I(i,k)$, we write the outstanding nominal for k defaults as $O(k) = 1 - \frac{k(1-R)}{n}$, where R is the recovery rate and n the number of names. According to standard market rules, the premium leg of the credit default swap index needs a slight adaptation since the premium payments are based only upon the number of non-defaulted names and do not take into account recovery rates. As a consequence, the outstanding nominal to be used in the recursion equations (40) and (41) providing $P^I(i,k)$ is such that $O(k) = 1 - \frac{k}{n}$.

3.5.3 Computation of Credit Deltas in the Tree Nodes

As usual in binomial trees, $\delta(i,k)$ is the ratio of the difference of the option value (at time t_{i+1}) in the upper state ($k+1$ defaults) and lower state (k defaults) and the corresponding difference for the underlying asset. In our case, both the CDO tranche and the credit default swap index are "dividend-baring." For instance, when the number of defaults switches from k to $k+1$, the default leg of the CDO tranche is associated with a default payment of $O(k) - O(k+1)$. Similarly, given the above discussion, when the number of defaults switches from k to $k+1$, the premium leg of the CDO tranche is associated with an accrued premium payment equal to[30]

$$-\kappa 1_{t_{i+1} \notin \{T_1, \ldots, T_p\}} \left(O(k) - O(k+1)\right) (t_{i+1} - T_l). \tag{42}$$

Thus, when a default occurs the change in value of the CDO tranche is the outcome of a capital gain of $V(i+1,k+1) - V(i+1,k)$ and of a cash-flow of

$$D(i,k) = \left(O(k) - O(k+1)\right) \left(1 - \kappa 1_{t_{i+1} \notin \{T_1, \ldots, T_p\}} (t_{i+1} - T_l)\right). \tag{43}$$

Similarly, when a default occurs the change in value of the credit default swap index is the outcome of a capital gain of $V^I(i+1,k+1) - V^I(i+1,k)$ and a cash-flow of

$$D^I(i,k) = \frac{1-R}{n} - \frac{1}{n} \kappa^I(0,0) 1_{t_{i+1} \notin \{T_1, \ldots, T_p\}} (t_{i+1} - T_l) \tag{44}$$

The credit delta of the CDO tranche at node (i,k) with respect to the credit default swap index is thus given by:

$$\delta(i,k) = \frac{V(i+1,k+1) - V(i+1,k) + D(i,k)}{V^I(i+1,k+1) - V^I(i+1,k) + D^I(i,k)}. \tag{45}$$

[29] This is an approximation of the index spread since, according to market rules, the first premium payment is reduced.

[30] If $t_{i+1} \in \{T_1, \ldots, T_p\}$, the premium payment is the same whether the number of defaults is equal to k or $k+1$. So, it does not appear in the computation of the credit delta.

Let us remark that using the previous credit deltas leads to a perfect replication of a CDO tranche within the tree, which is feasible since the approximating discrete market is complete.

In the next section, we compute CDO tranche credit deltas with respect to credit default swap index in two steps. We first calibrate loss intensities from a one factor Gaussian copula loss distribution. It allows us to examine how the correlation between defaults impact credit deltas. We then calibrate loss intensities from a loss distribution associated with a market base correlation structure and we compare our "default risk" deltas with some "credit spread" deltas computed on a basis of a bump of credit default swap index spread. We investigate in particular spread deltas computed from the standard market approach and spread deltas recently obtained in [1] and [26].

3.5.4 Model Calibrated on a Loss Distribution Associated with a Gaussian Copula

In this numerical illustration, the loss intensities λ_k are computed from a loss distribution generated from a one factor homogeneous Gaussian copula model.[31] The correlation parameter is equal to $\rho^2 = 30\%$, the credit spreads are assumed to be all equal to $\kappa = 20$ basis points per annum,[32] the recovery rate is such that $R = 40\%$ and the maturity is $T = 5$ years. The number of names is $n = 125$. Figure 9 shows the number of defaults distribution.

Loss intensities λ_k are calibrated up to $k = 49$ defaults according to the method proposed in Sect. 3.3.2. Under the Gaussian copula assumption, the default probabilities $p(5,k)$ are insignificant[33] for $k > 49$. To avoid numerical difficulties, we computed the remaining λ_k $(k > 49)$ by linear extrapolation.[34]

As can be seen from Fig. 10, loss intensities change almost linearly with respect to the number of defaults. Let us also remark that such rather linear behavior of loss intensities can be found in [48]. Our results can also be related to the analysis of Ding et al. [24] who deal with a dynamic model where the loss intensity is actually linear in the number of defaults.

---

[31] In the homogeneous Gaussian copula model, default times have the same marginal distribution, says F. In that model, default times are defined by $\tau_i = F^{-1}(\Phi(V_i))$, $i = 1, \ldots, n$, where Φ is the standard Gaussian cumulative distribution and $V_1, \ldots, V_n$ are some latent variables such that: $V_i = \rho V + \sqrt{1-\rho^2}\bar{V}_i$, $i = 1, \ldots, n$. The factors $V, \bar{V}_i$, $i = 1, \ldots, n$ are independent standard Gaussian random variables.

[32] Marginal default probabilities have been computed using the classical assumption, under which default times are exponentially distributed with parameter $\frac{\kappa}{1-R}$, i.e., the cumulative distribution of default times at time T is equal to $F(T) = \mathbb{Q}(\tau_1 \leq T) = 1 - \exp\left(-\frac{\kappa}{1-R}T\right)$.

[33] $\sum_{k \geq 50} p(5,k) \simeq 3 \times 10^{-4}$, $p(5,50) \simeq 3.2 \times 10^{-5}$ and $p(5,125) \simeq 4 \times 10^{-12}$.

[34] We checked that various choices of loss intensities for high number of defaults had no effect on the computation of deltas. Let us stress that this applies for the Gaussian copula case since the loss distribution has thin tails. For the market case example, we proceeded differently.

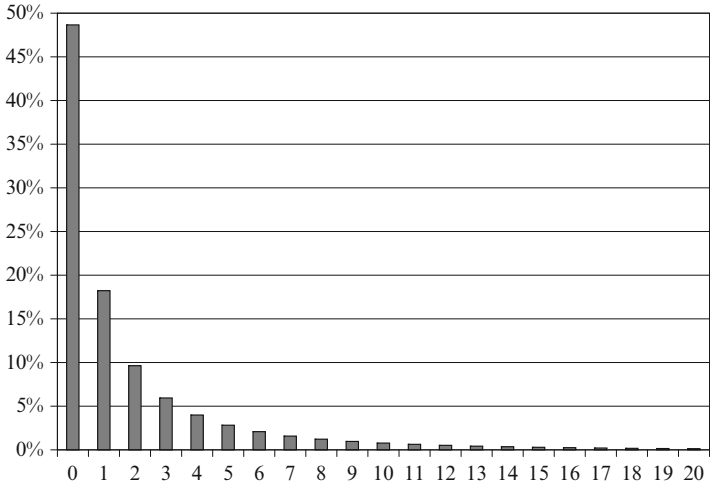


Fig. 9 Number of defaults distribution. Number of defaults on the x-axis. $\rho^2 = 30\%$: $p(5, k)$, $k = 0, \ldots, 20$

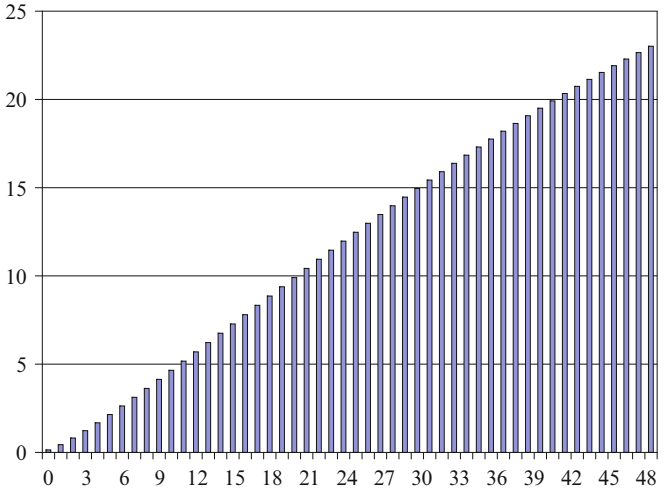


Fig. 10 Loss intensities λ_k, $k = 0, \ldots, 49$

Table 1 shows the dynamics of the credit default swap index spreads $\kappa^I(i, k)$ along the nodes of the tree. The continuously compounded default free rate is $r = 3\%$ and the time step is $\Delta = \frac{1}{365}$. It can be seen that default arrivals are associated with rather large jumps of credit spreads. For instance, if a (first) default occurs after a quarter, the credit default swap index spread jumps from 18 to 70 bps. An extra default by this time leads to an index spread of 148 bps.

The credit deltas with respect to the credit default swap index $\delta(i, k)$ have been computed for the (0–3%) and the (3–6%) CDO tranches (see Tables 2 and 3). As for

Table 1 Dynamics of credit default swap index spread $\kappa^I(i,k)$ in basis points per annum

	Weeks			
Nb defaults	0	14	56	84
0	20	18	14	13
1	0	70	54	46
2	0	148	112	93
3	0	243	182	150
4	0	350	261	215
5	0	466	347	285
6	0	589	437	359
7	0	719	531	436
8	0	856	630	516
9	0	997	732	598
10	0	1,142	839	683

Table 2 Delta of the [0–3%] equity tranche with respect to the credit default swap index

	Outstanding	Weeks			
Nb defaults	nominal	0	14	56	84
0	3.00%	0.538	0.591	0.755	0.859
1	2.52%	0	0.238	0.381	0.508
2	2.04%	0	0.074	0.137	0.212
3	1.56%	0	0.026	0.044	0.070
4	1.08%	0	0.011	0.017	0.024
5	0.60%	0	0.005	0.007	0.009
6	0.12%	0	0.001	0.001	0.001
7	0.00%	0	0	0	0

Table 3 Deltas of the [3–6%] with respect to the credit default swap index

	Outstanding	Weeks			
Nb defaults	nominal	0	14	56	84
0	3.00%	0.255	0.254	0.219	0.171
1	3.00%	0	0.280	0.349	0.357
2	3.00%	0	0.167	0.294	0.389
3	3.00%	0	0.068	0.158	0.265
4	3.00%	0	0.026	0.065	0.128
5	3.00%	0	0.014	0.027	0.053
6	3.00%	0	0.010	0.016	0.025
7	2.64%	0	0.008	0.011	0.015
8	2.16%	0	0.006	0.008	0.010
9	1.68%	0	0.004	0.005	0.007
10	1.20%	0	0.003	0.003	0.004
11	0.72%	0	0.002	0.002	0.002
12	0.24%	0	0.001	0.001	0.001
12	0.00%	0	0	0	0

the equity tranche, it can be seen that the credit deltas are positive anddecrease up to zero. This is not surprising given that a buy protection equity tranche involves a short put position over the aggregate loss with a 3% strike. This is associated with positive deltas, negative gammas and thus decreasing deltas. When the number of defaults is above 6, the equity tranche is exhausted and the deltas obviously are equal to zero.

At inception, the credit delta of the equity tranche is equal to 54% whilst it is only equal to 25% for the [3–6%] tranche which is deeper out of the money (see Table 3). Moreover, the [3–6%] CDO tranche involves a call spread position over the aggregate loss. As a consequence the credit deltas are positive and firstly increase (positive gamma effect) and then decrease (negative gamma) up to zero as soon as the tranche is fully amortized.

Given the recovery rate assumption of 40%, the outstanding nominal of the [3–6%] is equal to 3% for six defaults and to 2.64% for seven defaults. One might thus think that at the sixth default the [3–6%] tranche should behave almost like an equity tranche. However, as can be seen from Table 3, the credit delta of the [3–6%] tranche is much lower: around 1% instead of 60%. This is due to dramatic shifts in credit spreads when moving from the no-defaults to the six defaults state (see Table 1). In the latter case, the expected loss on the tranche is much larger, which is consistent with smaller deltas given the call spread payoff.

3.5.5 Dependence of Hedging Strategies upon the Correlation Parameter

Let us recall that the recombining tree is calibrated on a loss distribution over a given time horizon. The shape of the loss distribution depends critically upon the correlation parameter which was set up to now to $\rho^2 = 30\%$. Decreasing the dependence between default events leads to a thinner right-tail of the loss distribution and smaller contagion effects. We detail here the effects of varying the correlation parameter on the hedging strategies. For simplicity, we firstly focus the analysis on the equity tranche and shift the correlation parameter from 30 to 10%. It can be seen from Tables 2 and 4 that the credit deltas are much higher in the latter case. After 14 weeks, prior to the first default, the credit delta is equal to 59% for a 30% correlation and to 96% when the correlation parameter is equal to 10%.[35]

To further investigate how changes in correlation levels alter credit deltas, we computed the market value of the default leg of the equity tranche at a 14 weeks horizon as a function of the number of defaults under different correlation assumptions (see Fig. 11). The market value of the default leg, on the y-axis, is computed

[35] Let us remark that credit deltas can be above one in the no default case. This is due to the amortization scheme of the premium leg. We detail in the next section the impact of the premium leg on credit deltas.

Table 4 Deltas of the [0–3%] *equity* tranche with respect to the credit default swap index, $\rho^2 = 10\%$

Nb defaults	Outstanding nominal	Weeks 0	14	56	84
0	3.00%	0.931	0.960	1.009	1.058
1	2.52%	0	0.694	0.785	0.910
2	2.04%	0	0.394	0.485	0.645
3	1.56%	0	0.179	0.233	0.352
4	1.08%	0	0.072	0.092	0.145
5	0.60%	0	0.027	0.032	0.046
6	0.12%	0	0.004	0.005	0.007
7	0.00%	0	0	0	0

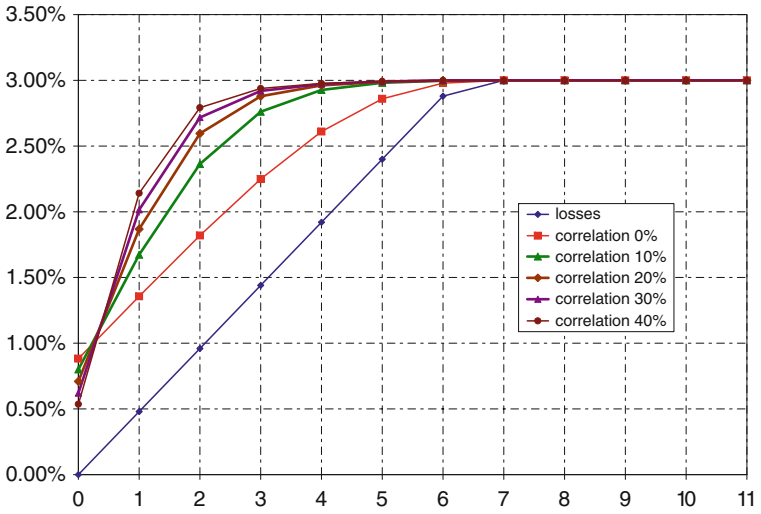


Fig. 11 Market value of equity default leg under different correlation assumptions. Number of defaults on the x-axis

as the sum of expected discounted cash-flows posterior to this 14 weeks horizon date and the accumulated defaults cash-flows paid before.[36] We also plotted the accumulated losses which represent the intrinsic value of the equity tranche default leg. Unsurprisingly, we recognize some typical concave patterns associated with a short put option payoff.

As can be seen from Fig. 11, prior to the first default, the value of the default leg of the equity tranche decreases as the correlation parameter increases from 0 to 40%. However, after the first default the ordering of default leg values is reversed. This can be easily understood since larger correlations are associated with larger

[36] For simplicity, we neglected the compounding effects over this short period.

jumps in credit spreads at default arrivals and thus larger changes in the expected discounted cash-flows associated with the default leg of the equity tranche.[37]

Therefore, varying the correlation parameter is associated with two opposite mechanisms:

- The first one is related to a typical negative vanna effect.[38] Increasing correlation lowers loss "volatility" and leads to smaller expected losses on the equity tranche. In a standard option pricing framework, this should lead to an increase in the credit delta of the short put position on the loss.
- This is superseded by the shifts due to contagion effects. Increasing correlation is associated with bigger contagion effects and thus larger jumps in credit spreads at the arrival of defaults. This, in turn leads to a larger jump in the market value of the credit index default swap. Let us recall that the default leg of the equity tranche exhibits a concave payoff and thus a negative gamma. As a consequence the credit delta, i.e., the ratio between the change in value of the option and the change in value of the underlying, decreases.

3.5.6 Model Calibrated on a Loss Distribution Associated with CDO Tranche Quotes

Up to now, the probabilities of number of defaults were computed thanks to a Gaussian copula and a single correlation parameter. In this example, we use a steep upward sloping *base correlation curve* for the iTraxx, typical of June 2007, as an input to derive the distribution of the probabilities of number of defaults (see Table 5). The maturity is still equal to 5 years, the recovery rate to 40% and the credit spreads to 20 bps. The default-free rate is now equal to 4%.

Rather than spline interpolation of base correlations, we used a parametric model of the 5 year loss distribution to fit the market quotes and compute the probabilities

Table 5 Base correlations with respect to attachment points (iTraxx June 2007)

Base tranches	[0–3%]	[0–6%]	[0–9%]	[0–12%]	[0–22%]
Base correlations	18%	28%	36%	42%	58%

[37] Let us remark that the larger the correlation the larger the change in market value of the default leg of the equity tranche at the arrival of the first default. Indeed, in a high correlation framework, this default means relatively higher default likelihood for the surviving names. This is not inconsistent with the previous results showing a decrease in credit deltas when the correlation parameter increases. The credit delta is the ratio of the change in value in the equity tranche and of the change in value in the credit default swap index. For a larger correlation parameter, the change in value in the credit default swap index is also larger due to magnified contagion effects.

[38] We recall that in option pricing, the vanna is the sensitivity of the delta to a unit change in volatility.

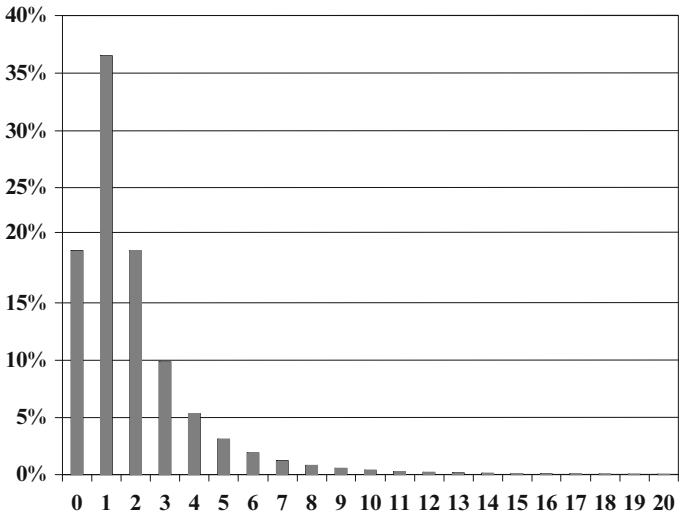


Fig. 12 Number of defaults distribution obtained from the *base correlation* structure described in Table 5. Number of defaults on the x-axis

of the number of defaults. This produces arbitrage free and smooth distributions that ease the calculation of the loss intensities.[39] Figure 12 shows the number of defaults distribution. This is rather different from the 30% flat correlation Gaussian copula case both for small and large losses. For instance, the probability of no defaults dropped from 48.7 to 19.5% while the probability of a single default rose from 18.2 to 36.5%. Let us stress that these figures are for illustrative purpose. The market does not provide direct information on first losses and thus the shape of the left tail of the loss distribution is a controversial issue. As for the right-tail, we have $\sum_{k\geq 50} p(5,k) \simeq 1.4\times 10^{-3}$ and $p(5,50) \simeq 3.3\times 10^{-6}$, $p(5,125) \simeq 1.38\times 10^{-3}$. The cumulative probabilities of large number of defaults are larger, compared with the Gaussian copula case. The probability of the names defaulting altogether is also quite large, corresponding to some kind of Armageddon risk. Once again these figures need to be considered with caution, corresponding to high senior and super-senior tranche premiums and disputable assumptions about the probability of all names defaulting.

Figure 13 shows the loss intensities calibrated onto market inputs compared with the loss intensities based on Gaussian copula inputs up to 39 defaults. As can be seen, the loss intensity increases much quicker with the number of defaults as compared with the Gaussian copula approach. The average relative change in the

[39] We also computed the number of defaults distribution using entropic calibration. Although we could still compute loss intensities, the pattern with respect to the number of defaults was not monotonic. Depending on market inputs, direct calibration onto CDO tranche quotes can lead to shaky figures.

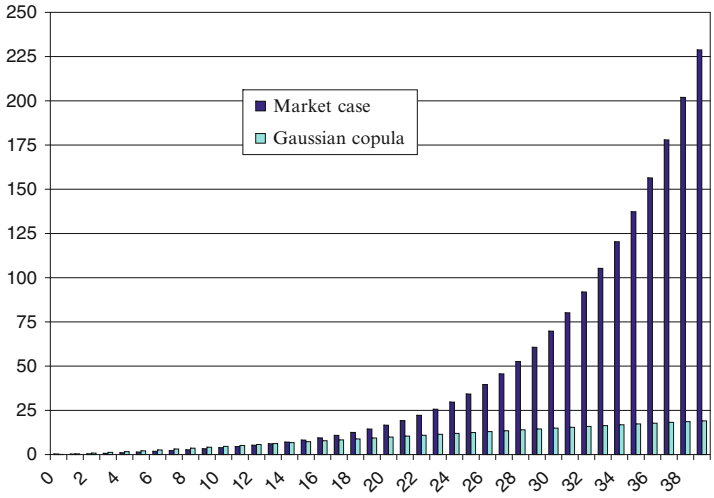


Fig. 13 Loss intensities for the Gaussian copula and market case examples. Number of defaults on the x-axis

loss intensities is equal to 19% when it is equal to 16% when computed under the Gaussian copula assumption. Unsurprisingly, a steep base correlation curve is associated with fatter upper tails of the loss distribution and magnified contagion effects.

Table 7 shows the dynamics of the credit default swap index spreads $\kappa^I(i,k)$ along the nodes of the tree. As for tree implementation, the time step is still $\Delta = \frac{1}{365}$. Let us remark that up to 12 defaults, loss intensities calibrated from market inputs are on the whole smaller than in the Gaussian copula case. Then, the contagion effect is smaller than in the flat 30% correlation Gaussian copula in low default states and greater for high default states. Unsurprisingly, market quotes lead to smaller index spreads up to 2 defaults at 14 weeks (see Tables 1 and 6). This is also coherent with Fig. 14 where the conditional expected losses in the two approaches cross each other at the third default. However, as mentioned above, this detailed pattern has to be considered with caution, since it involves the probabilities of 0, 1 and 2 defaults which are not directly observed in the market. After 2 defaults, credit spreads become definitely larger when calibrated from market inputs.

Thanks to Fig. 14 we can investigate the credit spread dynamics when using market inputs. We plotted the conditional (with respect to the number of defaults) expected loss $E\left[L_T \,|N_t\right]$ for $T = 5$ years and $t = 14$ weeks for the previous market inputs and for the 30% flat correlation Gaussian copula case. The conditional expected loss is expressed as a percentage of the nominal of the portfolio.[40] We also

[40] Thus, given a recovery rate of 40%, the maximum expected loss is equal to 60%.

Table 6 Dynamics of credit default swap index spread $\kappa^I(i,k)$ in basis points per annum

	Weeks			
Nb defaults	0	14	56	84
0	20	19	17	16
1	0	31	23	20
2	0	95	57	43
3	0	269	150	98
4	0	592	361	228
5	0	1,022	723	490
6	0	1,466	1,193	905
7	0	1,870	1,680	1,420
8	0	2,243	2,126	2,423
9	0	2,623	2,534	2,423
10	0	3,035	2,939	2,859

Table 7 Delta of the [0–3%] equity tranche with respect to the credit default swap index

	Outstanding	Weeks			
Nb defaults	nominal	0	14	56	84
0	3.00%	0.645	0.731	0.953	1.038
1	2.52%	0	0.329	0.584	0.777
2	2.04%	0	0.091	0.197	0.351
3	1.56%	0	0.023	0.045	0.090
4	1.08%	0	0.008	0.011	0.018
5	0.60%	0	0.004	0.003	0.004
6	0.12%	0	0.001	0.001	0.001
7	0.00%	0	0	0	0

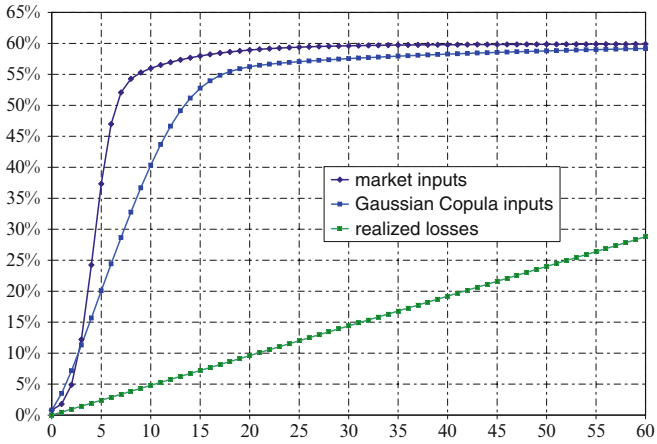


Fig. 14 Expected losses on the credit portfolio after 14 weeks over a 5 year horizon (y-axis) with respect to the number of defaults (x-axis) using market and Gaussian copula inputs

plotted the accumulated losses on the portfolio. The expected losses are greater than the accumulated losses due to positive contagion effects. There are some dramatic differences between the Gaussian copula and the market inputs examples. In the

Gaussian copula case, the expected loss is almost linear with respect to the number of defaults in a wide range (say up to 15 defaults). The pattern is quite different when using market inputs with huge non linear effects. This shows large contagion effects after a few defaults as can also be seen from Table 6 and Fig. 13. This rather explosive behavior was also observed by Herbertsson [35], Tables 3 and 4 and by Cont and Minca [13], Figs. 1 and 3. In Lopatin and Misirpashaev [48], the contagion effects are also magnified when using market data, compared with Gaussian copula inputs.

Table 7 shows the dynamic deltas associated with the equity tranche. We notice that the credit deltas drop quite quickly to zero with the occurrence of defaults. This is not surprising given the surge in credit spreads and dependencies after the first default (see Fig. 14): after only a few defaults the equity tranche is virtually exhausted.

It is noteworthy that the credit deltas $\delta(i,k)$ can be decomposed into a default leg delta $\delta_d(i,k)$ and a premium leg delta $\delta_p(i,k)$ as follows:

$$\delta(i,k) = \delta_d(i,k) - \kappa\delta_p(i,k), \tag{46}$$

where:

$$\delta_d(i,k) = \frac{D\left(i+1,k+1\right) - D\left(i+1,k\right) + O(k) - O(k+1)}{V^I\left(i+1,k+1\right) - V^I\left(i+1,k\right) + D^I(i,k)} \tag{47}$$

and

$$\delta_p(i,k) = \frac{P\left(i+1,k+1\right) - P\left(i+1,k\right) + \left(O(k) - O(k+1)\right) 1_{t_{i+1}\notin\{T_1,\ldots,T_p\}}\left(t_{i+1} - T_l\right)}{V^I\left(i+1,k+1\right) - V^I\left(i+1,k\right) + D^I(i,k)}. \tag{48}$$

Tables 8 and 9 detail the credit deltas associated with the default and premium legs of the equity tranche. As can be seen from Table 7, credit deltas for the equity tranche may be slightly above one when no default has occurred. Table 9 shows that this is due to the amortization scheme of the premium leg which is associated with significant negative deltas. Let us recall that premium payments are based on the outstanding nominal. Arrival of defaults thus reduces the commitment to pay. Furthermore, the increase in credit spreads due to contagion effects involves a decrease in the expected outstanding nominal. When considering the default leg only, we are led to credit deltas that actually remain within the standard 0–100% range. The default leg delta of the equity tranche with respect to the credit default swap index is initially equal to 54.1%. Let us also remark that credit deltas of the default leg gradually increase with time which is consistent with a decrease in time value.

Table 8 Delta of the default leg of the [0–3%] equity tranche with respect to the credit default swap index ($\delta_d(i,k)$)

Nb defaults	Outstanding nominal	Weeks 0	14	56	84
0	3.00%	0.541	0.617	0.823	0.910
1	2.52%	0	0.279	0.510	0.690
2	2.04%	0	0.072	0.166	0.304
3	1.56%	0	0.016	0.034	0.072
4	1.08%	0	0.004	0.006	0.012
5	0.60%	0	0.002	0.002	0.002
6	0.12%	0	0.001	0.000	0.000
7	0.00%	0	0	0	0

Table 9 Deltas of the premium leg of the [0–3%] equity tranche with respect to the credit default swap index ($\kappa\delta_p(i,k)$)

Nb defaults	Outstanding nominal	Weeks 0	14	56	84
0	3.00%	−0.104	−0.113	−0.130	−0.128
1	2.52%	0	−0.050	−0.074	−0.087
2	2.04%	0	−0.018	−0.031	−0.047
3	1.56%	0	−0.007	−0.011	−0.018
4	1.08%	0	−0.004	−0.004	−0.006
5	0.60%	0	−0.002	−0.002	−0.002
6	0.12%	0	−0.001	0.000	0.000
7	0.00%	0	0	0	0

Table 10 Market delta spreads and model deltas (a default event) at inception

Tranches	[0–3%]	[3–6%]	[3–9%]	[9–12%]	[12–22%]
Market deltas	27	4.5	1.25	0.6	0.25
Model deltas	21.5	4.63	1.63	0.9	0.6

3.5.7 Comparison with Standard Market Practice

We further examine the credit deltas of the different tranches at inception. These are compared with the deltas as computed by market participants under the previous base correlation structure assumption (see Table 10). These market deltas are calculated by bumping the credit curves by 1 basis point and computing the changes in present value of the tranches and of the credit default swap index. Once the credit curves are bumped, the moneyness varies, but the market practice is to keep constant the base correlations when recalculating the CDO tranches. This corresponds to the so-called "sticky strike" rule. The delta is the ratio of the change in present value of the tranche to the change in present value of the credit default swap index divided by the tranche's nominal. For example, a credit delta of an equity tranche previously equal to one would now lead to a figure of 33.33.

First of all we can see that the outlines are roughly the same, which is already noticeable since the two approaches are completely different. Then, we can remark that the model deltas are smaller for the equity tranche as compared with the market deltas, while there are larger for the other tranches.

These discrepancies can be understood from the dynamics of the dependence between defaults embedded in the Markovian contagion model. Figure 15 shows the base correlation curves at a 14 weeks horizon, when the number of defaults is

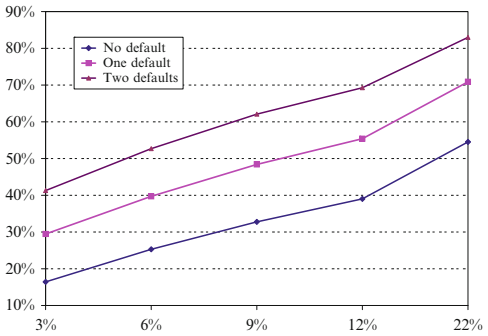


Fig. 15 Dynamics of the base correlation curve with respect to the number of defaults. Detachment points on the x-axis. Base correlations on the y-axis

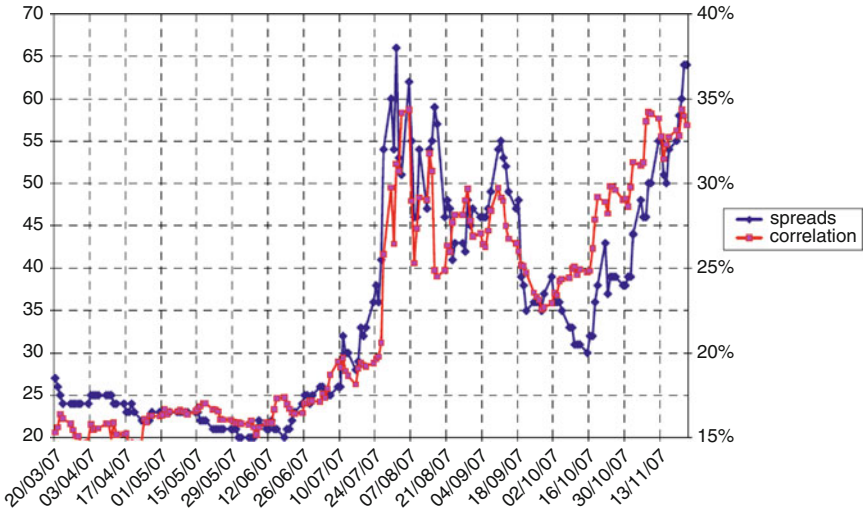


Fig. 16 Credit spreads on the 5 years iTraxx index (Series 7) in bps on the *left* axis. Implied correlation on the equity tranche on the *right* axis

equal to zero, one or two. We can see that the arrival of the first defaults is associated with parallel shifts in the base correlation curves. This increase in dependence counterbalances the increase of credit spreads and expected losses on the equity tranche and lowers the credit delta. The model deltas can be thought of as the "sticky implied tree" model deltas of Derman [22]. These are suitable in a regime of fear corresponding to systematic credit shifts.

The summer 2007 credit crisis provides some evidence that implied correlations tend to increase with credit spreads and thus with expected losses. Figure 16 shows the dynamics of the 5 year iTraxx credit spread and of the implied correlation of the equity tranche. Over this period the correlation between the two series was equal to 91%. This clearly favors the contagion model and once again suggests a flaw in the "sticky strike" market practice.

3.5.8 Comparison with Deltas Computed in Other Dynamic Credit Risk Models

We also thought that it was insightful to compare our model deltas and the results provided by Arnsdorf and Halperin [1], Fig. 7 (see Table 11).

The market conditions are slightly different since the computations were done in March 2007, thus the maturity is slightly smaller than 5 years. The market deltas are quoted deltas provided by major trading firms. We can see that these are quite close to the previous market deltas since the computation methodology involving Gaussian copula and base correlation is quite standard. The BSLP[41] model deltas (corresponding to "model B" in [1]) have a different meaning from ours: there are related to credit spread deltas rather that then default risk deltas and are not related to a dynamic replicating strategy. However, it is noteworthy that the model deltas in [1] are quite similar to ours, and thus rather far away from market deltas. Though this is not a formal proof, it appears from Fig. 11, that (systemic) gammas are rather small prior to the first default. If we could view a shock on the credit spreads as a small shock on the expected loss while a default event induces a larger shock (but not so large given the risk diversification at the index level) on the expected loss, the similarity between the different model deltas are not so surprising. As above, model deltas are lower for the equity tranche and larger for the other tranches, when compared with market deltas.

We also compare our model deltas with credit deltas obtained by Eckner [26], Table 5. Eckner model relies on an affine specification of default intensities (AJD model). Conditionally on the path of default intensities, default times are independent (i.e., there are no contagion effects at default times). The model is parametric with respect to the term structure of credit spreads and to CDO tranches. Calibration of the model parameters to credit spreads and liquid tranche quotes on the CDX NA IG5 index in December 2005 is provided and hedge ratios with respect to the credit default swap index are then computed. The sensitivities of CDO tranche and index prices are computed with respect to a uniform and relative shift of individual intensities. The approach can be extended in order to compute different hedge ratios with respect to the single name default swaps. However, the overall procedure, including the calibration and the computation of individual hedge ratios is likely to be rather involved.

In Table 12, the deltas obtained in the AJD intensity model can be compared with those computed from the Gaussian copula model and those computed within a contagion model calibrated to the same data set.

Table 11 Market and model deltas as in Arnsdorf and Halperin [1]

Tranches	[0–3%]	[3–6%]	[3–9%]	[9–12%]	[12–22%]
Market deltas	26.5	4.5	1.25	0.65	0.25
BSLP model deltas	21.9	4.81	1.64	0.79	0.38

[41] Bivariate spread-loss portfolio model.

Table 12 Market deltas, "intensity" model credit deltas in Eckner [26] and contagion model deltas

Tranches	[0–3%]	[3–7%]	[7–10%]	[10–15%]	[15–30%]
Market deltas	18.5	5.5	1.5	0.8	0.4
AJD model deltas	21.7	6.0	1.1	0.4	0.1
Contagion model deltas	17.9	6.3	2.5	1.3	0.8

Even though the approaches are completely different, once again the outlines are quite similar. Moreover, we can remark that the equity tranche deltas computed by Eckner are higher compared with the market deltas.

Another empirical comparison of various hedging strategies has recently been proposed by Cont and Kan [12]. This study provides several interesting observations related to the hedging of index CDO tranches, extending the ones presented in this section.

4 Conclusion

In this section, we were able to show that a CDO tranche payoff can be perfectly replicated with a self-financed strategy based on the underlying credit default swaps. This extends to any payoff which depends only upon defaults arrivals, such as basket default swaps, but does not address the issue of tranche options for instance. Clearly, the previous replication result is model dependent and relies on two critical assumptions. First, we preclude the possibility of simultaneous defaults. In other words, default times can be ordered from the first to the last default time. Hedging against simultaneous defaults would require trading credit default swaps contingent on several defaults, which are not currently traded in the market. The other important assumption, which is likely to be more questionable, is that credit default swap premiums are adapted to the filtration of default times, denoted $\mathbb{H}$, which therefore can be seen as the relevant information set on economic grounds. As a consequence, default swap premiums are deterministic between two default times. Our framework corresponds to a pure contagion model, where the arrival of defaults leads to jumps in the credit spreads of survived names, the magnitude of which depending upon the considered names and the whole history of defaults up to the current time. These jumps can be related to the derivatives of the joint survival function of default times. The dynamics of replicating prices of CDO tranches follows the same way. In other words, we only deal with default risks and not with spread risks. At a given point in time, there are only p sources of risk, related to the default occurrence of the p non defaulted names and we can trade the corresponding p credit default swaps. This provides the intuition of the completeness of the market following the rule of the thumb, "as many hedging instruments as sources of risk." The hedging strategy deals thus with default risks only and not with credit spread risks. Even though the underlying assumptions may look too restrictive, given the risk management and regulatory issues related to CDOs, we think that it may prove useful to rely

on benchmark models where the hedging can be fully described and analyzed in a dynamical way.

Unsurprisingly, the possibility of perfect hedging is associated with a martingale representation theorem under the filtration of default times. Subsequently, we exhibit a new probability measure under which the short term credit spreads (up to some scaling factor due to positive recovery rates) are the intensities associated with the corresponding default times. For the ease of presentation, we introduced first some instantaneous default swaps as a convenient basis of hedging instruments. Eventually, we can exhibit a replicating strategy of a CDO tranche payoff with respect to actually traded credit default swaps, for instance, with the same maturity as the CDO tranche. Let us note that no Markovian assumption is required for the existence of such a replicating strategy. Therefore the aggregate loss may not be a Markov process either. Since we dealt first with the dynamics of individual defaults, we are typically in a bottom-up model and no homogeneity assumption, such as equal credit spreads across names is required.

However, when going to implementing actual hedging strategies, one needs extra assumptions, both for the implementation to be feasible and to cope with quoted CDO tranches. We therefore consider the simplest way to specialize the above model: we assume that all pre-default intensities are equal and only depend on the current number of defaults. We also assume that all recovery rates are constant across names and time. In that framework, it can be shown that the aggregate loss process is a homogeneous Markov chain, more precisely a pure death process (thanks to the no simultaneous defaults assumption). The intensity associated with the Markov chain is simply the pre-default intensity times the number of non-defaulted names. Thanks to these restrictions, the model involves as many unknown parameters as the number of underlying names. On the other hand, the knowledge of upfront premiums of equity CDO tranches with different maturities and detachment points (and given some recovery rate) is equivalent to the knowledge of marginal distributions of the number of defaults at different time horizons. Thanks to the forward Kolmogorov equations, one can then perfectly compute the intensities of the aggregate loss process or the pre-default intensities of the names. Such fully calibrated and Markov model is also known as the local intensity model, the simplest form of aggregate loss models. As in local volatility models in the equity derivatives world, there is a perfect match of unknown parameters from a complete set of CDO tranches quotes. In other words, the model is fully specified from market inputs, which is clearly a desirable property, since given some market inputs, we deal with a single model and not with a family of parameterized models. The numerical implementation can be achieved through a binomial tree, well-known to finance people or by means of Markov chain techniques. We provide some examples and show that the market quotes of CDOs are associated with pronounced contagion effects. We can therefore explain the dynamics of the amount of hedging CDS and relate them to deltas computed by market practitioners. The figures are hopefully roughly the same, the discrepancies being mainly explained by contagion effects leading to an increase of dependence between default times after some defaults.

However, one cannot unfortunately observe a complete set of CDO tranche premiums. The set of local intensities consistent with the actually CDO tranches quotes is not a singleton. For a complete specification, one needs to introduce some extra assumptions: either, one can constrain the shape of intensities, for instance assume that there are piecewise constant with respect to the number of defaults with shifts associated to detachment points of traded tranches. Otherwise, as an intermediary step, we may think of fitting some marginal distributions of aggregate losses to CDO tranche quotes or use interpolation techniques consistent with the increase and concavity of the expected loss on equity tranches. Numerical examples in this section are constructed under the second approach. Unfortunately, for practical purpose, the computed deltas and thus hedging performance seem rather sensitive to the calibration technique.

One may compare the proposed framework with the standard structural approach, where default time of a given name is the first hitting time of a barrier by a Brownian motion associated with the asset process of the corresponding name. In that structural approach, dependence between default times stems from the correlation between the Brownian motions. In the latter framework, quite similar to a multivariate Black–Scholes setting, CDS are barrier-options and it is also possible to replicate a CDO tranche payoff by dynamically trading the CDS. While the former Markov chain approach focused on default risk, neglecting credit spread risk, the structural approach only deals with credit spread risk. Defaults are predictable and do not constitute an extra source of risk. On the other hand, a structural model can be well approximated in most cases by a one period structural model, where crossing the default barrier is only considered at maturity. This is known to be equivalent to the Gaussian copula model commonly used by practitioners. As mentioned above, an interesting feature is that the deltas with respect to underlying credit default swaps have the same order of magnitude in the two approaches.

However, extending the scope of the approach would result in adding extra complexity, both on mathematical grounds and regarding the specification of credit spreads dynamics. For instance, if we were to introduce some Brownian risks on top of jump to default risks, it is not clear how defaults would drive the volatility of credit spreads. The uncertainty with respect to this substantial model risk is likely to offset the benefit of dealing with credit spread and default risk altogether. At the time being, extra-complexity conveys the risk of darkening the risk management picture and providing a false sense of security. A better understanding of the multivariate dynamics of defaults and credit spreads is required before going any further. Another, more down to earth issue, but of practical importance is related to the set of hedging instruments. Given n names, one can think of using two credit default swaps of different maturities for each underlying name to cope both with default and credit spread risks. This induces extra complexity in implementing hedging strategies.

A more easy to reach extension of the previous framework consists in relaxing the homogeneity of names assumption, while remaining in a pure default setting. For instance, one could think of two homogeneous groups of names, say belonging to two different geographical regions, the intensities depending both upon the number of survived names in each group. This results in a two dimensional Markov

chain, since the portfolio state is characterized by the number of survived names in each group. We should then be able to discriminate CDS deltas for names within each subgroup. Let us note that given that we rely upon a bottom-up approach, once calibrated onto liquid CDO tranche quotes, one would be able to consistently price CDO tranches on any sub-portfolio, thus solving the difficult issue of bespoke tranche pricing.

Another possible and easy to implement extension of our setting consists in using a recovery rate depending upon the number of defaults. The easiest way to proceed is to assume some linear (and most likely negative) dependence with respect to the number of defaults in the portfolio. Such assumption will tend to raise the probability of large losses and ease the calibration to the senior tranches.

Eventually, we would like to stress that the approach described in this section should be fruitful in computing so called values on defaults. These assess the magnitude of losses on a portfolio (possibly including CDO tranches) after a default occurs. Usually, market practitioners do not take into account credit contagion effects associated with shifts of credit spreads of survived names, which can lead to gross misestimation of credit risk reserves. This can be easily dealt with in our framework.

Acknowledgements This research was supported by CRIS (Credit Risk Service) collaborative project for which we are very grateful. The authors also thank Marek Rutkowski for useful discussions and comments.

References

1. Arnsdorf, M., Halperin, I.: BSLP: Markovian bivariate spread-loss model for portfolio credit derivatives. Working paper, JP Morgan (2007)
2. Arvanitis, A., Laurent, J.-P.: On the edge of completeness. RISK **10**, 61–65 (1999)
3. Assaf, D., Langberg, N.A., Savits, T.H., Shaked M.: Multivariate phase-type distributions. Oper. Res. **32**(3), 688–702 (1984)
4. Azizpour, S., Giesecke, K.: Self-exciting corporate defaults: Contagion vs. frailty. Working paper, Stanford University (2008)
5. Beumee, J., Brigo, D., Schiemert, D., Stoyle, G.: Charting a course through the CDS Big Bang. Quantitative research, FitchSolutions (2009)
6. Bielecki, T.R., Rutkowski, M.: Credit Risk: Modeling, Valuation and Hedging. Springer, Berlin (2004)
7. Bielecki, T.R., Jakubowski, J., Vidozzi, A., Vidozzi, L.: Study of dependence for some stochastic processes. Working paper (2007)
8. Bielecki, T.R., Jeanblanc, M., Rutkowski, M.: Hedging of basket credit derivatives in default swap market. J. Credit Risk **3**(1), 91–132 (2007)
9. Blanchet-Scalliet, C., Jeanblanc, M.: Hazard rate for credit risk and hedging defaultable contingent claims. Finance Stochast. **8**, 145–159 (2004)
10. Brémaud, P.: Point Processes and Queues: Martingale Dynamics. Springer, Berlin (1981)
11. Brigo, D., Pallavicini, A., Torresetti, R.: Calibration of CDO tranches with the dynamical generalized-Poisson loss model. Working paper, Banca IMI (2007)
12. Cont, R., Kan, Y.H.: Dynamic hedging of portfolio credit derivatives in top-down models. Financial Engineering Report 2008-08, Columbia University (2008)

13. Cont, R., Minca, A.: Recovering portfolio default intensities implied by CDO quotes. Financial Engineering Report 2008-01, Columbia University (2008)
14. Cont, R., Deguest, R., Kan, Y.H.: Default intensities implied by CDO spreads: Inversion formula and model calibration. SIAM Journal on Financial Mathematics **1**, 555–585 (2010)
15. Das, S.: Credit Derivatives: CDOs and Structured Credit Products. Wiley, NY (2005)
16. Das, S.R., Duffie, D., Kapadia, N., Saita, L.: Common failings: How corporate defaults are correlated. J. Finance **62**(1), 93–117 (2007)
17. Davis, M., Lo, V.: Infectious defaults. Quant. Finance **1**, 382–387 (2001)
18. De Koch, J., Kraft, H.: CDOs in chains. Working paper, University of Kaiserslautern (2007)
19. De Servigny, A., Jobst, N.: The Handbook of Structured Finance. McGraw-Hill, NY (2007)
20. Deguest, R., Kan, Y.H.: A closed-form formula for the local intensity and its application to calibration. Working paper (2009)
21. Dellacherie, C.: Capacités et processus stochastiques. Springer, Berlin (1972)
22. Derman, E.: Regimes of volatility. Quantitative Strategies Research Notes, Goldman Sachs (1999)
23. Derman, E., Kani, I.: Riding on a smile. RISK **7**(2), 32–39 (1994)
24. Ding, X., Giesecke, K., Tomecek, P. Time-changed birth processes and multi-name credit. Oper. Res. **57**(4), 990–1005 (2009)
25. Dupire, B.: Pricing with a smile. RISK **7**(1), 18–20 (1994)
26. Eckner, A.: Computational techniques for basic affine models of portfolio credit risk. J. Comput. Finance **13**, 63–97 (2009)
27. Elliott, R.J., Jeanblanc, M., Yor, M.: On models of default risk. Math. Finance **10**, 179–196 (2000)
28. Epple, F., Morgan, S., Schloegl, L.: Joint distributions of portfolio losses and exotic portfolio products. Int. J. Theor. Appl. Finance **10**(4), 733–748 (2007)
29. Errais, E., Giesecke, K., Goldberg, L.R.: Pricing credit from the top down with affine point processes. Working paper, Stanford University (2007)
30. Frey, R., Backhaus, J.: Dynamic hedging of synthetic CDO tranches with spread and contagion risk. J. Econ. Dynam. Contr. **34**, 710–724 (2010)
31. Frey, R., Backhaus, J.: Pricing and hedging of portfolio credit derivatives with interacting default intensities. Int. J. Theor. Appl. Finance **11**(6), 611–634 (2008)
32. Galiani, S., Kakodkar, A., Shchetkovskiy, M., Jónsson, J.G., Gallo, A.: Base correlation and beyond: Interpolation and extrapolation. Working paper, Credit Derivatives Strategy, Merrill Lynch (2006)
33. Gibson, M.S.: Credit derivatives and risk management. Finance and Economics Discussion Series, Federal Reserve Board, Washington, DC (2007)
34. Giesecke, K., Goldberg, L.R.: A top down approach to multi-name credit. Working paper, Stanford University (2006)
35. Herbertsson, A.: Default contagion in large homogeneous portfolios. In: Gregoriou, G.N., Ali, P.U. (eds.) The Credit Derivatives Handbook. McGraw-Hill, NY (2007)
36. Herbertsson, A.: Pricing synthetic CDO tranches in a model with default contagion using the matrix-analytic approach. J. Credit Risk **4**(4), 3–35 (2008)
37. Herbertsson, A., Rootzén, H.: Pricing k-th to default swaps under default contagion, the matrix-analytic approach. J. Comput. Finance **12**(1), 49–78 (2006)
38. Jarrow, R.A., Yu, F.: Counterparty risk and the pricing of defaultable securities. J. Finance **56**, 1765–1799 (2001)
39. Kakodkar, A., Galiani, S., Jónsson, J.G., Gallo, A.: Credit Derivatives Handbook 2006, vol. 2. Merrill Lynch, NY (2006)
40. Karlin, S., Taylor, H.M.: A First Course in Stochastic Processes, 2nd edn. Academic Press, NY (1975)
41. Krekel, M., Partenheimer, J.: The implied loss surface of CDOs. Working paper, University of Ulm (2006)
42. Kusuoka, S.: A remark on default risk models. Adv. Math. Econ. **1**, 69–82 (1999)
43. Lando, D.: Credit Risk Modeling and Applications. Princeton University Press, NJ (2004)

44. Lando, D., Nielsen, M.S.: Correlation in corporate defaults: Contagion or conditional independence. J. Financ. Intermed. **19**(3), 355–372 (2010)
45. Laurent, J.-P.: A note on the risk management of CDOs. Working paper, ISFA actuarial School, University of Lyon and BNP Paribas (2006)
46. Laurent, J.-P., Leisen, D.: Building a consistent pricing model from observed option prices. In: Quantitative Analysis in Financial Markets: Collected Papers of the New York University Mathematical Finance Seminar, vol. II, pp. 216–238. World Scientific, Singapore (2000)
47. Laurent, J.-P., Cousin, A., Fermanian, J.-D.: Hedging default risk of CDOs in Markovian contagion models. Quant. Finance (2010)
48. Lopatin, A.V., Misirpashaev, T.: Two-dimensional Markovian model for dynamics of aggregate credit loss. Working paper, NumeriX (2007)
49. Meissner, G.: The Definitive Guide to CDOs. Riskbooks, London (2008)
50. Meyer-Dautrich, S., Wagner, C.K.J.: Minimum entropy calibration of CDO tranches. Working paper, Unicredit (2007)
51. Moler, C., Van Loan, C.: Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. SIAM Rev. **45**(1), 3–49 (2003)
52. Morgan, S., Mortensen, A.: CDO hedging anomalies in the base correlation approach, pp. 49–61. Lehman Brothers, Quantitative Credit Research Quarterly, October (2007)
53. Parcell, E., Wood, J.: Wiping the smile off your base (correlation curve). Working paper, Derivative Fitch (2007)
54. Parker, E., Brown, M.: The 2003 ISDA credit derivatives definitions. PLC Finance, NY (2003)
55. Petrelli, A., Zhang, J., Jobst, N., Kapoor, V.: A practical guide to CDO trading risk management. In: de Servigny, A., Jobst, N. (eds.) The Handbook of Structured Finance, pp. 339–371. McGraw Hill, NY (2007)
56. Protter, Ph.: Stochastic Integration and Differential Equations, 2nd edn. Springer, Berlin (2005)
57. Putyatin, P., Prieul, D., Maslova, S.: A Markovian approach to modelling correlated defaults. RISK **18**(5), 76–81 (2005)
58. Rubinstein, M.: Implied binomial trees. J. Finance **49**(3), 771–818 (1994)
59. Schönbucher, P.: Credit Derivatives Pricing Models. Wiley, Chichester (2003)
60. Schönbucher, P.J.: Portfolio losses and the term-structure of loss transition rates: A new methodology for the pricing of portfolio credit derivatives. Working paper, ETH Zürich (2006)
61. Schönbucher, P., Schubert, D.: Copula dependent default risk in intensity models. Working paper, Bonn University (2001)
62. Torresetti, R., Brigo, D., Pallavicini, A.: Implied expected tranched loss surface from CDO data. Working paper, Banca IMI (2007)
63. van der Voort, M.: An implied loss model. Working paper, ABN Amro and Erasmus University (2006)
64. Walker, M.B.: Simultaneous calibration to a range of portfolio credit derivatives with a dynamic discrete-time multi-step Markov loss model. Int. J. Theor. Appl. Finance **12**(5), 633–662 (2009)
65. Yu, F.: Correlated defaults in intensity-based models. Math. Finance **17**(2), 155–173 (2007)

About the Pricing Equations in Finance

Stéphane Crépey

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 of a Markovian model made of a jump-diffusion – like component X interacting with a continuous-time Markov chain – like component N. 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 establish the *Markov property* of the 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 the 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 the 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 related integro-differential obstacle problems. We further

S. Crépey (✉)
Département de Mathématiques, Université d'Évry Val d'Essonne, 91025 Évry Cedex, France
e-mail: stephane.crepey@univ-evry.fr

A.R. Carmona et al. (eds.), *Paris-Princeton Lectures on Mathematical Finance 2010*, Lecture Notes in Mathematics 2003, DOI 10.1007/978-3-642-14660-2_2,

© Springer-Verlag Berlin Heidelberg 2011

establish a *comparison principle* for semi-continuous viscosity solutions to these problems, which implies in particular the *uniqueness* of the viscosity solutions. This comparison principle is subsequently used for proving the convergence of *stable, monotone and consistent* approximation schemes to the 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 a financial derivative or on the underlying asset, as well as various forms of *discrete path-dependence*.

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, partial integro-differential equations or variational inequalities (PIDEs or PDEs for short).

In Parts II–IV, we tackle the resulting Markovian BSDE and PDE problems. In Crépey and Matoussi [38], 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 *Markovian reflected BSDEs*, which is used in Part III for studying the associated partial integro-differential systems of 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 for these problems. 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 30 and 31, which synthesize the major findings of Part II and III, respectively.

This paper lays the mathematical foundation of a large body of work in credit risk and financial modeling [15, 16, 20, 39]. Even if rather expected in their statement, many of the mathematical results derived in Parts II–IV are innovative. In particular, doubly reflected BSDEs with a delayed or an even more general intermittent upper barrier (RDBSDEs and RIBSDEs, see Definitions 9(ii) and 16), have not been considered elsewhere in the literature (if not for the preliminary results of Crépey and Matoussi [38]). Also, the Markovian model which is considered in detail in Parts II and III was already considered and some of the results of the present paper were

already announced and used in [16, 20, 38]. But the possibility to construct a model with all the required properties was taken for granted there. The mathematical construction of the model in Sect. 7 is non-trivial, and was not done elsewhere before. The treatment of the Markovian BSDEs with jumps and of their PDE interpretation in Parts II and III, including the proof of convergence of a numerical deterministic scheme to the viscosity solution of a system of integro-differential variational inequalities, is quite technical too.

As for Part I, we believe that, beyond providing the motivation for the mathematical results of Parts II–IV, it also has the merit of giving a unified, cross market perspective (see Sects. 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–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–IV whenever they are used therein (see Propositions 5, 6, 8, 14 and 16 in particular). Likewise readers mainly interested by the mathematical results of Parts II–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. [46] for a general reference on BSDEs in finance). The central result, Proposition 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, relatively 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 Sect. 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 Sect. 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 relatively to an arbitrary *numeraire* B (positive primary asset price process) which may be used for discounting other price processes, rather than a savings account (riskless asset) 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. [20], refines the risk-neutral martingale modeling approach of Sects. 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 Sect. 4.3 we deal with the issue of callability and *call protection* (*intermittent* call protection vs. call protection *before a stopping time*).

In Part I, well-posedness of the pricing BSDEs and PDEs is taken for granted. The following sections of the paper (Parts II–IV) are devoted to the mathematics of these pricing equations.

In Sect. 5 we recall the general set-up of [38] and the general form of the BSDEs we are interested in.

In Sect. 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 *portfolio credit risk* (see [16]).

But the construction of a model with such mutual dependence is a non-trivial issue, and we treat it in detail in Sect. 7, resorting to a suitable *Markovian change of probability measure*.

This model may also be viewed as a generalization of the interacting 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. This leads us to study in Sect. 8 the related *reflected and doubly reflected Markovian BSDEs* (see [20, 46, 47]), showing that they are *well-posed* in the sense that they have *unique solutions, which depend continuously on their input data.*

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

In Sect. 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 Sect. 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 Sect. 12 a *semi-continuous viscosity solutions comparison principle*, which implies in particular *uniqueness* of viscosity solutions for these problems.

This comparison principle is subsequently used in Sect. 13 for proving the convergence of *stable, monotone and consistent* approximation schemes (cf. Barles and Souganidis; see also [8] Briani et al. [28], Cont and Voltchkova [36] or Jakobsen et al. [64]) to the viscosity solutions of the equations. These results thus extend to models with regimes (whence *systems* of PDEs [9, 60]) the results of [8, 28], among others.

In Sects. 14–16 we provide extensions of the previous results to a factor process model (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 Sect. 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, PDEs. 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 $\mathcal{F}_T = \mathcal{F}$. 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.

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 Sect. 4.1), such that the prices of primary assets, once properly discounted and adjusted for dividends, are $\mathbb{P}$ – local martingales.

As we shall now see, under mild technical conditions, existence of such a martingale measure $\mathbb{P}$ is equivalent to a suitable notion of no-arbitrage.

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 $\beta_0 = 1$ and $\beta = B^{-1}$), 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 [42]). 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 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, where ζ^0 and the row-vector ζ represent the number of units held in the savings account and in each of the primary risky assets. 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

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

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 $t \in [0, T]$:

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

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

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

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

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), made of an initial wealth $w \in \mathbb{R}$ and an $\mathbb{R}^{1\otimes d}$-valued predictable locally bounded primary strategy in the risky assets ζ, with related wealth process $\mathcal{W}$ defined by (5).

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 below, some or all of the following cash flows (or payoffs):

- A bounded variation cumulative *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 $L \leq U$ and $L_T \leq \xi \leq U_T$.

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 (only 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. Sect. 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 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 1. 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 Sect. 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 (*put or maturity* time, to be precise) τ, and the call (or maturity) time σ, represent stopping times at the holder's and at the issuer's convenience, respectively.

Definition 2. **(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 1. 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 $\xi = (S_T - K)^{\pm}$.

Note 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

[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.

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 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 Sect. 4.2.1), 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 Sect. 4.3, building on the mathematical results of Sect. 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.

By classic arbitrage theory (see, e.g., [18, 32, 42]), 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.

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 primary market and the derivative*. The notion of arbitrage price process of a financial derivative referred to in the next result is the classical notion of No Free Lunch with Vanishing Risk condition of Delbaen and Schachermayer [42] in the case of European claims, subsequently extended to game (including American) claims by Kallsen and Kühn [67]. The proof of this result is based on a rather straightforward application of Theorem 2.9 in Kallsen and Kühn [67] (see Bielecki et al. [18] for the details).

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 $(\sigma, \tau) \in \bar{\mathcal{T}}_t \times \mathcal{T}_t$.

Proposition 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*

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

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

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

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

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

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 $\beta L = -c$, where $-c$ is a strict minorant of $\int_t^T \beta_u \, dD_u + \beta_T\xi$. Indeed, in view of Propositions 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 [68], and subsequently followed by El Karoui and Quenez [45], Kifer [69], Ma and Cvitanić [76], Hamadène [55], and, in the context of defaultable derivatives examined in Sect. 4.2, Bielecki et al. [20, 23] (see also Schweizer [85]). This

definition will be later shown to be consistent with the concept of arbitrage pricing of Proposition 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 3. 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 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.

Note 2. **(i)** The process Q is to be interpreted as the cumulative *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 *super-hedges*.
(iii) In relation with admissibility issues (see the end of Definition 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 [50] and *mean self-financing hedge* in Schweizer [85]). *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 *martingale*, or *compensator*, 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. [46] for a general reference in connection with finance and El Karoui et al. [47] for a seminal reference on reflected BSDEs) 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 [38, 56, 57]), 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. It is important to note for applications that it is also possible to deal with discrete dividends: see [20] and Sect. 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 line above,

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

Definition 4. *(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 continuous process null at time 0, 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 processes $K^{\pm}$ null at 0, defining mutually singular random measures on $[0,T]$.

Remark 3. 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 process $\int_{[0,\cdot]} \beta_t^{-1} d(\beta_t \widehat{\Pi}_t)$.

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 [38, 56, 57], 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, Sect. 16.2.2) such that $L \le Y \le U$ on $[0,T]$. Existence and uniqueness of a solution to (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 [38] as well as Note 8(v) and Proposition 30 in Part II). This covers, for instance, the put payment process L of an American vanilla option, or of a convertible bond (see Definition 7 and Bielecki et al. [18, 19]). 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 1 Equation (15) admits a solution (Π, M, K), with K equal to zero in the special case of an European derivative.

Proposition 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 18 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 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 Sect. 2.2. Let us set

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

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

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

Note that for any primary strategy ζ, the issuer's *Profit and Loss* (or *Tracking Error*) process $(e_t)_{t \in [0,T]}$ relative to the price process Π of Proposition 2 is given for $t \in [0,T]$ by:

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

where $\widehat{\Pi}$ is defined by (17), so that, in view of Proposition 2, $\widehat{\Pi}$ can be interpreted as the $\mathbb{P}$ – *cumulative price* of the option (cf. (2)). Observe in view of (18) that the tracking error process e 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 [75] (in the specific contexts of the Cox–Ross–Rubinstein or Black–Scholes models, analogous results can also be found in Kifer [69]).

Proposition 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:

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

or equivalently, using (22):

$$\begin{aligned} &\Pi_0 + \int_0^{\sigma^* \wedge \tau} \beta_u dM_u \\ &\geq \int_0^{\sigma^* \wedge \tau} \beta_u C_u du + \beta_{\sigma^* \wedge \tau} \left(\mathbb{1}_{\{\sigma^* \wedge \tau = \tau < T\}} L_\tau + \mathbb{1}_{\{\sigma^* < \tau\}} U_{\sigma^*} + \mathbb{1}_{\{\sigma^* = \tau = T\}} \xi \right) \end{aligned} \tag{25}$$

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 from (20) and from the following relations, which are valid by the terminal and put conditions in (15):

$$\Pi_T = \xi\,,\ \Pi_\tau \geq L_\tau.$$

(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 $t \in [0, T]$:

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

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:

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

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. □

Note 3. **(i)** Proposition 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 [50] or Schweizer [85].
(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 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 Sect. 4 of [46] and Part II).

Definition 5. 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 is $\bar{\sigma}$ is the first time of entry, capped at T, of the process (t, X), into a given closed subset of $[0, T] \times \mathbb{R}^q$.

Remark 4. 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 11 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 intimately, though non-trivially, related 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 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.

In particular $\delta(t, X_{t-})dN_t$ 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. 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 Sect. 3.3 for specific examples and Definition 10 in Part II for more precise statements.

Remark 5. 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 "mark" of the jump of X 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 vector-valued function u on $\mathbb{R}^q$ and for every $t \in [0, T]$,

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

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

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

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

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

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

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

Using the short-hand $\delta u(t, X_{t-})dN_t = \int_{x \in \mathbb{R}^q} \delta u(t, X_{t-}, x)N(dt, dx)$, note that one has (cf. (30)),

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

The Itô formula (32) may thus be rewritten as

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

where we set

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

The process X is thus a Markov process with generator $\mathcal{G}$ (see Proposition 29 in Part III for a more formal derivation).

Remark 6. 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 [84]). 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, then (38) yields, in view of the Itô formula (35):

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

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

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

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

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

Remark 7. 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, (37) yields by application of the Itô formula (35) to the functions u, v and uv, "$\triangleq$" standing for "equality up to a local martingale term":

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

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

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

Remark 8. The bilinear operator

$$(u,v) \mapsto \Gamma(u,v) = \mathcal{G}(uv) - u\mathcal{G}v - v\mathcal{G}u$$

which appears in the right-hand-side of (41) is known as the *carré du champ operator* associated to $\mathcal{G}$ (see, for instance, Sects. XV.20–26 of Dellacherie and Meyer [43]). In particular, formula (41) above corresponds to formula (22.1) on page 244 of [43].

We are now ready to prove the following,

Proposition 4. *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," one has in probability, for almost every t,*

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

where the subscript t stands for "conditional on $\mathcal{F}_t$."

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

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

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):

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

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 Sect. 6 in Part II), the jump process X is supported by a finite set which can be identified with $E = \{1, \ldots, n\}$, without loss of generality, and X is a continuous-time E-valued Markov chain X such that (cf. (31))

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

The generator $\mathcal{G}$ of X us then given by, 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 yet 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–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 written as (compare with (36))

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

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

Remark 9. 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 [35]):

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

where the $\bar{t}_l$s stand for the successive jump times of the process $t \mapsto N(\bar{B}_1 \times [0,t])$, in which $\bar{B}_1$ 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)):

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

3.3.3 Applications

With such versatile specifications ranging from pure diffusions, or (resorting to unbounded jump measures as explained in Sect. 3.3.2) Lévy processes, to

[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 $y \in \mathbb{R}^q$.

continuous-time Markov chains, the jump-diffusion model factor process model (28) offers a flexible setting which is rich enough 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*, or the most common forms of *Lévy models* used in finance for pricing purposes.

As will be explained in Sect. 4.1, the risk-neutral modeling approach can be readily extended to a martingale modeling approach relatively to an arbitrary *numeraire*, rather than the savings account 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.

Moreover, as we shall see in Sect. 4.2, one can easily accommodate in the risk-neutral (or in a more general martingale) modeling approach *defaultable derivatives* with terminal payoffs of the form $\mathbb{1}_{T<\theta}\varphi(X_T)$ (or $\mathbb{1}_{\nu<\theta}\varphi(X_\nu)$ 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 Sect. 4.2.1). A model X as of (28) is then typically used in the mode of a *pre-default factor process model* (see Sect. 4.2 and [20]).

Finally continuous-time Markov chains, or continuous-time Markov chains modulated by diffusions, which, as illustrated in Sect. 3.3.1 and made precise in Part II (see Sects. 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 et al. [74], Cont and Minca [34] or Herbertsson [59],
- A more general *homogeneous groups model* considered for different purposes by various authors in [22, 30, 53], among others,
- An even more general *basket credit migrations model* of Bielecki et al. [15, 16] in which the dynamics of the credit ratings of reference entities are modulated by the evolution of macro-economic factors, or another generation of *Markovian copula models* of Bielecki et al. [17] 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,

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

with terminal condition $u(T,x)=\xi(x)$. An application of the results of Part III (see Proposition 31(i) therein) yields,

Proposition 5. *Under mild conditions, the variational inequality (double obstacle 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 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 10. 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., [4, 5, 7, 11, 12]),

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

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 3 in Part II).

A further application of the results of Part III (Proposition 31 therein) then yields,

Proposition 6. (i) (Post-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 $\mathbb{P}$-price process Π can be represented as $\Pi_t=\bar u(t,X_t)$, where the function $\bar u$ is the unique viscosity solution of the following variational inequality (lower obstacle problem):*

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

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

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

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

Remark 12. Under more specific assumptions on the structure of X (see, e.g., Sect. 6 in Part II), the generic cascade of two PDEs (51), (55) must be suitably amended. 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, (51) and (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,

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

with the generator $\mathcal{G}$ therein given by (45).

In this article we refer to a decoupled system of partial integro-differential equations or obstacle problems, as to a *cascade* of PDEs. In particular (but not only, cf. above), this terminology will be used for systems consisting of equations defined over successive time intervals $[T_{l-1},T_l]$, in which 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 6, the first line of (15) takes the following form (cf. (18)):

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

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

Let us assume the same structure (without the barriers) on the primary market price process P, so $P_t = v(t,X_t)$ for a deterministic function $v(t,x)$, and

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

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

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

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

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

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

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

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 12), 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

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

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 13. In the context of credit derivatives (see also Sect. 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*, relatively 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 [84, 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):

$$\zeta_t^{va} = \frac{d< \Pi, P >_t}{dt} \left(\frac{d< P >_t}{dt} \right)^{-1} = (\nu, v)((v, v))^{-1}\big(t, X_{t-}\big). \tag{67}$$

Note 4. **(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 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 relatively 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 relatively 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 (4) (see, e.g., Protter [83]), 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 relatively to an arbitrary numeraire. Note that the pricing BSDE (15) (with β therein as mentioned above) to be solved in (Π, M, K), is 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)):

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

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

The conclusions of Propositions 2 and 3 are still valid in this context, provided "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:

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

We then have the following analog to Propositions 5 and 6.

Proposition 8. *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 $t \le \bar{\sigma}$.

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

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

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

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

One then has the following analog to Proposition 7.

Proposition 9. *$\Pi_0 = B_0\nu(t, X_0)$ (with $B_0 = 1$, in our set-up) 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)–(23) may be rewritten as, respectively (with $\rho_0 = 0$):*

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

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

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

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

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

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

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 further 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 7), 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 now consider defaultable derivatives with terminal payoffs of the form $\mathbb{1}_{T<\theta}\varphi(S_T)$ (or $\mathbb{1}_{\nu<\theta}\varphi(S_\nu)$ 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 [73] or Jarrow and Turnbull [65], subsequently generalized in many ways in the credit risk literature (see for instance Bielecki and Rutkowski [14]), and extended in particular to American and game claims in Bielecki et al. [18–20, 23], on which the material of this section is based.

We shall give hardly no proofs in this section, referring the interested reader to [18–20, 23].

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 so-called *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 $[0,T]\cup\{+\infty\}$-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(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 6. 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 $\tau \in \mathcal{T}_t$,

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

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

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

Remark 14. One can also cope with the case of *discrete coupons* (see [18–20, 23] and Sect. 14 in Part IV).

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 convertible bond, we need to introduce some additional notation:

$\bar{N}$: the nominal,
S: the price process of the asset underlying the bond,

$\bar{R}$: the recovery rate process on the bond upon default of the issuer,
η: the loss given default on the underlying asset,
κ : the bond's conversion factor,
$\bar{P}$, $\bar{C}$: the put and call nominal payments, with by assumption $\bar{P} \leq \bar{N} \leq \bar{C}$.

Definition 7. 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

$$R_t^{cb} = (1-\eta)\kappa S_{t-} \vee \bar{R}_t \,,\ \xi^{cb} = \bar{N} \vee \kappa S_T \tag{81}$$
$$L_t^{cb} = \bar{P} \vee \kappa S_t \,,\ U_t^{cb} = \bar{C} \vee \kappa S_t. \tag{82}$$

See [18] 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 1 yields (see Bielecki et al. [23]),

Proposition 10. *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 $t \in [0,T]$:*

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

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., [14]), we assume further that $\mathbb{F} = \mathbb{H} \vee \widetilde{\mathbb{F}}$, where the filtration $\mathbb{H}$ is generated by the *default indicator process* $I_t = \mathbb{1}_{\{\theta \leq t\}}$ and $\widetilde{\mathbb{F}}$ is some *reference filtration*. Moreover, we assume that the optional projection of J, defined by, for $t \in [0,T]$,

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

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

Note 5. **(i)** If Q is continuous, θ is a *totally inaccessible* $\mathbb{F}$ – stopping time (see, e.g., Dellacherie and Meyer [43]). 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. [33]).
(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, whereas 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 [77]).

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 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 1 (see, e.g., Bielecki et al. [23]). (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. $J_-\Pi = J_-\widetilde{\Pi}$ 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 relatively 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 and 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 2 (see Bielecki et al. [23]). *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 $\nu = \tau \wedge \sigma$,

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

in which we set $f = C + \gamma R$.

As a corollary to the previous results, we have,

Proposition 11 (see Bielecki et al. [23]). *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 $t \in [0,T]$,*

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

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

Hence, by Proposition 10, Π is an arbitrage price for the option, with 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 4 for the definition of a solution to (85)):

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

Hence, by Proposition 2, the $\widetilde{\mathbb{F}}$-semimartingale $\widetilde{\Pi}$ solves the $\widetilde{\mathbb{F}}$-Dynkin game with payoff $\widetilde{\pi}$. Thus, by Proposition 11, $\Pi := J\widetilde{\Pi}$ 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 3).

Lemma 3. *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 2), for any $0 \leq t \leq u \leq T$,

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

□

Let

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

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 12 can be seen as an extension of Proposition 3 to the defaultable case, in which two filtrations are involved. Note that our assumptions here are made relatively 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 12 (see Bielecki et al. [20, 23]). (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.*

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 process $\int_{[0,\cdot]} \beta_t^{-1} d(\beta_t \widehat{\Pi}_t)$.

Lemma 4. *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. [23] for the detail). This follows by computations similar to those of the proof of Kusuoka's Theorem 2.3 in [72] (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 recalled at Note 5(i), according to which $\mathbb{P}(\theta = \tau) = 0$ 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 decomposition is the analog, relatively to the primary market, of (90) for the game option.

Lemma 5 (see Bielecki et al. [20]). *Process $\alpha\bar{P}$ 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 13. *Under the previous assumptions, for any primary strategy ζ, the related cost $\rho = \rho(\zeta)$ in Proposition 12 satisfies, for every $t \in [0, T \wedge \theta]$,*

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

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 $\mu = r + \gamma, f = C + \gamma R, \xi, L, U$ of (85) are given as Borel-measurable functions of an $(\widetilde{\mathbb{F}}, \mathbb{P})$-Markov factor 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 pair $(u, \bar{u})$ of the *pre-default no protection pricing function* u and of the *pre-default protection pricing function* $\bar{u}$, namely (cf. (51), (55) or (71) above; see also [20]):

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

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

Proposition 14. *The variational inequality cascade* (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* $t \leq \bar{\sigma}$*.*

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

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

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

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

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

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

Exploiting (96) and (97) in (92), one gets, letting for notational convenience $R_t = R(t, X_{t-})$, $\mathcal{R}_t = \mathcal{R}(t, X_{t-})$,

Proposition 15. *For* $t \in [0, T \wedge \theta]$,

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

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

As in Sect. 3.5 (see also Bielecki et al. [19]), this decomposition of the hedging cost ρ can then be used for devising practical hedging schemes of a defaultable game option, like super-hedging ($\rho = 0$), hedging only the market (spread) risk B, hedging only the default risk H, or min-variance hedging.

Note 6. **(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 12 and Part III.
(ii) Analogous developments regarding defaultable derivatives can also be made relatively to a more general numeraire, cf. Sect. 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:

$$\bar{U}_t = \Omega_t^c \infty + \Omega_t U_t, \tag{99}$$

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

The identification between the *arbitrage*, or *infimal super-hedging*, $\mathbb{P}$*-price process* of a game option with intermittent call protection, and the state-process Π of a solution (Π, M, K), assumed to exist, to the BSDE (15) with $\bar{U}$ given by (99) therein, can be established by a straightforward adaptation of the arguments developed in Sect. 2 (See also Remark 29 in Part IV).

[4] *Boolean-valued* processes.

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

$$\Omega_t = \Omega(t, X_t, N_t) \tag{100}$$

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 $\Pi_t = u(t, X_t, N_t)$ on $[0, T]$ for a suitable pricing function u.

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 Sect. 16, in case of a call protection *discretely monitored* at the dates of a finite time grid $\mathfrak{T} = \{T_0, T_1 \ldots, T_m\}$.

As standing examples of such discretely monitored call protections, one can mention the following clauses, which are commonly found in convertible bonds contracts on an underlying stock S.

Let S_t be given by X_t^1, the first component of our factor process X_t.

Example 2. 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 $u = u(t, x)$,

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

One thus has by application of the results of Sect. 16 (cf. in particular (265)–(266)),

Proposition 16. In the situation of Example 2(i), *the BSDE (15) with $\bar{U}$ given by (99) admits a unique solution (Π, M, K), and one has $\Pi_t = u(t, X_t, N_t)$ on $[0, T]$, for a pricing function $u = u(t, x, k) = u_k(t, x)$ 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 u_ks to every set $[T_{l-1}\, T_l) \times [0, +\infty)$ are continuous, and $u_k(T_l-, x)$ 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 u solves the following cascade of variational inequalities:*

For l decreasing from m to 1,

- *At $t = T_l$, for $k \in \mathbb{N}_\imath$,*

$$u_k(T_l-, x) = \begin{cases} u_{k+1}(T_l, x)\,, \text{ or } u_k(T_l, x) \text{ if } k = \imath\,, & \text{on } \{S > \bar{S}\} \times \mathbb{R}^{q-1} \\ u_0(T_l, x), \text{ or } \min(u_0(T_l, x), U(T_l, x)) \text{ if } k = \imath\,, & \text{on } \{S < \bar{S}\} \times \mathbb{R}^{q-1}, \end{cases} \tag{102}$$

Or, in case $l = m$, $u_k(T_l-, x) = \xi(x)$ on $\mathbb{R}^q$,

- *On the time interval* $[T_{l-1}, T_l)$,

$$\max\left(\mathcal{G}u_k + C - ru_k\,,\ L - u_k\right) = 0\,,\ k = 0\ldots\imath - 1$$
$$\min\left(\max\left(\mathcal{G}u_\imath + C - ru_\imath\,,\ L - u_\imath\right), U - u_\imath\right) = 0\,.$$

In the situation of Example 2(ii), *the BSDE (15) with* $\bar{U}$ *given by (99) admits a unique solution* (Π, M, K), *and one has* $\Pi_t = u(t, X_t, N_t)$ *on* $[0, T]$, *for a suitable pricing function* $u = u(t, S, k) = u_k(t, S)$ *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* u_k*s to every set* $[T_{l-1}\, T_l) \times [0, +\infty)$ *are continuous, and the limit* $u_k(T_l-, x)$ *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* u *solves the following cascade of variational inequalities, with*

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

For l *decreasing from* m *to 1,*

- *At* $t = T_l$, *for* $k \in \{0,1\}^\jmath$, *on* $\{S \neq \bar{S}\}$,

$$u_k(T_\imath-, x) = \begin{cases} \min(u_{k_+}(T_\imath, x), U(T_\imath, x)), & \text{if } |k| \geq \imath \text{ and } |k_+| < \imath\,, \\ u_{k_+}(T_\imath, x), & \text{else} \end{cases} \tag{103}$$

Or, in case $l = m$, $u_k(T_l-, x) = \xi(x)$ *on* $\mathbb{R}^q$,

- *On the time interval* $[T_{l-1}, T_l)$, *for* $k \in \{0,1\}^\jmath$,

$$\max\left(\mathcal{G}u_k + C - ru_k\,,\ L - u_k\right) = 0\,,\ |k| < \imath$$
$$\min\left(\max\left(\mathcal{G}u_\imath + C - ru_\imath\,,\ L - u_\imath\right), U - u_\imath\right) = 0\,,\ |k| \geq \imath. \tag{104}$$

Note 7. **(i)** Existence of the limits $u_k(T_l-, x)$ in (102) or (103) for x in the hyperplane $\{S \neq \bar{S}\}$ of $\mathbb{R}^q$ follows in view of Remark 34.
(ii) Note that the system (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 [31, 39]).

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

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

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–IV will be mainly mathematical.

In this part (see Sect. 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 (28) therein.

Using the general results of Crépey and Matoussi [38], we then show that related Markovian reflected and doubly reflected BSDEs, covering the ones considered in Part I (see Definition 9, Note 8(v) and Definition 11), 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 [38]. Let us thus be given 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 "mark 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 and Shiryaev [62, Definition II.1.13, p. 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 [13, 62] 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 $t \in [0, T]$:

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

• $\mathcal{B}(\mathcal{O})$, 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

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

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

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

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

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

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

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

so in particular (cf. (105))

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

• $\mathcal{A}^2$, the space of finite variation continuous processes K with continuous Jordan components $K^\pm \in \mathcal{S}^2$, where by *the Jordan decomposition* of $K \in \mathcal{A}^2$, we mean the unique decomposition $K = K^+ - K^-$ of K as the difference of two non-decreasing processes $K^\pm$ null at 0 and defining mutually singular random measures on $[0, T]$;
• $\mathcal{A}^2_i$, the space of non-decreasing processes in $\mathcal{A}^2$.

Remark 15. By a slight abuse of notation we shall also write $\|X\|_{\mathcal{H}^p}$ for $\left(\mathbb{E}\left[\int_0^T X_t^2\, dt\right]^{\frac{p}{2}}\right)^{\frac{1}{p}}$ 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 [26]).

Proposition 17. **(i)** *The processes* $\int_0^{\cdot} Z_t dB_t$ *and* $\int_0^{\cdot}\int_E V_t(e)\widetilde{\mu}(dt,de)$ *are martingales, for any* $Z \in \mathcal{H}_d^p$ *and* $V \in \mathcal{H}_\mu^p$;
(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:*

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

for any $V \in \mathcal{H}_\mu^p$. □

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:
$(\mathbf{H.0})$ $\xi \in \mathcal{L}^2$;
$(\mathbf{H.1.i})$ $g_\cdot(y,z,v)$ is a progressively measurable process, and $\|g_\cdot(y,z,v)\|_{\mathcal{H}^2} < \infty$, for any $y \in \mathbb{R}$, $z \in \mathbb{R}^{1\otimes d}$, $v \in \mathcal{M}_\rho$;
$(\mathbf{H.1.ii})$ 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:

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

Remark 16. 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:
$(\mathbf{H.2.i})$ L and U are càdlàg processes in $\mathcal{S}^2$;
$(\mathbf{H.2.ii})$ $L_t \leq U_t$, $t \in [0,T)$ and $L_T \leq \xi \leq U_T$, $\mathbb{P}$-a.s.

Definition 8. **(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**) $Y \in \mathcal{S}^2, Z \in \mathcal{H}_d^2, V \in \mathcal{H}_\mu^2, K \in \mathcal{A}^2$,

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

(**iii**) $L_t \leq Y_t \leq U_t$ 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**) $Y \in \mathcal{S}^2, Z \in \mathcal{H}_d^2, V \in \mathcal{H}_\mu^2, K \in \mathcal{A}_i^2$

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

(**iii**) $L_t \leq Y_t$ 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 $(\mathbb{1}_{\cdot \leq \tau} g, \xi, L_{\cdot \wedge \tau}, , U_{\cdot \wedge \tau})$ satisfies (H.0), (H.1) and (H.2), like (g, ξ, L, U). One can thus state the following

Definition 9. 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 $(\mathbb{1}_{\cdot \leq \tau} g, \xi, L_{\cdot \wedge \tau}, U_{\cdot \wedge \tau})$, 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 $(\mathbb{1}_{\cdot \leq \tau} g, \xi, L_{\cdot \wedge \tau})$, 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

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

in Definition 8(a)(iii), with the convention that $0 \times \pm\infty = 0$.

Note 8. **(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 [38], reflected BSDEs stopped at a random time were introduced and presented as *reflected BSDEs with random terminal time* (only defined over the time interval $[0,\tau]$) as of Darling and Pardoux [41]. 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 [38] for preliminary general results on stopped RBSDEs and on RDBSDEs.
(iii) In the special case where $\tau = 0$, 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

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

is a solution to the stopped RBSDE with data $(g, Y_{\cdot\wedge\tau}, L, \tau)$.
(v) It will come out from the results of this part (Theorem 2; see also [38]) that the solution of an RDBSDE is essentially given as the solution of a stopped RBSDE before τ, appropriately pasted at τ with the solution of an 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 Sect. 16 in Part IV we shall consider doubly reflected BSDEs with an 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 [75], Cvitanić and Karatzas [40]). 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

[5] Boolean-valued processes.

following *Verification Principle*, which was used in the proof of Proposition 2 in Part I. We state it for an RDBSDE as of Definition 9(ii), which in view of Note 8(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 18. *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 [75]; or see also Bielecki et al. [23] for a proof of an analogous result in a context of 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, which is in fact standard since it does not involve τ, can be shown similarly. □

Remark 17. For general well-posedness (in the sense of existence, uniqueness and a priori estimates) and comparison results on the different variants of reflected BSDEs (RBSDEs, R2BSDEs and RDBSDEs) above, we refer the reader to Crépey and Matoussi [38]. We do not reproduce explicitly these results here, since we will state in Sect. 16.2 extensions of these results to more general RIBSDEs (see Note 8(vi)).

5.2 General Forward SDE

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

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

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 $s \geq 0$ and $e \in E$;
- $\widetilde{b}_s(0)$, $\widetilde{\sigma}_s(0)$ and $\widetilde{\delta}_s(0, e)$ are bounded in $s \geq 0$ 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 [54, Lemma 2.1, p. 84] for analogous results with proofs).

Proposition 19. *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:*

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

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

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 [16, 20, 38], 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 $I = \{1,\dots,k\}$, 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:

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

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 2 In (113) and (114), $m(dy)$ is a *finite jump measure* without atom at the origin 0_d of $\mathbb{R}^d$, 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 $n^{i,i}(t,x) = 0$.

Remark 18. We shall often find convenient to denote $v(t,x,i,\dots)$ rather than $v^i(t,x,\dots)$ for a function v of $(t,x,i,\dots)$, and $n^j(t,x,i)$ rather than $n^{i,j}(t,x)$. For instance, with $\mathcal{X}_t = (X_t, N_t)$ below, the notations $f(t,\mathcal{X}_t,y)$ and $n^j(t,\mathcal{X}_t)$

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

will typically be used rather than $f^{N_t}(t, X_t, y)$ and $n^{N_t,j}(t, X_t)$. 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 $[0,T] \times \mathbb{R}^d$.

6.2 Model Dynamics

Definition 10. 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 an initial condition $(t,x,i) \in \mathcal{E}$, such that $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$ is a stochastic basis on $[t,T]$, relatively 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 $s \in [t,T]$:

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

and the $\mathbb{P}^t$-compensatrices $\widetilde{\nu}^t$ and $\widetilde{\chi}^t$ of ν^t and χ^t are such that

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

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

Thus in particular $\nu_s^t(j)$ counts the number of transitions of N^t to state j between times t and s, and $\chi^t((0,s] \times A)$ counts the number of jumps of X^t with mark $y \in A$ between times t and s, for every $s \in [t,T]$ and $A \in \mathcal{B}(\mathbb{R}^d)$.

By an application of Jacod [61, Theorem 3.89, p. 109], the following variant of the *Itô formula* then holds (cf. formula (35) in Part I).

Proposition 20. *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 $s \in [t,T]$,*

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

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 and Kurtz [48, Sects. 7.A and 7.B]).

Note 9. **(i)** Of course, once the related semi-group and Markov properties will be established (see in particular Proposition 26, 28 and 29 as well as Theorems 3 and 4), in applications one can restrict attention to a "single" process $\mathcal{X}$, corresponding in practice to the "true" initial condition (t, x, i) of interest (cf. for instance the last section of [38] in which some of the results of this part were announced without proof). In the context of pricing problems in finance this "true initial condition of interest" corresponds to the current values of the underlyings and to the values of the model parameters calibrated to the current market data, see Part I.

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 $(t, x, i) \in \mathcal{E}$. We shall thus in effect be considering *Markov families* indexed by $(t, x, i) \in \mathcal{E}$.

(ii) 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 Sect. 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. [9], complemented by Barles and Imbert [6]).

(iii) 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 Sect. 7, resorting to a suitable *Markovian change of probability measure*. It should be noted that more specific sub-cases or related models were frequently considered in the literature. So (see also Sect. 6.6 for more comments about financial applications of this model):

- Barles et al. [9] consider jumps in X without regimes N, for a Lévy jump measure m (cf. point (i) above);
- Pardoux et al. [79] 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 Poisson process (instead of a Markov chain in our case, cf. Sect. 7.1);
- 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 Sect. 5, with (note that the initial time is t here instead of 0 therein; superscripts ^{t} are therefore added below to the notation of Sect. 5 where need be):

• E (the "mark space"), 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 $\zeta_s^t(e)$ respectively given by, for any $s \in [t, T]$ and $e = (y, j) \in E$:

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

• 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 with mark $y \in A$ and the jumps of N to state j between t and s, for any $s \geq t$, $A \in \mathcal{B}(\mathbb{R}^d)$ and $j \in I$.

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 $A \oplus \{j\} := \big(A \times \{0\}\big) \cup \big(\{0_d\} \times \{j\}\big)$:

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

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)):

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

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

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 q 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:
$(\mathbf{M.0})$ Φ lies in $\mathcal{P}_q$;
$(\mathbf{M.1.i})$ The mapping $(t,x,i) \mapsto g^i(t,x,u,z,r)$ lies in $\mathcal{P}_q$, for any $(u,z,r) \in \mathbb{R}^k \times \mathbb{R}^{1\otimes d} \times \mathbb{R}$;
$(\mathbf{M.1.ii})$ 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 $(u,z,r),(u',z',r') \in \mathbb{R}^k \times \mathbb{R}^{1\otimes d} \times \mathbb{R}$:

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

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

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

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

(see Remark 18 regarding our notational conventions such as "$g^i(t,x,\ldots) \equiv g(t,x,i,\ldots)$"), where $\widetilde{r}_s^t = \widetilde{r}_s^t(v)$ and $\widetilde{u}_s^t = \widetilde{u}_s^t(y,w)$ are defined by

$$\widetilde{r}_s^t = \int_{\mathbb{R}^d} v(y) f(s,\mathcal{X}_s^t,y) m(dy) \ , \ (\widetilde{u}_s^t)^j = \begin{cases} y, & j = N_s^t \\ y + w_j, & j \neq N_s^t \end{cases}. \tag{120}$$

Note 10. **(i)** The driver coefficient $\widetilde{g} = \widetilde{g}(s,\mathcal{X}_s^t,y,z,\widetilde{v})$ only depends on the functions v, w in $\widetilde{v} = (v,w)$ through their integral (or summation, in case of the "discrete function" w) with respect to the jump and regime switching intensities $f(s,\mathcal{X}_s^t,y)m(dy)$ and $n^j(s,\mathcal{X}_s^t)$. Indeed it is only under this structural assumption on the driver coefficient that one is able to derive a comparison theorem for a BSDE with jumps (see [38] or Sect. 16.2.3). As will be apparent in the proof of Theorem 11 in Part III, such a comparison theorem is key in establishing the connection between a BSDE and the related PDE problem.

(ii) The motivation to define $\widetilde{g}$ as g minus a regime switching related term in (119), is to get a related PDE of the simplest possible form in Part III (variational inequality problems $(\mathcal{V}2)$ and $(\mathcal{V}1)$ involving the operator $\widetilde{\mathcal{G}}$ defined by (188) rather than the "full generator" $\mathcal{G}$ of $\mathcal{X}$).
(iii) In the financial interpretation, one can think of the mute variables y and w in (119)–(120) as representing the *price* and the *regime switching deltas* (cf. (123), (124) or (126), (127) in Definition 11 below). Consequently $\widetilde{u}_s^t$ in (120) can be interpreted as the vector of the prices corresponding to the different possible regimes of the Markov chain component N_s^t, given the current time s and X_s^t. As for the mute variable z, it represents as usual the delta with respect to the continuous-space variable x.

Given the previous ingredients and an $\mathbb{F}^t$ – stopping time τ^t, where the parameter ^{t} stands in reference to an initial condition $(t,x,i) \in \mathcal{E}$ for $\mathcal{X}$, 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 11. (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

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

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

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

(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

$$\widetilde{g}(s, \mathcal{X}_s^t, y, z, \widetilde{v}),\ Y_{\tau^t}^t,\ \ell(s, \mathcal{X}_s^t),\ \tau^t \tag{122}$$

where Y^t is the state-process of $\mathcal{Y}^t$ in (ii).
(b) The solution is said to be *Markovian*, 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.:

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

$$\text{For any } j \in I:\ W_s^t(j) = u^j(s, X_{s-}^t) - u(s, \mathcal{X}_{s-}^t)\ ,\ s \in [t,T] \tag{124}$$

$$\int_t^s \widetilde{g}(\zeta, \mathcal{X}_\zeta^t, Y_\zeta^t, Z_\zeta^t, \mathcal{V}_\zeta^t)d\zeta = \int_t^s \Big[g(\zeta, \mathcal{X}_\zeta^t, u(\zeta, X_\zeta^t), Z_\zeta^t, \widetilde{r}_\zeta^t) - \sum_{j\in I} n^j(\zeta, \mathcal{X}_\zeta^t)\big(u^j(\zeta, X_\zeta^t) - u(\zeta, \mathcal{X}_\zeta^t)\big)\Big] \times d\zeta \ , \ s \in [t,T] \tag{125}$$

with in (125):

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

(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.:

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

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

$$\int_t^s \widetilde{g}(\zeta, \mathcal{X}_\zeta^t, \bar{Y}_\zeta^t, \bar{Z}_\zeta^t, \bar{\mathcal{V}}_\zeta^t)d\zeta = \int_t^s \Big[g(\zeta, \mathcal{X}_\zeta^t, v(\zeta, X_\zeta^t), \bar{Z}_\zeta^t, \bar{r}_\zeta^t) - \sum_{j\in I} n^j(\zeta, \mathcal{X}_\zeta^t)\big(v^j(\zeta, X_{\zeta-}^t) - v(\zeta, \mathcal{X}_{\zeta-}^t)\big)\Big] \times d\zeta \ , \ s \in [t, \tau^t] \tag{128}$$

with in (128):

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

(cf. (120)).

Remark 19. The terminology *Markovian solution* in part (b) of these definitions stands in reference to the fact that, as we will see in Part III, the Markovian consistency conditions (123)–(125) or (126)–(128) are the keys in establishing the bridge between the BSDE perspective and a PDE perspective, as well as in making the connection with applications (see, e.g., (131)).

6.5 Markovian Verification Principle

The following proposition is a Markovian counterpart to the general verification principle of Proposition 18 in Sect. 5.1.2.

Proposition 21. *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 Markovian solution to the Markovian decoupled forward backward stochastic differential equation with data* $\mathcal{G}, \mathcal{C}, \tau$, *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:*

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

with

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

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

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

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 from the definition of the barriers in (121), given the general verification principle of Proposition 18.
(ii) By (126) and the fact that $\bar{\mathcal{Y}}^t$ is stopped at τ^t, it comes,

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

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 18 (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 Sect. 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 model $\mathcal{X}$ is indeed required in order to cover the variety of situations encountered in financial modeling. So:

- In Bielecki et al. [20], this model is presented as a flexible risk-neutral pricing model in finance, for *equity and equity-to-credit (defaultable, cf. Sect. 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. [16], 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 a 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 $c^i(t,x) - \mu^i(t,x)y$, for *dividend and interest-rate* related functions c and μ (or dividends and interest-rates *adjusted for credit spread* in a more general context of defaultable contingent claims, cf. Sect. 4.2). Observe that in order for a Markovian solution $\mathcal{Z}^t$ to the Markovian FBSDE of Definition 11 to satisfy

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

for given functions c and μ on $\mathcal{E}$, it suffices in view of identities (123)–(125) to set

$$g^i(t,x,u,z,r) = c^i(t,x) - \mu^i(t,x)u^i + \sum_{j\in I} n^{i,j}(t,x)(u^j - u^i). \qquad (131)$$

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 the risk-neutral probability in Part I) would lead to a (z,r)-*dependent* driver coefficient function g.

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

Moreover, 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 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. Sect. 3.1).

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

Remark 20. In Sect. 16 in Part IV (see also Sect. 4.3 in Part I), we consider products with more general 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

Sections 7–9, which culminate in Proposition 30 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 Markovian 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 in Sect. 7.1 from a model with independent components. We shall then apply in Sect. 7.2 a Markovian change of probability measure in order to get a model with mutual dependence under the changed measure.

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 the related generator $\widehat{\mathcal{G}}$ such that

Assumption 3 **(i)** $\widehat{f} = 1$, $\widehat{n}^{i,j}(t,x) = \widehat{n}^{i,j} \geq 0$ for any $i, j \in I$, and $\widehat{n}^{i,i} = 0$ for any $i \in I$;
(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, 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[7]:

- A d-dimensional standard Brownian motion B;
- A compound Poisson process P with jump measure $m(dy)$;
- A continuous-time Markov chain Q on the state space I^2 with jump intensity from (l,j) to (l',j') given by $\widehat{n}^{l',j'}$, for every $(l,j) \neq (l',j')$ (and a given law at time 0, the nature of which plays no role in the sequel).

Remark 21. 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 with mark 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) between times 0 and s.

Lemma 6. *The $\mathbb{P}$-compensatrices $\widetilde{\chi}$ of χ and $\widetilde{\nu}$ of ν are respectively given by*

$$\widetilde{\chi}(ds,dy) = \chi(ds,dy) - m(dy)ds\ ,\ d\widetilde{\nu}_s(l,j) = d\nu_s(l,j) - \widehat{n}^{l,j}\,ds. \quad (132)$$

Moreover, for every $(l,j) \in I^2$, $\nu(l,j)$ is a Poisson process with intensity $\widehat{n}^{l,j}$.

Proof. That $m(dy)ds$ $\mathbb{P}$-compensates χ directly results from the definition of χ. Let us thus prove the results regarding ν. The chain Q is a bi-dimensional $\mathbb{F}$ – Markov chain with the generic state denoted as (j,l). Let $\lambda_s(q',q)$ denote the measure that counts the number of jumps of the chain Q from state $q' = (j',l')$ to state $q = (j,l)$ $(q' \neq q)$ on the time interval $(0,s]$. By the characterization of Markov chains in Bielecki et al. [21, Lemma 5.1], the $\mathbb{F}$-compensator ℓ of the measure λ is given as

$$d\ell_s, (q',q) = \mathbb{1}_{Q_s=q'}\widehat{n}^q ds.$$

Thus, the $\mathbb{F}$-compensator of the measure $\nu_s(q)$ counting the jumps of Q to the state $q = (j,l)$ on the time interval $(0,s]$, is given as

$$\sum_{q'} d\ell_s(q',q) = \widehat{n}^q \sum_{q'} \mathbb{1}_{Q_s=q'} ds = \widehat{n}^q ds.$$

[7] I thank Tomasz R. Bielecki for interesting discussions regarding the construction of this section, and in particular, for the proof of Lemma 6.

Consequently, $\widetilde{\nu}_s(j,l)$ is an $\mathbb{F}$-martingale. In view of Watanabe characterization of a Poisson process (see, e.g., Brémaud [27, Chap. II, Sect. 2, t5, p. 25]), $\nu(l,j)$ is thus a Poisson process. □

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

$$\begin{cases} dN_s^t = \sum\limits_{j\in I} (j - N_{s-}^t)\, d\nu_s(N_{s-}^t, j) \\ dX_s^t = \widehat{b}(s, \mathcal{X}_s^t)ds + \sigma(s, \mathcal{X}_s^t)\, dB_s + \int_{\mathbb{R}^d} \delta(s, \mathcal{X}_{s-}^t, y)\, \widetilde{\chi}(ds, dy). \end{cases} \tag{133}$$

Proposition 22. *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*[8] *$\mathcal{X}^t = (X^t, N^t)$. For any $p \in [2, +\infty)$, one has:*

$$\|X^t\|_{\mathcal{S}_d^p}^p \leq C_p\,(1 + |x|^p) \tag{134}$$

$$\|\mathbb{1}_{(s,r)}(X_\cdot^t - X_s^t)\|_{\mathcal{S}_d^p}^p \leq C_p\,(1 + |x|^p)\,(r-s). \tag{135}$$

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

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

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

where we set $\bar{x} = |x| \vee |x'|$.

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

$$dN_s^t = \sum_{(l,j)\in I^2} (j-l)\, \mathbb{1}_{\{l = N_{s-}^t\}} d\nu_s(l,j) \tag{138}$$

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

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

$$\begin{cases} dN_s^t = \sum\limits_{(l,j)\in I^2} (j-l)\, \mathbb{1}_{s<t} \mathbb{1}_{\{l=N_s^t\}} \widehat{n}^{l,j} ds + \sum\limits_{(l,j)\in I^2} (j-l)\, \mathbb{1}_{s<t} \mathbb{1}_{\{l=N_{s-}^t\}} d\widetilde{\nu}_s(l,j) \\ dX_s^t = \mathbb{1}_{s>t} \widehat{b}(s, \mathcal{X}_s^t)ds + \mathbb{1}_{st} \sigma(s, \mathcal{X}_s^t)\, dB_s + \int_{\mathbb{R}^d} \mathbb{1}_{st} \delta(s, \mathcal{X}_{s-}^t, y)\, \widetilde{\chi}(ds, dy)\ . \end{cases} \tag{140}$$

[8] Defined over $[0,T]$.

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, p. 297], the stochastic differential equation (140) with initial condition (x,i) at time t, 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) and (112) then yield, under Assumption 3:

$$\|X^t\|^p_{\mathcal{S}^p_d}\le C_pC_p^t \tag{141}$$

$$\|\mathbb{1}_{(s,r)}(X^t_\cdot-X^t_s)\|^p_{\mathcal{S}^p_d}\le C_pC_p^t(r-s) \tag{142}$$

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

with

$$C_p^t=|x|^p+\mathbb{E}\left[\sup_{[t,T]}|\widehat{b}(\cdot,0,N^t_\cdot)|^p+\sup_{[t,T]}|\sigma(\cdot,0,N^t_\cdot)|^p+\sup_{[t,T]}\int_{\mathbb{R}^d}|\delta(\cdot,0,N^t_\cdot,y)|^pm(dy)\right]$$

$$C_p^{t,t'}=|x-x'|^p+\mathbb{E}\left[\int_{t\wedge t'}^T|\widehat{b}(s,X^t_s,N^t_s)-\widehat{b}(s,X^t_s,N^{t'}_s)|^pds+\int_{t\wedge t'}^T|\sigma(s,X^t_s,N^t_s)-\sigma(s,X^t_s,N^{t'}_s)|^pds+\int_{t\wedge t'}^T\int_{\mathbb{R}^d}|\delta(s,X^t_s,N^t_s,y)-\delta(s,X^t_s,N^{t'}_s,y)|^pm(dy)ds\right]$$

The bound estimates (134) and (135) result from (141) and (142) by the boundedness Assumption 3(iii) on the coefficients. As for the error estimates (136)–(137), note that by construction of N via Q in (133), one has (assuming $t\le t'$, w.l.o.g.):

$$N^t\not\equiv N^{t'}\Longrightarrow\sum_{j\in I\setminus\{i\}}\nu_{t'}(i,j)>\sum_{j\in I\setminus\{i\}}\nu_t(i,j) \tag{144}$$

(which in words means, "at least one jump of ν on $(t,t']$"). Now, in view of Lemma 6, the probability of at least one jump of $\nu(i,j)$ on $(t,t']$ is $1-e^{-\widehat{n}^{i,j}|t-t'|}$, and therefore,

$$\mathbb{P}(N^t\not\equiv N^{t'})\le\sum_{j\in I\setminus\{i\}}\left(1-e^{-\widehat{n}^{i,j}|t-t'|}\right)\le\left(\sum_{j\in I\setminus\{i\}}\widehat{n}^{i,j}\right)|t-t'|\,,$$

which proves (136). Thus

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

where by (134) and the properties of b:

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

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

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

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

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

Hence (137) follows, in view of (143). □

Note 11. **(i)** Given the definition of N^t in the first line of (133), an application of Lemma 5.1 in Bielecki et al. [21] yields that N^t is an $\mathbb{F}$ – Markov chain (and therefore, a Markov chain with respect to its own filtration). The Markov property of N^t will be recovered independently in Proposition 29, as a by-product of Theorem 3 (cf. Note 9(i)). Note however that one of the messages of the present paper is that Markov properties are not really needed if one works in a SDE set-up. Indeed, SDE uniqueness results are then enough for most related purposes. In fact one of the keys of Theorem 3 precisely consists in SDE uniqueness results which underlie the SDE and BSDE semi-group properties of Sect. 8.
(ii) The reason why we introduce N^t indirectly via Q through (133) is the following. Defining a process N^t for every initial condition (t,x,i), and getting a "Markov family" $\mathcal{X}^t=(X^t,N^t)$ parameterized by ^{t} standing for a generic initial condition $(t,x,i)\in\mathcal{E}$, is not enough for the purpose of establishing the connection between an SDE and a PDE perspective in Part III. For this purpose, one also needs this "Markov family" to be "well-behaved" in the sense of bound and error estimates like (135) and (137) to be available. This is why we resort to the above "bi-dimensional construction" of N^t via Q, which implies (144), and yields as a consequence the error estimate (136) on N^t.
(iii) 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 $[t,T]$:

$$B^t = B - B_t\ ,\ \chi^t = \chi - \chi_t\ ,\ \nu^t(j) = \nu(N_{\cdot-}^t,j) - \nu_t(i,j). \tag{145}$$

Note that ν^t is a random measure on $[0,T]\times I$, such that for every $j\in I$, $\nu^t_s(j)$ counts the number of jumps of N^t to regime j on $(t,s]$. By contrast, ν is a random measure on $[0,T]\times I^2$, such that for every $(j,l)\in I^2$, $\nu_s(j,l)$ counts the jumps of the bi-dimensional Markov Chain Q to state (j,l) on $(0,s]$.

Remark 22. Of course $dB^t_s = dB_s$ and $\chi^t(ds,dy)=\chi(ds,dy)$, so the introduction of B^t and χ^t is not really necessary. The reason why we introduce B^t and χ^t is for notational consistency with ν^t (also note that B^t, χ^t and ν^t are defined over $[t,T]$, whereas B, χ and ν live over $[t,T]$).

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 23. **(i)** *Let $\mathcal{X}^t$ be defined as in Proposition 22. The stochastic differential equation (133), or equivalently (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*

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

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

Proof. **(i)** is straightforward, given Proposition 22 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 in view 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 [62, Theorem 4.34(a) Chap. III, p. 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

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

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. [66]), 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 Remark 23 below), under less restrictive conditions than in the previous section, with *state-dependent intensities*. Towards this end we shall apply to the model of Sect. 7.1 a *Markovian change of probability measure* (see Kunita and Watanabe [70], Palmowski and Rolski [78]; cf. also Bielecki et al. [16] or Becherer and Schweizer [10]).

Let thus a *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 in $\widehat{\mathcal{G}}$), for $(t, x, i) \in \mathcal{E}$:

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

(where we recall that $\widehat{f} \equiv 1$).

Lemma 7. ***(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 $\Gamma_t^t = 1$ and, 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 3(ii) and the regularity assumptions on γ.
(ii) By application of Bichteler [13, Proposition 5.2.25, p. 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. In particular the local martingale Γ^t is a genuine martingale. □

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

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

and we let finally $\mathbb{P}^t = \mathbb{P}_T^t$. Note that Γ_s^t is the $\mathcal{F}_s^t$-measurable version of the Radon–Nikodym density of $\mathbb{P}^t$ with respect to $\mathbb{P}$ on $\mathcal{F}_s^t$, for every $s \in [t,T]$.

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

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

The proof of the following lemma is classical and therefore deferred to Appendix 17.1. Note that this result does not depend on the special form of b in (147). Recall (145) for the definition of B^t.

Lemma 8. 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 24. **(i)** *The restriction to* $[t,T]$ *of* $\mathcal{X}^t = (X^t, N^t)$ *in Propositions 22 and 23(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*:*

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

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

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

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

Proof. **(i)** In view of (147) and (151), $\mathcal{X}^t$ 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 23(i).
(ii) The local martingale predictable representation property is preserved by equivalent changes of probability measures (see, e.g., Jacod–Shiryaev [62, Theorem 5.24, p. 196]), so the result follows from Proposition 23(ii). □

Note 12. **(i)** One might work with the following variant of (148):

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

As compared with (148), the change of probability measure defined by (154), which is used for instance in [16], would have the additional effect to further change the Brownian motion into

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

in (151), and to modify accordingly the coefficient of the first-order term in the generator of $\mathcal{X}$.
(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 Sect. 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 Sect. 4.2 and [20]. 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. [16], using (154) instead of (148)), is that of a *change of numeraire* (cf. Sect. 4.1).

7.3 Synthesis

In Sects. 8 and 9, we shall work with the models $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, χ^t, ν^t), $\mathcal{X}^t = (X^t, N^t)$ with generator $\mathcal{G}$ thus constructed, for initial conditions (t,x,i) varying in $\mathcal{E}$.

Remark 23. 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" $\widehat{n}, \widehat{f} \equiv 1, \widehat{b}$ 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 Sects. 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 Sect. 6.3 in relation to process $\mathcal{X}^t$ considered relatively to the stochastic basis $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, we define likewise, in relation with the process $\mathcal{X}^t$ considered relatively to $(\Omega, \mathbb{F}, \mathbb{P})$:

- F, the subset $\left(\mathbb{R}^d \times \{0_2\}\right) \cup \left(\{0_d\} \times I^2\right)$ 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 $t \in [0,T]$ and $e = (y,(l,j)) \in F$:

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

• μ, the $(\Omega, \mathbb{F}, \mathbb{P})$ – integer-valued random measure on $\big([0,T] \times F, \mathcal{B}([0,T]) \otimes \mathcal{B}_F\big)$ counting the jumps of χ with mark $y \in A$ and the jumps of ν to (l,j) between 0 and t, for any $t \geq 0$, $A \in \mathcal{B}(\mathbb{R}^d)$, $(l,j) \in I^2$.

We denote for short (cf. Sect. 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 BSDE assumptions (M.0)–(M.2) introduced in Sect. 6.4, as well as
$(\mathbf{M.3})$ $\ell = \varphi \vee c$ for a $\mathcal{C}^{1,2}$-function φ on $\mathcal{E}$ such that

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

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

Note 13. **(i)** The standing example for φ in (M.3) (see [38]) is $\varphi = x_1$, the first component of $x \in \mathbb{R}^d$ (assuming $d \geq 1$ in our model), whence $\mathcal{G}\varphi = b_1$. In this case (157) reduces to

$$b_1,\ \sigma_1,\ (t,x,i) \mapsto \int_{\mathbb{R}^d} |\delta^i_1(t,x,y)| m(dy) \ \in\ \mathcal{P}.$$

(ii) Alternatively to (M.3), one might work with the symmetric assumptions regarding h, namely $h = \varphi \wedge c$ 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.

In part (i) of the following theorem, building in particular upon the $(\mathbb{F}^t, \mathbb{P}^t; B^t, \chi^t, \nu^t)$ – martingale representation property of Proposition 24(ii), one establishes existence and uniqueness of an $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t) – solution $\mathcal{Y}^t$ of the R2BSDE on $[t, T]$ with data (121). This result is then "translated" in part (ii) in terms of an $(\Omega, \mathbb{F}, \mathbb{P})$, (B, μ) – solution $\widetilde{\mathcal{Y}}^t$ to another, auxiliary R2BSDE. The interest of the auxiliary R2BSDE is that the solutions $\widetilde{\mathcal{Y}}^t$ as (t, x, i) varies in $\mathcal{E}$ are defined with respect to the common stochastic basis $(\Omega, \mathbb{F}, \mathbb{P})$ (cf. Sect. 7.3). One can then proceed by application of the general estimates of [38] for deriving in Proposition 25 Markovian stability results regarding the $\widetilde{\mathcal{Y}}^t$s. These estimates are then used in Part III of this article for establishing the analytic interpretation of Y^t, the first component of $\mathcal{Y}^t$, which essentially coincides with that of $\widetilde{\mathcal{Y}}^t$ (see part (ii) below).

Theorem 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* $[0, T]$:

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

Then $\widetilde{\mathcal{Y}}^t = (Y^t, Z^t, \widetilde{\mathcal{V}}^t, K^t)$ *is an* $(\Omega, \mathbb{F}, \mathbb{P})$, (B, μ) *– 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

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

with

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

Proof. **(i)** Given assumptions (M.0)–(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

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

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.2) are satisfied by the data (158) relatively to $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, μ^t). Given the local martingale predictable representation property of Proposition 24(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 [38].
(ii) By the previous R2BSDE, one thus has for $s \in [t,T]$:

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

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 $s \ge t$:

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

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 [38] 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 $(\Omega,\mathbb{F},\mathbb{P})$, one then has the following stability result, whose proof is deferred to Appendix 17.2.

Proposition 25. **(i)** *One has the following estimate on $\widetilde{\mathcal{Y}}^t$ in Theorem 1:*

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

(ii) *Moreover, ^{t_n} referring to a perturbed initial condition $(t_n,x_n,i)\in\mathcal{E}$ with $(t_n,x_n)\to(t,x)$ as $n\to\infty$, then $\widetilde{\mathcal{Y}}^{t_n}$ converges in $\mathcal{S}^2\times\mathcal{H}^2_d\times\mathcal{H}^2_\mu\times\mathcal{A}^2$ 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 22 and Theorem 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 Sects. 7.1 and 7.2, with t' instead of t therein. Note in particular that $\widetilde{\mathbb{F}}^{t'}$ is embedded into the restriction $\mathbb{F}^t_{|[t',T]}$ of $\mathbb{F}^t$ to $[t',T]$.

Proposition 26. **(i)** *The stochastic differential equation (133), or equivalently (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*

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

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:

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

Proof. **(i)** By Bichteler [13, Proposition 5.2.25, p. 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 embedded into $\mathbb{F}^t_{|[t',T]}$. Now, $(N^t_r)_{t' \leq r \leq T}$ is an $\mathbb{F}^t_{|[t',T]}$-adapted process satisfying the first line of (133) on $[t', T]$. $(X^t_r)_{t' \leq r \leq T}$ is then in turn an $\mathbb{F}^t_{|[t',T]}$-adapted process satisfying the second line of (133) on $[t', T]$. Therefore $\mathcal{X}^{t'} = (\mathcal{X}^t_r)_{t' \leq r \leq T}$, by uniqueness relatively to $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P})$.
(ii) Note that the bound estimate (134) on X^t is also valid for solutions of stochastic differential equations *with random initial condition* like $X^{t'}$ in part (i) above, by application of Proposition 19 (cf. proof of Proposition 22). One thus has for any $p \in [2, +\infty)$, with $X^{t'}$ extended by $X^{t'} = X^t_{t'}$ on $[0, t']$:

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

where the last inequality comes from (134). Consequently, (H.0)'–(H.2)' in the proof of Theorem 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 24(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 [38]. 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.2)' as above. Since $\widetilde{\mathbb{F}}^{t'}$ is embedded into $\mathbb{F}^t_{|[t',T]}$, $\mathcal{Y}^{t'} = (Y^{t'}, Z^{t'}, \mathcal{V}^{t'}, K^{t'})$ 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 $(Y^t_r, Z^t_r, \mathcal{V}^t_r, K^t_r - K^t_{t'})_{t' \leq r \leq T}$ 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 relatively 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 2. **(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 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* $[0,T]$:

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

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,μ) *– 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 [38], 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 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 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,μ) – 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 Note 8(iv). □

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

Assumption 4 τ^t is an almost surely continuous random function of (t,x,i) on $\mathcal{E}$.

Example 3. Let τ^t denote the minimum of T and of the first exit time by $\mathcal{X}^t$ of an open domain $D \subseteq \mathbb{R}^d \times I$, that is:

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

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 4 is typically satisfied under a suitable *uniform ellipticity condition* on the diffusion coefficient σ of X. For related results, see, e.g., Darling and

Pardoux [41], Dynkin [44, Theorem 13.8], Freidlin [51], or Assumption A2.2 and the related discussion in Kushner–Dupuis [71, p. 281]. See also [31] for a precise statement and proof in case of a diffusion X (case $\chi \equiv 0$).

Under Assumption 4, one has the following stability results on $\overline{\mathcal{Y}}^t = (\bar{Y}^t, \bar{Z}^t, \overline{\mathcal{V}}^t, \bar{K}^t)$ in Theorem 2(ii). The proof is deferred to Appendix 17.3.

Proposition 27. **(i)** *The following bound estimate holds:*

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

(ii) *Moreover,* ^{t_n} *referring to a perturbed initial condition* $(t_n, x_n, i) \in \mathcal{E}$ *with* $(t_n, x_n) \to (t, x)$ *as* $n \to \infty$, *then* $\overline{\mathcal{Y}}^{t_n}$ *converges in* $\mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu \times \mathcal{A}^2$ *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 Sect. 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 2(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 $\bar{\mathbb{F}}^{t'} = (\bar{\mathcal{F}}^{t'}_r)_{r \geq t'}$ be defined by, for $r \in [t', T]$:

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

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

Assumption 5 τ' is an $\bar{\mathbb{F}}^{t'}$-stopping time.

Note that since we took D open in (167), Assumption 5 is satisfied in the case of Example 3.

Remark 24. Assumption 5 would not satisfied if the domain D had been taken closed instead of open in (167), for instance with $\{\psi^i \geq 0\}$ instead of $\{\psi^i > 0\}$ in (168).

Proposition 28. **(i)** *The following stochastic differential equation on* $[t', T]$*:*

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

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:

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

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

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

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

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

Proof. **(i)** By Bichteler [13, Proposition 5.2.25, p. 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 embedded into $\mathbb{F}^t_{|[t',T]}$. Now, $(\bar{N}^t_r)_{t'\le r\le T}$ is an $\mathbb{F}^t_{|[t',T]}$-adapted process satisfying the first line of (170) on $[t', T]$. $(\bar{X}^t_r)_{t'\le r\le T}$ is then in turn an $\mathbb{F}^t_{|[t',T]}$-adapted process satisfying the second line of (170) on $[t', T]$. Therefore $\bar{\mathcal{X}}^{t'} = (\bar{\mathcal{X}}^t_r)_{t'\le r\le T}$, by uniqueness relative to $(\Omega, \mathbb{F}^t_{|[t',T]}, \mathbb{P})$.
(ii) One has as in the proof of Proposition 26(ii):

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

Consequently the data

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

satisfy the general assumptions (H.0), (H.1), and the assumptions regarding L in (H.2), relatively 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 24(ii)) and the form postulated in (M.3) for ℓ, the general results of [38] imply 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 $[t', T]$. Besides, by Theorem 2(ii), $(\bar{Y}^t_r, \bar{Z}^t_r, \bar{\mathcal{V}}^t_r, \bar{K}^t_r)_{t\le r\le T}$ 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 $\bar{\mathcal{V}}^t = \mathbb{1}_{.\leq\tau^t}\widehat{\mathcal{V}}^t$ 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 9(i):

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

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

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

where $\ell(s, \bar{\mathcal{X}}^{t'}_s) \leq \bar{Y}^t_s$ for s in $(t', \tau']$ in the third line implies that the last inequality also holds at $s = t'$, by right-continuity. So $(\bar{Y}^t_r, \bar{Z}^t_r, \bar{\mathcal{V}}^t_r, \bar{K}^t_r - \bar{K}^t_{t'})_{t'\leq r\leq T}$ 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 9(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 Markovian forward SDE and the solutions $\mathcal{Y}$, $\bar{\mathcal{Y}}$ of our Markovian reflected backward SDEs.

Theorem 3. *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 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)–(125)):*

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

$$\textit{For any } j \in I: \ W_s^t(j) = u^j(s, X_{s-}^t) - u(s, \mathcal{X}_{s-}^t) \ , \ s \in [t,T] \tag{176}$$

$$\int_t^s \widetilde{g}(\zeta, \mathcal{X}_\zeta^t, Y_\zeta^t, Z_\zeta^t, \mathcal{V}_\zeta^t) d\zeta = \int_t^s \Bigg[g(\zeta, \mathcal{X}_\zeta^t, u(\zeta, X_\zeta^t), Z_\zeta^t, \widetilde{r}_\zeta^t) - \sum_{j \in I} n^j(\zeta, \mathcal{X}_\zeta^t) \big(u^j(\zeta, X_\zeta^t) - u(\zeta, \mathcal{X}_\zeta^t) \big) \Bigg] \times d\zeta \ , \ s \in [t,T] \tag{177}$$

with in (177):

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

(cf. (120)).

Proof. Letting $r = t' = s$ in the semi-group property (163) of $\mathcal{Y}$ yields:

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

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 25(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 25(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 (25) on $\widetilde{\mathcal{Y}}^t$. So $u^i(t_n, x_n) \to u^i(t,x)$ 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) $u(\cdot, \mathcal{X}^t_\cdot)$ are càdlàg processes. One then has on $\{\Delta N^t \neq 0\}$ (set on which $\Delta X^t = 0$), using also the continuity of u:

$$\begin{aligned}\Delta Y^t_s &= u(s, \mathcal{X}^t_s) - u(s, \mathcal{X}^t_{s-}) = \sum_{j\in I} \big(u^j(s, X^t_{s-}) - u(s, \mathcal{X}^t_{s-})\big)\, \Delta\nu^t_s(j) \\ &= \sum_{j\in I} W_s(j)\Delta\nu^t_s(j),\end{aligned}$$

where the last equality comes from the R2BSDE with data (121) satisfied by $\mathcal{Y}^t$. The last equality also trivially holds on $\{\Delta N^t = 0\}$. Denoting $\mathcal{W}^t_s(j) = u^j(s, X^t_{s-}) - u(s, \mathcal{X}^t_{s-})$, one thus has on $[t, T]$:

$$\begin{aligned}0 &= \sum_{j\in I} \big(\mathcal{W}^t_s(j) - W^t_s(j)\big)\Delta\nu^t_s(j) \\ &= \sum_{j\in I} \big(\mathcal{W}^t_s(j) - W^t_s(j)\big)\Delta\widetilde{\nu}^t_s(j) + \sum_{j\in I} \big(\mathcal{W}^t_s(j) - W_s(j)\big)n^j(s, \mathcal{X}^t_s)ds\end{aligned}$$

(recall (151) for the definition of $\widetilde{\nu}^t$), $\mathbb{P}^t$ – almost surely. Therefore $\mathcal{W}^t_s(j) = W^t_s(j)$ 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 $(y, z, \widetilde{v}) = (Y^t_s, Z^t_s, \mathcal{V}^t_s)$ in (120):

$$\widetilde{u}^t_s(N^t_s) = Y^t_s = u(s, \mathcal{X}^t_s),$$

by (175), and then for $j \neq N^t_s$:

$$(\widetilde{u}^t_s)^j = Y^t_s + W^t_s(j) = u(s, \mathcal{X}^t_s) + \big(u^j(s, X^t_{s-}) - u(s, \mathcal{X}^t_{s-})\big),$$

by (176). Therefore $\widetilde{u}^t_{s-} = u(s, X^t_{s-})$, so that by definition (119) of $\widetilde{g}$:

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

which gives (177).

□

As a by-product of Theorem 3, one has the following

Proposition 29. *$\mathcal{X}^t$ is an $(\mathbb{F}^t, \mathbb{P}^t)$-Markov process.*

Proof. In the case of a classical BSDE (without barriers) with

$$g^i(t,x,u,z,r) = \sum_{j\in I} n^{i,j}(t,x)(u_j - u_i),$$

using the Verification Principle of Proposition 18, identities (175) and (177) give:

$$Y_s^t = \mathbb{E}^t\big[\Phi(\mathcal{X}_T^t)\big|\mathcal{F}_s^t\big] = u(s,\mathcal{X}_s^t),$$

for a continuous bounded function u in $\mathcal{P}$. Therefore

$$\mathbb{E}^t\big[\Phi(\mathcal{X}_T^t)\big|\mathcal{F}_s^t\big] = \mathbb{E}^t\big[\Phi(\mathcal{X}_T^t)\big|\Sigma(\mathcal{X}_s^t)\big], \tag{180}$$

where $\Sigma(\mathcal{X}_s^t)$ denotes the sigma-field generated by $\mathcal{X}_s^t$. By the monotone class theorem, identity (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

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 $\bar{\mathcal{V}}^t = \mathbb{1}_{.\le\tau^t}\widehat{\mathcal{V}}^t$ 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 2(ii).

Theorem 4. **(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)–(128)):*

$$\bar{Y}_s^t = v(s,\mathcal{X}_s^t)\ ,\ s\in[t,\tau^t] \tag{181}$$

$$v(\tau^t, \mathcal{X}_{\tau^t}^t) = u(\tau^t, \mathcal{X}_{\tau^t}^t) \tag{182}$$

$$\text{For any } j\in I:\ \widehat{W}_s^t(j) = v^j(s, X_{s-}^t) - v(s,\mathcal{X}_{s-}^t)\ \ ,\ s\in[t,\tau^t] \tag{183}$$

$$\begin{aligned}\int_t^s \widetilde{g}(\zeta,\mathcal{X}_\zeta^t,\bar{Y}_\zeta^t,\bar{Z}_\zeta^t,\bar{\mathcal{V}}_\zeta^t)d\zeta = \int_t^s \Big[&g(\zeta,\mathcal{X}_\zeta^t, v(\zeta,X_\zeta^t),\bar{Z}_\zeta^t,\bar{r}_\zeta^t)\\ &-\sum_{j\in I} n^j(\zeta,\mathcal{X}_\zeta^t)\big(v^j(\zeta,X_\zeta^t) - v(\zeta,\mathcal{X}_\zeta^t)\big)\Big]\\ &\times d\zeta\ ,\ s\in[t,\tau^t]\end{aligned} \tag{184}$$

with in (184):

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

(cf. (120) for the definition of $\widetilde{r}^t$).

Proof. Letting $r = t' = s$ in the semi-group property (173) of $\bar{\mathcal{Y}}$ yields:

$$\bar{Y}^t_s = v(s, \bar{\mathcal{X}}^t_s)\ ,\ \mathbb{P}^t\text{−}a.s. \tag{185}$$

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 27(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 27(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. Since $\bar{Y}^t_{\tau^t} = Y^t_{\tau^t}$ (cf. Theorem 2(ii)), (181) and (175) in turn imply (182). One has further on $\{(\omega, s)\,;\ s \in [t, \tau^t],\ N^t_s \neq N^t_{s-}\}$ (on which $\Delta X^t = 0$), using also the continuity of v:

$$\begin{aligned}\Delta\bar{Y}^t_s &= v(s, \mathcal{X}^t_s) - v(s, \mathcal{X}^t_{s-}) = \sum_{j\in I} \big(v^j(s, X^t_{s-}) - v(s, \mathcal{X}^t_{s-})\big)\Delta\nu^t_s(j) \\ &= \sum_{j\in I} \widehat{W}_s(j)\Delta\nu^t_s(j)\end{aligned}$$

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 $[t,\tau^t]$:

$$\begin{aligned}0&=\sum_{j\in I}\big(\mathcal{W}^t_s(j)-\widehat{W}^t_s(j)\big)\Delta\nu^t_s(j)\\&=\sum_{j\in I}\big(\mathcal{W}^t_s(j)-\widehat{W}^t_s(j)\big)\Delta\widetilde{\nu}^t_s(j)+\sum_{j\in I}\big(\mathcal{W}^t_s(j)-\widehat{W}_s(j)\big)n^j(s,\mathcal{X}^t_s)ds\end{aligned}$$

(recall (151) for the definition of $\widetilde{\nu}^t$), $\mathbb{P}^t$ – 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. Hence (182) follows. Finally (184) derives from (181) and (182) like (177) from (175) and (176) (cf. proof of (177)). □

In summary, one has established in Sects. 7–9 the following proposition relatively to the Markovian FBSDE of Definition 11.

Proposition 30. *Under the assumptions of Sects. 7–9, the Markovian FBSDE with generator $\mathcal{G}$, cost functions $\mathcal{C}$ and (parameterized) stopping time τ has a Markovian 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 (147) and (148) (see Remark 23), are admittedly technical and involved, and by no means minimal. *In the sequel we shall give up all these specific assumptions, merely postulating instead that the Markovian FBSDE with data $\mathcal{G}$, $\mathcal{C}$ and τ has a Markovian solution* (as is for instance the case under the assumptions of Sects. 7–9).

Part III
Main PDE Results

In this part (see Sect. 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 of this part are used in Part I for giving a constructive and computational counterpart to the theoretical BSDE results of Sect. 2, in the Markovian factor process set-ups of Sects. 3, 4.1 or 4.2.4. We refer the reader to [31,39] for an alternative, simulation-based, computational approach.

As announced at the end of Part II, we give up all the specific assumptions made in Sect. 7–9. We make instead the following standing

Assumption 6 The Markovian FBSDE with data $\mathcal{G}, \mathcal{C}, \tau$ has a Markovian solution $\mathcal{Z}^t = (\Omega, \mathbb{F}^t, \mathbb{P}^t), (B^t, \chi^t, \nu^t), (\mathcal{X}^t, \mathcal{Y}^t, \bar{\mathcal{Y}}^t)$.

As illustrated in the previous sections, Assumption 6 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 ${}^t \equiv (t, x, i)$ varies in $\mathcal{E}$, 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 [4,5,7,11,12]).

We postulate from now on in this part that

Assumption 7 **(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 3 in Part II.

Let $\mathcal{D} = [0, T] \times \bar{D}$, where $\bar{D}$ denotes the closure[9] of D in $\mathbb{R}^d \times I$. Let also

[9] 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 $\mathbb{R}^d$.

$$\begin{aligned}\text{Int}\,\mathcal{E} &= [0,T)\times\mathbb{R}^d\times I\ ,\ \partial\mathcal{E} := \mathcal{E}\setminus\text{Int}\,\mathcal{E} = \{T\}\times\mathbb{R}^d\times I\\ \text{Int}\,\mathcal{D} &= [0,T)\times D\ ,\ \partial\mathcal{D} := \mathcal{E}\setminus\text{Int}\,\mathcal{D}\end{aligned}\tag{187}$$

stand for the *parabolic interior* and the *parabolic boundary* of $\mathcal{E}$ and $\mathcal{D}$, respectively.

Remark 25. The use of the "thick" boundary $\partial\mathcal{D}$ is motivated 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) and (114)):

$$\begin{aligned}\widetilde{\mathcal{G}}(\varphi,\varphi)^i(t,x) = \partial_t\varphi^i(t,x) &+ \frac{1}{2}Tr\big[a^i(t,x)\mathcal{H}\varphi^i(t,x)\big]\\ &+\nabla\varphi^i(t,x)\beta^i(t,x) + \mathcal{I}\varphi^i(t,x)\end{aligned}\tag{188}$$

with

$$\mathcal{I}\varphi^i(t,x) := \int_{\mathbb{R}^d}\Big(\varphi^i(t,x+\delta^i(t,x,y)) - \varphi^i(t,x)\Big)f^i(t,x,y)m(dy).\tag{189}$$

Let also $\widetilde{\mathcal{G}}\varphi$ stand for $\widetilde{\mathcal{G}}(\varphi,\varphi)$. So in particular (cf. (113)):

$$\widetilde{\mathcal{G}}\varphi^i(t,x) + \sum_{j\in I} n^{i,j}(t,x)\Big(\varphi^j(t,x) - \varphi^i(t,x)\Big) = \mathcal{G}\varphi^i(t,x).\tag{190}$$

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 11 (cf. Assumption 6), 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:

$$\begin{aligned}\max\Big(\min\Big(&-\widetilde{\mathcal{G}}u^i(t,x) - g^i(t,x,u(t,x),(\nabla u\sigma)^i(t,x),\mathcal{I}u^i(t,x)),\\ &u^i(t,x)-\ell^i(t,x)\Big), u^i(t,x)-h^i(t,x)\Big) = 0\end{aligned}$$

on $\text{Int}\,\mathcal{E}$, supplemented by the terminal condition Φ (the terminal cost function in the cost data $\mathcal{C}$) at T. We also consider the problem $(\mathcal{V}1)$ obtained by formally replacing h by $+\infty$ in $(\mathcal{V}2)$, that is

$$\begin{aligned}\min\Big(&-\widetilde{\mathcal{G}}u^i(t,x) - g^i(t,x,u(t,x),(\nabla u\sigma)^i(t,x),\\ &\times\mathcal{I}u^i(t,x))\ ,\ u^i(t,x)-\ell^i(t,x)\Big) = 0\end{aligned}$$

on $\text{Int}\,\mathcal{D}$, supplemented by a continuous boundary condition Ψ extending Φ on $\partial\mathcal{D}$.

The following *continuity property of the integral term* $\mathcal{I}$ in $\widetilde{\mathcal{G}}$ (cf. (189)) is key in the theory of viscosity solutions of nonlinear integro-differential equations (see for instance Alvarez–Tourin [1, p. 297]).

Lemma 9. *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

$$\begin{aligned}
&\mathcal{I}\psi^i(t_n,x_n) - \mathcal{I}\psi^i(t,x)\\
&\quad = -\int_{\mathbb{R}^d} \Big(\psi^i(t_n,x_n)f^i(t_n,x_n,y) - \psi^i(t,x)f^i(t,x,y)\Big)m(dy)\\
&\qquad + \int_{\mathbb{R}^d} \Big(\psi^i(t_n,x_n+\delta^i(t_n,x_n,y))f^i(t_n,x_n,y)\\
&\qquad -\psi^i(t,x+\delta^i(t,x,y))f^i(t,x,y)\Big)m(dy),
\end{aligned}$$

where

$$\begin{aligned}
&\int_{\mathbb{R}^d} \Big(\psi^i(t_n,x_n+\delta^i(t_n,x_n,y))f^i(t_n,x_n,y)\\
&\qquad -\psi^i(t,x+\delta^i(t,x,y))f^i(t,x,y)\Big)m(dy)\\
&= \int_{\mathbb{R}^d} \Big(\psi^i(t_n,x_n+\delta^i(t_n,x_n,y)) - \psi^i(t,x+\delta^i(t,x,y))\Big)f^i(t_n,x_n,y)m(dy)\\
&\quad + \int_{\mathbb{R}^d} \psi^i(t,x+\delta^i(t,x,y))\Big(f^i(t_n,x_n,y) - f^i(t,x,y)\Big)m(dy) \qquad (191)
\end{aligned}$$

goes to 0 as $\mathcal{E} \ni (t_n,x_n) \to (t,x)$, by Assumption 7(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. [37] or Fleming and Soner [49]), adapting further the definitions to finite activity jumps and *systems* of PIDEs as in [1, 9, 28, 60, 79].

Definition 12. **(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 $(t,x,i) \in \operatorname{Int}\mathcal{E}$, 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,

$$\max\Big(\min\Big(-\widetilde{\mathcal{G}}(u,\varphi)^i(t,x)-g^i(t,x,u(t,x),(\nabla\varphi\sigma)^i(t,x),\mathcal{I}u^i(t,x)),\\ u^i(t,x)-\ell^i(t,x)\Big),u^i(t,x)-h^i(t,x)\Big)\le 0,\ \text{resp. } \ge 0.$$

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

$$-\widetilde{\mathcal{G}}(u,\varphi)^i(t,x)-g^i(t,x,u(t,x),(\nabla\varphi\sigma)^i(t,x),\mathcal{I}u^i(t,x))\le 0,\ \text{resp. } \ge 0, \quad (192)$$

or inequality (192) 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 $(t,x,i)\in\operatorname{Int}\mathcal{E}$, 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 $\operatorname{Int}\mathcal{E}$, and such that $u\le\Phi$, resp. $u\ge\Phi$ pointwise at T.

(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 $u=\Phi$ at T.

(c) The notions of viscosity subsolutions, supersolutions and solutions of $(\mathcal{V}1)$ at $(t,x,i)\in\operatorname{Int}\mathcal{D}$, and, given a continuous boundary condition Ψ extending Φ on $\partial\mathcal{D}$, $\mathcal{P}$ – 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, $\operatorname{Int}\mathcal{D}$ to $\operatorname{Int}\mathcal{E}$, $\mathcal{C}^0(\mathcal{E})\cap\mathcal{C}^{1,2}(\mathcal{D})$ to $\mathcal{C}^{1,2}(\mathcal{E})$, "on $\partial\mathcal{D}$" to "at T" and Ψ to Φ therein.

Note 14. **(i)** We thus consider boundary conditions in the classical sense, rather than in the weak viscosity sense (cf. the proof of Lemma 13(ii) for more on this issue, see also Crandall et al. [37]).

(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 $\ell\le u\le h$. However a viscosity subsolution (resp. supersolution) u of $(\mathcal{V}2)$ does not need to satisfy $u\ge\ell$ (resp. $u\le h$). Likewise a viscosity solution v of $(\mathcal{V}1)$ necessarily satisfies $\ell\le u$, however a viscosity subsolution v of $(\mathcal{V}1)$ does not need to satisfy $u\ge\ell$.

(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 (192), or in the analogous inequalities regarding $(\mathcal{V}1)$, can be shown by an immediate adaptation to the present set-up of Barles et al. [9, Lemma 3.3, p. 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 11, under Assumption 6. This result establishes that u and v are viscosity solutions of the related obstacle problems, with u as boundary Dirichlet condition for v on $\partial\mathcal{D}$.

Theorem 5. **(i)** *The value function u is a $\mathcal{P}$-solution of $(\mathcal{V}2)$ on $\mathcal{E}$ for the terminal condition Φ at T.*
(ii) *The value function v is a $\mathcal{P}$-solution of $(\mathcal{V}1)$ on $\mathcal{E}$ for the boundary condition u on $\partial\mathcal{D}$.*

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 $\mathcal{X}$:

$$\begin{aligned} u^i(T,x) &= Y_T^T = \Phi^i(x) \\ \ell^i(t,x) \le Y_t^t &= u^i(t,x) \le h^i(t,x). \end{aligned}$$

So $u=\Phi$ pointwise at T and $\ell \le u \le h$ on $\mathcal{E}$. Let us show that u is a subsolution of $(\mathcal{V}2)$ on $\operatorname{Int}\mathcal{E}$. We let the reader check likewise that u is a supersolution of $(\mathcal{V}2)$ on $\operatorname{Int}\mathcal{E}$. Let thus $(t,x,i)\in\operatorname{Int}\mathcal{E}$ 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

$$-\widetilde{\mathcal{G}}\varphi^i(t,x) - g^i(t,x,u(t,x),(\nabla\varphi\sigma)^i(t,x),\mathcal{I}\varphi^i(t,x)) \le 0, \tag{193}$$

where it is further assumed that $u^i(t,x) > \ell^i(t,x)$ and $u^i(t,x)=\varphi^i(t,x)$ (cf. Definition 12(a)(i)). Suppose by contradiction that (193) does not hold. Then by a continuity argument using in particular Lemma 9:

$$\psi(s,y) := \widetilde{\mathcal{G}}\varphi^i(s,y) + g^i(s,y,u(s,y),(\nabla\varphi\sigma)^i(s,y),\mathcal{I}\varphi^i(s,y)) < 0 \tag{194}$$

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

$$\theta = \inf\big\{s\ge t\,;\,|X_s^t-x|\ge\alpha,\ N_s^t\ne i,\ Y_s^t=\ell^i(s,X_s^t)\big\}\wedge(t+\alpha) \tag{195}$$

$$(\widehat{Y}^t,\widehat{Z}^t,\widehat{V}^t,\widehat{K}^t) = \Big(\mathbb{1}_{\cdot<\theta}Y^t + \mathbb{1}_{\cdot\ge\theta}u^i(\theta,X_\theta^t),\ \mathbb{1}_{\cdot\le\theta}Z^t,\ \mathbb{1}_{\cdot\le\theta}V^t,\ K^t_{\cdot\wedge\theta}\Big) \tag{196}$$

$$\begin{aligned}(\widetilde{Y}^t,\widetilde{Z}^t,\widetilde{V}^t) = \Big(&\varphi^i(\cdot X^t_{\cdot\wedge\theta}),\ \mathbb{1}_{\cdot\le\theta}(\nabla\varphi\sigma)^i(\cdot,X^t_\cdot), \\ &\mathbb{1}_{\cdot\le\theta}\big(\big[\varphi^i(\cdot,X^t_{\cdot-}+\delta^i(\cdot,X^t_{\cdot-},y)) - \varphi^i(\cdot,X^t_{\cdot-})\big]\big)_{y\in\mathbb{R}^d}\Big).\end{aligned} \tag{197}$$

Note that $\theta > t$, $\mathbb{P}^t$ – almost surely. Thus, using also the continuity of u^i:

$$\widehat{Y}_t^t = Y_t^t = u^i(t,x) = \varphi^i(t,x) = \widetilde{Y}_t^t \ , \ \mathbb{P}^t\text{-a.s.} \tag{198}$$

Also observe that $K^{t,+} = 0$ on $[t,\theta]$ by the related minimality condition in the R2BSDE equation for $\mathcal{Y}^t$, given that $\ell^i(s,X_s^t) < Y_s^t$ on $[t,\theta)$. Let us now show that one has, for $s \in [t,\theta]$:

$$\begin{aligned}\widehat{Y}_s^t = u^i(\theta, X_\theta^t) + \int_s^\theta g^i(\zeta, X_\zeta^t, u(\zeta, X_\zeta^t), \widehat{Z}_\zeta^t, \widehat{r}_\zeta^t)d\zeta - (\widehat{K}_\theta^{t,-} - \widehat{K}_s^{t,-}) \\ - \int_s^\theta \widehat{Z}_\zeta^t dB_\zeta^t - \int_s^\theta \int_{\mathbb{R}^d} \widehat{V}_\zeta^t(y)\widetilde{\chi}^t(dy,d\zeta).\end{aligned} \tag{199}$$

Indeed this holds true on $\{s = \theta\}$ by definition of $\widehat{Y}^t$ in (195). Furthermore, on $\{s < \theta\}$:

- Either χ^t, whence X^t, do not jump at θ, and identity (199) with θ replaced by $r < \theta$ follows from the R2BSDE equation for $\mathcal{Y}^t$ (in which $K^{t,+} = 0$ on $[t,\theta]$), so that (199) itself holds by passage to the limit as $r \uparrow\uparrow \theta$,
- Or (cf. Definition 10(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 (199).

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 $s \in [t,\theta]$:

$$\begin{aligned}d\varphi^i(s,X_s^t) &= \mathcal{G}\widetilde{\varphi}(s,\mathcal{X}_s^t)ds + (\nabla\varphi\sigma)(s,\mathcal{X}_s^t)dB_s^t \\ &\quad + \int_{\mathbb{R}^d} \big(\varphi^i(s,X_{s-}^t + \delta(s,\mathcal{X}_{s-}^t,y)) - \varphi^i(s,\mathcal{X}_{s-}^t)\big)\widetilde{\chi}^t(ds,dy) \\ &= \widetilde{\mathcal{G}}\widetilde{\varphi}(s,\mathcal{X}_s^t)ds + (\nabla\varphi\sigma)(s,\mathcal{X}_s^t)dB_s^t \\ &\quad + \int_{\mathbb{R}^d} \big(\varphi^i(s,X_{s-}^t + \delta(s,\mathcal{X}_{s-}^t,y)) - \varphi^i(s,\mathcal{X}_{s-}^t)\big)\widetilde{\chi}^t(ds,dy) \\ &= \widetilde{\mathcal{G}}\varphi^i(s,X_s^t)ds + (\nabla\varphi\sigma)^i(s,X_s^t)dB_s^t \\ &\quad + \int_{\mathbb{R}^d} \big(\varphi^i(s,X_{s-}^t + \delta^i(s,X_{s-}^t,y)) - \varphi^i(s,\mathcal{X}_{s-}^t)\big)\widetilde{\chi}^t(ds,dy),\end{aligned}$$

where the second equality uses (190) 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. (197)):

$$\begin{aligned}\widetilde{Y}_s^t &= \varphi^i(\theta,X_\theta^t) - \int_s^\theta \widetilde{\mathcal{G}}\varphi^i(r,X_r^t)dr - \int_s^\theta \widetilde{Z}_r^t dB_r^t - \int_s^\theta \int_{\mathbb{R}^d} \widetilde{V}_r^t(y)\widetilde{\chi}^t(dy,dr) \\ &= \varphi^i(\theta,X_\theta^t) - \int_s^\theta \Big(\psi(r,X_r^t) - lg^i(r,X_r^t,u(r,X_r^t),(\nabla\varphi\sigma)^i(r,X_r^t),\mathcal{I}\varphi^i(r,X_r^t))\Big)dr \\ &\quad - \int_s^\theta \widetilde{Z}_r^t dB_r^t - \int_s^\theta \int_{\mathbb{R}^d} \widetilde{V}_r^t(y)\widetilde{\chi}^t(dy,dr),\end{aligned}$$

by definition (194) of ψ. One thus has for $s \in [t, \theta]$:

$$\widetilde{Y}_s^t = \varphi^i(\theta, X_\theta^t) - \int_s^\theta \Big(\psi(\zeta, X_\zeta^t) - g^i(\zeta, X_\zeta^t, u(\zeta, X_\zeta^t), \widetilde{Z}_\zeta^t, \mathcal{I}\varphi^i(\zeta, X_\zeta^t))\Big) d\zeta$$
$$- \int_s^\theta \widetilde{Z}_\zeta^t dB_\zeta^t - \int_s^\theta \int_{\mathbb{R}^d} \widetilde{V}_\zeta^t(y) \widetilde{\chi}^t(dy, d\zeta) \tag{200}$$

Note that in (199) and (200), one has by definitions (120) of $\widetilde{r}_\zeta^t = \widetilde{r}_\zeta^t(V_\zeta^t)$, (189) of $\mathcal{I}$ and (197) of $\widetilde{V}$:

$$\int_s^\theta g^i\Big(\zeta, X_\zeta^t, u(\zeta, X_\zeta^t), \widehat{Z}_\zeta^t, \widetilde{r}_\zeta^t\Big) d\zeta$$
$$= \int_s^\theta g^i\Big(\zeta, X_\zeta^t, u(\zeta, X_\zeta^t), \widehat{Z}_\zeta^t, \int_{\mathbb{R}^d} \widehat{V}_\zeta(y) f^i(\zeta, X_\zeta^t, y) m(dy)\Big) d\zeta$$
$$\int_s^\theta g^i\Big(\zeta, X_\zeta^t, u(\zeta, X_\zeta^t), \widetilde{Z}_\zeta^t, \mathcal{I}\varphi^i(\zeta, X_\zeta^t)\Big) d\zeta$$
$$= \int_s^\theta g^i\Big(\zeta, X_\zeta^t, u(\zeta, X_\zeta^t), \widetilde{Z}_\zeta^t, \int_{\mathbb{R}^d} \widetilde{V}_\zeta(y) f^i(\zeta, X_\zeta^t, y) m(dy)\Big) d\zeta \ .$$

In conclusion, (199) and (200) respectively mean that:

- $(\widehat{Y}^t, \widehat{Z}^t, \widehat{V}^t)$ is an $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, χ^t) – solution to the stopped BSDE on $[t, t+\alpha]$ with driver (cf. Definition 8(d) and Note 8(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;

- $(\widetilde{Y}^t, \widetilde{Z}^t, \widetilde{V}^t)$ is an $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, χ^t) – solution to 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 [38]):

$$\Gamma_t^t \delta Y_t^t = \mathbb{E}^t\Big[\Gamma_\theta^t \delta Y_\theta^t + \int_t^\theta \Gamma_s^t dA_s^t\Big] \tag{201}$$

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 (199) and (200) and since $u^i \leq \varphi^i$;

- $dA^t_s = \psi(r, X^t_s)ds - d\widehat{K}^{t,-}_s$, so that A^t is decreasing on $[t, \theta]$, by (194);
- Γ^t is a positive process, the so-called *adjoint* of δY^t (see, for instance, [38]).

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 (201). But this contradicts (198).

(ii) v is a continuous function on $\mathcal{E}$, by definition. Moreover by definitions of u, v, $\mathcal{Y}$ and $\bar{\mathcal{Y}}$ (with τ defined as in Example (3)), we have, for $(t, x, i) \in \partial\mathcal{D}$:

$$v^i(t,x) = \bar{Y}^t_t = Y^t_t = u^i(t,x),$$

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 $\partial\mathcal{D}$ and $\ell \le v$ on $\mathcal{E}$. We now show that v is a subsolution of $(\mathcal{V}1)$ on $\operatorname{Int}\mathcal{D}$. We let the reader check likewise that v is a supersolution of $(\mathcal{V}1)$ on $\operatorname{Int}\mathcal{D}$. Let then $(t, x, i) \in \operatorname{Int}\mathcal{D}$ 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

$$-\widetilde{\mathcal{G}}\varphi^i(t,x) - g^i(t,x,v(t,x),(\nabla\varphi\sigma)^i(t,x),\mathcal{I}\varphi^i(t,x)) \le 0, \tag{202}$$

where it is further assumed that $v^i(t,x) > \ell^i(t,x)$ and $v^i(t,x) = \varphi^i(t,x)$ (cf. Definition 12(a)(i)). Suppose by contradiction that (202) does not hold. One then has by continuity,

$$\psi(s,y) := \widetilde{\mathcal{G}}\varphi^i(s,y) + g^i(s,y,v(s,y),(\nabla\varphi\sigma)^i(s,y),\mathcal{I}\varphi^i(s,y)) < 0 \tag{203}$$

for any (s, y) such that $(s, y, i) \in \operatorname{Int}\mathcal{D}$, $s \in [t, t+\alpha]$ and $|y - x| \le \alpha$, for some small enough $\alpha > 0$. Let

$$\theta = \inf\{s \ge t\,;\, |X^t_s - x| \ge \alpha,\, N^t_s \ne i,\, \bar{Y}^t_s = \ell^i(s, X^t_s)\} \wedge (t+\alpha) \wedge \tau^t \tag{204}$$

$$(\widehat{Y}^t, \widehat{Z}^t, \widehat{V}^t, \widehat{K}^t) = \left(\mathbb{1}_{.<\theta}\bar{Y}^t + \mathbb{1}_{.\ge\theta}v^i(\theta, X^t_\theta)\ ,\ \mathbb{1}_{.\le\theta}\bar{Z}^t\ ,\ \mathbb{1}_{.\le\theta}\bar{V}^t\ ,\ \bar{K}^t_{.\wedge\theta}\right) \tag{205}$$

$$(\widetilde{Y}^t, \widetilde{Z}^t, \widetilde{V}^t) = \Big(\varphi^i(\cdot X^t_{.\wedge\theta})\ ,\ \mathbb{1}_{.\le\theta}(\nabla\varphi\sigma)^i(\cdot, X^t_.)\ , \mathbb{1}_{.\le\theta}\big(\big[\varphi^i(\cdot, X^t_{.-} + \delta^i(\cdot, X^t_{.-}, y)) - \varphi^i(\cdot, X^t_{.-})\big]\big)_{y\in\mathbb{R}^d}\Big). \tag{206}$$

Using in particular the fact that D is open in (167), one has that $\theta > t$, $\mathbb{P}^t$ – almost surely. Thus, using also the continuity of v^i:

$$\widehat{Y}^t_t = \bar{Y}^t_t = v^i(t,x) = \varphi^i(t,x) = \widetilde{Y}^t_t\ ,\ \mathbb{P}^t\text{-a.s.} \tag{207}$$

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 $\ell^i(s, X_s^t) < \bar{Y}_s^t$ on $[t, \theta)$ and $\theta \leq \tau^t$. By using the stopped RBSDE equation for $\bar{\mathcal{Y}}^t$, one can then show like (199) in part (i) that one has, for any $s \in [t, \theta]$:

$$\widehat{Y}_s^t = v^i(\theta, X_\theta^t) + \int_s^\theta g^i(\zeta, X_\zeta^t, v(\zeta, X_\zeta^t), \widehat{Z}_\zeta^t, \bar{r}_\zeta^t) d\zeta - \int_s^\theta \widehat{Z}_\zeta^t dB_\zeta^t - \int_s^\theta \int_{\mathbb{R}^d} \widehat{V}_\zeta^t(y) \widetilde{\chi}^t(d\zeta, dy) \tag{208}$$

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, $(\widehat{Y}^t, \widehat{Z}^t, \widehat{V}^t)$ is an $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, χ^t) – solution to the stopped BSDE on $[t, t + \alpha]$ with driver (cf. Note 8(i))

$$g^i\Big(s, X_s^t, v(s, X_s^t), z, \int_{\mathbb{R}^d} \nu(y) f^i(s, X_s^t, y) m(dy)\Big) ds$$

(where, to avoid confusion with the value function $v = v^i(t, x)$ in $v(\zeta, X_\zeta^t)$, $\nu(y)$ here, usually denoted by $v(y)$ elsewhere, refers to a generic function $\nu \in \mathcal{M}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m(dy); \mathbb{R})$), and terminal condition $v^i(\theta, X_\theta^t)$ at θ. Besides, one can show as in part (i) above that $(\widetilde{Y}^t, \widetilde{Z}^t, \widetilde{V}^t)$ is an $(\Omega, \mathbb{F}^t, \mathbb{P}^t)$, (B^t, χ^t) – solution to the stopped BSDE on $[t, t + \alpha]$ with driver

$$g^i\Big(s, X_s^t, v(s, X_s^t), z, \int_{\mathbb{R}^d} \nu(y) f^i(s, X_s^t, y) m(dy)\Big) ds - \psi(s, X_s^t) ds$$

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, 9], Pardoux et al. [79], Pham [82], Harraj et al. [58], Amadori [2, 3] and Ma and Cvitanic [76], among others.

Assumption 8 **(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

$$|g^i(t, x, u, z, r) - g^i(t, x', u, z, r)| \leq \eta_R(|x - x'|(1 + |z|))$$

for any $t \in [0, T]$, $i \in I$, $z \in \mathbb{R}^{1\otimes d}$, $r \in \mathbb{R}$ and $x, x' \in \mathbb{R}^d$, $u \in \mathbb{R}^k$ with $|x|, |x'|, |u| \leq R$;
(iii) The function g^i is non-decreasing with respect to u^j, for any $(i, j) \in I^2$ with $i \neq j$.

Note 15. **(i)** By Assumption 8(i), one has in particular

$$|b| \vee |\sigma| \vee |\delta| < C(1 + |x|) \tag{209}$$

on $\mathcal{E}$.
(ii) The monotonicity Assumption 8(iii) on g means that we deal with a *cooperative* system of PIDEs (see, for instance, Busca and Sirakov [29]).

We are now in position to establish the following

Theorem 6. *One has $\mu \leq \nu$ 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 at T, respectively of $(\mathcal{V}1)$ on $\mathcal{E}$ with boundary condition u on $\partial\mathcal{D}$.*

As we first show, one can reduce attention, for the sake of establishing Theorem 6, to the special case where 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 8(iii). Note 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 8(iii). Thus,

Lemma 10. *If Theorem 6 holds in the special case where g^i is non-decreasing with respect to u^j for any $(i, j) \in I^2$, then Theorem 6 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. Indeed, under the general assumptions of Theorem 6, $e^{-Rt}\mu$ and $e^{-Rt}\nu$ are respectively $\mathcal{P}$-subsolution and $\mathcal{P}$-supersolution of the following transformed problem, for $(\mathcal{V}2)$,

$$\begin{aligned}\max\Big(\min\Big(&-\widetilde{\mathcal{G}}\varphi^i(t, x) - e^{-Rt}g^i(t, x, e^{Rt}\varphi(t, x), e^{Rt}(\nabla\varphi\sigma)^i(t, x),\\ &e^{Rt}\mathcal{I}\varphi^i(t, x)) - R\varphi^i(t, x),\\ &\varphi^i(t, x) - e^{-Rt}\ell^i(t, x)\Big), \varphi^i(t, x) - e^{-Rt}h^i(t, x)\Big) = 0\end{aligned}$$

on $\mathrm{Int}\,\mathcal{E}$, supplemented by the terminal condition $\varphi = e^{-Rt}\Phi$ at T (and likewise with $h = +\infty$ for $(\mathcal{V}1)$ on $\mathrm{Int}\,\mathcal{D}$, supplemented by the boundary condition $\varphi = e^{-Rt}\Psi$ on $\partial\mathcal{D}$). Now, for R large enough, Assumption 8(iii) and the Lipschitz continuity property of g with respect to the last variable imply that $e^{-Rt}g(t,x,e^{Rt}u,e^{Rt}z,e^{Rt}r) + Ru^i$ is non-decreasing with respect to u. One thus concludes by an application of the assumed restricted version of Theorem 6. □

Given Lemma 10, one may and do reduce attention, in order to prove Theorem 6, to the case where the function g is non-decreasing with respect to u. The statement regarding $(\mathcal{V}2)$ in Theorem 6 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 6 would be analogous, substituting $(\mathcal{V}1)$ to $(\mathcal{V}2)$, $+\infty$ to h, $\mathrm{Int}\,\mathcal{D}$ to $\mathrm{Int}\,\mathcal{E}$ and $\mathcal{C}^0(\mathcal{E}) \cap \mathcal{C}^{1,2}(\mathcal{D})$ to $\mathcal{C}^{1,2}(\mathcal{E})$ in Lemma 11 below and its proof.

Let $\Lambda_1 = k\Lambda$ where Λ is the Lipschitz constant of g (cf. Assumption (M.1.ii) in Sect. 6.4).

Lemma 11. *Given a $\mathcal{P}$-subsolution μ and a $\mathcal{P}$-supersolution ν of $(\mathcal{V}2)$ on $\mathcal{E}$, assuming g non-decreasing with respect to u :*
(i) *$\omega = \mu - \nu$ is a $\mathcal{P}$-subsolution of*

$$\min\Big(w, -\widetilde{\mathcal{G}}\omega \; - \Lambda_1\big(\max_{j\in I}(\omega^j)^+ + |\nabla\omega\sigma| + (\mathcal{I}\omega)^+\big)\Big) = 0$$

on $\mathcal{E}$ with null boundary condition at T, in the sense that:
- *$\omega \le 0$ at T, and*
- *$\omega^i(t,x) > 0$ implies*

$$-\widetilde{\mathcal{G}}\varphi^i(t,x) \; - \Lambda_1\Big(\max_{j\in I}(\omega^j(t,x))^+ + |\nabla\varphi^i(t,x)\sigma^i(t,x)| + (\mathcal{I}\varphi^i(t,x))^+\Big) \le 0 \tag{210}$$

for any $(t,x,i) \in \mathrm{Int}\,\mathcal{E}$ and $\varphi \in \mathcal{C}^{1,2}(\mathcal{E})$ such that $\omega^i - \varphi^i$ reaches a global null maximum at (t,x).
(ii) *For every $q_1 > 0$, 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*

$$\min\Big(\chi, -\widetilde{\mathcal{G}}\chi - \Lambda_1\big(\chi + |\nabla\chi\sigma| + (\mathcal{I}\chi)^+\big)\Big) = 0$$

on $\mathcal{E}$, in the sense that $\chi > 0$ and

$$-\widetilde{\mathcal{G}}\chi - \Lambda_1\big(\chi + |\nabla\chi\sigma| + (\mathcal{I}\chi)^+\big) > 0 \tag{211}$$

on $\mathcal{E}$.

(iii) *For q_1 in part (ii) greater than q_2 such that $\mu, \nu \in \mathcal{P}_{q_2}$, where q_2 is provided by our assumption that $\mu, \nu \in \mathcal{P}$, one has $\max_{i \in I}(\omega^i)^+ \leq \alpha\chi$ 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. [9] (see also Pardoux et al. [79] and Harraj et al. [58]). Here are the main differences (our assumptions are fitted to financial applications, cf. Part I):

- We consider a model with jumps in X and regimes represented by N, whereas [9] or [58] only consider jumps in X, and [79] only considers regimes;
- 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 [9] or [58] consider general Levy measures, bounded jumps, and continuous solutions with sub-exponential (strictly including polynomial) growth in x;
- [9] deals with classical BSDEs (without barriers);
- We consider time-dependent coefficients b, σ, δ whereas [9] considers homogeneous dynamics.

Because of these differences we provide a detailed proof in Appendix 18.1.

To conclude this section we can state the following proposition, which sums-up the results of Theorems 5 and 6.

Proposition 31. **(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 at T;*
(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 $\partial\mathcal{D}$.* □

13 Approximation

An important feature of semi-continuous viscosity solutions comparison principles like Theorem 6 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 [8]. See also Briani, La Chioma and Natalini [28], Cont and Voltchkova [36] or Jakobsen et al. [64] for various extensions of these results to PIDEs.

The following results thus extend to models with regimes, thus *systems* of PIDEs, the results of [8, 28], among others.

The following lemma is standard and elementary, and thus stated without proof.

Lemma 12. *Let $(\mathcal{E}_h)_{h>0}$ denote a family of rectangular time-space meshes of step h over $\mathcal{E}$, the time mesh including in particular the maturity time T, for every $h > 0$.*

Let $(u_h)_{h>0}$ be a family of uniformly locally bounded real-valued 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:*

$$\lim_{n\to+\infty} u^i_{h_n}(t_n,x_n) \text{ with } h_n \to 0 \text{ and } \mathcal{E}^{h_n} \ni (t_n,x_n,i) \to (t,x,i) \text{ as } n \to \infty, \tag{212}$$

is non empty and compact in $\mathbb{R}$. It admits as such a smallest and a greatest element: $\underline{u}^i(t,x) \leq \overline{u}^i(t,x)$ 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. Let us be given families of operators

$$\widetilde{\mathcal{G}}_h = \widetilde{\mathcal{G}}_h u^i(t_h,x_h)\,,\ \nabla_h = \nabla_h u^i(t_h,x_h),\ \mathcal{I}_h = \mathcal{I}_h u^i(t_h,x_h)$$

devoted to approximate $\widetilde{\mathcal{G}}u^i(t_h,x_h)$, $\nabla u^i(t_h,x_h)$ and $\mathcal{I}u^i(t_h,x_h)$ on $\mathcal{E}_h$ for real-valued functions u on $\mathcal{E}$, respectively. For $\mathcal{L} = \nabla$, $\mathcal{I}$ or $\widetilde{\mathcal{G}}$, we say that:
(i) The discretized operator $\mathcal{L}_h = \nabla_h$, $\mathcal{I}_h$ or $\widetilde{\mathcal{G}}_h$ is *monotone*, if

$$\mathcal{L}_h u^i_1(t_h,x_h) \leq \mathcal{L}_h u^i_2(t_h,x_h) \tag{213}$$

for any functions $u_1 \leq u_2$ on $\mathcal{E}_h$ with $u^i_1(t_h,x_h) = u^i_2(t_h,x_h)$;
(ii) The 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:

$$\mathcal{L}_h(\varphi+\xi_h)^i(t_h,x_h) \to \mathcal{L}\varphi^i(t,x) \tag{214}$$

whenever $h \to 0+$, $\mathcal{E}_h \ni (t_h,x_h,i) \to (t,x,i) \in \mathcal{E}$ and $\mathbb{R} \ni \xi_h \to 0$.

Moreover we also assume g to be monotone in the following sense.

Assumption 9 The function

$$\mathbb{R}^k \times \mathbb{R}^{1\otimes d} \times \mathbb{R} \ni (u,p,r) \mapsto g^{\cdot}(\cdot,\cdot,u,p\sigma,r) \in \mathbb{R}^{\mathcal{E}} \tag{215}$$

is non-decreasing, in the sense that for any $(u,p,r) \leq (u',p',r')$ coordinate by coordinate in $\mathbb{R}^k \times \mathbb{R}^{1\otimes d} \times \mathbb{R}$, 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 $(t,x,i) \in \mathcal{E}$.

Note 16. **(i)** The "abstract" monotonicity and consistency conditions of Definition 13 need to be verified carefully on a case-by-case basis for any concrete approximation scheme under consideration (e.g., finite difference schemes). We refer the reader to Cont and Voltchkova [36] (see also Jakobsen et al. [64]) for the complete analysis of specific schemes under various sets of assumptions.
(ii) The monotonicity of g with respect to p, which is the most stringent condition in Assumption 9, 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 Sect. 6.6);
• σ is equal to zero, which corresponds to the situation of pure jump models; note however that our continuity Assumption 4 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 Example 3 (cf. Assumption 7(iii));
• The dimension d 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 $\nabla\varphi\sigma$, by non-negativity of σ in the scalar case (see, for instance, Kushner and Dupuis [71]).
(iii) 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, for R large enough, the mapping $u^i(t,x) \mapsto \widetilde{u}^i(t,x) := e^{-Rt}u^i(t,x)$ transforms the problem into one in which Assumption 9 holds (see the proof of Lemma 10). 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 part (a) of the following lemma, we mean that u_h is bounded by $C(1+|x|^q)$ for some C and q independent of h.

Lemma 13. *Let us be given monotone and consistent approximation schemes*

$$(\widetilde{\mathcal{G}}_h)_{h>0},\ (\nabla_h)_{h>0} \text{ and } (\mathcal{I}_h)_{h>0}$$

for $\widetilde{\mathcal{G}}$, ∇ and $\mathcal{I}$ respectively, g satisfying the monotonicity Assumption 9.

(a) *Let* $(u_h)_{h>0}$ *be uniformly polynomially bounded and satisfy*

$$\max\Big(\min\Big(-\widetilde{\mathcal{G}}_h {u_h}^i(t_h,x_h) - g^i(t_h,x_h,u_h(t_h,x_h),(\nabla_h u_h\sigma)^i(t_h,x_h), \times\mathcal{I}_h u_h^i(t_h,x_h)), \tag{216}$$

$$u_h^i(t_h,x_h) - \ell^i(t_h,x_h)\Big), u_h^i(t_h,x_h) - h^i(t_h,x_h)\Big) = 0 \tag{217}$$

on $\mathrm{Int}\,\mathcal{E}\cap\mathcal{E}_h$ *and* $u_h = \Phi$ *on* $\partial\mathcal{E}\cap\mathcal{E}_h$ *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* $\mathrm{Int}\,\mathcal{E}$;
(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*

$$\min\Big(-\widetilde{\mathcal{G}}_h v_h^i(t_h,x_h) - g^i(t_h,x_h,v_h(t_h,x_h),(\nabla_h v_h\sigma)^i(t_h,x_h), \times\mathcal{I}_h v_h^i(t_h,x_h)), \tag{218}$$

$$v_h^i(t_h,x_h) - \ell^i(t_h,x_h)\Big) = 0 \tag{219}$$

on $\mathrm{Int}\,\mathcal{D}\cap\mathcal{E}_h$ *and* $v_h = u$ *on* $\partial\mathcal{D}\cap\mathcal{E}_h$ *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* $\mathrm{Int}\,\mathcal{D}$;
(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 11). *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* $\partial\mathcal{D}$; *see comments in part (ii) below.*
(i) We prove that $\overline{u}$ is a viscosity subsolution of $(\mathcal{V}2)$ on $\mathrm{Int}\,\mathcal{E}$. The fact that $\underline{u}$ is a viscosity supersolution of $(\mathcal{V}2)$ on $\mathrm{Int}\,\mathcal{E}$ can be shown likewise. First note that $\overline{u}\le h$, by (216) on $\mathrm{Int}\,\mathcal{E}\cap\mathcal{E}_h$, inequality $\Phi\le h$ at T (cf. Assumption (M.2.ii) in Sect. 6.4) and continuity of h and Φ. Let then $(t^\star,x^\star,i)\in\mathrm{Int}\,\mathcal{E}$ 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. (192)):

$$-\widetilde{\mathcal{G}}\varphi^i(t^\star,x^\star) - g^i\left(t^\star,x^\star,\overline{u}(t^\star,x^\star),(\nabla\varphi\sigma)^i(t^\star,x^\star),\mathcal{I}\varphi^i(t^\star,x^\star)\right)\le 0. \tag{220}$$

By a classical argument in the theory of viscosity solutions (see, e.g., Barles and Souganidis [8]), 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):

$$u_h^i\le\varphi^i + (u_h-\varphi)^i(t,x) \tag{221}$$

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 (216):

$$-\widetilde{\mathcal{G}}_h u_h^i(t,x) - g^i(t,x,u_h(t,x),(\nabla_h u_h\sigma)^i(t,x),\mathcal{I}_h u_h^i(t,x)) \le 0. \quad (222)$$

Given (221), one thus has by monotonicity of the scheme and of g (Assumption 9):

$$\begin{aligned}
-\widetilde{\mathcal{G}}_h(\varphi+\xi_h)^i(t,x) &\le g^i\left(t,x,u_h(t,x),(\nabla_h(\varphi+\xi_h)\sigma)^i(t,x),\mathcal{I}_h(\varphi+\xi_h)^i(t,x)\right)\\
&\le g^i\left(t^\star,x^\star,\overline{u}(t^\star,x^\star),(\nabla\varphi\sigma)^i(t^\star,x^\star),\mathcal{I}\varphi^i(t^\star,x^\star)\right)\\
&\quad+\eta(|t-t^\star|)+\eta_R(|x-x^\star|(1+|(\nabla\varphi\sigma)^i(t^\star,x^\star)|))\\
&\quad+\Lambda_1\max_{j\in I}(u_h^j(t,x)-\overline{u}^j(t^\star,x^\star))^+ + \Lambda|(\nabla_h(\varphi+\xi_h)\sigma)^i(t,x)\\
&\quad-(\nabla\varphi\sigma)^i(t^\star,x^\star)| + \Lambda(\mathcal{I}_h(\varphi+\xi_h)^i(t,x)-\mathcal{I}\varphi^i(t^\star,x^\star))^+,
\end{aligned}$$

where in the last inequality (cf. the proof of Lemma 11(i) in Appendix 18.1):

- η is a modulus of continuity of g^i on a "large" compact set around

$$(t^\star,x^\star,\overline{u}(t^\star,x^\star),(\nabla\varphi\sigma)^i(t^\star,x^\star),\mathcal{I}\varphi^i(t^\star,x^\star)),$$

Λ_1 stands for $k\Lambda$, and η_R is the modulus of continuity standing in Assumption 8(ii);
- The three last terms come from the Lipschitz continuity and monotonicity properties of g.

Inequality (220) follows by sending h to zero in the previous inequality, using the consistency (214) 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* at T, namely in the case of $\overline{u}$ (the related statement and proof are similar in the case of $\underline{u}$): Inequality (220) holds for any $(t^\star = T, x^\star, i)$ and $\varphi \in \mathcal{C}^{1,2}(\mathcal{E})$ such that

$$\overline{u}^i(t^\star,x^\star) > \ell^i(t^\star,x^\star) \vee \Phi^i(t^*,x^*) \quad (223)$$

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 (221) 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 (223) implies that

$$u_h^i(t,x) > \ell^i(t,x) \vee \Phi^i(t,x)$$

for h small enough, whence $(t,x,i) \in \operatorname{Int}\mathcal{E}$ and by (216):

$$-\widetilde{\mathcal{G}}_h u_h^i(t,x) - g^i(t,x,u_h(t,x),(\nabla_h u_h\sigma)^i(t,x),\mathcal{I}_h u_h^i(t,x)) \le 0\,. \quad (224)$$

Inequality (220) 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 $\partial\mathcal{D}$, 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 $\operatorname{Int}\mathcal{E}$ satisfying the boundary condition in the weak viscosity sense at T, satisfies it pointwise at T. So, in our case, suppose for instance by contradiction that

$$\overline{u}^i(T,x^\star) > \Phi^i(T,x^*) \tag{225}$$

for some $x^\star \in \mathbb{R}^d$. Let us then introduce the function

$$\varphi^i_\varepsilon(t,x) = \overline{u}^i(t,x) - \frac{|x^*-x|^2}{\varepsilon} - C_\varepsilon(T-t) \tag{226}$$

in which

$$\begin{aligned} C_\varepsilon > \sup_{(t,x)\in[t-\eta,T]\times\bar{B}_1(x^*)} & \widetilde{\mathcal{G}}\left(\frac{|y-x^*|^2}{\varepsilon}\right)^i(t,x) + g^i\left(t,x,\overline{u}(t,x),\left(\frac{2(y-x^*)\sigma}{\varepsilon}\right)^i(t,x),\right. \\ & \left.\mathcal{I}\left(\frac{|y-x^*|^2}{\varepsilon}\right)^i(t,x)\right) \end{aligned} \tag{227}$$

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. (226))

$$0 \le \frac{|x^*-x|^2}{\varepsilon} + C_\varepsilon(T-t) \le \overline{u}^i(t,x) - \overline{u}^i(T,x^\star) \tag{228}$$

so

$$\overline{u}^i(T,x^\star) \le \overline{u}^i(t,x). \tag{229}$$

Since $\overline{u}$ is locally bounded, (228) implies that $(t,x) \to (T,x^\star)$ as $\varepsilon \to 0$, which, joint to the upper semi-continuity of $\overline{u}$ and to (229), implies that $\overline{u}^i(t,x) \to \overline{u}^i(T,x^\star)$ as $\varepsilon \to 0$.

Now one has $\ell \leq \Phi$ pointwise at T, therefore (225) 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

$$C_\varepsilon - \widetilde{\mathcal{G}}\left(\frac{|y-x^*|^2}{\varepsilon}\right)^i(t,x) - g^i\left(t,x,\overline{u}(t,x),\left(\frac{2(y-x^*)\sigma}{\varepsilon}\right)^i(t,x),\right.$$
$$\left.\mathcal{I}\left(\frac{|y-x^*|^2}{\varepsilon}\right)^i(t,x)\right) \leq 0,$$

which for ε small enough contradicts (227). □

Proposition 32. *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(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 $\partial\mathcal{D} \cap \{t < T\}$.*

Proof. **(a)** By Lemma 13(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 $\overline{u} \leq \underline{u}$, by Theorem 6. Moreover $\underline{u} \leq \overline{u}$ by Lemma 12(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 12(iii).
(b) By Lemma 13(b)(i), $\overline{v}$ and $\underline{v}$ are respectively viscosity subsolutions and supersolutions of $(\mathcal{V}1)$ on $\operatorname{Int}\mathcal{D}$. Moreover, they satisfy $\overline{v} \leq u \leq \underline{v}$ at T, by Lemma 13(b)(ii). If, in addition, $v_h \to v(=u)$ on $\partial\mathcal{D} \cap \{s < T\}$, then $\overline{v} \leq u \leq \underline{v}$ on $\partial\mathcal{D}$, 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 26. The convergence result regarding v in Proposition 32(b) can only be considered as a partial result, since one only gets the convergence on $\mathcal{E}$ conditionally on the convergence on $\partial\mathcal{D} \cap \{t < T\}$, 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 $\partial\mathcal{D}$ in the approximation scheme for v (cf. the boundary condition "$v_h = u$ on $\partial\mathcal{D} \cap \mathcal{D}_h$" in Lemma 13(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 practical issues like *discrete dividends* or *discrete path-dependence* in the context of pricing problems in finance.

Let us thus be given a set $\mathfrak{T} = \{T_0, T_1 \ldots, T_m\}$ 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. We set, for $l = 1, \ldots, m$,

$$\mathcal{E}_l = [T_{l-1}, T_l] \times \mathbb{R}^d \times I \ , \ \mathcal{D}_l = [T_{l-1}, T_l] \times D$$

and we define $\mathrm{Int}\,\mathcal{E}_l$, $\partial\mathcal{E}_l$, $\mathrm{Int}\,\mathcal{D}_l$ and $\mathrm{Int}\,\mathcal{D}$ as the *parabolic interiors and boundaries* of $\mathcal{E}_l$ and $\mathcal{D}_l$ as of (187). Note that the sets $\mathrm{Int}\,\mathcal{E}_l$s and $\partial\mathcal{E} = \{T\} \times \mathbb{R}^d \times I$, partition $\mathcal{E}$.

Discrete dividends on a financial derivative or on an underlying asset (component of the factor process $\mathcal{X}$) motivate separate developments presented in Sects. 14 and 15, respectively. Sect. 16 deals with the issue of discretely monitored call protection (discretely monitored and *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 Sect. 4.2.1), 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. Now, discrete coupons imply predictable jumps, by the coupon amounts, of the related financial derivatives arbitrage price processes at the T_ls. But 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 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. One might thus think that pricing problems with discrete coupons are not amenable to the methods of this paper.

However, as demonstrated in [18–20, 23], this apparent difficulty can be handled by working with a suitable notion of *clean* (instead of *ex-dividend*) price process for a financial derivative. Here clean price means (ex-dividend) 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.

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 $\mathrm{Int}\,\mathcal{E}_l$s of functions g_ls satisfying our usual assumptions relatively to the $\mathcal{E}_l$s. Definition 12 for viscosity solutions of $(\mathcal{V}2)$ and $(\mathcal{V}1)$ then needs to be amended as follows.

Definition 14. **(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 $(t,x,i)\in\mathrm{Int}\,\mathcal{E}$, if and only if the restriction of u to $\mathcal{E}_l$ with $(t,x,i)\in\mathrm{Int}\,\mathcal{E}_l$ is a viscosity subsolution, resp. supersolution, resp. resp. solution, of $(\mathcal{V}2)$ at (t,x,i), relatively to $\mathcal{E}_l$ (cf. Definition 12(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 formally defined as in Definition 12(b), with the embedded notions of viscosity subsolution, resp. supersolution, resp. resp. solution, of $(\mathcal{V}2)$ at any (t,x,i) in $\mathrm{Int}\,\mathcal{E}$ defined as in (i) above.
(iii) The notions of viscosity subsolutions, supersolutions and solutions of $(\mathcal{V}1)$ at $(t,x,i)\in\mathrm{Int}\,\mathcal{D}$, and, given a further continuous boundary condition Ψ on $\partial\mathcal{D}$ such that $\Psi=\Phi$ at T, those of $\mathcal{P}$ – viscosity subsolutions, supersolutions and solutions of $(\mathcal{V}1)$ on $\mathcal{E}$, are defined similarly (cf. Definition 12(c)).

Proposition 33. *Using Definition 14 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 $\widetilde{g}(s,\mathcal{X}^t_s,Y^t_s,Z^t_s,\mathcal{V}^t_s)$ (cf. (119)) for any $(Y^t,Z^t,\mathcal{V}^t)\in\mathcal{S}^2\times\mathcal{H}^2_d\times\mathcal{H}^2_{\mu^t}$, and second, for the stability results of Propositions 25(ii) and 27(ii). But it can be checked by inspection of the related proofs that these stability results are still true under the currently relaxed assumption on g. Moreover the process $\widetilde{g}(s,\mathcal{X}^t_s,Y^t_s,Z^t_s,\mathcal{V}^t_s)$ is obviously still well-defined under the current assumption on g, for any $(Y^t,Z^t,\mathcal{V}^t)\in\mathcal{S}^2\times\mathcal{H}^2_d\times\mathcal{H}^2_{\mu^t}$.

In Part III, Theorem 5 still holds true, by immediate inspection of its proof. Moreover, under an "l by l" version of Assumption 8(ii) on the g_ls, Lemma 11 and Theorem 6 (hence Proposition 31 also follows) can be proven together iteratively on l as we now show. Let thus 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 11 relative to $\mathcal{E}_m$ is proven in exactly the same way as before. We thus have (cf. Theorem 6) $\mu\le\nu$ on $\mathcal{E}_m$. One can then establish likewise the version of Lemma 11 relative to $\mathcal{E}_{m-1}$ (note that $\mu-\nu\le 0$ on $\partial\mathcal{E}_{m-1}$, by the first step of the proof). So $\mu\le\nu$ on $\mathcal{E}_{m-1}$, and so on until $l=1$. Lemma 12 is of course not affected by the relaxation of the assumption on g. Finally, given Definition 14, Lemma 13(a)(i) can be proven exactly as before, on each $\mathrm{Int}\,\mathcal{E}_l$, and the proof of Lemma 13(a)(ii) does not change. Lemma 13(a) is thus still true, and so is likewise Lemma 13(b), hence Proposition 32 holds 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 Sect. 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 Sect. 14), specifically,

$$X_{T_l} = \theta_l(\mathcal{X}_{T_l-}),$$

where the *jump function* θ is given as a system of Lipschitz functions $y \mapsto \theta_l^j(y)$ from $\mathbb{R}^d$ into itself, for every $i \in I$ and $l = 1, \ldots, m$.

Definition 15. **(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 $l = 1, \ldots, m$:

$$u_l^i(T_l, x) = u_{l+1}^i(T_l, \theta_l^i(x)) \tag{230}$$

where, in case $l = m$, u_{l+1}^i is to be understood as Φ in the right-hand-side of (230). 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 Int $\mathcal{E}_l$s of the u_ls, along with the terminal condition Φ at T.

The formal analogue of Definition 11 for a Markovian 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 then be formulated, where :
• A "model $\mathcal{X}$ with generator $\mathcal{G}$" in Definition 11(a) is to be understood here 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 11(b):
– The deterministic value function u in Definition 11(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 11(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 $l = 1,\dots,m$, and $\Lambda \le \Phi \le \Upsilon$, whence in particular

$$\ell^i_m(T,x) = \Lambda^i(T,\theta^i_m(x)) \le \Phi^i(T,\theta^i_m(x)) \le \Upsilon^i(T,\theta^i_m(x)) = h^i_m(T,x). \quad (231)$$

Note that $\ell(s,\mathcal{X}^t_s)$ and $h(s,\mathcal{X}^t_s)$ are then quasi-left continuous processes satisfying our standing assumption (H.2) in Sect. 5.1, as should be in view of application of general reflected BSDE results (see, e.g., [38]).

Suitable semi-group properties analogous to Propositions 26 and 28 in Part II, and existence of a Markovian solution in the above sense to the Markovian decoupled forward backward SDE with data $\mathcal{G}$, $\mathcal{C}$ and τ (cf. Theorems 3, 4 and Proposition 30 in Part II), can then be established like in Part II (see also Theorem 11 in Part IV below).

Remark 27. The fact that the value functions u and v are defined by continuous Cauchy cascades can be established much like Theorem 11 below (see also Chassagneux and Crépey [31]). Since the proof is simpler here, we do not provide it, referring the reader to the proof of Theorem 11 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 obstacles partial integro-differential problems.

Reasoning as in Part III (cf. the proof of Proposition 33 for a review of the main arguments), one can thus show,

Proposition 34. *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 $l = 1,\dots,m$,*

- *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 $u^i_{l+1}(T_l,\theta^i_l(x))$ on $\partial\mathcal{E}_l$ – with u_{l+1} in the sense of Φ, in case $l = m$,*
- *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 $\partial\mathcal{D}_l$.*

Part (ii) of this Proposition is thus the generalization to the present set-up of Proposition 31 in Part III. As for the approximation arguments of Sect. 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 $\partial\mathcal{D}_l$ in the approximation scheme for v_l, and provided $v_{l,h} \to v_l (= u_l)$ on $\partial\mathcal{D}_l \cap \{t < T_l\}$.

Of course, in practice (cf. also Remark 26):

- 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 $\partial\mathcal{D}^l$ when it comes to approximating v_l on $\mathcal{E}_l$, using the already computed function $u_{l,h}$ on $\partial\mathcal{D}_l$ as input data.

There is thus clearly room 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 Sect. 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 found in *cliquet options*, *volatility and variance swaps*, or *discretely monitored Asian options*. As is well known, these can often be efficiently priced by PDE methods after an appropriate extension of the state space. We refer the reader to Windcliff et al. [86, 87] for illustrations in the cases of cliquet options and volatility and variance swaps, respectively.

Provided one works with a suitably extended state space, the methods and results of the present paper are 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 $X_t = (S_t, S_t^0, \ldots, S_t^{m-1})$, 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 Sect. 15.1 (case of a degenerate model $\mathcal{X} = (X, N) = X$ therein), which provides a second motivation for the developments of Sect. 15.1.

But this state space extension is not the only possible one. Exploiting 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. [86, 87]).

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.

One would thus like to be able to address the issue of a *discretely monitored call protection* τ, like for instance (cf. Examples 2 and 5),

Example 4. 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 Sect. 16 (cf. Sect. 16.3.5), it is actually possible to deal with such forms of path-dependence, resorting to a "degenerate variant" $\mathcal{X} = (X, N)$ of the general jump-diffusion with regimes of this paper, in which X is a Markovian jump-diffusion not depending on N, and where the I-valued pure jump marker process N is constant except for deterministic jumps at the T_ls, from $N^t_{T_l-}$ to

$$N^t_{T_l} = \theta_l(\mathcal{X}^t_{T_l-}), \tag{232}$$

for a suitable jump function θ.

Note 17. In this set-up:
(i) In the notation of Sect. 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 23(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 25 are valid, independently of the error estimate (136) on N. Incidentally note that (136) does not hold anymore, since N now depends on X via (232), even under the original measure $\mathbb{P}$ (before the change of measure to $\mathbb{P}^t$).

15.3 General Case

The situations of Sects. 15.1 and 15.2 are both special cases, covering many practical pricing applications, of deterministic jumps of the factor process $\mathcal{X}$ at fixed times T_ls. 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:

- 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 25(ii) and 27(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 Sect. 15.2 where X does not depend on N;
- Second, the martingale representation property of Proposition 23(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 24(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 Sect. 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

$$\bar{U}_t = \Omega_t^c \infty + \Omega_t U_t \tag{233}$$

for given càdlàg event-indicator[10] processes $\Omega_t, \Omega_t^c = 1 - \Omega_t$, rather than more specifically (cf. (107))

$$\bar{U}_t = \mathbb{1}_{\{t<\tau\}} \infty + \mathbb{1}_{\{t\geq\tau\}} U_t \tag{234}$$

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 $t \in [0, T]$,

$$\Omega_t = \mathbb{1}_{\{l_t \text{ odd}\}}, \tag{235}$$

where l_t is the index l of the random time interval $[T_l, T_{l+1})$ containing t.

Remark 28. Considering sequences τ such that $\tau_0 = \tau_1 = 0$ and $\tau_2 > 0$ almost surely, observe that this formalism includes the case where the protection is inactive on the first non-empty time interval.

[10] Boolean-valued.

In the special case of a doubly reflected BSDE of the form (15) with a generalized effective call payoff process $\bar{U}$ as of (233), (235) 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 Sect. 16.2.1).

Remark 29. 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 4 of a solution to a reflected BSDE (see Definition 16 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. [23, Long Preprint Version] or Chassagneux et al. or [31] for more about this.

However doubly reflected BSDEs with a generalized upper barrier as of (233), (235) 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 Sect. 16.2 extend to more general RIBSDEs (see Definition 16 and Remark 17) the abstract RDBSDE results of Crépey and Matoussi [38]: general well-posedness (in the sense of existence, uniqueness and a priori estimates) and comparison results. In order to recover the results of [38], simply consider in Sect. 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 $l \geq 2$. Also note that the component K of the solution is continuous in case of an RDBSDE.

We then deal with the Markovian case in Sect. 16.3.

16.2 General Set-Up

In this section one works in the general set-up and under the assumptions of Sect. 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 an 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 8(a)(iii), with for $t \in [0,T]$,

$$\bar{U}_t = \mathbb{1}_{\{l_t \text{ even}\}}\infty + \mathbb{1}_{\{l_t \text{ odd}\}}U_t \tag{236}$$

where l_t is defined by $\tau_{l_t} \leq t < \tau_{l_t+1}$. 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 (possibly discontinuous) Jordan components denoted as usual by $K^{\pm}$.

Definition 16. 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:

$$
\begin{aligned}
&(\mathbf{i})\ Y \in \mathcal{S}^2, Z \in \mathcal{H}_d^2, V \in \mathcal{H}_\mu^2, K \in A^2, \\
&(\mathbf{ii})\ Y_t = \xi + \int_t^T g_s(Y_s, Z_s, V_s)ds + K_T - K_t \\
&\qquad\qquad - \int_t^T Z_s dB_s - \int_t^T \int_E V_s(e)\widetilde{\mu}(ds, de)\ ,\ t \in [0, T]\ , \\
&(\mathbf{iii})\ L \leq Y \text{ on } [0, T]\ ,\ Y \leq \bar{U} \text{ on } [0, T] \\
&\text{and } \int_0^T (Y_{t-} - L_{t-})dK_t^+ = \int_0^T (\bar{U}_{t-} - Y_{t-})dK_t^- = 0,
\end{aligned}
$$

where $\bar{U}$ is defined by (236), and with the convention that $0 \times \pm\infty = 0$ in (iii).

Remark 30. In the special case where $\tau_2 = T$ a.s. (so $\tau_l = \tau_2 = T$ for $l \geq 2$), the RIBSDE with data (g, ξ, L, U, τ) reduces to the RDBSDE with data (g, ξ, L, U, τ_1) (see Definition 9(ii)). If moreover $\tau_1 = 0$, one then 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 18 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 29).

Proposition 35. *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:*

$$
\begin{aligned}
\rho_t^\varepsilon &= \inf\Big\{ s \in [t, T]\,;\, Y_s \geq \bar{U}_u - \varepsilon \Big\} \wedge T\ ,\ \theta_t^\varepsilon \\
&= \inf\Big\{ s \in [t, T]\,;\, Y_s \leq L_u + \varepsilon \Big\} \wedge T\ .
\end{aligned}
$$

So, for any $\rho, \theta \in \times\, \mathcal{T}_t$,

$$
\mathbb{E}\big[J^t(\tau, \theta)\big|\mathcal{F}_t\big] - \varepsilon \leq Y_t \leq \mathbb{E}\big[J^t(\rho, \theta)\big|\mathcal{F}_t\big] + \varepsilon. \tag{237}
$$

Of course, given the definition of $\bar{U}$ in (236), this Dynkin game effectively reduces to a "constrained Dynkin game" with upper payoff process U (instead of $\bar{U}$ in Proposition 35),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,

$$\rho_t^{\varepsilon} = \inf\left\{ s \in \cup_{l\geq 0}[\tau_{2l+1} \vee t, \tau_{2l+2} \vee t)\,;\, Y_s \geq U_u - \varepsilon \right\} \wedge T.$$

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 [43] (see also Protter [84]).

Lemma 14. (i) (See Sect. VI.40 of [43]) *Among the 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 [43]) *Any quasimartingale X* belonging to $\mathcal{S}^2$ *is a* special semimartingale *with canonical decomposition*

$$X_t = X_0 + M_t + A_t\ ,\ t \in [0,\, T] \tag{238}$$

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 [38].

Theorem 7. *We consider a sequence of RIBSDEs with data and solutions indexed by n, but for a* common *sequence τ of stopping times, with lower barriers L_n given as quasimartingales in $\mathcal{S}^2$, and with predictable finite variation components denoted by A_n (cf. (238)). The data are assumed to be bounded in the sense that the driver coefficients $g^n = g_t^n(y, z, v)$ are uniformly Λ – Lipschitz continuous in (y, z, v), and one has for some constant c_1:*

$$\|\xi^n\|_2^2 + \|g_\cdot^n(0,0,0)\|_{\mathcal{H}^2}^2 + \|L^n\|_{\mathcal{S}^2}^2 + \|U^n\|_{\mathcal{S}^2}^2 + \|A^{n,-}\|_{\mathcal{S}^2}^2 \leq c_1. \tag{239}$$

Then one has for some constant $c(\Lambda)$:

$$\|Y^n\|_{\mathcal{S}^2}^2 + \|Z^n\|_{\mathcal{H}_d^2}^2 + \|V^n\|_{\mathcal{H}_\mu^2}^2 + \|K^{n,+}\|_{\mathcal{S}^2}^2 + \|K^{n,-}\|_{\mathcal{S}^2}^2 \leq c(\Lambda)c_1. \tag{240}$$

Indexing by ^{n,p} *the differences* $\cdot^n - \cdot^p$, *one also has:*

$$\|Y^{n,p}\|^2_{\mathcal{S}^2} + \|Z^{n,p}\|^2_{\mathcal{H}^2_d} + \|V^{n,p}\|^2_{\mathcal{H}^2_\mu} + \|K^{n,p}\|^2_{\mathcal{S}^2} \leq c(\Lambda)c_1\Big(\|\xi^{n,p}\|_2^2 + \|g_\cdot^{n,p}(Y_\cdot^n, Z_\cdot^n, V_\cdot^n)\|^2_{\mathcal{H}^2} + \|L^{n,p}\|_{\mathcal{S}^2} + \|U^{n,p}\|_{\mathcal{S}^2}\Big). \tag{241}$$

Assume further $dA^{n,-} \leq \alpha_t^n dt$ *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 (239) and (241).*

Suppose additionally that $\|\alpha^n\|_{\mathcal{H}^2}$ *is bounded over* $\mathbb{N}$ *and that when* $n \to \infty$:

- $g_\cdot^n(Y_\cdot, Z_\cdot, V_\cdot)$ *converges in* $\mathcal{H}^2$ *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) *converges in* $\mathcal{L}^2 \times \mathcal{H}^2 \times \mathcal{S}^2$ *to* (ξ, L, U).

Then (Y^n, Z^n, V^n, K^n) *converges in* $\mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu \times \mathcal{S}^2$ *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 (240)–(241) "with* $n = \infty$*" therein.*

Moreover, in the special case $L^{n,p} = U^{n,p} = 0$, one has like for R2BSDEs that estimate (241) holds with $L^{n,p} = U^{n,p} = 0$ therein (cf. Appendix A of [38]), irrespectively of the specific assumptions on the L_ns in Theorem 7. In particular,

Proposition 36. *Uniqueness holds for an RIBSDE satisfying the standing assumptions (H.0)–(H.2).*

16.2.3 Comparison

In this section we specialize the general assumption (H.1) in Sect. 5.1 to the case where (cf. Sect. 4 of [38])

$$g_t(y, z, v) = \widetilde{g}_t\Big(y, z, \int_E v(e)\eta_t(e)\zeta_t(e)\,\rho(de)\Big), \tag{242}$$

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)' $|\widetilde{g}_t(y, z, r) - \widetilde{g}_t(y', z', r')| \leq \Lambda\big(|y - y'| + |z - z'| + |r - r'|\big)$, 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

$$\Big|\int_E (v(e) - v'(e))\eta_t(e)\zeta_t(e)\,\rho(de)\Big| \leq |v - v'|_t|\eta_t|$$

with $|\eta_t|_t$ uniformly bounded, so g defined by (242) satisfies (H.1).

The following RIBSDE comparison result is then an easy generalization of the R2BSDE comparison result of Crépey and Matoussi [38].

Theorem 8. *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.2). We assume further that g satisfies (H.1)'. Then $Y \leq Y'$, $d\mathbb{P} \otimes dt$ – almost everywhere, whenever:*
(i) $\xi \leq \xi'$, $\mathbb{P}$ – almost surely,
(ii) $g_\cdot(Y'_\cdot, Z'_\cdot, V'_\cdot) \leq g'_\cdot(Y'_\cdot, Z'_\cdot, V'_\cdot)$, $d\mathbb{P} \otimes dt$ – almost everywhere,
(iii) $L \leq L'$ and $\bar{U} \leq \bar{U}'$, $d\mathbb{P} \otimes dt$ – almost everywhere, where $\bar{U}$ is defined by (236) and $\bar{U}'$ is the analogous process relative to τ'.

Remark 31. The inequality $\bar{U} \leq \bar{U}'$ which is assumed in part (iii) implies in particular that

$$(\tau_{2l}, \tau_{2l+1}) \subseteq (\tau'_{2l}, \tau'_{2l+1}), \ l \geq 0.$$

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

$$M_t = M_0 + \int_0^t Z_s \, dB_s + \int_0^t \int_E V_s(e) \widetilde{\mu}(ds, de), \quad t \in [0, T] \tag{243}$$

for some $Z \in \mathcal{H}_d^2$ and $V \in \mathcal{H}_\mu^2$.

We also strengthen Assumption (H.2.i) into:
(**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 ${}^pX = X_-$ (see, e.g., Jacod–Shiryaev [62, Propositions I.2.26, p. 22 and I.2.35, p. 25]). We thus work in this section under assumptions (H) and (H.0)–(H.2)', where (H.2)' denotes (H.2) with (H.2.i) replaced by (H.2.i)' therein.

Finally we postulate the so-called *Mokobodski condition* (see [38]), namely the existence of a quasimartingale X with Rao components in $\mathcal{S}^2$ and such that $L \leq X \leq U$ over $[0, T]$. In view of Lemma 14, This is tantamount to the existence of non-negative supermartingales X^1, X^2 belonging to $\mathcal{S}^2$ and such that $L \leq X^1 - X^2 \leq U$ over $[0, T]$. 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 7.

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 is that of an RDBSDE (or RIBSDE with $\tau_2 = T$, see Remark 30).

Lemma 15. *Assuming (H), (H.0)–(H.2)' and the Mokobodski condition, then, in the special case where $\tau_2 = T$ almost surely, the RIBSDE with data (g, ξ, L, U, τ) has a (unique) solution (Y, Z, V, K). Moreover the reflecting process K is continuous.*

Proof. Under the present assumptions, existence of a solution to an RDBSDE was established in Crépey and Matoussi [38] (in which continuity of the reflecting 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 28).

Let $[\![\theta]\!]$ denotes the graph of a stopping time θ.

Lemma 16. *Assuming (H), (H.0)–(H.2)' and the Mokobodski condition, then, in the special case where $0 = \tau_1 \leq \tau_2 \leq \tau_3 = T$ 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*

$$\{(\omega, t)\,;\, \Delta K_t^- \neq 0\} \subseteq [\![\tau_2]\!]\,,\ \Delta Y_{\tau_2} = \Delta K_{\tau_2}^- = (Y_{\tau_2} - U_{\tau_2})^+\,.$$

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 an RDBSDE in [38], 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 9(i) below. In particular, in case $Y_{\tau_2} U_{\tau_2}$, the jump $\Delta K_{\tau_2}^-$ of the reflecting process K^- at time τ_2 is set to the effect that

$$Y_{\tau_2 -} = U_{\tau_2} = U_{\tau_2 -} = \bar{U}_{\tau_2 -},$$

so that the upper obstacle related conditions are satisfied in Definition 16(iii). 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 τ_2 either. □

Iterated and alternate applications of Lemmas 15 and 16 yield the following existence result for an RIBSDE,

Theorem 9. *Let us be given an RIBSDE with data (g, ξ, L, U, τ). We assume (H), (H.0)–(H.2)' and 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*

$$\begin{cases} g\ ,\ Y^{l+1}_{\tau_{l+1}}\ ,\ L\ ,\ \tau_{l+1} & (l \text{ even}) \\ g\ ,\ \min(Y^{l+1}_{\tau_{l+1}}, U_{\tau_{l+1}})\ ,\ L\ ,\ U\ ,\ \tau_{l+1} & (l \text{ odd}) \end{cases} \tag{244}$$

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* $l = 0, \ldots, m$ *:*

- $(Y, Z, V) = (Y^l, Z^l, V^l)$ *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})*,*

$$\Delta K_{\tau_l} = (Y^l_{\tau_l} - U_{\tau_l})^+ = \Delta Y_{\tau_l}\ (= 0 \textit{ for } l \textit{ odd})$$

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*

$$\{(\omega, t)\,;\, \Delta K^- \neq 0\} \subseteq \bigcup_{\{l \textit{ even}\}} [\![\tau_l]\!]\ ,\ \Delta Y = \Delta K^- = (Y - U)^+ \textit{ on } \bigcup_{\{l \textit{ even}\}} [\![\tau_l]\!]\ .$$

Remark 32. We conjecture that one does not need the condition that $\tau_{m+1} = T$ for some fixed index m in Theorem 9. 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 [81]. 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 [81], which are based on Peng [80], 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 Sect. 5.1). Moreover it is apparent that the penalization approach and the related results of Peng [80] and Peng and Xu [81] 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 9 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 Sect. 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

$$N^t_{T_l} = \theta_l(\mathcal{X}^t_{T_l-}), \tag{245}$$

for a suitable jump function θ. Again (see Remark 17), 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 valid.

Let us set, for a regular function u over $[0,T] \times \mathbb{R}^d$ (cf. (113) and the related comments):

$$\mathcal{G}u(t,x) = \partial_t u(t,x) + \tfrac{1}{2}\mathrm{Tr}[a(t,x)\mathcal{H}u(t,x)] + \nabla u(t,x)\widetilde{b}(t,x) + \int_{\mathbb{R}^d} \Big(u(t,x+\delta(t,x,y)) - u(t,x)\Big) f(t,x,y) m(dy) \tag{246}$$

with

$$\widetilde{b}(t,x) = b(t,x) - \int_{\mathbb{R}^d} \delta(t,x,y) f(t,x,y) m(dy). \tag{247}$$

In the present set-up, the operator $\mathcal{G}$ defined by (246) 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 (246), and the specification of the jump size function θ of N in (245),
- Cost data $\mathcal{C}$ as of Sect. 6.4, assumed here not to depend on $i \in I$,
- The parameterized sequence of stopping times τ^t defined by $\tau_0^t = t$ and, for every $l \geq 0$ (to be compared with the stopping time τ of Example 3/Hypothesis 7(iii) in Part III):

$$\tau_{2l+1}^t = \inf\{s > \tau_{2l}^t \,;\, N_s^t \notin \Delta\} \wedge T \;,\; \tau_{2l+2}^t = \inf\{s > \tau_{2l+1}^t \,;\, N_s^t \in \Delta\} \wedge T \;, \tag{248}$$

for a given subset Δ of I, resulting in an effective upper payoff process $\bar{U}$ of the Markovian form (233) corresponding to the event-process

$$\Omega_s^t = \mathbb{1}_{N_s^t \notin \Delta}. \tag{249}$$

Observe that since the cost data do not depend on i, the only impact of the marker process N^t is via its influence on τ^t. Also note that the τ_l^ts effectively reduce to $\mathfrak{T}$-valued stopping times, and that one almost surely has $\tau_{m+1}^t = T$.

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 2 in Part I.

Example 5. Given a constant *trigger level* $\bar{S}$ and an integer $\imath \leq m$, τ^t of the form (248) 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, capped at $\imath$, of consecutive monitoring dates T_l with $S_{T_l}^t \geq \bar{S}$ from time s backwards since the initial time t. Call is possible whenever $N_s^t = \imath$, 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,
(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

$$\theta_l^i(x) = (\mathbb{1}_{S \geq \bar{S}}, i_1, \ldots, i_{d-1})$$

(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 $S_{T_l}^t \geq \bar{S}$ 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 $v \in \mathcal{M}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m(dy); \mathbb{R})$:

$$\begin{aligned} &\widetilde{g}(s, \mathcal{X}_s^t, y, z, v) = g(s, X_s^t, y, z, \widetilde{r}_s^t) \text{ with } \widetilde{r}_s^t = \widetilde{r}_s^t(v) = \int_{\mathbb{R}^d} v(y) f(s, X_s^t, y) m(dy) \\ &\widehat{g}(s, \mathcal{X}_s^t, y, z, \widehat{v}) = g(s, X_s^t, y, z, \widetilde{r}_s^t) + (\widetilde{r}_s^t - \widehat{r}) \text{ with } \widehat{r} = \widehat{r}(v) = \int_{\mathbb{R}^d} v(y) m(dy). \end{aligned} \tag{250}$$

Accordingly, the V^t-component of a solution to any Markovian BSDE (cf. Theorem 37) lives in $\mathcal{H}_{\mu^t}^2 = \mathcal{H}_{\chi^t}^2$.

Proposition 37. **(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*

$$\begin{cases} \mathbb{1}_{\{s>t\}}\widetilde{g}(s, X_s^t, y, z, v),\ Y^{l+1,t}_{\tau_{l+1}^t},\ \ell(s \vee t, X^t_{s\vee t}),\ \tau^t_{l+1} & \textit{(l even)} \\ \mathbb{1}_{\{s>t\}}\widetilde{g}(s, X_s^t, y, z, v),\ \min(Y^{l+1,t}_{\tau_{l+1}^t}, h(\tau^t_{l+1}, X^t_{\tau^t_{l+1}})), & \\ \ell(s \vee t, X^t_{s\vee t}),\ \ h(s \vee t, X^t_{s\vee t}),\ \tau^t_{l+1} & \textit{(l odd)} \end{cases} \tag{251}$$

where, in case $l = m$, $Y^{l+1,t}_{\tau_{l+1}^t}$ is to be understood as $\Phi(X_T^t)$.

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 9(ii). So in particular $Y^t = Y^{l,t}$ on $[\tau_l^t, \tau_{l+1}^t)$, for every $l = 0, \ldots, m$, and

$$Y_t^t = \begin{cases} Y_t^{0,t},\ i \in \Delta \\ Y_t^{1,t},\ i \notin \Delta. \end{cases} \tag{252}$$

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

$$\widetilde{g}(s, X_s^t, y, z, v),\ \Phi(X_T^t),\ \ell(s, X_s^t),\ h(s, X_s^t),\ \tau^t. \tag{253}$$

(ii) *For every $l = 0, \ldots, m$, we extend $Y^{l,t}$ by $Y_t^{l,t}$, and $K_t^{l,t}$, $Z^{l,t}$ 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 (251), with $\widehat{g}$ instead of $\widetilde{g}$ therein.*

Proof. Part (i) follows by application of Proposition 9. Identity (252) simply results from the fact that, since $Y^t = Y^{l,t}$ on $[\tau_l^t, \tau_{l+1}^t)$,

$$\begin{cases} Y_t^t = Y_t^{0,t},\ N_t^t \in \Delta \\ Y_t^t = Y_t^{1,t},\ N_t^t \notin \Delta \end{cases} \tag{254}$$

with $N_t^t = i$. Part (ii) then follows from part (i) as in the proof of Theorem 2. □

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_l^ts continuous, which would be the "naive analog" of Assumption 4, would be too strong in regard to applications. This is for instance typically not satisfied in the situations of Example 5. One is thus led to introduce the following weaker

Assumption 10 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 $t \notin \mathfrak{T}$, and right-continuous at (t, x, i) if $t \in \mathfrak{T}$,
(ii) Left-limited at (t, x, i) if $t = T_l \in \mathfrak{T}$ 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 $t \notin \mathfrak{T}$, or, for $t = T_l \in \mathfrak{T}$, if $\mathcal{E}_{l+1} \ni (t_n, x_n, i) \to (T_l, x, i)$;
- If $\text{Int}\,\mathcal{E}_l \ni (t_n, x_n, i) \to (t = T_l, x, i)$ 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 $l' \geq m+1$.

Observe that since the τ^t_ls are in fact $\mathfrak{T}$-valued stopping times:

- The continuity assumption on τ^t effectively means that $\tau^{t_n}_l = \tau^t_l$ for n large enough, almost surely, for every $l = 1, \ldots, m+1$ and $\mathcal{E} \ni (t_n, x_n, i) \to (t, x, i) \in \mathcal{E}$ with $t \notin \mathfrak{T}$;
- The right-continuity, resp. left-limit assumption, effectively means that for n large enough $\tau^{t_n}_l = \tau^t_l$, resp. $\widetilde{\tau}^t_l$, almost surely, for every $l = 1, \ldots, m+1$ and $\mathcal{E}_{l'+1} \ni$, resp. $\mathrm{Int}\,\mathcal{E}_{l'} \ni (t_n, x_n, i) \to (T_{l'}, x, i) \in \mathcal{E}$.

Remark 33. It is intuitively clear, though we shall not try to prove this in this article, that Assumption 10 is satisfied in the situations of Example 5, in case the jump-diffusion X is uniformly elliptic in the direction of its first component S (cf. Example 3). We refer the reader to [31] 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 11 h is Lipschitz in (t, x).

One denotes by $\widetilde{\mathcal{Y}}^t = (\widetilde{\mathcal{Y}}^{l,t})_{0 \leq l \leq m}$, 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 $l = 0, \ldots, m$, 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 37(i).

Theorem 10. *For every $l = m, \ldots, 0$:*
(i) *One has the following estimate on $\mathcal{Y}^{l,t}$,*

$$\|Y^{l,t}\|^2_{\mathcal{S}^2} + \|Z^{l,t}\|^2_{\mathcal{H}^2_d} + \|V^{l,t}\|^2_{\mathcal{H}^2_\mu} + \|K^{l,t}\|^2_{\mathcal{S}^2} \leq C(1 + |x|^{2q}). \tag{255}$$

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 $t \notin \mathfrak{T}$, $\mathcal{Y}^{l,t_n}$ converges in $\mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu \times A^2$ to $\mathcal{Y}^{l,t}$ as $\mathcal{E} \ni (t_n, x_n, i) \to (t, x, i)$;*
- *In case $t = T_{l'} \in \mathfrak{T}$,*
  - *$\mathcal{Y}^{l,t_n}$ converges in $\mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu \times A^2$ 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 $\mathcal{Y}^{l,t_n}$ converges in $\mathcal{S}^2 \times \mathcal{H}^2_d \times \mathcal{H}^2_\mu \times A^2$ to $\widetilde{\mathcal{Y}}^{l,t}$ as $\mathrm{Int}\,\mathcal{E}_{l'} \ni (t_n, x_n, i) \to (t, x, i)$.*

Proof. Under Assumption 10, these results can be established, recursively on l decreasing from m to 0, by easy amendments to the proof of Proposition 27 in Appendix 17.3, using Assumption 11 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 (276). □

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 Sect. 16.3.1 and Theorem 37, respectively. Given $t'\geq t$, let $\widetilde{\mathbb{F}}^{t'}$ stand for $(\widetilde{\mathcal{F}}^{t'}_r)_{r\geq t'}$ with for $r\geq t'$

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

Let $\tau'=t'\vee\tau^t$, in the sense that $\tau'_l=t'\vee\tau^t_l$, for $l=1,\ldots,m+1$. As for $\mathbb{F}^{t'}=(\mathcal{F}^{t'}_r)_{r\geq t'}$, $\mathbb{P}^{t'}$, $B^{t'}$ and $\mu^{t'}$, they are defined as usual as in Sects. 7.1–7.2, with t' instead of t therein. Note in particular that $\widetilde{\mathbb{F}}^{t'}$ is embedded into 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 26, 28 in Part II.

Proposition 38. **(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:*

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

(ii) *For t and t' in the same monitoring time strip, so $T_{l-1}\leq t<t'<T_l$ for some $l\in\{1,\ldots,m\}$, then $\tau'=t'\vee\tau^t$ is an $\widetilde{\mathbb{F}}^{t'}$ – stopping time, and the RIBSDE on $[t',T]$ with data*

$$\widetilde{g}(s,X^{t'}_s,y,z,\widetilde{v}),\ \Phi(X^{t'}_T),\ \ell(s,X^{t'}_s),\ h(s,X^{t'}_s),\tau' \tag{256}$$

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 $\mathcal{Y}^t=(Y^t_r,Z^t_r,V^t_r,K^t_r)_{t\leq r\leq T}$ defined as in Theorem 37:

$$\mathcal{Y}^{t'}=(Y^{t'}_r,Z^{t'}_r,V^{t'}_r,K^{t'}_r)_{t'\leq r\leq T}=(Y^t_r,Z^t_r,V^t_r,K^t_r-K^t_{t'})_{t'\leq r\leq T}. \tag{257}$$

Proof. Part (i) can be shown much like Proposition 26(i). It implies in particular that whenever $T_{l-1}\leq t<t'<T_l$ for some $l\in\{1,\ldots,m\}$, then $N^t_r=N^{t'}_r=i$ for $r\in[t',T_l)$. In view of (248) one thus has $\tau'_0=t'$ and, for every $l\geq 0$:

$$\tau'_{2l+1}=\inf\{s>\tau'_{2l};\ N^{t'}_s\notin\Delta\}\wedge T,\ \tau'_{2l+2}=\inf\{s>\tau'_{2l+1};\ N^{t'}_s\in\Delta\}\wedge T. \tag{258}$$

This shows that τ' is an $\widetilde{\mathbb{F}}^{t'}$ – stopping time, namely the analog of τ^t relatively to $N^{t'}$. Knowing this, part (ii) can then be established much like Proposition 26(ii) or 28(ii) in Part II. □

In the present set-up the suitable notion of a Cauchy cascade (cf. Definition 15) takes the following form.

Definition 17. **(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 $\nu = (v_l)_{1\le l\le m}$ 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:

$$v_l^i(T_l,x) = \begin{cases} \min(v_{l+1}(T_l,x,\theta_l^i(x)), h(T_l,x)) & \text{if } i \notin \Delta \text{ and } \theta_l^i(x) \in \Delta, \\ v_{l+1}(T_l,x,\theta_l^i(x)) & \text{else} \end{cases} \tag{259}$$

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 by the concatenation of the v_ls on the Int $\mathcal{E}_l$s, and by the terminal condition Φ at T.

Remark 34. 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 4 or 5 / 2, 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 35.

Theorem 11. **(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 37. 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.,*

$$Y_s^t = v(s, \mathcal{X}_s^t),\ s \in [t,T]. \tag{260}$$

(iii) *For every $\varepsilon > 0$, an ε – saddle-point of the related Dynkin game at time t is given by,*

$$\rho_t^\varepsilon = \inf\left\{ s \in \cup_{l\ge 0}[\tau_{2l+1}^t, \tau_{2l+2}^t);\ (s, \mathcal{X}_s^t) \in \mathcal{E}_\epsilon^- \right\} \wedge T,$$
$$\theta_t^\varepsilon = \inf\left\{ s \in [t,T];\ (s, \mathcal{X}_s^t) \in \mathcal{E}_\epsilon^+ \right\} \wedge T$$

with

$$\mathcal{E}_\epsilon^- = \{(t,x,i) \in \mathcal{E};\, v^i(t,x) \geq h^i(t,x) - \varepsilon\},$$
$$\mathcal{E}_\epsilon^+ = \{(t,x,i) \in \mathcal{E}\,;\, v^i(t,x) \geq \ell^i(t,x) + \varepsilon\}.$$

Proof. Let us prove parts (i) and (ii), which immediately imply (iii) by an application of Proposition 35. By taking $r = t'$ in the semi-group property (257) of $\mathcal{Y}$, one gets, for every $l = 1, \ldots, m$ and $T_{l-1} \leq t \leq r < T_l$,

$$Y_r^t = v_l(r, \mathcal{X}_r^t),\ \mathbb{P}^t\text{–}a.s. \tag{261}$$

for a deterministic function v_l on $\mathrm{Int}\,\mathcal{E}_l$. In particular,

$$Y_t^t = v^i(t,x), \text{for any } (t,x,i) \in \mathcal{E}, \tag{262}$$

where v is the function defined on $\mathcal{E}$ by the concatenation of the v_ls and of the terminal condition Φ. In view of (252), the fact that v is of class $\mathcal{P}$ then directly follows from the bound estimates (255) on $\mathcal{Y}^{0,t}$ and $\mathcal{Y}^{1,t}$.

Let us show that the v_ls are continuous on the $\mathrm{Int}\,\mathcal{E}_l$s. Given $\mathcal{E} \ni (t_n, x_n, i) \to (t,x,i)$ with $t \notin \mathfrak{T}$ or $t_n \geq T_l = t$, one decomposes by (252):

$$|u^i(t,x) - u^i(t_n,x_n)| = |Y_t^t - Y_{t_n}^{t_n}| \leq$$
$$\begin{cases} |\mathbb{E}(Y_t^{0,t} - Y_{t_n}^{0,t})| + \mathbb{E}|Y_{t_n}^{0,t} - Y_{t_n}^{0,t_n}|,\ i \in \Delta \\ |\mathbb{E}(Y_t^{1,t} - Y_{t_n}^{1,t})| + \mathbb{E}|Y_{t_n}^{1,t} - Y_{t_n}^{1,t_n}|,\ i \notin \Delta. \end{cases}$$

In either case we conclude as in the proof of Theorem 4(i), using Proposition 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, and that the jump condition (259) is satisfied (except maybe at the boundary points (T_l, x, i) such that θ_l^i is discontinuous at x).

Given $\mathrm{Int}\,\mathcal{E}_l \ni (t_n, x_n, i) \to (t = T_l, x, i)$ 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 $v_l^i(T_l,x)$ here is defined by (259). We distinguish four cases.

• In case $i \notin \Delta$ and $\theta_l^i(x) \in \Delta$, one has, denoting $\widetilde{v}^j(s,y) = \min(v(s,y,\theta_l^j(y)), h(y)), \widehat{v}^j(s,y) = \min(v(s,y,j), h(y))$,

$$|\widetilde{u}^i(t,x) - u^i(t_n,x_n)|^2 = |\widetilde{u}^i(t,x) - Y_{t_n}^{1,t_n}|^2$$
$$\leq 2\mathbb{E}|\widetilde{v}^i(t,x) - \widehat{v}(t, \mathcal{X}_t^{t_n})|^2 + 2|\mathbb{E}\big(\widehat{v}(t, \mathcal{X}_t^{t_n}) - Y_{t_n}^{1,t_n}\big)|^2. \tag{263}$$

By continuity of θ_l at (x,i), it comes that $\theta_l(\mathcal{X}_t^{t_n}) = \theta_l^i(x) \in \Delta$ for $\mathcal{X}_t^{t_n}$ close enough to x, say $||\mathcal{X}_t^{t_n} - x|| \le c$. In this case $t = \tau_2^{t_n}$, therefore (cf. (251)) $Y_t^{1,t_n} = \widehat{v}(t, \mathcal{X}_t^{t_n})$. So

$$\mathbb{E}|\mathbb{1}_{||\mathcal{X}_t^{t_n} - x|| \le c}\big(\widehat{v}(t, \mathcal{X}_t^{t_n}) - Y_{t_n}^{1,t_n}\big)|^2 \le \mathbb{E}|Y_t^{1,t_n} - Y_{t_n}^{1,t_n}|^2,$$

which converges to zero as $n \to \infty$, by the R2BSDE satisfied by Y^{1,t_n} and the convergence of $\mathcal{Y}^{1,t_n}$ to $\widetilde{\mathcal{Y}}^{1,t}$. Moreover $\mathbb{E}|\mathbb{1}_{||\mathcal{X}_t^{t_n} - x|| > c}\big(\widehat{v}(t, \mathcal{X}_t^{t_n}) - Y_{t_n}^{1,t_n}\big)|^2$ goes to zero as $n \to \infty$ by the a priori estimates on X and Y^{1,t_n} and the continuity of $\widehat{v}$ already established over $\mathrm{Int}\,\mathcal{E}_{l+1}$. Finally by this continuity and the a priori estimates on X the first term in (263) also goes to zero as $n \to \infty$. So, as $n \to \infty$,

$$v^i(t_n, x_n) \to \widetilde{v}^i(t,x) = \min(v(t,x,\theta_l^i(x)), h(t,x)) = v_l^i(T_l, x).$$

- In case $i \in \Delta$ and $\theta_l^i(x) \notin \Delta$, one can show likewise, using $\breve{v}^j(s,y) := v(s,y, \theta_l^j(y))$ instead of $\widetilde{v}^j(s,y)$, $v(t, \mathcal{X}_t^{t_n})$ instead of $\widehat{v}(t, \mathcal{X}_t^{t_n})$ and Y^0 instead of Y^1 above, that

$$v^i(t_n, x_n) \to \breve{v}^i(t,x) = v_l^i(T_l, x) \tag{264}$$

as $n \to \infty$.
- If $i, \theta_l^i(x) \notin \Delta$, it comes,

$$\begin{aligned}
|\breve{u}^i(t,x) - u^i(t_n, x_n)|^2 &= |\breve{u}^i(t,x) - Y_{t_n}^{1,t_n}|^2 \\
&\le 2\mathbb{E}|\breve{v}^i(t,x) - v(t, \mathcal{X}_t^{t_n})|^2 + 2|\mathbb{E}\big(v(t, \mathcal{X}_t^{t_n}) - Y_{t_n}^{1,t_n}\big)|^2 \\
&\le 2\mathbb{E}|\breve{v}^i(t,x) - v(t, \mathcal{X}_t^{t_n})|^2 + 2|\mathbb{E}\big(Y_{T_l}^{1,t_n} - Y_{t_n}^{1,t_n}\big)|^2,
\end{aligned}$$

which goes to zero as $\to \infty$ by an analysis similar to (actually simpler than) that of the first bullet point. Hence (264) follows.
- If $i, \theta_l^i(x) \in \Delta$, (264) 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 $\nu = (v_l)_{1 \le l \le m}$, 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 $x \in \mathbb{R}^d$ with θ_l^i continuous at x,

$$v_l^i(T_l, x) = \begin{cases} \min(v_{l+1}(T_l, x, \theta_l^i(x)), h(T_l, x)), & i \notin \Delta \text{ and } \theta_l^i(x) \in \Delta \\ v_{l+1}(T_l, x, \theta_l^i(x)), & \text{else} \end{cases} \tag{265}$$

with v_{l+1} in the sense of Φ in case $l = m$;

- On the time interval $[T_{l-1}, T_l)$, for every $i \in I$,

$$\begin{cases} \min\Big(-\mathcal{G}v_l^i - g^{v_l^i}, v_l^i - \ell \Big) = 0, \ i \in \Delta \\ \max\Big(\min\Big(-\mathcal{G}v_l^i - g^{v_l^i}, v_l^i - \ell \Big), v_l^i - h \Big) = 0, \ i \notin \Delta \end{cases} \tag{266}$$

where $\mathcal{G}$ is given by (246) and where we set, for any function $\varphi = \varphi(t,x)$,

$$g^{\varphi} = g^{\varphi}(t,x) = g(t,x,\varphi(t,x),(\nabla\varphi\sigma)(t,x),\mathcal{I}\varphi(t,x)). \tag{267}$$

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 11. Using the notions of viscosity solutions introduced in Definition 14, one then has in virtue of arguments already used in Part III (cf. also Proposition 34(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 (265)–(266), 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 5, 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 10, 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 (265) and (266) 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 (265) and (266).

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 (249) reduces to,

$$\Omega_s^t = \mathbb{1}_{N_s^t \notin \Delta} = \mathbb{1}_{s \geq \tau_1^t}. \tag{268}$$

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 4, which corresponds to the "stopped" version of Example 5/2.

From a mathematical point of view one is back to an RDBSDE as of Definition 9(ii) (cf. (107), (16)). But this is for a stopping time, τ_1^t, which falls outside the

scope of Example 3/Assumption 6 in Part III, so that the PDE results of Part III cannot be applied directly. However, assuming (268), one can check by inspection in the arguments of Sects. 16.3.2–16.3.4, that:

- For $i \notin \Delta$, the $\mathcal{Y}^{l,t}$s do not depend on i, and $\mathcal{Y}^t$ in Theorem 37 coincides with $\mathcal{Y}^t$ in Theorem 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 two),
- Theorem 10 is true independently of Assumption 11 (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 11 is in fact not required in this section.

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 Sects. 16.3.2–16.3.4,

Proposition 39. **(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*

$$\max\Big(\min\Big(-\mathcal{G}u - g^u, u - \ell\Big), u - h\Big) = 0 \tag{269}$$

on $\mathcal{E}$ with boundary condition Φ at T, where $\mathcal{G}$ is given by (246) and where g^u is defined by (267).
(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 of Theorem 11, which for $i \notin \Delta$ does not depend on i (v^i is constant in i outside Δ, assuming (268)).

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 (265) and (266) 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 (260) in Theorem 11, $(v_l^i)_{1\le l\le m}^{i\in\Delta}$ 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 $x \in \mathbb{R}^d$ with θ_l^i continuous at x,

$$v_l^i(T_l, x) = \begin{cases} u(T_l, x), & l = m \text{ or } \theta_l^i(x) \notin \Delta \\ v_{l+1}(T_l, x, \theta_l^i(x)), & \text{else,} \end{cases} \tag{270}$$

- On the time interval $[T_{l-1}, T_l)$, for every $i \in \Delta$,

$$\min\Big(-\mathcal{G} v_l^i - g^{v_l^i}, v_l^i - \ell \Big) = 0. \tag{271}$$

Given a pertaining notion of discontinuous viscosity solution of (270) and (271), $(v_l^i)_{1\le l\le m}^{i\in\Delta}$ could then be characterized as the unique solution in this sense to (270) and (271).

Remark 35. The Cauchy–Dirichlet cascade (269)–(271) involves less equations than the Cauchy cascade (265) and (266). However "less" here is still often far too much (see for instance Example 4(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.

17 Proofs of Auxiliary BSDE Results

17.1 Proof of Lemma 8

Recall that a càdlàg process Z^t is a $\mathbb{P}^t$ – local martingale if and only if $\Gamma^t Z^t$ is a $\mathbb{P}$ – local martingale (see, e.g., Proposition III.3.8 in Jacod–Shiryaev [62]). Now for

$$Z^t = B^t\,,\ \text{resp.}\ \int_t^{\cdot}\int_{\mathbb{R}^d} V_s^t(y)\,\widetilde{\chi}^t(ds,dy)\,,\ \text{resp. resp.}\ \sum_{j\in I}\int_t^{\cdot} W_s^t(j)\,d\widetilde{\nu}_s^t(j)$$

with V^t, W^t in the related spaces of predictable integrands, we have, "$\triangleq$" standing for "equality up to an $(\mathbb{F}^t, \mathbb{P})$ – local martingale term":

$$d(\Gamma^t Z^t)_s \triangleq \Gamma_{s-}^t dZ_s^t + \Delta\Gamma_s^t \Delta Z_s^t$$

where

$$\Delta Z_s^t = 0\,,\ \text{resp.}\ \textstyle\int_{\mathbb{R}^d} V_s^t(y)\,\chi(ds,dy)\,,\ \text{resp. resp.}\ \sum_{j\in I} W_s^t(j)\,d\nu_s^t(j).$$

• In case $Z^t = B^t$, $\Gamma^t Z^t$ is obviously a $\mathbb{P}$ – local martingale. Process 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 $Z^t = \int_t^{\cdot}\int_{\mathbb{R}^d} V_s^t(y)\,\widetilde{\chi}^t(ds,dy)$, since χ and ν cannot jump together (see Remark 21), one has by (148):

$$\Delta\Gamma_s^t \Delta Z_s^t = \Delta Z_s^t \Gamma_{s-}^t \int_{\mathbb{R}^d}\left(\frac{f(s,\mathcal{X}_{s-}^t,y)}{\widehat{f}(s,\mathcal{X}_{s-}^t,y)}-1\right)\chi(ds,dy).$$

So

$$\begin{aligned}
d(\Gamma^t Z^t)_s &\triangleq \Gamma_{s-}^t \int_{\mathbb{R}^d} V_s^t(y)\,\widetilde{\chi}^t(ds,dy)\\
&\quad + \Gamma_{s-}^t \int_{\mathbb{R}^d} V_s^t(y)\left(\frac{f(s,\mathcal{X}_{s-}^t,y)}{\widehat{f}(s,\mathcal{X}_{s-}^t,y)}-1\right)\chi\,(ds,dy)\\
&= -\Gamma_{s-}^t \int_{\mathbb{R}^d} V_s^t(y)\,f(s,\mathcal{X}_s^t,y)m(dy)ds\\
&\quad + \Gamma_{s-}^t \int_{\mathbb{R}^d} V_s^t(y)\,\frac{f(s,\mathcal{X}_{s-}^t,y)}{\widehat{f}(s,\mathcal{X}_{s-}^t,y)}\,\chi\,(ds,dy)\\
&= \Gamma_{s-}^t \int_{\mathbb{R}^d} V_s^t(y)\frac{f(s,\mathcal{X}_{s-}^t,y)}{\widehat{f}(s,\mathcal{X}_{s-}^t,y)}\,\widetilde{\chi}\,(ds,dy)
\end{aligned}$$

and $\Gamma^t Z^t$ is also a $\mathbb{P}$ – local martingale.
• In case $Z^t = \sum_{j\in I}\int_t^{\cdot} W_s^t(j)\,d\widetilde{\nu}_s^t(j)$ one gets likewise

$$d(\Gamma^t Z^t)_s \triangleq \Gamma_{s-}^t \sum_{j\in I} W_s^t(j)\,\frac{n^j(s,\mathcal{X}_{s-}^t)}{\widehat{n}^j(N_{s-}^t)}d\widetilde{\nu}_s(j)$$

and $\Gamma^t Z^t$ is again a $\mathbb{P}$ – local martingale.

17.2 Proof of Proposition 25

First we have, using the facts that f (cf. Lemma 7(i)) and $\widehat{n}$ are bounded, with f positively bounded for (H.1.ii)":
(H.1.i)" $\mathbb{1}_{\{\cdot>t\}}\widehat{g}(\cdot,\mathcal{X}_{\cdot}^t,y,z,\widehat{v})$ is a progressively measurable process with

$$\|\mathbb{1}_{\{\cdot>t\}}\widehat{g}(\cdot,\mathcal{X}_{\cdot}^t,y,z,\widehat{v})\|_{\mathcal{H}^2} < \infty \text{ for any } y\in\mathbb{R},\ z\in\mathbb{R}^{1\otimes d},\ \widehat{v}\in\mathcal{M}_\pi;$$

(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 $s\in[0,T]$, $y,y'\in\mathbb{R}$, $z,z'\in\mathbb{R}^{1\otimes d}$, $\widehat{v},\widehat{v}'\in\mathcal{M}_\pi$:

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

(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.2), relatively to $(\Omega,\mathbb{F},\mathbb{P}),(B,\mu)$.
(i) By the general results of [38], one thus has the following bound estimate on $\widetilde{\mathcal{Y}}^t$:

$$\|Y^t\|^2_{\mathcal{S}^2}+\|Z^t\|^2_{\mathcal{H}^2_d}+\|\widetilde{\mathcal{V}}^t\|^2_{\mathcal{H}^2_\mu}+\|K^{t,+}\|^2_{\mathcal{S}^2}+\|K^{t,-}\|^2_{\mathcal{S}^2}\leq c(\Lambda)c_1$$

with

$$\begin{aligned}c_1:=&\,\|\Phi(\mathcal{X}^t_T)\|^2_2+\|\mathbb{1}_{\{\cdot>t\}}\widehat{g}(\cdot,\mathcal{X}^t_\cdot,0,0,0)\|^2_{\mathcal{H}^2}\\&+\|\ell(\cdot\vee t,\mathcal{X}^t_{\cdot\vee t})\|^2_{\mathcal{S}^2}+\|h(\cdot\vee t,\mathcal{X}^t_{\cdot\vee t})\|^2_{\mathcal{S}^2}+\|\int_{\cdot\wedge t}^{\cdot}\mathcal{G}\varphi(r,\mathcal{X}^t_r)dr\|^2_{\mathcal{S}^2},\end{aligned}$$

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 [38], we also have the following error estimate in which c_1 is as above:

$$\begin{aligned}&\|Y^t-Y^{t_n}\|^2_{\mathcal{S}^2}+\|Z^t-Z^{t_n}\|^2_{\mathcal{H}^2_d}+\|\widetilde{V}^t-\widetilde{V}^{t_n}\|^2_{\mathcal{H}^2_\mu}+\|K^t-K^{t_n}\|^2_{\mathcal{S}^2}\\&\leq c(\Lambda)c_1\Big(\|\Phi(\mathcal{X}^t_T)-\Phi(\mathcal{X}^{t_n}_T)\|^2_2+\|\mathbb{1}_{\{\cdot>t\}}\widehat{g}(\cdot,\mathcal{X}^t_\cdot,Y^t_\cdot,Z^t_\cdot,\widetilde{\mathcal{V}}^t_\cdot)-\mathbb{1}_{\{\cdot>t_n\}}\\&\times\widehat{g}(\cdot,\mathcal{X}^{t_n}_\cdot,Y^t_\cdot,Z^t_\cdot,\widetilde{\mathcal{V}}^t_\cdot)\|^2_{\mathcal{H}^2}+\|\ell(\cdot\vee t,\mathcal{X}^t_{\cdot\vee t})-\ell(\cdot\vee t_n,\mathcal{X}^{t_n}_{\cdot\vee t_n})\|_{\mathcal{S}^2}\\&+\|h(\cdot\vee t,\mathcal{X}^t_{\cdot\vee t})-h(\cdot\vee t_n,\mathcal{X}^{t_n}_{\cdot\vee t_n})\|_{\mathcal{S}^2}\Big).\end{aligned}\tag{272}$$

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 (272). 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

$$\begin{aligned}\Omega^{m,n}_s:=\{s\geq t\vee t_n\}\cap\{N^t_s=N^{t_n}_s\}\cap\{|X^t_s|\vee|X^{t_n}_s|\\\vee|Y^t_s|\vee|Z^t_s|\vee r^t_s\leq R_m\},\end{aligned}$$

with $r_s^t := |\widehat{r}_s^t| \vee |\widetilde{r}_s^t| \vee |\widetilde{r}_s^{t_n}|$, where

$$\widehat{r}_s^t = \int_{\mathbb{R}^d} V_s^t(y)m(dy),\ \widetilde{r}_s^t = \int_{\mathbb{R}^d} V_s^t(y)f(s,X_s^t,N_s^t,y)m(dy),$$
$$\widetilde{r}_s^{t_n} = \int_{\mathbb{R}^d} V_s^t(y)f(s,X_s^{t_n},N_s^t,y)m(dy), \tag{273}$$

and let $\bar{\Omega}_s^{m,n}$ denote the complement of the set $\Omega_s^{m,n}$. One has for any m,n:

$$\begin{aligned}
&\|\mathbb{1}_{\{\cdot>t\}}\widehat{g}(\cdot,\mathcal{X}_\cdot^t,Y_\cdot^t,Z_\cdot^t,\widetilde{\mathcal{V}}_\cdot^t) - \mathbb{1}_{\{\cdot>t_n\}}\widehat{g}(\cdot,\mathcal{X}_\cdot^{t_n},Y_\cdot^t,Z_\cdot^t,\widetilde{\mathcal{V}}_\cdot^t)\|_{\mathcal{H}^2}^2\\
&= \mathbb{E}\int_{t\wedge t_n}^T \Big[\mathbb{1}_{\{s>t\}}\widehat{g}(s,\mathcal{X}_s^t,Y_s^t,Z_s^t,\widetilde{\mathcal{V}}_s^t) - \mathbb{1}_{\{s>t_n\}}\widehat{g}(s,\mathcal{X}_s^{t_n},Y_s^t,Z_s^t,\widetilde{\mathcal{V}}_s^t)\Big]^2 ds\\
&= \mathbb{E}\int_{t\wedge t_n}^T \Big[\mathbb{1}_{\{s>t\}}\widehat{g}(s,\mathcal{X}_s^t,Y_s^t,Z_s^t,\widetilde{\mathcal{V}}_s^t)\\
&\quad -\mathbb{1}_{\{s>t_n\}}\widehat{g}(s,\mathcal{X}_s^{t_n},Y_s^t,Z_s^t,\widetilde{\mathcal{V}}_s^t)\Big]^2 \mathbb{1}_{\bar{\Omega}_s^{m,n}} ds\\
&\quad + \mathbb{E}\int_{t\wedge t_n}^T \Big[\mathbb{1}_{\{s>t\}}\widehat{g}(s,\mathcal{X}_s^t,Y_s^t,Z_s^t,\widetilde{\mathcal{V}}_s^t)\\
&\quad - \mathbb{1}_{\{s>t_n\}}\widehat{g}(s,\mathcal{X}_s^{t_n},Y_s^t,Z_s^t,\widetilde{\mathcal{V}}_s^t)\Big]^2 \mathbb{1}_{\Omega_s^{m,n}} ds\\
&\le 2\mathbb{E}\int_{t\wedge t_n}^T \Big[\widehat{g}(s,\mathcal{X}_s^t,Y_s^t,Z_s^t,\widetilde{\mathcal{V}}_s^t)^2 + \widehat{g}(s,\mathcal{X}_s^{t_n},Y_s^t,Z_s^t,\widetilde{\mathcal{V}}_s^t)^2\Big] \mathbb{1}_{\bar{\Omega}_s^{m,n}} ds\\
&\quad + \mathbb{E}\int_0^T \Big[\widehat{g}(s,\mathcal{X}_s^t,Y_s^t,Z_s^t,\widetilde{\mathcal{V}}_s^t) - \widehat{g}(s,\mathcal{X}_s^{t_n},Y_s^t,Z_s^t,\widetilde{\mathcal{V}}_s^t)\Big]^2\\
&\quad \times \mathbb{1}_{\Omega_s^{m,n}} ds =: I_{m,n} + II_{m,n}.
\end{aligned}$$

Now,

$$\begin{aligned}
&\widehat{g}(s,\mathcal{X}_s^t,Y_s^t,Z_s^t,\widetilde{V}_s^t)^2 + \widehat{g}(s,\mathcal{X}_s^{t_n},Y_s^t,Z_s^t,\widetilde{V}_s^t)^2\\
&\quad \le C\Big(1 + |X_s^t|^{2q} + |X_s^{t_n}|^{2q} + |Y_s^t|^2 + |Z_s^t|^2 + |\widetilde{\mathcal{V}}_s^t|^2\Big).
\end{aligned} \tag{274}$$

Note that $|X_s^{t_n}|^{2q}$ is equi – $d\mathbb{P}\otimes dt$-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 (274), since $\widetilde{\mathcal{Y}}^t \in \mathcal{S}^2 \times \mathcal{H}_d^2 \times \mathcal{H}_\mu^2 \times \mathcal{A}^2$. Besides, one has that

$$\mathbb{E}\int_{t\wedge t_n}^T \mathbb{1}_{\bar{\Omega}_s^{m,n}} ds \le T|t-t_n| + \mathbb{E}\int_{t\vee t_n}^T \mathbb{1}_{\bar{\Omega}_s^{m,n}} ds, \tag{275}$$

where for $s \ge t \vee t_n$:

$$\bar{\Omega}_s^{m,n} \subseteq \{N_s^t \ne N_s^{t_n}\} \cup \{|X_s^t| \vee |X_s^{t_n}| \vee |Y_s^t| \vee |Z_s^t| \vee |r_s^t| \ge R_m\}.$$

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:

$$\begin{aligned}
&\mathbb{E}\int_{t\vee t_n}^T \mathbb{1}_{\bar{\Omega}_s^{m,n}} ds\\
&\quad \le C|t-t_n| + \mathbb{E}\int_{t\vee t_n}^T \big(\mathbb{1}_{\{|X_s^t|\ge R_m\}} + \mathbb{1}_{\{|X_s^{t_n}|\ge R_m\}} + \mathbb{1}_{\{|Y_s^t|\ge R_m\}}\\
&\qquad + \mathbb{1}_{\{|Z_s^t|\ge R_m\}} + \mathbb{1}_{\{|r_s^t|\ge R_m\}}\big) ds \le C\big(|t-t_n| + \tfrac{1}{R_m^2}\big).
\end{aligned}$$

Therefore, given (275), $\mathbb{E}\int_{t\wedge t_n}^T \mathbb{1}_{\bar{\Omega}_s^{m,n}} ds$ 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 – $d\mathbb{P} \otimes dt$-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:

$$\begin{aligned} II_{m,n} &= \mathbb{E}\int_0^T \left[\widehat{g}(s, X_s^t, N_s^t, Y_s^t, Z_s^t, \widetilde{\mathcal{V}}_s^t) - \widehat{g}(s, X_s^{t_n}, N_s^t, Y_s^t, Z_s^t, \widetilde{\mathcal{V}}_s^t)\right]^2 \mathbb{1}_{\Omega_s^{m,n}} ds \\ &\le \mathbb{E}\int_0^T \eta_m \left(|X_s^t - X_s^{t_n}| + |\widetilde{r}_s^t - \widetilde{r}_s^{t_n}|\right) ds \end{aligned}$$

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 $I_{m_\epsilon,\, n} \le \frac{\varepsilon}{2}$ for $n \ge n_\epsilon$. Let further μ_ϵ be such $\eta_{m_\epsilon}(\rho) \le \varepsilon$ for $\rho \le \mu_\epsilon$. C_ϵ denoting an upper bound on η_{m_ϵ}, it comes, for every n:

$$\begin{aligned} II_{m_\epsilon,n} &\le \mathbb{E}\int_0^T \eta_{m_\epsilon} \left(|X_s^t - X_s^{t_n}| + |\widetilde{r}_s^t - \widetilde{r}_s^{t_n}|\right) ds \\ &\le \mathbb{E}\int_0^T \left(\varepsilon + C_\epsilon \mathbb{1}_{\{|X_s^t - X_s^{t_n}| \ge \mu_\epsilon\}} + C_\epsilon \mathbb{1}_{\{|\widetilde{r}_s^t - \widetilde{r}_s^{t_n}| \ge \mu_\epsilon\}}\right) ds \\ &\le T\left(\varepsilon + C_\epsilon \mathbb{P}[\sup_{[0,T]} |X_\cdot^t - X_\cdot^{t_n}| \ge \mu_\epsilon]\right) + C_\epsilon \mathbb{E}\int_0^T \mathbb{1}_{\{|\widetilde{r}_s^t - \widetilde{r}_s^{t_n}| ds \ge \mu_\epsilon\}}. \end{aligned}$$

Now, given estimate (137), one has that $\mathbb{P}[\sup_{[0,T]} |X_\cdot^t - X_\cdot^{t_n}| \ge \mu_\epsilon] \to 0$ as $n \to \infty$, by Markov's inequality. Moreover (cf. (273))

$$|\widetilde{r}_s^t - \widetilde{r}_s^{t_n}| \le \int_{\mathbb{R}^d} |V_s^t(y)||f(s, X_s^t, N_s^t, y) - f(s, X_s^{t_n}, N_s^t, y)| m(dy),$$

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(i). Thus by Markov's inequality:

$$\mathbb{E}\int_0^T \mathbb{1}_{\{|\widetilde{r}_s^t - \widetilde{r}_s^{t_n}| ds \ge \mu_\epsilon\}} \le \frac{\|\widetilde{r}^t - \widetilde{r}^{t_n}\|_{\mathcal{H}^2}^2}{\mu_\epsilon^2}$$

converges to 0 as $n \to \infty$.

In conclusion $I_{m_\epsilon,n} + II_{m_\epsilon,n} \le \varepsilon$ for $n \ge n_\epsilon \vee n'_\epsilon$, for any $\varepsilon > 0$, which proves that

$$\|\mathbb{1}_{\{\cdot > t\}} \widehat{g}(\cdot, \mathcal{X}_\cdot^t, Y_\cdot^t, Z_\cdot^t, \widetilde{\mathcal{V}}_\cdot^t) - \mathbb{1}_{\{\cdot > t_n\}} \widehat{g}(\cdot, \mathcal{X}_\cdot^{t_n}, Y_\cdot^t, Z_\cdot^t, \widetilde{\mathcal{V}}_\cdot^t)\|_{\mathcal{H}^2}^2 \to 0 \text{ as } n \to \infty.$$

17.3 Proof of Proposition 27

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 25 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), and the assumptions regarding L in (H.2) relatively to $(\Omega,\mathbb{F},\mathbb{P})$, (B,μ).

(i) By the general results of [38], 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 [38]:

$$\begin{aligned}
&\|\bar{Y}^t-\bar{Y}^{t_n}\|^2_{\mathcal{S}^2}+\|\bar{Z}^t-\bar{Z}^{t_n}\|^2_{\mathcal{H}^2_d}+\|\overline{\mathcal{V}}^t-\overline{\mathcal{V}}^{t_n}\|^2_{\mathcal{H}^2_\mu}\\
&+\|\bar{K}^t-\bar{K}^{t_n}\|^2_{\mathcal{S}^2}\le c(\Lambda)c_1\times\Big(\|Y^t_{\tau^t}-Y^{t_n}_{\tau^{t_n}})\|_2^2\\
&+\|\mathbb{1}_{\{t<\cdot<\tau^t\}}\widehat{g}(\cdot,\mathcal{X}^t_\cdot,\bar{Y}^t_\cdot,\bar{Z}^t_\cdot,\overline{\mathcal{V}}^t_\cdot)-\mathbb{1}_{\{t_n<\cdot<\tau^{t_n}\}}\widehat{g}(\cdot,\mathcal{X}^{t_n}_\cdot,\bar{Y}^t_\cdot,\bar{Z}^t_\cdot,\overline{\mathcal{V}}^t_\cdot)\|^2_{\mathcal{H}^2}\\
&+\|\ell(t\vee\cdot\wedge\tau^t,\mathcal{X}^t_{t\vee\cdot\wedge\tau^t})-\ell(t_n\vee\cdot\wedge\tau^{t_n},\mathcal{X}^{t_n}_{t_n\vee\cdot\wedge\tau^{t_n}})\|_{\mathcal{H}^2}\Big)
\end{aligned}\tag{276}$$

(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. [38]). 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 (276). 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 thus $\Omega^{m,n}_s$ and $\bar{\Omega}^{m,n}_s$ be defined as in the proof of Proposition 25(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:

$$\|\mathbb{1}_{\{t<s\cdot<\tau^t\}}\widehat{g}(\cdot,\mathcal{X}^t_\cdot,\bar{Y}^t_\cdot,\bar{Z}^t_\cdot,\overline{\mathcal{V}}^t_\cdot)-\mathbb{1}_{\{t_n<\cdot<\tau^{t_n}\}}\widehat{g}(\cdot,\mathcal{X}^{t_n}_\cdot,\bar{Y}^t_\cdot,\bar{Z}^t_\cdot,\overline{\mathcal{V}}^t_\cdot)\|^2_{\mathcal{H}^2}$$

$$=\mathbb{E}\int_0^T\left[\mathbb{1}_{\{t<s<\tau^t\}}\widehat{g}(s,\mathcal{X}^t_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)-\mathbb{1}_{\{t_n<s<\tau^{t_n}\}}\widehat{g}(s,\mathcal{X}^{t_n}_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)\right]^2 ds$$

$$\leq 2\mathbb{E}\int_0^T\left[\widehat{g}(s,\mathcal{X}^t_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)^2+\widehat{g}(s,\mathcal{X}^{t_n}_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)^2\right]\mathbb{1}_{\bar{\Omega}^{m,n}_s}ds$$

$$+\mathbb{E}\int_0^T\left[\mathbb{1}_{\{t<s<\tau^t\}}\widehat{g}(s,\mathcal{X}^t_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)\right.$$

$$\left.-\mathbb{1}_{\{t_n<s<\tau^{t_n}\}}\widehat{g}(s,\mathcal{X}^{t_n}_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)\right]^2\mathbb{1}_{\Omega^{m,n}_s}ds$$

$$=: I_{m,n}+II_{m,n}.$$

As in the proof Proposition 25(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

$$II_{m,n}=\mathbb{E}\int_0^T\left[\mathbb{1}_{\{t<s<\tau^t\}}\widehat{g}(s,X^t_s,N^t_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)\right.$$

$$\left.-\mathbb{1}_{\{t_n<s<\tau^{t_n}\}}\widehat{g}(s,X^{t_n}_s,N^t_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)\right]^2\mathbb{1}_{\Omega^{m,n}_s}ds$$

$$\leq 2\mathbb{E}\int_0^T\left[\widehat{g}(s,X^t_s,N^t_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)-\widehat{g}(s,X^{t_n}_s,N^t_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)\right]^2$$

$$\mathbb{1}_{\Omega^{m,n}_s}ds+2\mathbb{E}\int_0^T|\mathbb{1}_{\{t<s<\tau^t\}}-\mathbb{1}_{\{t_n<s<\tau^{t_n}\}}|\widehat{g}(s,X^t_s,N^t_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)^2ds,$$

where in the last inequality:

- $\mathbb{E}\int_0^T\left[\widehat{g}(s,X^t_s,N^t_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)-\widehat{g}(s,X^{t_n}_s,N^t_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)\right]^2\mathbb{1}_{\Omega^{m,n}_s}ds\leq\mathbb{E}\int_0^T\eta_m\left(|X^t_s-X^{t_n}_s|\right)ds$ for a non-negative bounded function η_m continuous and null at 0 (cf. the proof of Proposition 25(ii));
- $\mathbb{E}\int_0^T|\mathbb{1}_{\{t<s<\tau^t\}}-\mathbb{1}_{\{t_n<s<\tau^{t_n}\}}|(\widehat{g}(s,X^t_s,N^t_s,\bar{Y}^t_s,\bar{Z}^t_s,\overline{\mathcal{V}}^t_s)^2ds$ goes to 0 as $n,m\to\infty$, by $d\mathbb{P}\otimes dt$-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 4).

We conclude the proof as for Proposition 25(ii).

18 Proofs of Auxiliary PDE Results

18.1 Proof of Lemma 11

(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 (210) 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

$$\varphi^i_{\varepsilon,\alpha}(t,x,s,y) = \mu^i(t,x) - \nu^i(s,y) - \frac{|x-y|^2}{\varepsilon^2} - \frac{|t-s|^2}{\alpha^2} - \varphi^i(t,x) \quad (277)$$

on $[0,T] \times \mathbb{R}^d$, in which $\varepsilon,\ \alpha$ 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. [37] or Fleming and Soner [49]), 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,

$$\begin{aligned} &\mu^i(t^\star,x^\star) - \nu^i(t^\star,x^\star) - \varphi^i(t^\star,x^\star) = \varphi^i_{\epsilon,\alpha}(t^\star,x^\star,t^\star,x^\star) \\ &\le \varphi^i_{\varepsilon,\alpha}(t,x,s,y) = \mu^i(t,x) - \nu^i(s,y) - \frac{|x-y|^2}{\varepsilon^2} \\ &\quad - \frac{|t-s|^2}{\alpha^2} - \varphi^i(t,x); \end{aligned} \quad (278)$$

- $(t,x), (s,y) \to (t^\star, x^\star)$ 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 [37, Theorem 8.3] that there exists symmetric matrices $X, Y \in \mathbb{R}^{d \otimes d}$ such that

$$\begin{gathered} (a + \partial_t \varphi(t,x), p + \nabla \varphi^i(t,x), X) \in \bar{\mathcal{P}}^{2,+} \mu^i(t,x) \\ (a,p,Y) \in \bar{\mathcal{P}}^{2,-} \nu^i(s,y) \\ \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \le \tfrac{4}{\varepsilon^2} \begin{pmatrix} \mathrm{Id}_d & -\mathrm{Id}_d \\ -\mathrm{Id}_d & \mathrm{Id}_d \end{pmatrix} + \begin{pmatrix} \mathcal{H}\varphi(t,x) & 0 \\ 0 & 0 \end{pmatrix} \end{gathered} \quad (279)$$

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 [37,49]), and

$$a = \frac{2(t-s)}{\alpha^2},\ p = \frac{2(x-y)^{\mathsf{T}}}{\varepsilon^2}. \tag{280}$$

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 (278) 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

$$\begin{aligned}
&-a - \partial_t\varphi^i(t,x) - \frac{1}{2}\mathrm{Tr}(a^i(t,x)X) - pb^i(t,x) - \nabla\varphi^i(t,x)\Big(b^i(t,x)\\
&\quad - \int_{\mathbb{R}^d}\delta^i(t,x,z)f^i(t,x,z)m(dz)\Big)\\
&\quad - \int_{\mathbb{R}^d}\Big(\mu^i(t,x+\delta^i(t,x,z)) - \mu^i(t,x) - p\delta^i(t,x,z)\Big)f^i(t,x,z)m(dz)\\
&\quad - g^i(t,x,\mu(t,x),(p+\nabla\varphi^i(t,x))\sigma^i(t,x),\mathcal{I}\mu^i(t,x)) \leq 0\\
&\quad - a - \frac{1}{2}\mathrm{Tr}(a^i(s,y)) - pb^i(s,y)\\
&\quad - \int_{\mathbb{R}^d}\Big(\nu^i(s,y+\delta^i(s,y,z)) - \nu^i(s,y) - p\delta^i(s,y,z)\Big)f^i(s,y,z)m(dz)\\
&\quad - g^i(s,y,\nu(s,y),p\sigma^i(s,y),\mathcal{I}\nu^i(s,y)) \geq 0
\end{aligned}$$

Note 18. **(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 [37]) as done here. In the case of unbounded Levy measures however, Barles and Imbert [6] (see also Jakobsen and Karlsen [63]) recently established that this Lemma (and thus the related uniqueness proofs in Barles et al. [9], and then in turn in Harraj et al. [58]) has to be amended in a rather involved way.

By subtracting the previous inequalities, there comes:

$$
\begin{aligned}
&-\partial_t\varphi^i(t,x)-\frac{1}{2}\Big(\mathrm{Tr}(a^i(t,x)X)-\mathrm{Tr}(a^i(s,y))\Big)\\
&\quad-p\Big(b^i(t,x)-b^i(s,y))\Big)-\nabla\varphi^i(t,x)\Big(b^i(t,x)\\
&\quad-\int_{\mathbb{R}^d}\delta^i(t,x,z)f^i(t,x,z)m(dz)\Big)\\
&\quad-\int_{\mathbb{R}^d}\Big[(\mu^i(t,x+\delta^i(t,x,z))-\mu^i(t,x))-(\nu^i(s,y+\delta^i(s,y,z))-\nu^i(s,y))\\
&\quad-p(\delta^i(t,x,z)-\delta^i(s,y,z))\Big]f^i(t,x,z)m(dz)\\
&\quad+\int_{\mathbb{R}^d}\Big[\nu^i(s,y+\delta^i(s,y,z))-\nu^i(s,y)-p\delta^i(s,y,z))\Big]\\
&\qquad\times\Big[f^i(t,x,z)-f^i(s,y,z)\Big]m(dz)\\
&\quad-\Big(g^i(t,x,\mu(t,x),(p+\nabla\varphi^i(t,x))\sigma^i(t,x),\mathcal{I}\mu^i(t,x))\\
&\quad-g^i(s,y,\nu(s,y),p\sigma^i(s,y),\mathcal{I}\nu^i(s,y))\Big)\le 0
\end{aligned}
$$

Now, by straightforward computations analogous to those in [9, p. 76 and 77] (see also [79]) using the maximization property of (t,x,s,y), the definition of p (cf. (280)), the matrix inequality (279) and the Lipschitz continuity properties of the data (and accounting for the fact that we deal 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 [9, 79]), we have:

$$
\begin{aligned}
|p|(|t-s|+|x-y|)&\le C\frac{|t-s|^2+|x-y|^2}{\varepsilon^2}\\
\mathrm{Tr}(a^i(t,x)X)-\mathrm{Tr}(a^i(s,y)Y)&\le C\frac{|t-s|^2+|x-y|^2}{\varepsilon^2}+\mathrm{Tr}(a^i(t,x)\mathcal{H}\varphi^i(t,x))\\
|p\Big(b^i(t,x)-b^i(s,y))\Big)|&\le C\frac{|t-s|^2+|x-y|^2}{\varepsilon^2}(\mu^i(t,x+\delta^i(t,x,z))\\
&\quad-\mu^i(t,x))-(\nu^i(s,y+\delta^i(s,y,z))-\nu^i(s,y))\\
&\le(\varphi^i(t,x+\delta^i(t,x,z))-\varphi^i(t,x))\\
&\quad+\Big(\frac{|x+\delta^i(t,x,z)-y-\delta^i(s,y,z)|^2}{\varepsilon^2}\\
&\qquad-\frac{|x-y|^2}{\varepsilon^2}\Big)
\end{aligned}
$$

where in the last inequality

$$\begin{aligned}
&\frac{|x+\delta^i(t,x,z)-y-\delta^i(s,y,z)|^2}{\varepsilon^2}-\frac{|x-y|^2}{\varepsilon^2}\\
&\quad=\frac{1}{\varepsilon^2}\left[2(x-y)^{\mathsf{T}}(\delta^i(t,x,z)-\delta^i(s,y,z))+|\delta^i(t,x,z)-\delta^i(s,y,z)|^2\right]\\
&\quad=p(\delta^i(t,x,z)-\delta^i(s,y,z))+\frac{1}{\varepsilon^2}|\delta^i(t,x,z)-\delta^i(s,y,z)|^2\\
&\quad\le C\frac{|t-s|^2+|x-y|^2}{\varepsilon^2}
\end{aligned}$$

Therefore

$$\begin{aligned}
&-\partial_t\varphi^i(t,x)-\frac{1}{2}\mathrm{Tr}(a^i(t,x)\mathcal{H}\varphi^i(t,x))-\nabla\varphi^i(t,x)\Big(b^i(t,x)\\
&\quad-\int_{\mathbb{R}^d}\delta^i(t,x,z)f^i(t,x,z)m(dz)\Big)\\
&\quad-\int_{\mathbb{R}^d}\Big(\varphi^i(t,x+\delta^i(t,x,z))-\varphi^i(t,x)\Big)f^i(t,x,z)m(dz)\\
&\quad-\Big(g^i(t,x,\mu(t,x),(p+\nabla\varphi^i(t,x))\sigma^i(t,x),\mathcal{I}\mu^i(t,x))\\
&\quad-g^i(s,y,\nu(s,y),p\sigma^i(s,y),\mathcal{I}\nu^i(s,y))\Big)\\
&\qquad\le C\Big(|t-s|+|x-y|+\frac{|t-s|^2+|x-y|^2}{\varepsilon^2}\Big)\\
&\mathcal{I}\mu^i(t,x)-\mathcal{I}\nu^i(s,y))\\
&\quad=\int_{\mathbb{R}^d}\Big[\nu^i(s,y+\delta^i(s,y,z))-\nu^i(s,y)\Big]\Big[f^i(t,x,z)-f^i(s,y,z)\Big]m(dz)\\
&\qquad+\int_{\mathbb{R}^d}\Big[(\mu^i(t,x+\delta^i(t,x,z))-\mu^i(t,x))-(\nu^i(s,y+\delta^i(s,y,z))\\
&\qquad-\nu^i(s,y))\Big]f^i(t,x,z)m(dz)\\
&\quad\le\int_{\mathbb{R}^d}\Big[(\varphi^i(t,x+\delta^i(t,x,z))-\varphi^i(t,x))\Big]f^i(t,x,z)m(dz)\\
&\qquad+C\Big(|t-s|+|x-y|+\frac{|t-s|^2+|x-y|^2}{\varepsilon^2}\Big)\\
&\quad=\mathcal{I}\varphi^i(t,x)+C\Big(|t-s|+|x-y|+\frac{|t-s|^2+|x-y|^2}{\varepsilon^2}\Big)\\
&g^i(t,x,\mu(t,x),(p+\nabla\varphi^i(t,x))\sigma^i(t,x),\mathcal{I}\mu^i(t,x))\\
&\qquad-g^i(s,y,\nu(s,y),p\sigma^i(s,y),\mathcal{I}\nu^i(s,y))\\
&\quad\le\eta_\epsilon(|t-s|)+\eta_R(|x-y|(1+|p\sigma^i(s,y)|))+\Lambda_1\max_{j\in I}(\mu^j(t,x)-\nu^j(s,y))^+\\
&\qquad+\Lambda|p(\sigma^i(t,x)-\sigma^i(s,y))+(\nabla\varphi\sigma)^i(t,x)|+\Lambda(\mathcal{I}\mu^i(t,x)-\mathcal{I}\nu^i(s,y))^+
\end{aligned}$$

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 (280) is bounded independently of α, for given ε;
• η_R is the modulus of continuity standing in Assumption 8(ii);
• The three last terms come from the Lipschitz continuity and/or monotonicity properties of g with respect to its three last variables. Therefore

$$\begin{aligned}
-\widetilde{\mathcal{G}}\varphi^i(t,x) &= -\partial_t\varphi^i(t,x) - \frac{1}{2}\mathrm{Tr}(a^i(t,x)\mathcal{H}\varphi^i(t,x)) - \nabla\varphi^i(t,x)\Big(b^i(t,x) \\
&\quad - \int_{\mathbb{R}^d}\delta^i(t,x,z)f^i(t,x,z)m(dz)\Big) \\
&\quad - \int_{\mathbb{R}^d}\Big(\varphi^i(t,x+\delta^i(t,x,z)) - \varphi^i(t,x)\Big)f^i(t,x,z)m(dz) \\
&\le \Lambda_1\Big(\max_{j\in I}(\mu^j(t,x)-\nu^j(s,y))^+ + |(\nabla\varphi\sigma)^i(t,x)| + \mathcal{I}\varphi^i(t,x)^+\Big) \\
&\quad + \eta_\epsilon(|t-s|) + \eta_R(|x-y|(1+|p\sigma^i(s,y)|)) \\
&\quad + C\Big(|t-s| + |x-y| + \frac{|t-s|^2+|x-y|^2}{\varepsilon^2}\Big)
\end{aligned}$$

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:

$$\begin{aligned}
&-\widetilde{\mathcal{G}}\varphi^i(t^*,x^*) - \Lambda_1\Big(\max_{j\in I}(\mu^j(t,x)-\nu^j(s,y))^+ \\
&\qquad\qquad + |(\nabla\varphi\sigma)^i(t^*,x^*)| + \mathcal{I}\varphi^i(t^*,x^*)^+\Big) \\
&\le \rho + \eta_\epsilon(|t-s|).
\end{aligned}$$

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 $\alpha \le \alpha_\varepsilon (\le \varepsilon)$:

$$\begin{aligned}
&-\widetilde{\mathcal{G}}\varphi^i(t^*,x^*) - \Lambda_1\Big(\max_{j\in I}(\mu^j(t,x)-\nu^j(s,y))^+ \\
&\quad + |(\nabla\varphi\sigma)^i(t^*,x^*)| + \mathcal{I}\varphi^i(t^*,x^*)^+\Big) \le 2\rho
\end{aligned}$$

Sending $\rho, \varepsilon, \alpha$ to zero with $\varepsilon \le \varepsilon_\rho$ and $\alpha \le \alpha_\varepsilon$, inequality (210) 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 (210) holds at $(t^\star, x^\star, i)$ in case $t^\star > 0$.

Now in case $t^\star = 0$ let us introduce the function

$$\varphi^i_\epsilon(t,x) = \omega^i(t,x) - (\varphi^i(t,x) + \frac{\varepsilon}{t}) \tag{281}$$

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 (210) to the function $(s, y) \mapsto \varphi^i(s, y) + \frac{\varepsilon}{s}$ at (t, x, i), whence:

$$
\begin{aligned}
&- \widetilde{\mathcal{G}}\varphi^i(t,x) \; - \Lambda_1\Big(\max_{j\in I}(\omega^j(t,x))^+ + |(\nabla\varphi\sigma)^i(t,x)| + (\mathcal{I}\varphi^i(t,x))^+ \Big) \\
&\quad \le -\frac{\varepsilon}{t^2} \le 0.
\end{aligned}
$$

Sending ε to 0 in the left hand side we conclude by upper semi-continuity of $\max_{j\in I}(\omega^j)^+$ that (210) holds at $(t^\star = 0, x^\star, i)$.

(ii) Straightforward computations give:

$$
\begin{gathered}
-\partial_t\chi(t,x) = C_1\chi(t,x) \\
(1+|x|)|\nabla\chi(t,x)| \;\vee\; (1+|x|^2)|\mathcal{H}\chi(t,x)| \;\vee\; \chi(t, x+\delta^i(t,x,z))| \le C|\chi(t,x)|
\end{gathered}
$$

on $\mathcal{E}$, for a constant C independent of C_1. Therefore for $C_1 > 0$ large enough

$$
-\widetilde{\mathcal{G}}\chi - \Lambda_1\big(\chi + |\nabla\chi\sigma| + (\mathcal{I}\chi)^+\big) > 0
$$

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

$$
\sup_{(t,x,i)\in\mathcal{E}} \Big((\omega^i(t,x))^+ - \alpha\chi(t,x) \Big) e^{-\Lambda_1(T-t)} \le 0. \tag{282}
$$

Assume by contradiction that one has $>$ instead of $\le$ in (282). Then by upper semi-continuity of ω^+ the supremum is reached at some point $(t^*, x^*, i) \in \operatorname{Int}\mathcal{E}$ in the left hand side of (282), and

$$
(\omega^i(t^*,x^*))^+ \ge \omega^i(t^*,x^*))^+ - \alpha\chi(t^*,x^*) > 0. \tag{283}
$$

Therefore one has on $[0, T] \times \mathbb{R}^d$:

$$
\begin{aligned}
\Big(\omega^i(t,x) - \alpha\chi(t,x)\Big)e^{-\Lambda_1(T-t)} &\le \Big((\omega^i(t,x))^+ - \alpha\chi(t,x)\Big)e^{-\Lambda_1(T-t)} \\
&\le \Big((\omega^i(t^*,x^*))^+ - \alpha\chi(t^*,x^*)\Big)e^{-\Lambda_1(T-t^*)} \\
&= \Big(\omega^i(t^*,x^*) - \alpha\chi(t^*,x^*)\Big)e^{-\Lambda_1(T-t^*)}
\end{aligned}
$$

thus

$$
\omega^i(t,x) - \alpha\chi(t,x) \le \Big(\omega^i(t^*,x^*) - \alpha\chi(t^*,x^*)\Big)e^{-\Lambda_1(t-t^*)}.
$$

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 (283)):

$$\begin{aligned} &-\widetilde{\mathcal{G}}\varphi^i(t^*,x^*) \; - \Lambda_1\Big(\max_{j\in I}(\omega^j(t^*,x^*))^+ + |\nabla\varphi^i(t^*,x^*)\sigma^i(t^*,x^*)| \\ &+(\mathcal{I}\varphi^i(t^*,x^*))^+\Big) \le 0. \end{aligned} \tag{284}$$

But the left hand side in this inequality is nothing but

$$\begin{aligned} &-\alpha\widetilde{\mathcal{G}}\chi(t^*,x^*) + \Lambda_1\Big(\omega^i(t^*,x^*) - \alpha\chi(t^*,x^*)\Big) \\ &\quad -\Lambda_1\Big(\omega^i(t^*,x^*) + \alpha|\nabla\chi(t^*,x^*)\sigma^i(t^*,x^*)| + \alpha(\mathcal{I}\chi^i(t^*,x^*))^+\Big) \\ &= -\alpha\widetilde{\mathcal{G}}\chi(t^*,x^*) - \Lambda_1\Big(\alpha\chi(t^*,x^*) + \alpha|\nabla\chi(t^*,x^*)\sigma^i(t^*,x^*)| \\ &\quad +\alpha(\mathcal{I}\chi^i(t^*,x^*))^+\Big) \end{aligned}$$

which should be positive by (211) in (ii), in contradiction with (284).

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

References

1. Alvarez O., Tourin, A.: **13**(3), 293–317 (1996)
2. Amadori, A.L.: Nonlinear integro-differential evolution problems arising in option pricing: A viscosity solutions approach. J. Diff. Integr. Equat. **16**(7), 787–811 (2003)
3. Amadori, A.L.: The obstacle problem for nonlinear integro-differential operators arising in option pricing. Ricerche Matemat. **56**(1), 1–17 (2007)
4. Bally, V., Matoussi, A.: Weak solutions for SPDEs and Backward doubly stochastic differential equations. J. Theor. Probab. **14**(1), 125–164 (2001)
5. Bally, V., Caballero, E., Fernandez, B., El-Karoui, N.: Reflected BSDE's PDE's and variational inequalities. INRIA Technical Report (2002)
6. Barles, G., Imbert, C.: Second-order elliptic integro-differential equations: Viscosity solutions' theory revisited. Annales de l'IHP **25**(3), 567–585 (2008)
7. Barles, G., Lesigne, L.: SDE, BSDE and PDE. Backward stochastic differential equations. In: El Karoui, N., Mazliak, L. (eds.) Pitman Research Notes in Mathematics Series, vol. 364, pp. 47–80. Longman, Harlow (1997)
8. Barles, G., Souganidis, P.E.: Convergence of approximation schemes for fully nonlinear second order equations. Asymptotic Analysis **4**, 271–283 (1991)
9. Barles, G., Buckdahn, R., Pardoux, E.: Backward stochastic differential equations and integral-partial differential equations. Stoch. Stoch. Rep. **60**, 57–83 (1997)

10. Becherer, D., Schweizer, M.: Classical solutions to reaction-diffusion systems for hedging with interacting Itô and point processes. Ann. Appl. Probab. **15**, 1111–1144 (2005)
11. Bensoussan, A., Lions, J.L.: Applications of Variational Inequalities in Stochastic Control. North-Holland, Amsterdam (1982)
12. Bensoussan, A., Lions, J.L.: Impulse Control and Quasi-variational Inequalities. Gauthier-Villars, Paris (1984)
13. Bichteler, K.: Stochastic Integration with Jumps. Cambridge University Press, Cambridge (2002)
14. Bielecki, T.R., Rutkowski, M.: Credit Risk: Modeling, Valuation and Hedging. Springer, Berlin (2002)
15. Bielecki, T.R., Vidozzi, A., Vidozzi, L.: An efficient approach to valuation of credit basket products and ratings triggered step-up bonds. Working Paper (2006)
16. Bielecki, T.R., Crépey, S., Jeanblanc, M., Rutkowski, M.: Valuation of basket credit derivatives in the credit migrations environment. In: Birge, J., Linetsky, V. (eds.) Handbook of Financial Engineering. Elsevier, Amsterdam (2007)
17. Bielecki, T.R., Vidozzi, A., Vidozzi, L.: A Markov copulae approach to pricing and hedging of credit index derivatives and ratings triggered step-up bonds. J. Credit Risk **4**(1) (2008)
18. Bielecki, T.R., Crépey, S., Jeanblanc, M., Rutkowski, M.: Arbitrage pricing of defaultable game options with applications to convertible bonds. Quant. Finance **8**(8), 795–810 (2008) (Preprint version available online at `www.defaultrisk.com`)
19. Bielecki, T.R., Crépey, S., Jeanblanc, M., Rutkowski, M.: Convertible bonds in a defaultable diffusion model. Forthcoming In: Kohatsu-Higa, A., Privault, N., Sheu, S.J. (eds.) Stochastic Analysis with Financial Applications (36 pages). Birkhäuser Verlag (2010)
20. Bielecki, T.R., Crépey, S., Jeanblanc, M., Rutkowski, M.: Defaultable options in a Markovian intensity model of credit risk. Math. Finance **18**, 493–518 (2008) (Updated version available on `www.defaultrisk.com`)
21. Bielecki, T.R., Jakubowski, J., Vidozzi, A., Vidozzi, L.: Study of dependence for some stochastic processes. Stoch. Anal. Appl. **26**(4), 903–924 (2008)
22. Bielecki, T.R., Crépey, S., Jeanblanc, M.: Up and Down Credit Risk. Quantitative Finance, 1469–7696, April 2010
23. Bielecki, T.R., Crépey, S., Jeanblanc, M., Rutkowski, M.: Valuation and hedging of defaultable game options in a hazard process model. J. Appl. Math. Stoch. Anal. **2009**, 33 (2009); Article ID 695798 (Long Preprint version available on `www.defaultrisk.com`)
24. Boel, R., Varaiya, P., Wong, E.: Martingales on jump processes: I Representation results. SIAM JC **13**(5), 999–1021 (1975)
25. Boel, R., Varaiya, P., Wong, E.: Martingales on jump processes: II Applications. SIAM JC **13**(5), 1023–1061 (1975)
26. Bouchard, B., Elie R.: Discrete time approximation of decoupled Forward-Backward SDE with jumps. Stoch. Process. Appl. **118**, 53–75 (2008)
27. Brémaud, P.: Point Processes and Queues, Martingale Dynamics. Springer, New York (1981)
28. Briani, M., La Chioma, C., Natalini, R.: Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in financial theory. Numer. Math. **98**(4), 607–646 (2004)
29. Busca, J., Sirakov, B.: Harnack type estimates for nonlinear elliptic systems and applications. Annales de l'institut Henri Poincaré (C) Analyse non linéaire **21**(5), 543–590 (2004)
30. Carmona, R., Crépey, S.: Importance sampling and interacting particle systems for the estimation of markovian credit portfolios loss distribution. Int. J. Theor. Appl. Finance **13**(4), 577–602 (2010)
31. Chassagneux, J.-F., Crépey, S.: Doubly reflected BSDEs with call protection and their approximation. Working Paper (2009)
32. Cherny, A., Shiryaev, A.: Vector stochastic integrals and the fundamental theorems of asset pricing. Proceedings of the Steklov Institute of mathematics, vol. 237, pp. 6–49 (2002)
33. Coculescu, D., Nikeghbali, A.: Hazard processes and martingale hazard processes. Working Paper (2008)

34. Cont, R., Minca, A.: Recovering portfolio default intensities implied by CDO quotes. Working Paper (2008)
35. Cont, R., Tankov, P.: Financial Modelling with Jump Processes. Chapman & Hall, London (2003)
36. Cont, R., Voltchkova, K.: Finite difference methods for option pricing methods in exponential Levy models. SIAM J. Numer. Anal. **43**, 1596–1626 (2005)
37. Crandall, M., Ishii, H., Lions, P.-L.: User's guide to viscosity solutions of second order partial differential equations. Bull. Amer. Math. Soc. **27**, 1–67 (1992)
38. Crépey, S., Matoussi, A.: Reflected and doubly reflected BSDEs with jumps: A priori estimates and comparison principle. Ann. Appl. Probab. **18**(5), 2041–2069 (2008)
39. Crépey, S., Rahal, A.: Pricing convertible bonds with call protection. Forthcoming in J. Comput. Finance
40. Cvitanic, J., Karatzas, I.: Backward stochastic differential equations with reflection and Dynkin games. Ann. Probab. **24**(4), 2024–2056 (1996)
41. Darling, R., Pardoux, E.: Backward SDE with random terminal time and applications to semilinear elliptic PDE. Ann. Probab. **25**, 1135–1159 (1997)
42. Delbaen, F., Schachermayer, W.: The fundamental theorem of asset pricing for unbounded stochastic processes. Mathematische Annalen **312**, 215–250 (1997)
43. Dellacherie, C., Meyer, P.-A.: Probabilities and Potential. North-Holland, Amsterdam (1982)
44. Dynkin, E.B.: Markov Processes. Springer, Berlin (1965)
45. El Karoui, N., Quenez, M.-C.: Nonlinear pricing theory and backward stochastic differential equations. In: Runggaldier, W. (ed.) Financial Mathematics, Bressanone, 1996. Lecture Notes in Math., vol. 1656, pp. 191–246. Springer, Berlin (1997)
46. El Karoui, N., Peng, S., Quenez, M.-C.: Backward stochastic differential equations in finance. Math. Finance **7**, 1–71 (1997)
47. El Karoui, N., Kapoudjian, E., Pardoux, C., Peng, S., Quenez, M.-C.: Reflected solutions of backward SDEs, and related obstacle problems for PDEs. Ann. Probab. **25**, 702–737 (1997)
48. Ethier, H.J., Kurtz, T.G.: Markov Processes. Characterization and Convergence. Wiley, NY (1986)
49. Fleming, W., Soner, H.: Controlled Markov Processes and Viscosity Solutions, 2nd edn. Springer, Berlin (2006)
50. Föllmer, H., Sondermann, D.: Hedging of non-redundant contingent claims. In: Mas-Colell, A., Hildenbrand, W. (eds.) Contributions to Mathematical Economics, pp. 205–223. North-Holland, Amsterdam (1986)
51. Freidlin, M.: Markov Processes and Differential Equations: Asymptotic Problems. Birkhaüser, Basel (1996)
52. Frey, R., Backhaus, J.: Dynamic hedging of synthetic CDO-tranches with spread- and contagion risk. J. Econ. Dyn. Contr. **34**(4), 710–724 (2010)
53. Frey, R., Backhaus, J.: Pricing and hedging of portfolio credit derivatives with interacting default intensities. Int. J. Theor. Appl. Finance **11**(6), 611–634 (2008)
54. Fujiwara, T., Kunita, H.: Stochastic differential equations of jump type and Lévy processes in diffeomorphisms group. J. Math. Kyoto Univ. **25**, 71–106 (1985)
55. Hamadène, S.: Mixed zero-sum differential game and American game options. SIAM J. Contr. Optim. **45**, 496–518 (2006)
56. Hamadène, S., Hassani, M.: BSDEs with two reflecting barriers driven by a Brownian motion and an independent Poisson noise and related Dynkin game. Electr. J. Probab. **11**(5), 121–145 (2006)
57. Hamadène, S., Ouknine, Y.: Reflected backward stochastic differential equation with jumps and random obstacle. Electr. J. Probab. **8**, 1–20 (2003)
58. Harraj., N., Ouknine, Y., Turpin, I.: Double-barriers-reflected BSDEs with jumps and viscosity solutions of parabolic integrodifferential PDEs. J. Appl. Math. Stoch. Anal. **1**, 37–53 (2005)
59. Herbertsson, A.: Pricing portfolio credit derivatives. PhD Thesis, Göteborg University (2007)
60. Ishii, H., Koike, S.: Viscosity solutions for monotone systems of second-order elliptic PDEs. Comm. Part. Diff. Equat. **16**(6–7), 1095–1128 (1991)

61. Jacod, J.: Calcul Stochastique et Problèmes de Martingales, 2nd edn. Springer, Berlin (2003)
62. Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes. Springer, Berlin (2003)
63. Jakobsen, E.R., Karlsen, K.H.: A maximum principle for semi-continuous functions applicable to integro-partial differential equations. Nonlinear Differ. Equat. Appl. **13**, 137–165 (2006)
64. Jakobsen, E.R., Karlsen, K.H., La Chioma, C.: Error estimates for approximate solutions to Bellman equations associated with controlled jump-diffusions. Numer. Math. **110**(2), 221–255 (2008)
65. Jarrow, R.A., Turnbull, S.: Pricing derivatives on financial securities subject to credit risk. J. Finance **50**, 53–85 (1995)
66. Jeanblanc, M., Yor, M., Chesney, M.: Mathematical Methods for Financial Markets. Springer, Berlin (2009)
67. Kallsen, J., Kühn, C.: Convertible bonds: financial derivatives of game type. In: Kyprianou, A., Schoutens, W., Wilmott, P. (eds.) Exotic Option Pricing and Advanced Lévy Models, pp. 277–288. Wiley, NY (2005)
68. Karatzas, I.: On the pricing of American options. Appl. Math. Optim. **17**, 37–60 (1988)
69. Kifer, Y.: Game options. Finance Stoch. **4**, 443–463 (2000)
70. Kunita, H., Watanabe, T.: Notes on transformations of Markov processes connected with multiplicative functionals. Mem. Fac. Sci. Kyushu Univ. **A17**, 181–191 (1963)
71. Kushner, H., Dupuis, B.: Numerical Methods for Stochastic Control Problems in Continuous Time. Springer, Berlin (1992)
72. Kusuoka, S.: A remark on default risk models. Adv. Math. Econ. **1**, 69–82 (1999)
73. Lando, D.: Three Essays on Contingent Claims Pricing. Ph.D. Thesis, Cornell University (1994)
74. Laurent, J.-P., Cousin, A., Fermanian, J.-D.: Hedging default risks of CDOs in Markovian contagion models. Quantitative Finance, 1469–7696, April 2010
75. Lepeltier, J.-P., Maingueneau, M.: Le jeu de Dynkin en théorie générale sans l'hypothèse de Mokobodski. Stochastics **13**, 25–44 (1984)
76. Ma, J., Cvitanić, J.: Reflected Forward-Backward SDEs and obstacle problems with boundary conditions. J. Appl. Math. Stoch. Anal. **14**(2), 113–138 (2001)
77. Nikeghbali, A., Yor, M.: A definition and some characteristic properties of pseudo-stopping times. Ann. Probab. **33**, 1804–1824 (2005)
78. Palmowski, Z., Rolski, T.: A technique for exponential change of measure for Markov processes. Bernoulli **8**(6), 767–785 (2002)
79. Pardoux, E., Pradeilles, F., Rao, Z.: Probabilistic interpretation of systems of semilinear PDEs. Annales de l'Institut Henri Poincaré, série Probabilités-Statistiques **33**, 467–490 (1997)
80. Peng, S.: Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob-Meyer's type. Probab. Theor. Relat. Field. **113**(4), 473–499 (1999)
81. Peng, S., 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 **41**(3), 605–630 (2005)
82. Pham, H.: Optimal stopping of controlled jump-diffusion processes: A viscosity solution approach. J. Math. Syst. Estim. Contr. **8**, 1–27 (1998)
83. Protter, P.E.: A partial introduction to financial asset pricing theory. Stoch. Process. Appl. **91**(2), 169–203 (2001)
84. Protter, P.E.: Stochastic Integration and Differential Equations, 2nd edn. Springer, Berlin (2004)
85. Schweizer, M.: Option hedging for semimartingales. Stoch. Process. Appl. **37**, 339–363 (1991)
86. Windcliff, H.A., Forsyth, P.A., Vetzal, K.R.: Numerical methods for valuing Cliquet options. Appl. Math. Finance **13**, 353–386 (2006)
87. Windcliff, H.A., Forsyth, P.A., Vetzal, K.R.: Pricing methods and hedging strategies for volatility derivatives. J. Bank. Finance **30**, 409–431 (2006)

Mean Field Games and Applications

Olivier Guéant, Jean-Michel Lasry, and Pierre-Louis Lions

About this text:

This text is inspired from a "Cours Bachelier" held in January 2009 and taught by Jean-Michel Lasry. This course was based upon the articles of the three authors and upon unpublished materials they developed. Proofs were not presented during the conferences and are now available. So are some issues that were only rapidly tackled during class.

The content of this text is therefore far more important than the actual "Cours Bachelier" conferences, though the guiding principle is the same and consists in a progressive introduction of the concepts, methodologies and mathematical tools of mean field game theory.

Mean field game theory was created in 2006 by Jean-Michel Lasry and Pierre-Louis Lions and the first results and developments are given in the publications [34–36]: structures, concepts, definitions of equilibria, forward-backward Hamilton–Jacobi–Bellman/Kolmogorov equation systems, existence theorems in static and dynamic cases, links with Nash equilibria and dynamics in n-player games theory when n tends to infinity, variational principle for decentralization, etc. A number of developments were then implemented by Jean-Michel Lasry and Pierre-Louis Lions, several of them in collaboration with Olivier Guéant: notions

O. Guéant (✉)
UFR de Mathématiques, Université Paris-Diderot, 175, rue du Chevaleret, 75013 Paris, France
e-mail: olivier.gueant@gmail.com

J.-M. Lasry
Institut de Finance, Université Paris-Dauphine, Place du Maréchal de Lattre de Tassigny, Paris Cedex 16, 75775, France
e-mail: lasry@dauphine.fr

P.-L. Lions
Collège de France, rue d'Ulm 3, Paris CX 05, 75005, France
e-mail: pierre-louis.lions@college-de-france.fr

A.R. Carmona et al. (eds.), *Paris-Princeton Lectures on Mathematical Finance 2010*, Lecture Notes in Mathematics 2003, DOI 10.1007/978-3-642-14660-2_3,
© Springer-Verlag Berlin Heidelberg 2011

of stability of solutions, specific numerical methods, numerical eductive algorithms, and developments in 1/n for a better approximation to n-player games. These developments were presented in three successive courses at the Collège de France [38], in a Bachelier course, in various publications [23, 24] and in Olivier Guéant's PhD thesis [23]. Various applications, notably on the economics of scarce resources, were implemented or are ongoing (in collaboration: Pierre Noël Giraud, Olivier Guéant, Jean-Michel Lasry, Pierre-Louis Lions). Advances in population dynamics were made by Olivier Guéant [23]. Since 2008, several other authors have made further contributions, or are working on new applications and/or properties of MFG models [21, 33].

1 Introduction to Mean Field Games

Mean field game theory is a branch of game theory. It is therefore a set of concepts, mathematical tools, theorems, simulation methods and algorithms, which like all game theory, is intended to help specialists model situations where agents make decisions strategically. These specialists, as in other areas of game theory, will probably be economists (micro or macro), sociologists, engineers and even architects or urban planners. In any case, these applications emerge from the panorama created by the first "toy models" presented in this text.

We choose the term "toy models" to indicate the particular status of game theory and of its many "examples". Consider the famous "prisoner's dilemma". Nobody thinks of taking the story literally, nor that this example was created to be applied to the real-life situation it is supposed to evoke. In fact it is a fable intended to introduce an archetype of strategic interaction: an archetype that can thus be recognized in many negotiation situations in business life and elsewhere. Many of our examples have a similar status. "What time does the meeting start?" or the "Mexican wave equation" should not be taken literally, as a desire to scientifically model these situations in social life. Even if there is clearly an element of truth in our models for these two examples, we believe that the interest for the reader is primarily in the structure that is indicated through these "toy models". The Mexican wave equation, for example, shows how a sophisticated propagation phenomenon in a social setting can be constructed from non-cooperative individual behaviors in a rational-expectations context, once a certain taste for imitation is present in agents' utility function.

Introducing mean field games through these "toy models" is also a way of leading the reader to progressively discover the concepts and the mathematics of mean field game theory.

In this text we present a large number of results and mathematical proofs. Nevertheless, we cover only some parts of the large mathematical corpus built up since 2006. Thus for mathematicians this course can be seen as an introduction, or

a reading in parallel with the mean field games mathematical literature and with the three different courses held by Pierre-Louis Lions at the Collège de France (06–07, 07–08, 08–09), which present the whole mathematical corpus and which can be downloaded from the Collège de France website [38].

1.1 Three Different Avenues

There are three routes leading naturally to mean field game theory. Each route casts light on a particular aspect of mean field game theory, and the three complement each other.

1.1.1 First Route: From Physics to Mean Field Games

The first route takes its departure from particle physics.[1] In particle physics, mean field theory designates a highly effective methodology for handling a wide variety of situations in which there are too many particles to permit the dynamics or equilibrium to be described by modeling all the inter-particle interactions. The enormous number of all these interactions makes a detailed model ineffective – unreadable and unsuitable for both calculation and simulation, the model becomes unusable.

Nevertheless, in many situations of this kind, it is possible to construct an excellent approximation to the situation by introducing one or more "mean fields" that serve as mediators for describing inter-particle interactions. In this kind of model, one describes the contribution of each particle to the creation of a mean field and the effect of the mean field on each particle by conceiving each particle as infinitesimal, i.e. by carrying out a kind of limit process on the number n of particles ($n \to \infty$).

A large proportion of types of inter-particle interactions, though not all, lend themselves to this methodology: the inter-particle interactions must be sufficiently "weak" or "regular" in order for the statistical phenomena to emerge.

Mean field game theory provides an adaptation of this methodology to cases in which the particles are replaced by agents who mutually interact in socioeconomic and/or strategic situations. The main difference – indeed the challenge – is to take account not only of the ability of agents to make decisions, but also the capacity for strategic interaction, i.e. the capacity of each agent to construct strategies that involve thinking about his peers, who in turn and at the same time follow their own individual strategies.

[1] Several articles were written using the mean field notion of physicists and applying it to economic dynamics. One may see [15,16] or [27] as instances of such an approach. Our approach is different from the approach of the "econophysicists" since we are more influenced by control theory and hence more keen on mixing optimization and mean fields. As a consequence, the forward/backward structure of our approach is not present in most of the preceding works.

This new sophistication changes the nature of the mean field: it is no longer a statistic on the domain of particle states, but rather a statistic on the domain of agent states and hence in the domain of strategies and information.

This first route is certainly the one that sheds the most light on the operating characteristics of mean field game theory: since the methodology of mean fields works very well in particle physics and provides computable models in a priori complex situations, it offers a good prospect for transposition to the world of agents.

But this route is also the most demanding, and would probably be the most discouraging on any initial attempt. Mean field game theory has been able to emerge only because N-player game theory has long existed, thanks to the remarkable pioneering work carried out by Von Neumann and Morgenstern [47] sixty years ago and to the important developments made since then, notably by Nash [43–46] and then Aumann [9], and through the many applications that have been developed, particularly in most branches of economics.

What we want to provide is indeed a new branch of game theory for large games that relies on Nash equilibria and on the various concepts introduced in this field during the last 50 years. We are not applying, *mutatis mutandis*, the tools of statistical physics to economic problems. This is an important difference between mean field games and econophysics and we need to insist on it. Econophysicists only apply theories and methods originally rooted in physics to describe an economy and, although they often manage to have good models for the topics under scrutiny, these models are only descriptive. For instance, econophysicists manage to have good descriptive models for wealth distributions using only kinetic models and microscopic interactions (see [14]) but they never explain why people may want to exchange money as in their models (our last model can be seen as an attempt to model wealth distributions in a different way). Mean field game theory proposes to use the tools of physics but to use them within the classical economic paradigm, to explain (and not only to describe) phenomenon. Hence we will assign rationality to agents and not regard them as just gas particles, and not even as robots applying some predetermined behavioral strategy: strategic choices are endogenous in our models as they are in game theory.

1.1.2 Second Route: From Game Theory to Mean Field Games

This route is the most well-charted from a mathematical standpoint: it involves studying the limit of a large class of N-player games when N tends to infinity. Usually, differential games with N-players can be summed up by an HJB-Isaacs system of PDEs that turns out to be untractable. Fortunately things are simplified, at least for a wide range of games that are symmetrical in the players as their number increases, and for deep reasons. Indeed, interindividual complex strategies can no longer be implemented by the players, for each player is progressively lost in the crowd in the eyes of other players when the number of players increases.

More precisely, the class that proves to be best suited to this passage to the limit is that of games in which players of the same kind can be interchanged without altering the game: a form of anonymity of context where nothing is dependent on the individual. This hypothesis is particularly appropriate in the modeling of applications when there are a large number of players. From a mathematical standpoint this hypothesis of invariance through permutation is crucial in moving to the limit.

Moving to the limit causes a situation to emerge in which each player has become infinitesimal amidst the mass of other players, and constructs his strategies from his own state and from the state of the infinite mass of his co-players, who in turn simultaneously construct their strategies in the same way. It is this equilibrium of each player in the mass of co-players that we term the mean field approach.

Continuums of players are not new in the literature and they have been widely used since Robert Aumann and his seminal paper on general equilibrium with infinitely many players (see [9]). However, our approach is different in many ways from what has been studied up to now (see the literature on large games for instance – [29–32]). An example is the set of specific partial differential equations systems developed and studied in the seminal articles [34–36] and in [38]. Another instance is the approximation of a N-player game by a mean field game and the study on the error term in $\frac{1}{N}$ (see our first toy model). This strategy of approximation allows us to constrain the strategies of the players (since no complex strategies involving specific players can be played) and hence to reduce the finite dimension of the game to a granularity effect that leads to a common noise for the group of players.

1.1.3 Third Route: From Economics to Mean Field Games

In the theory of general economic equilibrium, agents have little concern for each other: everyone looks only to his own interest and to market prices. The only level at which the existence of others applies is found in the hypothesis of rational expectation. A theory is viewed as credible from the standpoint of rational expectations only if each agent can check whether, by putting himself in the place of others, he would find the behavior predicted by the theory. This is the only requirement that removes the agent of general equilibrium from his solipsism. In other words, in the theory of general equilibrium, prices mediate all social interactions. Yet we know that in many cases there are other economic effects which give rise to other interactions between agents: externality, public goods, etc. The incorporation of these effects when they are of a statistical nature, which is most often the case, leads to a "mean field"-type definition (in the sense given above) of equilibrium between agents. Similarly, the issues of industrial economics in which agents are involved in complex systems of signaling, entry barriers, positioning in relation to the competition, etc. can become mean field games equilibria as the size of the groups of agents concerned grows.

These interactions between agents are the main interests of economists. They want to understand how prices form through rational behaviors and the consequence of externality effects. Also, economists are interested in the evolution of an economy

and hence they have been spending a lot of time on anticipation and the way prices or, more generally, behaviors form in an intertemporal context. This field of economics is clearly untractable for econophysicists since econophysics only considers forward problems without anticipations except perhaps from a heuristical point of view that makes the backward dimension vanish. That's another difference between mean field game theory and the mean fields of econophysicists: mean field games have a forward/backward structure. In most mean field games models, we try not only to describe but also, and most importantly, to explain a phenomenon using the economic toolbox of utility maximization and rational expectations. Hence mean field game theory appears as a toolbox to be used by economists and not as a new competing paradigm in social sciences that avoids considering the major methodological advances made by economists in the last decades.

1.2 Fields of Application

1.2.1 Mean Field Games Versus N-Player Modeling

These three routes place the representation of agents in mean field game theory. They are more sophisticated than the agents of general equilibrium in economics, who as we have seen are largely indifferent to their co-agents and are concerned only with prices. Conversely, the agents of mean field game theory are less sophisticated than the players of N-player game theory since they base their strategies only on the statistical state of the mass of co-agents.

Nevertheless, this lesser sophistication of the mean field games agent compared to the N-player game theory player produces by way of compensation a wide variety of positive effects in two very different respects: in terms of efficiency on the one hand and of widening the field of application on the other.

As far as efficiency is concerned

A large part of this efficiency and readability comes from the possibility of deploying the power of differential calculus. This advantage is, moreover, one of the differences between mean field games and other prior developments already mentioned in games with an infinite number of players. These works, which follow from Robert Aumann's outstanding contribution, basically use measure theories, as we do, to represent the continuum of players, but they only use measure theory. From a mathematical standpoint, mean field game theory takes a completely new direction by opening the door to extensive use of differential calculus. Differential calculus has been one of the most powerful and productive tools for some 300 years and there have been major advances in the last decades in many applied fields outside physics: applications of partial differential equations (PDE) to control problems, Itô or Malliavin stochastic calculus, SPDE, advanced methods of functional analysis,

etc. Mean field game theory has moreover enabled a new and autonomous corpus to be developed in this mathematical field, including at the junction of differential calculus and measure theory, in the form of sophisticated developments in the geometry of Wasserstein spaces.

An enlargement of the field of application: two examples

1. A substantial gain in relation to N-player game theory derives from the ease with which questions of player entries and exits can be dealt with. Indeed, through the representation of players by a continuum, the modeling of the renewal of player generations is no longer a problem. Like time and player distribution in space, states are continuous variables, and entries and exits are simple flows whose technical treatment presents no special problems. One can thus implement overlapping generation models without pain.
2. The emergence of a social dimension to mean field games models, since, in these models, statistical data on other players emerge as fundamental constituents of individual strategies. From this point of view, the approximation of N-player games by the mean field games limit with (if necessary) the use of the corrective $1/N$ term, allows this approach to introduce a "social" dimension in regard to players, even in limited groups of, say, a few hundred agents.

In view of the benefits compared with N-player games, it seems quite natural to us to consider mean field games as "solutions" to problems of N-player games. Consider, for example, an N-player game where N is fairly small (of the order of a few dozen) and with player entries and exits. It is very likely that in a large number of cases the mean field games limit ($N \to \infty$) provides a good first approximation to the N-player solution and that the first term of the development in $1/N$ is sufficient to described with enough precision the effects due to granularity (produced by the fact that N is finite and rather small).

Thus there is a wide field of application for mean field games models. It ranges from general equilibrium with externality to the Mexican wave, and its center of gravity seems to us, from today's standpoint, to be socioeconomic modeling in a rational expectations context.

1.2.2 A Large Family of Examples

To illustrate the different aspects of mean field game theory, and to indicate something of its domains of application, we shall in the follow-up to this course present a series of "toy models". In other words, as we mentioned above, we generally present extremely stylized models, which are not to be taken literally and require being reworked by specialists, but which show the possible architecture of applications to various questions:

- Effects of coordination and interactions in the face of externality (meeting time)
- Production of a limited resource (peak oil, Hotelling's rule)

- Mimicry and propagation of behaviors in the social area (Mexican wave, forenames, fashion, etc.)
- Agoraphobia/agoraphilia, search for identity, etc. (quadratic-Gaussian population models)
- Distortion of financial management in the presence of a signaling problem (managers and classification)
- Effects of competition on the dynamics of human capital (Pareto-type distribution of salaries: an example of the positive effect of negative externality)

In the course of presenting these "toy models", the mathematical concepts and methods, indeed the whole mean field games toolbox, will become progressively apparent. The considerations mentioned above will thus acquire substance.

1.3 The Mathematical Tools of the Mean Field Approach

The implementation of the mean field game theory as a modeling methodology led to writing new types of systems of equation, then developing the mathematical apparatus required for handling these equations: theorems for the existence of solutions, numerical calculation algorithms, specific definition of stability, variational principles, etc.

We shall return in the conclusion of this course to the mathematical corpus which the reader will be able to discover through these examples.

2 A First Toy Model: When Does the Meeting Start?

We begin with a "toy model" constructed as a series of episodes, or rather as the old TV show "Double your Money", in which the dramatic tension progressively grows. We shall here adopt the serious true/false-type question: "What time does the meeting start?".

We recall what was previously said in the introduction on the role of "toy models", of which the prisoner's dilemma is archetypical. Nevertheless, we shall proceed as if it involved a scientific investigation of the subject, by tackling it in an increasingly sophisticated step-by-step manner. We therefore begin with a relatively simple framework, then we progressively add various difficulties to give the model greater depth and at the same time reveal the technological answers provided by the mean field games approach. As this "toy model" is presented, we hope the readers will come to feeling that it applies to real examples they are particularly familiar with. In the next section, we shall offer a stylized model of oil production over a long period of time that will show how our "toy model" can shed light on more serious contexts.

2.1 A First Simple Model

2.1.1 Introduction

A meeting scheduled for a certain time t very often only starts several minutes after the scheduled time. The actual time T when the meeting starts depends on the dynamics of the arrival of its participants. If a rule sets the start of the meeting at the point when a certain quorum is reached, this rule sets up a form of strategic interaction between agents. We shall construct a first mean field games approach to this situation.

We consider a meeting with a very large number of participants and we consider them as a continuum of agents (the justification will be provided further on). Our agents are rational and understand the situation. More precisely, all the data that we shall provide pertaining to the problem is common knowledge to the meeting participants.

Three moments will be important in this model:

- t the scheduled time of the meeting.
- τ^i the time at which agent i would like to arrive in view of the problem. In reality, we suppose that he will arrive at time $\tilde{\tau}^i = \tau^i + \sigma^i \tilde{\epsilon}^i$ where $\tilde{\epsilon}^i$ is a normal noise with variance 1, specific to agent i (hypothesis of idiosyncratic noise[2]). More precisely, τ^i is a variable controlled by the agent i and $\sigma^i \tilde{\epsilon}^i$ is an uncertainty the agent is subject to. These uncertainties and their intensity differ in the population of agents since some agents come a long way to participate in the meeting and others are very close. We will note m_0 the distribution of σ^i in the population.
- T the time which the meeting will start at (the rule which sets the meeting starting time T according to the arrival of participants is given further on).

To decide on his arrival time, or at least his intended arrival time τ^i, each agent will optimize a total cost that, to simplify things (since it is "toy model"), we assume is made up of three components:

- A cost (reputation effect) of lateness in relation to the scheduled time t:

$$c_1(t, T, \tilde{\tau}) = \alpha[\tilde{\tau} - t]_+$$

- A cost (personal inconvenience) of lateness in relation to the actual starting time of the meeting T:

$$c_2(t, T, \tilde{\tau}) = \beta[\tilde{\tau} - T]_+$$

- A waiting time cost that corresponds to the time lost waiting to reach time T:

$$c_3(t, T, \tilde{\tau}) = \gamma[T - \tilde{\tau}]_+$$

[2] This hypothesis of independence will simplify the equations to determine the equilibrium.

Let us note $c(t, T, \tilde{\tau})$ the sum of these three costs which is a convex function of $\tilde{\tau}$.

We can already make the model more complex by assuming different values for c_1, c_2, c_3 according to the agent, but our goal is precisely the opposite: to make it simple in order to explain the methodology.

2.1.2 Resolution

The optimization problem that each agent faces is therefore to minimize his expected total cost. Nash-MFG equilibrium, with rational expectations, presupposes that each agent optimizes by assuming T to be known. T is a priori a random variable but because we consider an infinite number of players, the "law of large numbers" will imply that T is deterministic and we consider a deterministic T from now.[3]

For agent i the problem is therefore:

$$\tau^i = argmin \mathbb{E}\left[c(t, T, \tilde{\tau}^i)\right], \qquad \tilde{\tau}^i = \tau^i + \sigma^i \tilde{\epsilon}^i$$

Here T is the mean field, the exhaustive summary for each agent of the behavior of the others.

The exercise is to show the existence of a fixed point T, i.e. to show that individual optimization behaviors, supposing T is known, fully generate the realization of this time T.

To show that this equilibrium exists, one must first examine more closely agents' individual choices, which is done simply by obtaining a first-order condition.

Proposition 1 (FOC). *The optimal τ^i of an agent having a σ equal to σ^i is implicitly defined by:*

$$\alpha \mathcal{N}\left(\frac{\tau^i - t}{\sigma^i}\right) + (\beta + \gamma)\mathcal{N}\left(\frac{\tau^i - T}{\sigma^i}\right) = \gamma$$

where $\mathcal{N}$ is the cumulative distribution function associated to a normal distribution.

Proof. The expression to minimize is:

$$\begin{aligned}
&\mathbb{E}\left[\alpha[\tilde{\tau}^i - t]_+ + \beta[\tilde{\tau}^i - T]_+ + \gamma[T - \tilde{\tau}^i]_+\right] \\
&\quad = \mathbb{E}\left[\alpha[\tilde{\tau}^i - t]_+ + (\beta + \gamma)[\tilde{\tau}^i - T]_+ - \gamma(\tilde{\tau}^i - T)\right] \\
&\quad = \alpha \mathbb{E}\left([\tau^i - t + \sigma^i \tilde{\epsilon}^i]_+\right) + (\beta + \gamma)\mathbb{E}\left([\tau^i - T + \sigma^i \tilde{\epsilon}^i]_+\right) - \gamma(\tau^i - T)
\end{aligned}$$

[3] Hence rational expectations are simply perfect expectations.

The first order condition of the problem is therefore given by:

$$\alpha\mathbb{P}\left(\tau^i - t + \sigma^i\tilde{\epsilon}^i > 0\right) + (\beta+\gamma)\mathbb{P}\left(\tau^i - T + \sigma^i\tilde{\epsilon}^i > 0\right) = \gamma$$

$$\alpha\mathcal{N}\left(\frac{\tau^i - t}{\sigma^i}\right) + (\beta+\gamma)\mathcal{N}\left(\frac{\tau^i - T}{\sigma^i}\right) = \gamma$$

Since $\mathcal{N}$ is a strictly monotonic cumulative distribution function and since the three parameters α, β and γ are positive, the existence and uniqueness of τ^i can be deduced easily. □

From this characterization of τ^i as a function of (t, T, σ^i) we can deduce the dynamics of agents' arrival. For this let us consider first of all the distribution m_0 of the σ^i within the continuum. Because of the continuum and the law of large numbers, this distribution is transported by the application $\sigma^i \mapsto \tilde{\tau}^i$.

If we therefore note F the (deterministic!) cumulative distribution function of the agents' real arrival times, it is natural to establish a rule on the real starting time T from the meeting, which depends on the function $F(\cdot)$. An example is that of a quorum: the meeting starts after the scheduled time and only when a proportion θ of the participants have arrived.

We then have to prove the existence and uniqueness of a fixed point. Starting from a value T, we obtain agents' optimal strategies $(\tau^i(\cdot;T))_i$. These optimal strategies are the targeted times but each person's arrival time is affected by a noise: we obtain the real arrival times $(\tilde{\tau}^i(\cdot;T))_i$. Then from the law of large numbers and the hypothesis of the independence of agents' uncertainties, these arrival times are distributed according to F, which is deterministic, and T is deduced from F by the meeting starting time rule $(T^*(F))$, in this case the quorum. This is straightforwardly summarized by the following scheme:

$$T^{**} : T \mapsto (\tau^i(\cdot;T))_i \mapsto (\tilde{\tau}^i(\cdot;T))_i \mapsto F = F(\cdot;T) \mapsto T^*(F)$$

The result we obtain is as follows:

Proposition 2 (Equilibrium T). *If $\alpha > 0, \beta > 0, \gamma > 0$ and if $0 \notin \overline{supp(m_0)}$ then T^{**} is a contraction mapping of $[t; +\infty[$, and there is a unique solution T to our problem.*

Proof. First, let's differentiate with respect to T the first order condition that defines τ^i.

$$\frac{d\tau^i}{dT}\left[\alpha\mathcal{N}'\left(\frac{\tau^i - t}{\sigma^i}\right) + (\beta+\gamma)\mathcal{N}'\left(\frac{\tau^i - T}{\sigma^i}\right)\right] = (\beta+\gamma)\mathcal{N}'\left(\frac{\tau^i - T}{\sigma^i}\right)$$

Since 0 is supposed not to be in the support of m_0, this leads to $\frac{d}{dT}\tau(t,\sigma;T) \leq k < 1$.

Hence, $\forall T, s, h > 0$,

$$\begin{aligned} F(s; T+h) &= \mathbb{P}(\tau^i(\sigma^i; T+h) + \sigma^i \epsilon^i \leq s) \geq \mathbb{P}(\tau^i(\sigma^i; T) + kh + \sigma^i \epsilon^i \leq s) \\ &= F(s - kh; T) \end{aligned}$$

Consequently,

$$\begin{aligned} & T^*(F(\cdot; T+h)) \leq T^*(F(\cdot - kh; T)) \leq T^*(F(\cdot; T)) + kh \\ & \Rightarrow T^{**}(T+h) - T^{**}(T) \leq kh \end{aligned}$$

and this proves the result through the contraction mapping theorem. □

It is interesting to notice that the quorum case is not special in the sense that the preceding proof only requires the T setting rule ($T^* : F(\cdot) \mapsto T$) to verify the following properties for the above result to be true.

- $\forall F(\cdot), T^*(F(\cdot)) \geq t$: the meeting never starts before t
- **(Monotony)** Let's consider two cumulative distribution functions $F(\cdot)$ and $G(\cdot)$. If $F(\cdot) \leq G(\cdot)$ then $T^*(F(\cdot)) \geq T^*(G(\cdot))$
- **(Sub-additivity)** $\forall s > 0, T^*(F(\cdot - s)) - T^*(F(\cdot)) \leq s$

In the more general case where the cost depends on F, strategic interaction no longer simply reduces to time T. It is very natural that the social cost for each agent depends on the proportion of participants who are already there when he arrives. In this more general case, F is the mean field: each person makes his decision according to F. In return, the decisions construct F. From a mathematical standpoint, the fixed point concerns F.

2.2 Variations

There are many possible ways of enriching this initial "toy model". For example, one variant involves considering a shared disturbance in addition to the idiosyncratic disturbances. This is an important variant as it is an example where the dynamics of the population is stochastic. Nonetheless, as it would lead us to too long developments we will not consider this variant here.

The variant we shall present is a "geographical" model, i.e. the agents are initially distributed in different places and must come to where the meeting is being held.

The interest of this variant is that it will show how coupled forward/backward PDEs, which are the core of mean field game theory (in continuous time, with a continuous state space), emerge.

2.2.1 The Framework

Thus let us suppose that the agents are distributed on the negative half-line according to distribution function $m_0(\cdot)$ (with compact support and such that $m_0(0) = 0$) and that they must go to the meeting held at 0. Suppose that in order to get to 0, an agent i moves according to the process $dX_t^i = a_t^i dt + \sigma dW_t^i$ where drift a is controlled in return for a quadratic cost $\frac{1}{2}a^2$ (here σ is the same for everyone). This distribution hypothesis may seem to be a rather artificial representation in this example of transport uncertainties. In practice, we shall see that it is relatively pertinent in other applications.

Each agent is thus faced with an optimization problem, written as:

$$Min_{a(\cdot)}\mathbb{E}\left[c(t,T,\tilde{\tau}^i) + \frac{1}{2}\int_0^{\tilde{\tau}^i} a^2(t)dt\right]$$

with $X_0^i = x_0, \quad dX_t^i = a_t^i dt + \sigma dW_t^i$ and the time to reach 0 is given by $\tilde{\tau}^i = min\{s/X_s^i = 0\}$.

If one looks for a Nash-MFG equilibrium, one will reason at a given T and each agent's problem is one of stochastic control. We thus have the following Hamilton–Jacobi–Bellman equation[4]:

$$0 = \partial_t u + \min_a \left(a\partial_x u + \frac{1}{2}a^2\right) + \frac{\sigma^2}{2}\partial_{xx}^2 u$$

This equation can be written:

$$(HJB) \qquad \partial_t u - \frac{1}{2}(\partial_x u)^2 + \frac{\sigma^2}{2}\partial_{xx}^2 u = 0$$

The condition at the limit is simply $\forall \tau, u(\tau, 0) = c(t, T, \tau)$, where T is deterministic for the same reason as before. This condition corresponds to the total cost on arrival at the meeting (we will assume that c has the same shape as in the preceding setup but we impose c not to be piecewise-linear but twice continuously differentiable).

The Hamilton–Jacobi–Bellman equation gives a Bellman function u and hence indicates the optimal behavior of agents for a fixed T. This equation is the same for all agents since they have the same cost criterion and differ only in their point

[4] As it is often the case in this text, we will consider that the solution of the Hamilton–Jacobi–Bellman equation is a solution of the optimization problem. In general, if we do not provide any verification result for the solutions of the mean field games partial differential equations, it should be noticed that verifying (u, m) indeed provides a solution of the optimization problem is like verifying u, solution of the Hamilton–Jacobi–Bellman equation with m fixed, provides a solution of the optimization problem with m fixed. Hence, there should not be specific tools of verification for mean field games.

of departure at the initial moment. Moreover, the solution here is Markovian as in most stochastic optimization problems and the strategy, i.e. the optimal drift $a(s,x) = -\partial_x u(s,x)$ therefore depends only on the place x and the time s. Thus, in particular, it does not depend on the agent concerned: the agents, whom unknown factors lead to the same point x at time s, adopt the same strategy, which is natural, since they have the same information, same transport costs and same final criterion. This property simplifies the expression of the problem.

The "law of large numbers" then gives us the distribution m of agents through the Kolmogorov equation.[5] This distribution corresponds to the distribution of players who have not yet arrived at 0 and therefore m loses mass (through 0), as agents gradually arrive at the meeting. The dynamics of m is:

$$(Kolmogorov) \qquad \partial_t m + \partial_x((-\partial_x u)m) = \frac{\sigma^2}{2}\partial^2_{xx} m$$

$m(0,\cdot) = m_0(\cdot)$ is obviously fixed, and we will try to find a solution with the following "smooth fit" condition: $m(\cdot,0) = 0$.

Moreover, as we have chosen to model the problem by the dynamics of Brownian diffusion, the model must be complemented and restricted to a compact domain. In the proof that follows, we suppose that the domain is $[0, T_{max}] \times [-X_{max}, 0]$ and the boundary conditions are

$$u(T_{max},\cdot) = c(t,T,T_{max}), \quad u(\cdot,-X_{max}) = c(t,T,T_{max}), \quad m(\cdot,-X_{max}) = 0$$

In this context, the flow reaching 0 (when the agents reach the meeting place) is $s \mapsto -\partial_x m(s,0)$. Thus the cumulative distribution function F of arrival times is defined by

$$F(s) = -\int_0^s \partial_x m(v,0)dv$$

Now, T is fixed by the quorum rule (with let's say $\theta = 90\%$) but we impose that it must be in the interval $[t, T_{max}]$. In other words:

$$T = \begin{cases} t, & \text{if } F^{-1}(\theta) \leq t \\ T_{max}, & \text{if } F(T_{max}) \leq \theta \\ F^{-1}(\theta), & \text{otherwise} \end{cases}$$

[5] Note that this is based not only on the independence hypothesis of noises but also on the simple structure of noises. For example, if volatility depends on state, the associated elliptic operator would replace the Laplace operator. Also, If noises were not independent, the deterministic partial differential equation would have to be replaced by a stochastic one. For all these developments, we refer to [38].

2.2.2 Existence of an Equilibrium for the Meeting Starting Time

As in the first simple case, we need to prove that there is a time T coherent with the (rational) expectations of the agents. We are going to use a fixed point theorem as before. Indeed, one goes from a given T and deduces u. The Kolmogorov equation then gives us m and therefore the arrival flow at 0. Since the time T in our example is given by the arrival of a proportion θ of all the agents, it clearly is a matter of fixed point.

Before going deeply in the mathematics, let's introduce some hypotheses:

- We suppose that $T \mapsto c(t, T, \tau)$ is a continuous function
- We suppose that $\tau \mapsto c(t, T, \tau)$ is a C^2 function
- We suppose that $m_0(0) = m_0(-X_{max}) = 0$. Also, we suppose that $|m_0'(0)| > 0$ and $|m_0'(-X_{max})| > 0$

Now, we consider the following scheme (the functional spaces involved in the scheme will be proved to be the right ones in what follows):

$$\begin{aligned} T \mapsto c(t, T, \cdot) \in C^2 \mapsto u \in C^2 \mapsto \partial_x u \in C^1 \mapsto m \in C^1 \\ \mapsto -\partial_x m(\cdot, 0) \in C^0 (\mapsto F) \mapsto T \end{aligned}$$

Since the scheme is from $[t, T_{max}]$ to $[t, T_{max}]$, to obtain a fixed point result, we just need to prove that the scheme is continuous.

The first part of the scheme ($T \mapsto c(t, T, \cdot) \in C^2$) is continuous and well defined by hypothesis. For the second part of the scheme ($c(t, T, \cdot) \in C^2 \mapsto u \in C^2$), we just state a lemma:

Lemma 1. *Let's consider the following PDE:*

$$(HJB) \qquad \partial_t u - \frac{1}{2}(\partial_x u)^2 + \frac{\sigma^2}{2}\partial^2_{xx} u = 0$$

with the boundary conditions

$$\begin{aligned} u(\cdot, 0) = c(t, T, \cdot) \quad u(T_{max}, \cdot) = c(t, T_{max}, T_{max}), \quad u(\cdot, -X_{max}) \\ = c(t, T_{max}, T_{max}) \end{aligned}$$

The solution u is in $C^2(]0, T_{max}[\times] - X_{max}, 0[)$ and $\exists K, \forall T \in [t, T_{max}], \partial_x u$ is a K-Lipschitz function.

Moreover the mapping $c(t, T, \cdot) \in C^2 \mapsto u \in C^2$ is continuous.

Now that we get u and then the control $-\partial_x u$ we can turn to the Kolmogorov equation. We state a lemma that is an application of Hopf's principle.

Lemma 2. *Let's consider the following PDE:*

$$(Kolmogorov) \qquad \partial_t m + \partial_x (am) = \frac{\sigma^2}{2}\partial^2_{xx} m$$

with a a C^1 (and hence Lipschitz) function and the boundary conditions $m(0,\cdot) = m_0(\cdot), \quad m(\cdot,0) = 0, \quad m(\cdot,-X_{max}) = 0$ where m_0 is supposed to verify the above hypotheses.

Then the solution m is in $C^1((0,T_{max}) \times (-X_{max},0))$ and

$$\exists \epsilon > 0, \inf |\partial_x m(\cdot,0)| \geq \epsilon$$

Moreover ϵ only depends on the Lipschitz constant of the function a.

Also the mapping $a \mapsto m \in C^1$ is continuous.

From these two lemmas, we can deduce a third one adapted to our problem. Indeed, since u is a C^2 function, $a = -\partial_x u$ is a Lipschitz function and hence we have a lower bound to the flow arriving at the meeting:

Lemma 3.

$$\exists \epsilon > 0, \forall T \in [t, T_{max}], \inf |\partial_x m(\cdot,0)| \geq \epsilon$$

Now, let's consider the mapping $\Psi : -\partial_x m(\cdot,0) \in C^0 \mapsto T$, defined above using (here) the quorum rule. We are going to prove that Ψ is continuous as soon as $-\partial_x m(\cdot,0)$ has a strictly positive lower bound.

Lemma 4. *Ψ is a Lipschitz function on $C^0([0,T_{max}], \mathbb{R}_+^*)$.*

Proof. Let's consider two functions ψ_1 and ψ_2 that stand for two possible flows of arrival and let's define ϵ a common lower bound to these two functions. Then, let's define $T_1 = \Psi(\psi_1)$ and $T_2 = \Psi(\psi_2)$. If T_1 and T_2 are both in $]t, T_{max}[$, then, assuming $T_1 \leq T_2$, we can write:

$$\begin{aligned} 0 &= \int_0^{T_1} \psi_1 - \int_0^{T_2} \psi_2 = \int_0^{T_1} (\psi_1 - \psi_2) - \int_{T_1}^{T_2} \psi_2 \\ &\Rightarrow \epsilon(T_2 - T_1) \leq \int_{T_1}^{T_2} \psi_2 = \int_0^{T_1} (\psi_1 - \psi_2) \leq T_{max} |\psi_1 - \psi_2|_\infty \end{aligned}$$

Hence, in this case, the function is Lipschitz.

In all other cases, still assuming $T_1 \leq T_2$, we have instead of an equality the following inequality:

$$\int_0^{T_1} \psi_1 - \int_0^{T_2} \psi_2 \geq 0$$

and the result follows from the same reasoning.

Thus, the function is Lipschitz and hence continuous. □

By now, we have proved that the scheme is continuous and therefore, using the Brouwer fixed point theorem we have existence of an equilibrium T.

Proposition 3 (Existence). *The scheme that defines the actual T as a function of the anticipated T is continuous and has at least one fixed point.*

Hence, there is at least one equilibrium T.

2.2.3 Comments

This variant is a good example of a mean field game in continuous time and the mechanism will often be the same. First agents anticipate what will be the dynamics of the community and hence anticipate m. Here, the relevant information was captured by T that is a function of m so that they had to anticipate m to anticipate T. From this anticipation agents use a backward reasoning described by the Hamilton–Jacobi–Bellman equation. Then, from this equation, individual actions can be plugged into the forward equation (the Kolmogorov equation) to know the actual dynamics of the community implied by individual behaviors. Finally, the rational expectation hypothesis implies that there must be coherence between the anticipated m and the actual m.

This forward/backward mechanism is the core of mean field game theory in continuous time and we will see it in action later on.

2.3 Mean Field Games Equilibrium as the N-Player Nash Limit Equilibrium

2.3.1 Introduction

Let us return for the sake of simplicity to our first model. It is now time to come back to the continuum hypothesis by considering the game with N players. To simplify the account, and because it involves a "toy model", we look at the same particular case as above (which is rather technical since the criterion is not regular but is very graphic) in which the meeting begins once a proportion θ (we shall assume $\theta = 90\%$ for the sake of simplicity) of the participants have arrived (but still we force T to be between times t and T_{max}). In addition, let us suppose that all the agents have the same σ. Various questions then naturally arise:

- Does the N-player game have Nash equilibria?
- Is there uniqueness of such equilibria?
- Do N-player equilibria tend towards the mean field games equilibrium when $N \to \infty$?
- If need be, is the rate of convergence known?

This case is simple, but it allows – since we shall answer the above questions in the affirmative (in the symmetrical case) – to pave the way for an approximation of an N-player game by MFG.

This example of approximation of a N-player game through a first order expansion "$G_0 + \frac{1}{N}G_1 + \ldots$", where (formally) G_0 is the mean field game and G_1 the first order correction coefficient, leads to a new type of solution of a N-player game equilibrium. The solution of "$G_0 + \frac{1}{N}G_1$" reflects a strategic world in which agents

do not care about other agents, individually at least, but only about the population dynamics and a world in which N, the number of players, is only entered to take into account the "granularity" of the game and the imperfectness of the continuum hypothesis.

2.3.2 Solution of the N-Player Games

To simplify, let us say that the number of players is $N = 10k$ ($k = 1, 2, 3, \ldots$) and thus that the meeting begins with the arrival of the $9k^{th}$ player. A given player (let us say player 1) will aim for an arrival time τ^* which should verify (symmetrical Nash equation):

$$\tau^* = argmin_{\tau^1}\mathbb{E}[C(\tau^1 + \sigma\tilde{\epsilon}^1, \tau^* + \sigma\tilde{\epsilon}^2, \ldots, \tau^* + \sigma\tilde{\epsilon}^N)]$$

This function C does not really depend on all the components of $(\tau^* + \sigma\tilde{\epsilon}^2, \ldots, \tau^* + \sigma\tilde{\epsilon}^N)$ but only on two statistics of order $\tau^* + \sigma\tilde{\epsilon}_{(9k-1)}$ and $\tau^* + \sigma\tilde{\epsilon}_{(9k)}$ where one has noted $\tilde{\epsilon}_{(r)}$ the r^{th} element, in the order, in $\{\tilde{\epsilon}^2, \ldots, \tilde{\epsilon}^N\}$. Indeed it is obvious that the 90-percentile of $(\tau^1 + \sigma\tilde{\epsilon}^1, \tau^* + \sigma\tilde{\epsilon}^2, \ldots, \tau^* + \sigma\tilde{\epsilon}^N)$ is hidden among $\tau^1 + \sigma\tilde{\epsilon}^1$, $\tau^* + \sigma\tilde{\epsilon}_{(9k-1)}$ and $\tau^* + \sigma\tilde{\epsilon}_{(9k)}$.

Thus the Nash equilibrium is characterized by:

$$\tau^* = argmin_{\tau^1}\mathbb{E}[G(\tau^1 + \sigma\tilde{\epsilon}^1, \tau^* + \sigma\tilde{y}, \tau^* + \sigma\tilde{z})]$$

where $(\tilde{y}, \tilde{z})$ are statistics of order corresponding to the $(9k-1)^{th}$ and $9k^{th}$ ordered elements of $\{\tilde{\epsilon}^2, \ldots, \tilde{\epsilon}^N\}$. Hence, the variables $(\tilde{y}, \tilde{z})$ are independent of $\tilde{\epsilon}^1$.

Taking up the initial model, the function G is defined by:

$$\forall a, \forall b, \forall c \geq b, \quad G(a, b, c) = G(a, t \vee b \wedge T_{max}, t \vee c \wedge T_{max})$$

$$\forall b \leq c \in [t, T_{max}], G(a, b, c) = \begin{cases} -\gamma(a-b) & a \leq t \\ -\gamma(a-b) + \alpha(a-t) & a \in (t, b] \\ \alpha(a-t) & a \in (b, c] \\ \alpha(a-t) + \beta(a-c) & a > c \end{cases}$$

We have the following property:

Lemma 5. *$\forall b, \forall c \geq b, a \mapsto G(a, b, c)$ is continuous, piecewise linear and convex.*

G is not practical for optimization purposes. Let's introduce H the function $(\tau^1, b, c) \mapsto \int_{-\infty}^{\infty} G(\tau^1 + \sigma x, b, c)\mathcal{N}'(x)dx$ where $\mathcal{N}$ still is the cumulative distribution function of a normal variable with variance 1.

Lemma 6. *$\forall b \leq c$, H is a strictly convex function of τ^1 that decreases and then increases. Moreover, we can compute some derivatives $(t \leq b \leq c \leq T_{max})$:*

$$\partial_1 H(\tau^1, b, c) = \left[-\gamma \mathcal{N}\left(\frac{b-\tau^1}{\sigma}\right) + \alpha\left(1 - \mathcal{N}\left(\frac{t-\tau^1}{\sigma}\right)\right) + \beta\left(1 - \mathcal{N}\left(\frac{c-\tau^1}{\sigma}\right)\right)\right]$$

$$\partial^2_{11} H(\tau^1, b, c) = \frac{1}{\sigma}\left[\gamma \mathcal{N}'\left(\frac{b-\tau^1}{\sigma}\right) + \alpha \mathcal{N}'\left(\frac{t-\tau^1}{\sigma}\right) + \beta \mathcal{N}'\left(\frac{c-\tau^1}{\sigma}\right)\right]$$

$$\partial^2_{12} H(\tau^1, b, c) = -\frac{1}{\sigma}\gamma \mathcal{N}'\left(\frac{b-\tau^1}{\sigma}\right)$$

$$\partial^3_{13} H(\tau^1, b, c) = -\frac{1}{\sigma}\alpha \mathcal{N}'\left(\frac{c-\tau^1}{\sigma}\right)$$

Proof.

$$\int_{-\infty}^{\infty} G(\tau^1 + \sigma x, b, c)\mathcal{N}'(x)dx = \frac{1}{\sigma}\int_{-\infty}^{\infty} G(t, b, c)\mathcal{N}'\Big(\frac{t-\tau^1}{\sigma}\Big)dt$$

Hence, we can differentiate with respect to τ^1 and we get:

$$\begin{aligned}\partial_1 H(\tau^1, b, c) &= -\frac{1}{\sigma^2}\int_{-\infty}^{\infty} G(t, b, c)\mathcal{N}''\Big(\frac{t-\tau^1}{\sigma}\Big)dt \\ &= -\frac{1}{\sigma}\int_{-\infty}^{\infty} G(\tau^1 + \sigma x, b, c)\mathcal{N}''(x)dx\end{aligned}$$

Then, using derivatives in the distribution sense, we get:

$$\partial_1 H(\tau^1, b, c) = \int_{-\infty}^{\infty} \partial_1 G(\tau^1 + \sigma x, b, c)\mathcal{N}'(x)dx$$

$$\partial_1 H(\tau^1, b, c) = \int_{-\infty}^{\infty} [-\gamma 1_{\tau^1+\sigma x \leq b} + \alpha 1_{\tau^1+\sigma x \geq t} + \beta 1_{\tau^1+\sigma x \geq c}]\mathcal{N}'(x)dx$$

$$\partial_1 H(\tau^1, b, c) = \left[-\gamma \mathcal{N}\left(\frac{b-\tau^1}{\sigma}\right) + \alpha\left(1 - \mathcal{N}\left(\frac{t-\tau^1}{\sigma}\right)\right) + \beta\left(1 - \mathcal{N}\left(\frac{c-\tau^1}{\sigma}\right)\right)\right]$$

We can differentiate once again and we get:

$$\partial^2_{11} H(\tau^1, b, c) = \frac{1}{\sigma}\left[\gamma \mathcal{N}'\left(\frac{b-\tau^1}{\sigma}\right) + \alpha \mathcal{N}'\left(\frac{t-\tau^1}{\sigma}\right) + \beta \mathcal{N}'\left(\frac{c-\tau^1}{\sigma}\right)\right]$$

This is strictly positive so that H is strictly convex as a function of the first variable. Since $H(\tau^1, b, c) = \mathbb{E}[G(\tau^1 + \sigma\tilde{\epsilon}^1, b, c)] \geq G(\mathbb{E}[\tau^1 + \sigma\tilde{\epsilon}^1], b, c) = G(\tau^1, b, c)$, H must be decreasing and then increasing.

Other derivatives are straightforwardly given by:

$$\partial^2_{12} H(\tau^1, b, c) = -\frac{1}{\sigma}\gamma \mathcal{N}'\left(\frac{b-\tau^1}{\sigma}\right)$$

$$\partial^3_{13} H(\tau^1, b, c) = -\frac{1}{\sigma}\alpha \mathcal{N}'\left(\frac{c-\tau^1}{\sigma}\right)$$

□

Let's now recall that we want to find a symmetrical Nash equilibrium and the condition is given by:

$$\tau^* = argmin_{\tau^1} \mathbb{E}[G(\tau^1 + \sigma\tilde{\epsilon}^1, \tau^* + \sigma\tilde{y}, \tau^* + \sigma\tilde{z})]$$

Clearly this can be rewritten using the function H and we get:

$$\tau^* = argmin_{\tau^1} \mathbb{E}[H(\tau^1, \tau^* + \sigma\tilde{y}, \tau^* + \sigma\tilde{z})]$$

Using the first order condition associated to the preceding minimization we see that we need to better understand the function $\partial_1 H$. The following lemma will be helpful in what follows because it introduces compactness in the problem:

Lemma 7.

$$B = \{\tau^1 | \exists b \leq c, \partial_1 H(\tau^1, b, c) = 0\}$$

is a bounded set.

Proof. The set we introduced corresponds to the set of points at which H reaches its minimum for all possible couples (b, c) with $b \leq c$. Because $\forall a, \forall b, \forall c \geq b, \quad G(a, b, c) = G(a, t \vee b \wedge T_{max}, t \vee c \wedge T_{max})$, the same type of properties applies for H and hence our set B is the same as

$$\{\tau^1 | \exists (b, c), t \leq b \leq c \leq T_{max}, \partial_1 H(\tau^1, b, c) = 0\}$$

Now, $\partial_1 H(\tau^1, b, c) = 0$ implicitly defines a function $\tau^1(b, c)$ that is continuous and hence the set B is compact (and then bounded) as the image of a bounded set $(\{(b, c), t \leq b \leq c \leq T_{max}\})$ by a continuous mapping. □

Let's introduce now the best response function of agent 1. This function Γ is defined as:

$$\Gamma(\tau^*) = argmin_{\tau^1}\mathbb{E}[H(\tau^1, \tau^* + \sigma\tilde{y}, \tau^* + \sigma\tilde{z})]$$

Another (though implicit) definition of this function is based on the first order condition:

$$\mathbb{E}[\partial_1 H(\Gamma(\tau^*), \tau^* + \sigma\tilde{y}, \tau^* + \sigma\tilde{z})] = 0 \qquad (*)$$

Lemma 8.

$$\forall \tau^*, \inf B \leq \Gamma(\tau^*) \leq \sup B$$

Proof. Since H is decreasing and then increasing as a function of the first variable, we clearly now that $\forall \xi < \inf B$:

$$\mathbb{E}[\partial_1 H(\xi, \tau^* + \sigma\tilde{y}, \tau^* + \sigma\tilde{z})] < 0$$

Hence $\inf B \leq \Gamma(\tau^*)$. The other inequality is obtained using the same reasoning. □

Since a Nash equilibrium simply is a fixed point of Γ, we can restrict Γ to the set $K = [\inf B, \sup B]$.

If we define $\Gamma_{|K} : \tau^* \in K \mapsto \Gamma(\tau^*)$, we see that any symmetrical Nash equilibrium must be a fixed point of $\Gamma_{|K}$.

Now we have our last lemma before the existence and uniqueness theorem that says:

Lemma 9. *$\Gamma_{|K}$ is a contraction mapping from K to K.*

Proof. Let's go back to the implicit definition of the function Γ given by $(*)$. Using the implicit function theorem we have:

$$\Gamma'(\tau^*) = -\frac{\mathbb{E}[\partial^2_{12} H(\Gamma(\tau^*), \tau^* + \sigma\tilde{y}, \tau^* + \sigma\tilde{z}) + \partial^2_{13} H(\Gamma(\tau^*), \tau^* + \sigma\tilde{y}, \tau^* + \sigma\tilde{z})]}{\mathbb{E}[\partial^2_{11} H(\Gamma(\tau^*), \tau^* + \sigma\tilde{y}, \tau^* + \sigma\tilde{z})]}$$

Since $0 < -\partial^2_{12} H - \partial^2_{13} H < \partial^2_{11} H$, we have $0 \leq \Gamma'(\tau^*) < 1$. Now because K is compact, there exists a constant $\varepsilon > 0$ so that $\forall \tau^* \in K, \Gamma'_{|K}(\tau^*) \leq 1 - \epsilon$. □

Now using a classical fixed point result we have:

Proposition 4 (Existence and Uniqueness). *There exists a unique symmetrical Nash equilibrium for the game with N players.*

Remark: We restrict ourselves to cases where $N = 10k$ and $\theta = 90\%$ for the sake of simplicity but the preceding result is still true for all N and θ.

2.3.3 Approximation in $1/N$

Before beginning the analysis, recall that the equilibrium is a Dirac measure in the mean field game case since all individuals have the same σ. We note this equilibrium τ^*_{MFG}, and the starting time for the meeting will be (except when a limit is reached) $\tau^*_{MFG} + \sigma F^{-1}(\theta)$ where F is here the cumulative distribution function of a normal distribution.

Thus, rather than being defined by:

$$\mathbb{E}[\partial_1 H(\tau^*_N, \tau^*_N + \sigma\tilde{y}, \tau^*_N + \sigma\tilde{z})] = 0$$

the mean field games equilibrium is defined by:

$$\partial_1 H(\tau^*_{MFG}, \tau^*_{MFG} + \sigma F^{-1}(\theta), \tau^*_{MFG} + \sigma F^{-1}(\theta)) = 0$$

We see that there is an advantage in introducing J defined by $J(t, y, z) = \partial_1 H(t, t+\sigma y, t+\sigma z)$ and that we can then carry out the following Taylor expansion:

$$\begin{aligned}
0 &= \mathbb{E}J(\tau^*_N, \tilde{y}, \tilde{z}) \\
&= \underbrace{J(\tau^*_{MFG}, F^{-1}(\theta), F^{-1}(\theta))}_{=0} + (\tau^*_N - \tau^*_{MFG})\partial_1 J(\tau^*_{MFG}, F^{-1}(\theta), F^{-1}(\theta)) \\
&\quad + \mathbb{E}(\tilde{y} - F^{-1}(\theta))\partial_2 J(\tau^*_{MFG}, F^{-1}(\theta), F^{-1}(\theta)) \\
&\quad + \mathbb{E}(\tilde{z} - F^{-1}(\theta))\partial_3 J(\tau^*_{MFG}, F^{-1}(\theta), F^{-1}(\theta)) \\
&\quad + \frac{1}{2}\mathbb{E}(\tilde{y} - F^{-1}(\theta))^2 \partial_{22} J(\tau^*_{MFG}, F^{-1}(\theta), F^{-1}(\theta)) \\
&\quad + \frac{1}{2}\mathbb{E}(\tilde{z} - F^{-1}(\theta))^2 \partial_{33} J(\tau^*_{MFG}, F^{-1}(\theta), F^{-1}(\theta)) \\
&\quad + \mathbb{E}(\tilde{z} - F^{-1}(\theta))(\tilde{y} - F^{-1}(\theta))\partial_{23} J(\tau^*_{MFG}, F^{-1}(\theta), F^{-1}(\theta)) \\
&\quad + o(\tau^*_N - \tau^*_{MFG}) + o(1/N)
\end{aligned}$$

Detailed study of the properties of order statistics (see [17]), i.e. the variables $\tilde{y}$ and $\tilde{z}$, show that the convergence of τ^*_N toward τ^*_{MFG} occurs in $1/N$.

Indeed, if we write[6]

$$\begin{aligned}
\xi &= \lim_{N\to\infty} N\mathbb{E}(\tilde{y} - F^{-1}(\theta)) \in \mathbb{R} \\
\zeta &= \lim_{N\to\infty} N\mathbb{E}(\tilde{z} - F^{-1}(\theta)) \in \mathbb{R} \\
v &= \lim_{N\to\infty} N\mathbb{E}(\tilde{y} - F^{-1}(\theta))^2 \\
&= \lim_{N\to\infty} N\mathbb{E}(\tilde{z} - F^{-1}(\theta))^2 = \lim_{N\to\infty} N\mathbb{E}(\tilde{z} - F^{-1}(\theta))^2 \in \mathbb{R}
\end{aligned}$$

[6] The fact that these constants exist is not obvious and relies on the properties of order statistics.

then we see that

$$\tau_N^* = \tau_{MFG}^* - \frac{1}{N}\frac{1}{\partial_1 J}\left[\xi\partial_2 J + \zeta\partial_3 J + \frac{v^2}{2}\left(\partial_{22}J + \partial_{33}J + 2\partial_{23}J\right)\right] + o\left(\frac{1}{N}\right)$$

The mean field games framework is therefore an approximation of an N-player game when N is large and we know the order of magnitude of the error occurring when we consider a mean field games model for solving an N-player game.

3 Application of Mean Field Game to Economics: Production of an Exhaustible Resource

A fairly typical example of mean field game is that of the production of an exhaustible resource by a continuum of producers. We know from Hotelling's work (see [28]) that there is a rent involved in the production of an exhaustible resource, but it is interesting to examine this in greater depth in a competitive situation and to understand the dynamics of exhaustion of a scarce resource. We therefore present a basic model onto which other models can be grafted. For instance, the model can be improved to take account of a Stackelberg-type competition, to consider the existence of big players (OPEC in the instance we have in mind), etc. It is also a basis for studying important problems such as the entry of new competitors, particularly those who are developing alternative energy sources (see [20] for a complete analysis of this question in a mean field game framework identical to the one developed here. This framework allows for instance to consider with powerful analytical tools the negative effect in terms of carbon emissions of a subsidy to alternative energy producers (see [25]) as in [18]).

This example will enable us to show the general character of mean field games PDEs when addressing Forward/Backward problems. It also offers a transparent way of dealing with externality.

3.1 Basis of the Model

We consider a large number of oil producers, which can be viewed either as wells or from a more macro standpoint as oil companies. The only assumption we make is that there is a sufficiently large number of them and that one can apply simple hypotheses such as that of the continuum (mean field games modeling) and perfect competition (price-taker behavior of agents).

Each of these oil producers initially has a reserve that is termed R_0. We assume that these reserves are distributed among producers according to an (initial) distribution $m(0,\cdot)$. These reserves will of course contribute to production q such that,

for any specific agent, we have $dR(t) = -q(t)dt + \nu R(t)dW_t$ where the brownian motion is specific to the agent considered.

Production choices will be made in order to optimize a profit criterion (the same for all agents) of the following form:

$$Max_{(q(t))_t}\mathbb{E}\int_0^\infty (p(t)q(t) - C(q(t)))e^{-rt}ds \quad s.t.\ q(t) \geq 0, R(t) \geq 0$$

where:

- C is the cost function which we will then write as quadratic: $C(q) = \alpha q + \beta\frac{q^2}{2}$.
- the prices p are determined according to the supply/demand equilibrium on the market at each moment, demand being given by a function $D(t, p)$ at instant t (that could be written $D(t, p) = We^{\rho t}p^{-\sigma}$ where $We^{\rho t}$ denotes the total wealth affected by a constant growth rate to model economic growth and where σ is the elasticity of demand that can be interpreted in a more general model as the elasticity of substitution between oil and any other good) and supply is naturally given by the total oil production of the agents.

Our model can be dealt with in the deterministic case or in the stochastic case depending on the value of ν.

We are going to start with the deterministic case where $\nu = 0$. In that case, a solution can be found without mean field methods. The mean field methods will be necessary in the stochastic case and the economic equilibrium will appear as a very special case of the PDE system, leaving an empty room to add externality effects and for more complex specifications.

3.2 The Deterministic Case

3.2.1 Characterization of the Equilibrium

Proposition 5 (Equilibrium in the deterministic case). *The equilibrium is characterized by the following equations where p, q and λ are unknown functions and R_0 the level of initial oil reserve.*

$$D(s, p(s)) = \int q(s, R_0)m_0(R_0)dR_0$$

$$q(s, R_0) = \frac{1}{\beta}\left[p(s) - \alpha - \lambda(R_0)e^{rs}\right]_+$$

$$\int_0^\infty q(s, R_0)ds = R_0$$

Proof. Let's consider the problem of an oil producer with an oil reserve equal to R_0.

The optimal production levels can be found using a Lagrangian:

$$\mathcal{L} = \int_0^\infty (p(s)q(s) - C(q(s)))e^{-rs}ds + \lambda\left(R_0 - \int_0^\infty q(s)ds\right)$$

The first order condition is:

$$p(s) = C'(q(s)) + \lambda e^{rs}$$

where λe^{rs} is the Hotelling rent. Noteworthy, if one considered a monopoly, the price would not be "marginal cost + rent" but "marginal cost + rent" multiplied by the usual markup. In other words, the actual rent is increasing with the market power.

Now, using our specification for the costs, we get, as long as $q(s)$ is positive:

$$p(s) - \alpha - \beta q(s) = \lambda e^{rs}$$

Hence, $q(s)$ is given by:

$$q(s) = \frac{1}{\beta}\left[p(s) - \alpha - \lambda e^{rs}\right]_+$$

In this equation λ depends on the initial oil stock (or reserve) and it will be denoted $\lambda(R_0)$. This lagrangian multiplier is given by the intertemporal constraint that equalizes the whole stream of production and the initial oil reserve:

$$\int_0^\infty q(s, R_0)ds = \frac{1}{\beta}\int_0^\infty \left(p(s) - \alpha - \lambda(R_0)e^{rs}\right)_+ ds = R_0$$

Now, we need to find the prices that were left unknown. This simply is given by the demand/supply equality.

$$D(s, p(s)) = \int q(s, R_0)m_0(R_0)dR_0$$

If we compile all these results we get the three equations that characterize the equilibrium. □

3.2.2 Computation of an Equilibrium

Since q only depends on $\lambda(\cdot)$ and $p(\cdot)$ we can totally separate the variables t and R_0. More precisely, if we consider an eductive algorithm (eductive algorithms will be used later to solve coupled PDEs) we can consider two "guesses" $\lambda(\cdot)$ and $p(\cdot)$ to compute $q(\cdot, \cdot)$ and then update $\lambda(\cdot)$ and $p(\cdot)$ using respectively the constraints $\int_0^\infty q(s, R_0)ds = R_0$ and $D(s, p(s)) = \int q(s, R_0)m_0(R_0)dR_0$.

More precisely, we consider a dynamical system indexed by the variable θ like the following[7]

$$\partial_\theta p(t,\theta) = D(t,p(t,\theta)) - \int q(t,R_0) m_0(R_0) dR_0$$
$$\partial_\theta \lambda(R_0,\theta) = \int_0^\infty q(t,R_0)dt - R_0$$

where

$$q(t,R_0) = \frac{1}{\beta}\left[p(t,\theta) - \alpha - \lambda(R_0,\theta)e^{rt}\right]_+$$

Once a dynamical system is chosen, the solution for $R_0 \mapsto \lambda(R_0)$ and $t \mapsto p(t)$ and hence the productions of all oil producers is obtained by:

$$\lim_{\theta \to +\infty} p(t,\theta) = p(t)$$
$$\lim_{\theta \to +\infty} \lambda(R_0,\theta) = \lambda(R_0)$$

As an example we can illustrate the evolution of total oil production in this model where we consider a CES demand function, namely $D(t,p) = We^{\rho t}p^{-\sigma}$.

We took the following values for the parameters: the interest rate considered by oil producers is $r = 5\%$, the average growth rate of the world economy is $\rho = 2\%$, the initial marginal cost of producing an oil barrel is $\alpha = 10$, $\beta = 100$ to model the importance of capacity constraints, $\sigma = 1.2$ because oil is not a highly elastic good and $W = 40$ to obtain meaningful values in the model. The problem is considered over 150 years and the initial distribution of reserves has the following form (Fig. 1):

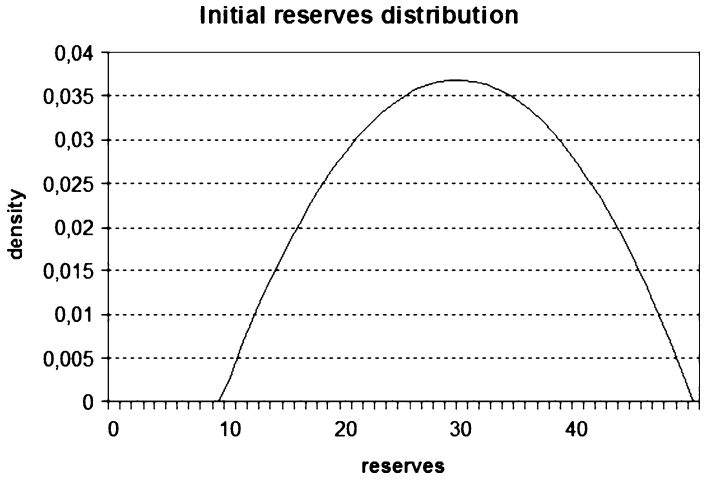


Fig. 1 m_0

[7] The system can be multiplied by the inverse of its Jacobian matrix. Different multiplying factors can also be added to the two equations.

If we consider the global production of oil producers, its evolution is given by the first graph below (Fig. 2) where the horizontal axis represents the years and the vertical one the global production at each date. The associated evolution of oil prices is also represented where we only plot the first 50 years to avoid ending up with very large values after too many decades and hence a graph that is unreadable (Fig. 3).

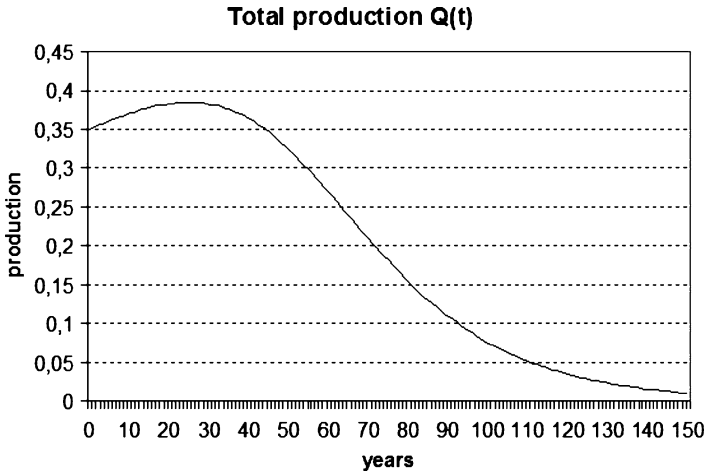


Fig. 2 Evolution of the total oil production

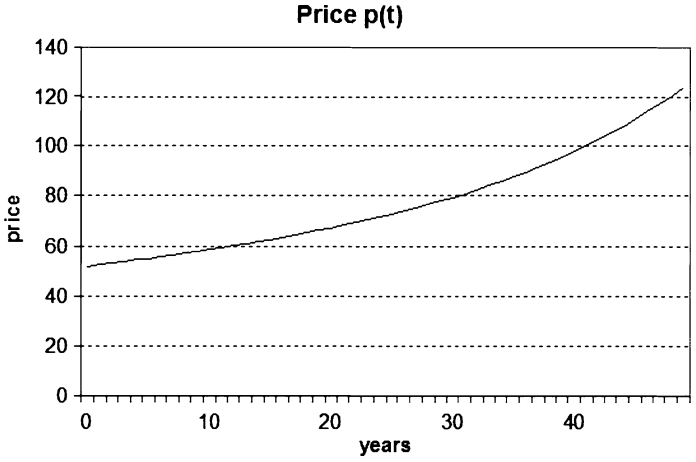


Fig. 3 Evolution of prices over 50 years

3.2.3 Comments on the Deterministic Model

This deterministic model will appear as classical to most readers. Though, some comments deserve to be made. First of all, we saw that prices were not equal to marginal cost since the optimal production (when positive) is given by:

$$p(t) = C'(q(t, R_0)) + \lambda(R_0)e^{rt}$$

Hence, the Hotelling rent ($\lambda(R_0)e^{rt}$) increases with time and differs among producers. Since λ measures the strength of the constraint associated to the exhaustible nature of oil, it is a decreasing function of R_0. As a consequence, the rent is higher when it comes to consider a smaller producer.

Another remarkable phenomenon is the shape of the curve. Oil production first increases and then decreases. It's a form of the so-called Hubbert peak even though we do not have the symmetry result associated to the usual Hubbert peak.[8]
Economic growth pushes oil producers to produce more (and especially producers with a large oil reserve) but the intrinsic exhaustibility of oil induces a decrease in the production after a certain period of time.

3.3 The Stochastic Case

The above model was a mean field game as any general equilibrium economic model. In the simple deterministic case developed above, the mean field games tools didn't need to be used and classical tools were sufficient, except perhaps when it came to find a numerical solution. However, when it comes to noise or externality in the model, the mean field games partial differential equations will be necessary. In contrast with the PDEs developed for the first toy model, the PDEs will now be completely coupled and not only coupled through boundary conditions.

3.3.1 The Mean Field Games PDEs

To start writing the equations, let's introduce $u(t, R)$ the Bellman function of the problem, namely:

$$u(t, R) = Max_{(q(s))_{s\geq t}, q\geq 0} \mathbb{E} \int_t^\infty (p(s)q(s) - C(q(s)))e^{-r(s-t)}ds$$

$$s.t.\ dR(s) = -q(s)ds + \nu R(s)dW_s, R(t) = R$$

[8] Our model not being suited for it since we do not focus on the discovery and exploitation of new wells.

The Hamilton–Jacobi–Bellman equation associated to this optimal control problem is:

$$(HJB) \qquad \partial_t u(t,R) + \frac{\nu^2}{2} R^2 \partial^2_{RR} u(t,R) - ru(t,R) + \max_{q \geq 0} (p(t)q - C(q) - q\partial_R u(t,R)) = 0$$

Now, let's denote $m(t,R)$ the distribution of oil reserves at time t. This distribution is transported by the optimal production decisions of the agents $q^*(t,R)$ where, now, R is the reserve at time t and not the initial reserve as in the deterministic case. The transport equation is:

$$(Kolmogorov) \qquad \partial_t m(t,R) + \partial_R(-q^*(t,R)m(t,R)) = \frac{\nu^2}{2} \partial^2_{RR} \left[R^2 m(t,R)\right]$$

with $m(0,\cdot)$ given.

Now, let's discuss the interdependence between u and m.
m is linked to u quite naturally since m is transported by the optimal decisions of the agents determined by the optimal control in the HJB equation. This optimal control is given by[9]:

$$q^*(t,R) = \left[\frac{p(t) - \alpha - \partial_R u(t,R)}{\beta}\right]_+$$

Now, u depends on m through the price $p(t)$ and this price can be seen as a function of m. Indeed, because $p(t)$ is fixed so that supply and demand are equal, $p(t)$ is given by:

$$p(t) = D(t,\cdot)^{-1} \left(-\frac{d}{dt} \int R m(t,R) dR\right)$$

If we want to conclude on this part and rewrite the equations to focus on the interdependence, we may write the following expressions:

$$\partial_t u(t,R) + \frac{\nu^2}{2} R^2 \partial^2_{RR} u(t,R) - ru(t,R) + \frac{1}{2\beta} \left[\left(D(t,\cdot)^{-1} \left(-\frac{d}{dt} \int R m(t,R) dR\right) - \alpha - \partial_R u(t,R)\right)_+\right]^2 = 0$$

[9] Once again we suppose that the solution of the (HJB) equation we consider is a solution of the underlying optimization problem. Verification results still need to be proved.

$$\partial_t m(t,R)+\partial_R\left(-\left[\frac{D(t,\cdot)^{-1}\left(-\frac{d}{dt}\int Rm(t,R)dR\right)-\alpha-\partial_R u(t,R)}{\beta}\right]_+ m(t,R)\right)$$
$$=\frac{\nu^2}{2}\partial^2_{RR}\left(R^2 m(t,R)\right)$$

These equations are the coupled equations associated to our optimization problem but it is still an interesting problem to establish a verification theorem that would prove a solution of the Hamilton–Jacobi–Bellman equation to be a solution of the optimization problem.

3.3.2 Generalization and Externality

The equations, as stated above to focus on the interdependence, are less practical and intuitive than the preceding forms of the equation. Though, they express something really important we want to insist upon: general equilibrium in its classical form can appear as a very special case of a mean field game. A natural consequence is that we can add other effects in a very simple manner while adding meaningful terms to the PDEs.

For example, it's widely thought amongst oil specialists that oil producers not only want to maximize profit but also want to avoid being in the last producers to produce because they do not know what is going to happen to oil companies at the end of the oil era.

This kind of effect would have been very hard to introduce with the first (usual) approach we presented. With the mean field games approach, the addition of such an effect is just another dependence on m in the HJB equation that defines u. One possibility is for example to introduce a ranking effect in the Hamilton–Jacobi–Bellman equation. The Hamilton–Jacobi–Belmman equation may become:

$$\partial_t u(t,R)+\frac{\nu^2}{2}R^2\partial^2_{RR}u(t,R)-ru(t,R)+H\left(\int_0^R m(t,\varphi)d\varphi\right)$$
$$+\frac{1}{2\beta}\left[\left(D(t,\cdot)^{-1}\left(-\frac{d}{dt}\int Rm(t,R)dR\right)-\alpha-\partial_R u(t,R)\right)_+\right]^2=0$$

where H is a decreasing function. In addition to the intertemporal profit optimization, the producer wants to have less oil reserve than its competitors.

This generalization is one amongst many. We just aim at convincing the reader about the variety of effects and particularly externality effects the mean field games approach allows to handle quite easily. To see how this mean field game can be adapted to the study of competition between oil producers and potential entrants that produce alternative energy, see [25] and [20].

4 The Mexican Wave

4.1 Introduction

Before moving on to more complex models, let us look at a "toy model" which is prompted by mean field games and models the mimicry responsible for the Mexican wave phenomenon in stadiums.

Mexican wave is called this way because it seems that it appeared for the first time in a stadium in Mexico. The goal of our model here is to understand how a Mexican wave can be one of the solution of a mean field game involving a (infinite) set of supporters and a taste for mimicry. Let's start with the description of the stadium. To simplify the study, we regard our stadium as a circle of length L (hence the stadium is a one-dimension object, though it wouldn't be difficult to generalize). Mathematically, the stadium will then be the interval $[0, L)$ regarded as a torus.

In the stadium, there is a continuum of individuals; each one being referenced by a coordinate $x \in [0, L)$. Each agent is free to behave and can be either seated ($z = 0$) or standing ($z = 1$) or in an intermediate position $z \in (0, 1)$. Some positions are less comfortable than others and we model this phenomenon using a utility function u.

Typically u will be of the following form to express that being standing or being seated is more comfortable than being in an intermediate position:

$$u(z) = -Kz^{\alpha}(1-z)^{\beta}$$

Now, let's describe the optimization function for any agent:

- An agent pays a price $h(a)dt$ to change his position from z to $z + adt$. $h(a)$ will simply be a quadratic cost function: $\frac{a^2}{2}$.
- An agent wants to behave as his neighbors. Mathematically an agent in x maximizes

$$-\frac{1}{\epsilon^2}\int (z(t,x) - z(t,x-y))^2 \frac{1}{\epsilon} g\left(\frac{y}{\epsilon}\right) dy$$

  where g is a gaussian kernel.
- An agent maximizes his comfort described by u.

The optimization criterion for an agent localized at x is then

$$\sup_{z(\cdot,x)} \liminf_{T\to+\infty} \frac{1}{T}\int_0^T \left\{ \left[-\frac{1}{\epsilon^2}\int (z(t,x) - z(t,x-y))^2 \frac{1}{\epsilon} g\left(\frac{y}{\epsilon}\right) dy \right] \right.$$
$$\left. + u(z(t,x)) - \frac{\dot{z}(t,x)^2}{2} \right\} dt$$

This ergodic control problem can be formally transformed in a differential way and we get:

$$-\frac{2}{\epsilon^2}\int (z(t,x)-z(t,x-y))\frac{1}{\epsilon}g\left(\frac{y}{\epsilon}\right)dy+u'(z(t,x))=-\partial^2_{tt}z(t,x)$$

If we let ϵ tends to 0, we get in the distribution sense that our problem is to solve the equation[10]:

$$\partial^2_{tt}z(t,x)+\partial^2_{xx}z(t,x)=-u'(z(t,x))$$

Before going on and solve the problem, we must notice that $z=0$ and $z=1$ should be solutions of the problem. Consequently, we must have $u'(0)=u'(1)=0$ and hence α and β have to be strictly greater than 1.

4.2 Mexican Wave as a Solution

A Mexican wave is, by definition, a wave. Hence we are going to look for a solution of the form $z(t,x)=\varphi(x-vt)$ where v is the speed of the wave. But what we call Mexican wave is usually a specific form of wave and we want to call Mexican wave a function φ with a compact support on $(0,L)$ that is first increasing from 0 to 1 and then decreasing form 1 to 0.

If we look for such a function φ, we can easily see that it must solve:

$$(1+v^2)\varphi''=-u'(\varphi)$$

[10] This equation doesn't seem to be of the mean field type but we can write the associated mean field equations.
Let's consider that agents are indexed by x. For each x, the Bellman function associated to the problem of an agent in x can be written as $J(x;\cdot)$ solving the Hamilton–Jacobi equation:

$$0=\partial_t J(x;t,z)+\frac{1}{2}\left(\partial_z J(x;t,z)\right)^2+u(z)-\frac{1}{\epsilon^2}\int (z-\tilde{z})^2 m(\tilde{x};t,\tilde{z})\frac{1}{\epsilon}g\left(\frac{x-\tilde{x}}{\epsilon}\right)d\tilde{z}d\tilde{x}$$

where $m(x;t,\cdot)$ is the probability distribution function of the position z of an agent situated in x. $m(x;\cdot,\cdot)$ solves a Kolmogorov equation that is:

$$\partial_t m(x;t,z)+div(\partial_z J(x;t,z)m(x;t,z))=0$$

with $m(x;0,z)=\delta_{z(0,x)}(z)$ Hence, the problem can be written as a set of Hamilton–Jacobi equations indexed by x with the associated Kolmogorov equations. Because the problem is purely deterministic, we can directly follow the position of each individual and consider an equation in $z(t,x)$ instead of this complex system.

Proposition 6 (Existence of Mexican waves for $\alpha, \beta \in (1;2)$). *Suppose that $\alpha, \beta \in (1;2)$. Then, for any v verifying*

$$\frac{\Gamma(1-\frac{\alpha}{2})\Gamma(1-\frac{\beta}{2})}{\Gamma(2-\frac{\alpha+\beta}{2})} < \sqrt{\frac{K}{2(1+v^2)}}L$$

there exists a Mexican wave φ solution of $(1+v^2)\varphi'' = -u'(\varphi)$.

Proof. We use an "energy method" to solve the equation $(1+v^2)\varphi'' = -u'(\varphi)$. First, let's multiply the equation by φ' and integrate. We get:

$$\frac{1+v^2}{2}\varphi'^2 = -u(\varphi) + C$$

Since $\varphi = 0$ must be a solution, the constant has to be zero. Consequently, we are left with an ODE:

$$\varphi' = \pm\sqrt{\frac{2K}{1+v^2}}\varphi^{\alpha/2}(1-\varphi)^{\beta/2}$$

If α were greater than 2 the Cauchy–Lipschitz theorem would apply using the boundary condition $\varphi(0) = 0$ or $\varphi(L) = 0$ and the unique solution would be $z = \varphi = 0$. Now because we supposed $\alpha \in (1;2)$, we can have a local non-uniqueness result.

Let's build a solution different from 0. First we can consider that φ is equal to zero in a neighborhood of 0 e.g. $\forall s \in [0, \eta]$. Now for $s > \eta$, we can integrate the ODE and define $\varphi(s)$ implicitly by:

$$\int_0^{\varphi(s)} w^{-\alpha/2}(1-w)^{-\beta/2}dw = \sqrt{\frac{2K}{1+v^2}}(s-\eta)$$

This definition holds as long as $\varphi(s) \leq 1$, i.e. as long as $s \leq M$ where M is defined as $B(1-\alpha/2, 1-\beta/2) = \sqrt{\frac{2K}{1+v^2}}(M-\eta)$ (B stands for the beta function). Now, for $s > M$, we build the solution in the same way and we can do so because $\beta \in (1;2)$. We define implicitly $\varphi(s)$ by:

$$\int_{\varphi(s)}^1 w^{-\alpha/2}(1-w)^{-\beta/2}dw = \sqrt{\frac{2K}{1+v^2}}(s-M)$$

as long as $\varphi(s)$ remains positive. This happens in $s = M'$ where $B(1-\alpha/2, 1-\beta/2) = \sqrt{\frac{2K}{1+v^2}}(M'-M)$. Now, φ is supposed to be 0 for $s \geq M'$.

We have built a differentiable function φ but we need to check that M' can be smaller than L for a sufficiently small η.

We have $2B(1-\alpha/2, 1-\beta/2) = \sqrt{\frac{2K}{1+v^2}}(M'-\eta)$. Hence M' can be smaller than L if and only if there exists η such that

$$\eta < L - \sqrt{\frac{2(1+v^2)}{K}} B(1-\alpha/2, 1-\beta/2)$$

Such a positive η exists if and only if $\sqrt{\frac{K}{2(1+v^2)}} L > B(1-\alpha/2, 1-\beta/2)$ and this is equivalent to our condition thanks to the link between the functions Γ and B. □

We can represent a solution φ as described above (supporters do not keep standing before going down to the seated position) (Fig. 4):

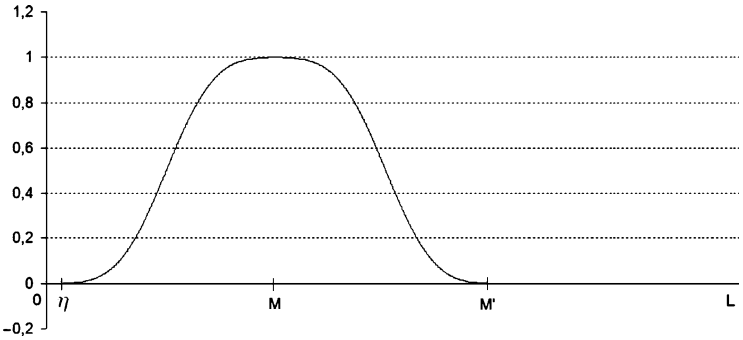


Fig. 4 $\alpha = \beta = 1.5$

Remark: This solution is not unique in general for two reasons. The first one is obvious: η in the preceding proof can be chosen in an interval. However, this non-uniqueness is only due to a translation invariance of the problem on the torus and is therefore meaningless. A second reason is that supporters may stand for a while before going back to the seated position.

4.3 Mean Field Games Versus Descriptive Models

All the models developed to represent the Mexican wave assume that the supporters behave like automata: they carry out actions according to the context with possibly some memory of the preceding moments. This logic of automatic functioning is perfectly adequate for producing a Mexican wave-type crowd movement, and even for producing an equation of the dynamics that is the same as what we have written: if it has not already been done, we would be able to do so.

The difference between our model and a model based on automata agents (see for instance [19]), lies in the meaning given to the agents' actions. While automata produce actions dictated by the context, our agents produce the same actions

as a consequence of a process of thinking about the behavior of other agents, the coherence of these behaviors, and the personal preferences in view of these behaviors. That this gives the same result as if agents were automata should not be cause for disappointment: the parsimony principle does not apply here; simply because agents behave as if they were automata is no reason for not giving another meaning to their actions. And not only for ethical reasons. Indeed, if one wishes to study the stability of the Mexican wave, and behaviors apart from equilibrium, it becomes necessary to return to the mechanism that has enabled the equation to be constructed. And hence, if the Mexican wave has been disturbed, if some rows of spectators have not moved for an exogenous reason, models based on automata generally predict erratic behavior in situations in which our mean field games agents, after a moment's thought, behave in such a way that the collective motion of the Mexican wave is re-established.

Thus the meaning given to behaviors sheds light on what one would expect in the event of disturbance to various processes.

5 A Model of Population Distribution

5.1 Introduction

Let's now consider a model of population distribution. This model is the archetype of a mean field games model in continuous time with a continuous state space. Many models can be derived from this one and most importantly the notion of stability introduced in what follows is arguably one of the most relevant one to deal with stability in forward/backward models such as intertemporal mean field games models. This stability notion called eductive stability turns out to be useful to circumvent the issues of the forward/backward structure, especially when it comes to find numerical solutions to mean field games PDEs.

In what follows we only present some aspects of the model. The readers may refer to [23] to go deeper into the different notions of stability or they may read [22] for an even more complete presentation with generalization to multi-population issues.

5.2 The Framework

We consider a large number of agents modeled as usual by a continuum. These agents have geographic, economic and social characteristics that we assume are represented by a finite number n of values. A simple example is the position of an agent represented by his coordinates in space. Another example is that of a technology used by an agent. In short, we assume that the agents have characteristics denoted by $X \in \mathbb{R}^n$.

Each individual will have control over his characteristics, and we choose the case in which agents wish to resemble their peers. To resemble others, an agent has to move in the state space $\mathbb{R}^n$. When an agent wants to make move of size α in the characteristics space (hereafter social space or state space) he will pay a cost of the quadratic form $\frac{|\alpha|^2}{2}$. Moreover, this control is not perfect, since we add Brownian noise. In mathematical terms, our problem is thus written (for an agent i):

$$\sup_{(\alpha_s)_{s\geq 0}, X_0^i = x} \mathbb{E}\left[\int_0^\infty \left(g(t, X_t^i, m_t) - \frac{|\alpha(s, X_s^i)|^2}{2}\right) e^{-\rho s} ds\right]$$

$$dX_t^i = \alpha(t, X_t^i)dt + \sigma dW_t^i$$

where m is the distribution of agents in the social space and where the function g will model the will to resemblance depending on the type of problem. Various specifications for g will produce our results:

$$g(t, x, m) = -\beta(x - \int y m(t, y) dy)^2$$

$$g(t, x, m) = -\beta \int (x - y)^2 m(t, y) dy$$

$$g(t, x, m) = \ln(m(t, x))$$

To simplify the exposition and to stick to the papers cited in the introduction, we consider the logarithmic case where g is a local function of m.

The control problem is re-written in differential way. We obtain the PDEs which are at the heart of mean field game theory:

$$(HJB) \qquad \partial_t u + \frac{\sigma^2}{2}\Delta u + \frac{1}{2}|\nabla u|^2 - \rho u = -\ln(m)$$

$$(Kolmogorov) \qquad \partial_t m + \nabla \cdot (m \nabla u) = \frac{\sigma^2}{2}\Delta m$$

and in this case the optimal control is written $\alpha(t, X_t) = \nabla u(t, X_t)$.

What is fundamental in this PDE system is the forward/backward dimension. The Hamilton–Jacobi–Bellman equation is backward like all Bellman equations (in finite time, there would be a final condition of type $u(T, x) = u_T(x)$) - this can also be seen in the diffusion part of the equation, which is a backward heat equation. Conversely, the transport equation is forward and transport an initial distribution $m(0, x)$ according to agents' optimal choices. We clearly see the emergence of agents' reasoning in this forward/backward aspect. They assume a dynamic for m and optimize as a result to get the optimal control ∇u. The behavior we obtained transports the distribution of agents. The coherence is finally found if we assume that the expectations are rational (and hence perfect) on the distribution m. This is the usual reasoning in mean field games.

5.3 Stationary Solutions

We are interested firstly in stationary solutions. The framework of quadratic costs and logarithmic utility allows us to have explicit quadratic solutions for u and Gaussian solutions for m, as in the following result:

Proposition 7 (Gaussian solutions). *Suppose that* $\rho < \frac{2}{\sigma^2}$.

There exist three constants, $s^2 > 0$, $\eta > 0$ *and* ω *such that* $\forall \mu \in \mathbb{R}^n$, *if* m *is the probability distribution function associated to a gaussian variable* $\mathcal{N}(\mu, s^2 I_n)$ *and* $u(x) = -\eta|x-\mu|^2 + \omega$, *then* (u, m) *is a solution of our problem.*

These three constants are given by:

- $s^2 = \frac{\sigma^4}{4-2\rho\sigma^2}$
- $\eta = \frac{1}{\sigma^2} - \frac{\rho}{2} = \frac{\sigma^2}{4s^2}$
- $\omega = -\frac{1}{\rho}\left[\eta n\sigma^2 - \frac{n}{2}\ln\left(\frac{2\eta}{\pi\sigma^2}\right)\right]$

Proof. First, let's note that the stationary equation for m (the Kolmogorov equation) can be rewritten as:

$$\nabla \cdot (m\nabla u - \frac{\sigma^2}{2}\nabla m) = 0$$

Hence, we can restrict ourselves to solutions of:

$$m\nabla u = \frac{\sigma^2}{2}\nabla m$$

Consequently, we just need to solve the Hamilton–Jacobi–Bellman equation if we replace m by $K\exp(\frac{2}{\sigma^2}u)$ where K is chosen to ensure that m is indeed a probability distribution function.

We are looking for a solution for u of the form:

$$u(x) = -\eta|x-\mu|^2 + \omega$$

If we put this form in the Hamilton–Jacobi–Bellman equation we get:

$$2\eta^2|x-\mu|^2 + \rho\eta|x-\mu|^2 - \rho\omega - \eta n\sigma^2 = -\ln(K) + \frac{2\eta|x-\mu|^2}{\sigma^2} - \frac{2\omega}{\sigma^2}$$

A first condition for this to be true is:

$$2\eta^2 + \rho\eta = \frac{2\eta}{\sigma^2}$$

$$\iff \eta = \frac{1}{\sigma^2} - \frac{\rho}{2}$$

A second condition, to find ω, is related to the fact that m is a probability distribution function. This clearly requires η to be positive but this is guaranteed by the hypothesis $\rho\sigma^2 < 2$. This also implies:

$$K \exp\left(\frac{2\omega}{\sigma^2}\right) \int_{\mathbb{R}^n} \exp\left(\frac{-2\eta}{\sigma^2}|x-\mu|^2\right) = K \exp\left(\frac{2\omega}{\sigma^2}\right) \left(\frac{\pi\sigma^2}{2\eta}\right)^{\frac{n}{2}} = 1$$

$$\Rightarrow \rho\omega + \eta n\sigma^2 = \frac{n}{2} \ln\left(\frac{2\eta}{\pi\sigma^2}\right)$$

and this last equation gives ω.

From this solution for u we can find a solution for m. We indeed know that m is a probability distribution function and that m is given by

$$m(x) = K \exp(\frac{2u(x)}{\sigma^2})$$

As a consequence, m is the probability distribution function of an n-dimensional gaussian random variable with variance equal to $s^2 I_n$ where $s^2 = \frac{\sigma^2}{4\eta}$ i.e. $s^2 = \frac{\sigma^4}{4-2\rho\sigma^2}$. □

A priori, nothing guaranteed that a solution exists insofar as the cases usually well treated (see [36]) correspond most often to a decreasing function g and not, as here, to an increasing one. On the other hand nothing shows there is a uniqueness result. First, there is invariance by translation and we must therefore localize the problem in order to address this question. This localization is done simply by replacing $\ln(m(t,x))$ by $\ln(m(t,x)) - \delta x^2$ ($\delta > 0$) and we obtain the same type of results. Even when localized, we do not guarantee uniqueness (though, localization will be helpful for other purposes). Although we do not prove uniqueness, we are nevertheless interested in the problem of the stability of solutions. Since we have a stationary result, a first step towards studying the dynamics is to study stability.

5.4 Stability Results

5.4.1 Two Notions of Stability

Two types of stability are relevant to our problem. We will call the first one "physical stability" and the second one "eductive stability". The physical stability concept might look more standard to the reader. The second, the eductive stability, refers to many papers by Roger Guesnerie and other authors (e.g. [26]) on stability in a rational expectation economic context. These papers inspired the mean field games eductive stability concept.

If we consider a problem on $[0,T]$ with conditions stationary solutions on each side ($u(T,\cdot) = u^*$ and $m(0,\cdot) = m^*$ given), we can look at what happens (as $T \to \infty$) when we disturb the boundary conditions (u^*, m^*). The stability associated with this perturbation in 0 for m and in T for u is the physical stability and we refer to [23] for a complete study.

A second possibility is to add a variable θ (virtual time) and to consider a different, purely forward, dynamic system, whose stationary equilibrium is the same as the one we are looking for. If there is convergence (when $\theta \to \infty$) in this new dynamic system where we reverse the time in the backward equation by imposing arbitrary conditions in $\theta = 0$, then we shall call this eductive stability.

In what follows, we focus on eductive stability, and more precisely on local eductive stability, because it helps a lot to develop and justify numerical methods.

5.4.2 Eductive Stability

The physical stability, briefly described above, is intrinsically linked to the forward/backward structure of the equations.

Here, we want to circumvent this forward/backward structure and we introduce a virtual time θ that will be purely forward in the sense that we consider the following new system of PDEs[11]:

$$\partial_\theta u = \frac{\sigma^2}{2}u'' + \frac{1}{2}u'^2 - \rho u + \ln(m)$$
$$\partial_\theta m = \frac{\sigma^2}{2}m'' - (mu')'$$

Let's consider two "initial guesses" ($u(\theta = 0, x)$ and $m(\theta = 0, x)$) that are not too far from the stationary equilibrium (u^*, m^*) associated to $\mu = 0$, as defined in Proposition 7:

$$m(0,x) = m^*(x)(1 + \varepsilon\psi(0,x))$$
$$u(0,x) = u^*(x) + \varepsilon\varphi(0,x)$$

We are going to linearize these equations. After easy computations we obtain the following linear PDEs:

$$\partial_\theta \varphi = \frac{\sigma^2}{2}\varphi'' - 2\eta x\varphi' - \rho\varphi + \psi$$
$$\partial_\theta \psi = \frac{\sigma^2}{2}\psi'' + 2\eta x\psi' - \varphi'' + \frac{x}{s^2}\varphi'$$

[11] We consider the problem in dimension 1 for the sake of simplicity but the problem in general is the same.

A more convenient way to write these linearized PDEs is to introduce the operator $\mathcal{L}$: $f \mapsto \mathcal{L}f = -\frac{\sigma^2}{2}f'' + 2\eta x f'$ and we get the following equations for the couple (φ, ψ):

Proposition 8.

$$\partial_\theta \varphi = -\mathcal{L}\varphi - \rho\varphi + \psi$$
$$\partial_\theta \psi = -\mathcal{L}\psi + \frac{2}{\sigma^2}\mathcal{L}\varphi$$

Proof. It simply is a consequence of the link between the variables, namely $s^2 = \frac{\sigma^2}{4\eta}$. □

Now, we are going to use the properties of the operator $\mathcal{L}$ we have just introduced. To do that we need to use some properties of the Hermite polynomials associated to the space $L^2(m^*(x)dx)$ (see [1] for more details).

Proposition 9 (Hermite polynomials). *We define the n^{th} Hermite polynomial of $L^2(m^*(x)dx)$ by:*

$$H_n(x) = s^n \frac{1}{\sqrt{n!}}(-1)^n \exp\left(\frac{x^2}{2s^2}\right)\frac{d^n}{dx^n}\exp\left(-\frac{x^2}{2s^2}\right)$$

The polynomials $(H_n)_n$ form an orthonormal basis of the Hilbert space $L^2(m^*(x)dx)$.

The Hermite polynomials H_n are eigenvectors of $\mathcal{L}$ and:

$$\mathcal{L}H_n = 2\eta n H_n$$

To study the linearized equations, we are going to consider the space $L^2(m^*(x)dx)$ and consider a decomposition on the Hermite polynomials basis. Because the problem is purely forward in θ we need to have, for each coordinate, two negative eigenvalues.

To this purpose, let's introduce the matrices $(B_n)_n$:

$$B_n = \begin{pmatrix} -(\rho + 2\eta n) & 1 \\ \frac{n}{s^2} & -2\eta n \end{pmatrix}$$

Lemma 10 (Eigenvalues of B_n). *Let's consider $n \geq 2$.*

The eigenvalues $\xi_n^1 < \xi_n^2$ of B_n are both negative with:

$$\xi_n^{1,2} = \frac{1}{2}\left[-\rho - 4\eta n \pm \sqrt{\rho^2 + \frac{4n}{s^2}}\right]$$

Proposition 10. *Let's suppose that the initial conditions $\varphi(0,\cdot)$ and $\psi(0,\cdot)$ are in the Hilbert space $H = L^2(m^*(x)dx)$.*

Let's consider for $n \geq 2$ the functions $\begin{pmatrix}\varphi_n\\ \psi_n\end{pmatrix}$ that verify:

$$\begin{pmatrix}\partial_\theta\varphi_n\\ \partial_\theta\psi_n\end{pmatrix} = B_n \begin{pmatrix}\varphi_n\\ \psi_n\end{pmatrix}$$

with $\varphi_n(0)$ equal to $\varphi(0,\cdot)_n =< H_n, \varphi_n(0) >$ and $\psi_n(0)$ equal to $\psi(0,\cdot)_n =< H_n, \psi_n(0) >$.

We have for a fixed θ and as n tends to ∞:

$$\varphi_n(\theta) = \mathcal{O}(|\varphi_n(0)|e^{\xi_n^2\theta})$$
$$\psi_n(\theta) = \mathcal{O}(\sqrt{n}|\varphi_n(0)|e^{\xi_n^2\theta})$$

In particular,

$$\forall\theta > 0, \forall k \in \mathbb{N}, (n^k\varphi_n(\theta))_n \in l^1(\subset l^2), (n^k\psi_n(\theta))_n \in l^1(\subset l^2)$$

Proof. After straightforward algebraic manipulations, we get:

$$\begin{pmatrix}\varphi_n(\theta)\\ \psi_n(\theta)\end{pmatrix} = A_n e^{\xi_n^1\theta}\begin{pmatrix}1\\ a_n\end{pmatrix} + B_n e^{\xi_n^2\theta}\begin{pmatrix}1\\ b_n\end{pmatrix}$$

where:

$$a_n = \rho + 2\eta n + \xi_n^1, \qquad b_n = \rho + 2\eta n + \xi_n^2$$

Now, to find the two constants we need to use the conditions on $\varphi_n(0)$ and $\psi_n(0)$:

$$\begin{cases}\varphi_n(0) = A_n + B_n\\ \psi_n(0) = a_n A_n + b_n B_n\end{cases}$$

Hence:

$$\begin{cases}A_n = \frac{b_n\varphi_n(0) - \psi_n(0)}{b_n - a_n}\\ B_n = \frac{a_n\varphi_n(0) - \psi_n(0)}{a_n - b_n}\end{cases}$$

Using the fact that $a_n \sim -\frac{\sqrt{\eta}}{\sigma}\sqrt{n}$ and $b_n \sim \frac{\sqrt{\eta}}{\sigma}\sqrt{n}$ we can deduce the asymptotic behavior of the constants as n goes to infinity.

$$A_n \sim_{n\to\infty} \frac{\varphi_n(0)}{2}, \qquad B_n \sim_{n\to\infty} \frac{\varphi_n(0)}{2}$$

Hence, since $\xi_n^1 < \xi_n^2$,

$$\varphi_n(\theta) = \mathcal{O}(|\varphi_n(0)|e^{\xi_n^2\theta})$$

$$\psi_n(\theta) = \mathcal{O}(\sqrt{n}|\varphi_n(0)|e^{\xi_n^2\theta})$$

These two estimations prove the results. □

These estimations show that the solutions will be far more regular than the initial conditions.

Proposition 11 (Resolution of the linearized PDEs). *Suppose that:*

- *The initial conditions $\varphi(0,\cdot)$ and $\psi(0,\cdot)$ are in the Hilbert space $H = L^2(m^*(x)dx)$*
- *$\int \psi(0,x)m^*(x)dx = 0$ (this is guaranteed if the initial guess for m is a probability distribution function)*
- *$\int x\varphi(0,x)m^*(x)dx = 0$ (this is guaranteed if the initial guess is even)*
- *$\int x\psi(0,x)m^*(x)dx = 0$ (this is guaranteed if the initial guess is even)*

Let's define $(\varphi_n)_n$ and $(\psi_n)_n$ by:

- *$\varphi_0(\theta) = \varphi_0(0)e^{-\rho\theta}$ and $\psi_0(\theta) = 0$.*
- *$\varphi_1(\theta) = \psi_1(\theta) = 0$.*
- *$\forall n \geq 2$, φ_n and ψ_n defined as in the preceding proposition.*

Then $\varphi(\theta,x) = \sum_{n=0}^{\infty} \varphi_n(\theta)H_n(x)$ and $\psi(\theta,x) = \sum_{n=0}^{\infty} \psi_n(\theta)H_n(x)$ are well defined in H, are in C^∞, are solutions of the PDEs and verify the initial conditions.

Proof. First of all, the above proposition ensures that the two functions φ and ψ are well defined, in C^∞, and that we can differentiate formally the expressions. Then, the first three conditions can be translated as $\psi_0(0,\cdot) = 0$, $\varphi_1(0,\cdot) = 0$ and $\psi_1(0,\cdot) = 0$ and so the conditions at time 0 is verified.

The fact that the PDEs are verified is due to the definition of φ_n and ψ_n and also to the fact that we can differentiate under the sum sign because of the estimates of the preceding proposition. □

Proposition 12 (Local eductive stability). *Suppose that:*

- *The initial guesses $\varphi(0,\cdot)$ and $\psi(0,\cdot)$ are in the Hilbert space $H = L^2(m^*(x)dx)$.*
- *$\int \psi(0,x)m^*(x)dx = 0$ (this is guaranteed if the initial guess for m is a probability distribution function)*
- *$\int x\varphi(0,x)m^*(x)dx = 0$ (this is guaranteed if the initial guess is even)*
- *$\int x\psi(0,x)m^*(x)dx = 0$ (this is guaranteed if the initial guess is even)*

Then the solution (φ,ψ) of the PDEs converges in the sense that:

$$\lim_{\theta\to\infty} ||\varphi(\theta,\cdot)||_{L^2(m^*(x)dx)} = 0 \qquad \lim_{\theta\to\infty} ||\psi(\theta,\cdot)||_{L^2(m^*(x)dx)} = 0$$

Proof. We basically want to show that:

$$\sum_{n=0}^{+\infty} |\varphi_n(\theta)|^2 \to_{\theta\to+\infty} 0, \qquad \sum_{n=0}^{+\infty} |\psi_n(\theta)|^2 \to_{\theta\to+\infty} 0$$

This is actually a pure consequence of the estimates proved earlier and of the Lebesgue's dominated convergence theorem. □

These stability results are interesting but the symmetry conditions to obtain them may seem cumbersome. Indeed, when it comes to apply this kind of methodology to find stationary solutions, we clearly need a result that is less sensitive to initial conditions. A good way to proceed is to consider the case introduced at the beginning where there is no translation invariance, that is the localized case in which $\delta > 0$.

5.4.3 Eductive Stability in the Localized Case

In the proof of the eductive stability, there was a need to impose symmetry conditions on the initial guesses. These conditions were necessary to ensure stability because B_1 was singular. If one wants to have stability results for more general initial guesses, the intuitive idea is to break the translation invariance of the problem.

Interestingly, we introduced localization earlier. This localization idea can be used once again, to have more general stability results. If we center the problem around 0 as before, we can see that the only relevant difference between the original problem and the problem with an additional term $-\delta x^2$, that localizes the problem around 0, is the positive constant η that depends on δ according to the equation:

$$2\eta^2 - \eta\left(\frac{2}{\sigma^2} - \rho\right) = \delta$$

Now, in this context we can prove that the eigenvalues of B_n are both negative for $n \geq 1$ (remember that we needed n to be larger than 2 to have these properties in the case where $\delta = 0$).

This result can be used to prove general stability results when $\delta > 0$. It is indeed straightforward that all our stability results can be rewritten exactly the same if one replaces the conditions

$$\begin{cases} \int x\psi(0,x)m^*(x)dx = 0 \\ \int x\varphi(0,x)m^*(x)dx = 0 \end{cases}$$

by $\delta > 0$.

Thus, in this localized context, (and up to a linear approximation) if we start from a couple (u, m) close to a stationary equilibrium, there will be convergence toward stationary equilibrium as $\theta \to \infty$ when using the purely forward PDEs system. Numerically, this is very interesting and the eductive methods give very good results, both for finding stationary equilibrium and for generalizing the approach for seeking dynamic equilibrium (see [23]).

5.5 Numerical Methods

The forward/backward structure of mean field games is quite an issue when it comes to find numerical solutions. One can try to find a fixed point (u, m) solving alternatively the backward equation and the forward equation but there is a priori no guarantee that a solution can be found in this way. The eductive stability property proved earlier, however, can be adapted to design a numerical method. Other authors have developed several methods and the interested reader may for instance see [3].

We are going to present our methods to find stationary solutions. Interestingly, if one replaces the Laplace operator by heat operators (forward or backward, depending on the context), the numerical recipes presented below still work to find dynamical solutions (see [22, 23]).

5.5.1 Stationary Equilibrium

First, let's recall the two equations that characterize a stationary equilibrium:

$$(HJB) \qquad \frac{\sigma^2}{2}\Delta u + \frac{1}{2}|\nabla u|^2 - \rho u = -g(x, m)$$

$$(Kolmogorov) \qquad \nabla \cdot (m \nabla u) = \frac{\sigma^2}{2}\Delta m$$

where, now, g is not anymore supposed to be the logarithm function.

The Hamilton–Jacobi–Bellman equation can be simplified using the change of variable[12] $\beta = \exp\left(\frac{u}{\sigma^2}\right)$ and we obtain:

$$(HJB)' \qquad \frac{\sigma^2}{2}\Delta\beta = \beta\left[\rho \ln(\beta) - \frac{1}{\sigma^2} g(x, m)\right]$$

$$(Kolmogorov)' \qquad \nabla \cdot \left[\sigma^2 \left(m \frac{\nabla \beta}{\beta}\right)\right] = \frac{\sigma^2}{2}\Delta m$$

The two equations $(HJB)'$ and $(Kolmogorov)'$ can be written in a more practical way for numerical resolutions by "inverting" the Δ operators. This can be done in the Kolmogorov equation by restricting the Laplace operator to probability

[12] This change of variable, combined with the change of variable $\alpha = m \exp\left(-\frac{u}{\sigma^2}\right)$, simplifies the PDEs. Easy calculations give in fact that the equations in (u, m) become:

$$\partial_t \beta + \frac{\sigma^2}{2}\Delta\beta = \beta h(\alpha, \beta)$$

$$\partial_t \alpha - \frac{\sigma^2}{2}\Delta\alpha = -\alpha h(\alpha, \beta)$$

where h simply is $h(\alpha, \beta) = \rho \ln(\beta) - \frac{1}{\sigma^2} g(\alpha\beta)$.

distribution functions (since in practice we restrict ourselves to Fourier series with only a finite number of harmonics) and we obtain:

$$(Kolmogorov)' \qquad -m + \left[\frac{\sigma^2}{2}\Delta\right]^{-1}\left(\sigma^2\nabla\cdot\left(m\frac{\nabla\beta}{\beta}\right)\right) = 0$$

This cannot be done in the case of the Hamilton–Jacobi–Bellman equation but we can invert an operator like $\frac{\sigma^2}{2}\Delta - \epsilon Id$ for any $\epsilon > 0$. This gives:

$$(HJB)' \qquad -\beta + \left[\frac{\sigma^2}{2}\Delta - \epsilon Id\right]^{-1}\left(\beta\left[\rho\ln(\beta) - \frac{1}{\sigma^2}g(x,m) - \epsilon\right]\right) = 0$$

Using these equations we can consider the ideas of eductive stability and try to obtain solutions by solving the following equations where we introduce the virtual time θ:

$$\begin{cases} \partial_\theta m = -m + \left[\frac{\sigma^2}{2}\Delta\right]^{-1}\left(\sigma^2\nabla\cdot\left(m\frac{\nabla\beta}{\beta}\right)\right) \\ \partial_\theta\beta = -\beta + \left[\frac{\sigma^2}{2}\Delta - \epsilon Id\right]^{-1}\left(\beta\left[\rho\ln(\beta) - \frac{1}{\sigma^2}g(x,m) - \epsilon\right]\right) \end{cases}$$

Numerically these equations are quite easy to solve. An example is shown below where $g(x,m) = \sqrt{m} - \delta x^2$ with $\sigma^2 = 0.4$, $\rho = 0.4$, $\delta = 0.5$ on the domain $[-1,1]$ (we took $\epsilon = \frac{\rho}{3}$) (Fig. 5).

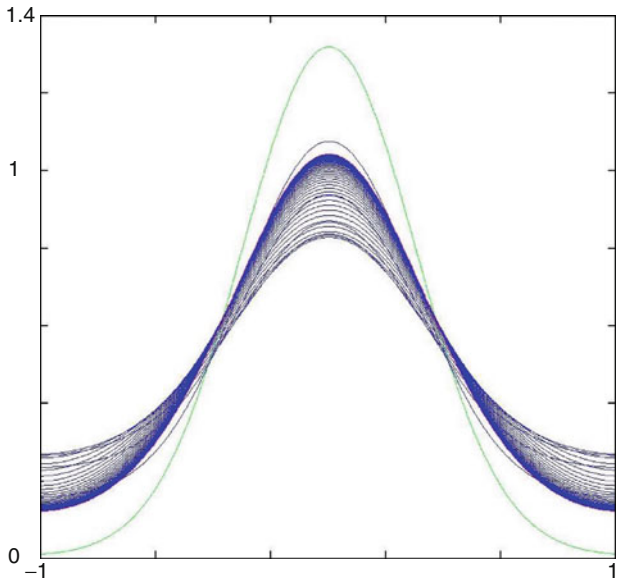


Fig. 5 Initial guess $\simeq \mathcal{N}(0, 0.3)$. Solution after 8,000 iterations with $d\theta \simeq 0.01$ (an iteration is drawn every 40 iterations). Only 15 harmonics are used.

We see that after a certain number of steps in θ, the distribution $m(\theta, \cdot)$ converges towards a limit distribution m that is a good candidate for being a stationary equilibrium.

5.5.2 Generalizations

This method works really well in practice for stationary solutions. In addition to be fast and effective, the eductive algorithm (as we term it) can be generalized to find not only stationary solutions but dynamical solutions of the mean field game PDEs. In short, the idea is simply to invert the heat operators instead of Laplace operators before introducing the virtual time θ. This is done in [22, 23].

6 Asset Managers and Ranking Effect

6.1 Introduction

When someone entrusts his saving to an asset manager, he does so according to his risk profile, i.e. he will try and find an asset manager whose management is as close as possible, in terms of return/risk for example, to what would be his own management strategy were it not delegated. However, the asset manager to whom he entrusts his savings does not have the sole aim of satisfying his current customers. He may wish to increase the number of his customers and therefore the assets under management or, as an individual, perform better in order to increase his bonus.

We offer a model which, starting off from the classic Markowitz model (see [41]) or the CARA-Gaussian model, adds a classification dimension among the asset managers: each asset manager will want to optimize, over and above his usual criterion, a function that depends on his classification (in terms of return) among all the asset managers.

6.2 The Model

Our model therefore considers a continuum of asset managers who at time 0 have the same unitary amount to manage. These managers will invest in risk-free and risky assets in creating their portfolio. A proportion θ of their portfolio will be invested in risky assets and a proportion $1 - \theta$ in risk-free assets with return r. The risky assets have a return which we denote $r + \tilde{\epsilon}$, where $\tilde{\epsilon}$ is a random variable that we will assume is distributed normally, with the mean and variance still be to specified.

To build their portfolio, managers will optimize a criterion of the following form:

$$\mathbb{E}[u(X) + \beta \tilde{C}]$$

where:

- $u(x) = -\exp(-\lambda x)$ is a CARA utility function.
- $X = 1 + r + \theta\tilde{\epsilon}$ is the fund value at date 1.
- β measures the relative importance of the additional criterion of competition among managers.
- $\tilde{C}$ is the random variable representing the classification. This variable $\tilde{C}$ has values in $[0, 1]$, with 0 corresponding to the worst performance and 1 to the best performance obtained by a manager.

It now remains to specify how managers differ. If they all have the same amount to invest, they nevertheless have different beliefs as regards the return on the risky asset, i.e. in relation to the variable $\tilde{\epsilon}$. These beliefs will concern the mean of $\tilde{\epsilon}$ (we assume that there is agreement on volatility), such that an agent will be of type ϵ if he thinks that $\tilde{\epsilon} \sim \mathcal{N}(\epsilon, \sigma^2)$. We will assume in what follows that the beliefs ϵ are distributed according to a probability distribution function f (even, for example).

6.3 Resolution

To solve this problem, let us consider an agent of type ϵ. The proportion θ of his portfolio placed in the risky asset is given by the following optimality condition:

Proposition 13 ($\mathbf{FOC}_\epsilon$).

$$(FOC_\epsilon) \quad -\lambda^2\sigma^2\left(\theta - \frac{\epsilon}{\lambda\sigma^2}\right)\exp\left(-\lambda(1+r) - \lambda\theta\epsilon + \frac{1}{2}\lambda^2\theta^2\sigma^2\right) + \beta m(\theta)C(\epsilon) = 0$$

where m is the distribution of the θ's in the population at equilibrium and where $C(\cdot) = 2\left[N\left(\frac{\cdot}{\sigma}\right) - \frac{1}{2}\right]$, N being the cumulative distribution function of a normal variable $\mathcal{N}(0, 1)$.

Proof. The asset manager maximizes:

$$\mathbb{E}_\epsilon\left[u(1 + r + \theta\tilde{\epsilon}) + \beta\tilde{C}\right]$$

It's easy to see that $\tilde{C} = 1_{\tilde{\epsilon}>0}M(\theta) + 1_{\tilde{\epsilon}\leq 0}(1 - M(\theta))$ where M stands for the cumulative distribution function of the weights θ.

Also,

$$\begin{aligned}\mathbb{E}_\epsilon\left[u(1 + r + \theta\tilde{\epsilon})\right] &= -\mathbb{E}_\epsilon\left[\exp\left(-\lambda\left(1 + r + \theta\tilde{\epsilon}\right)\right)\right] \\ &= -\exp\left(-\lambda\left(1 + r + \theta\epsilon\right) + \frac{1}{2}\lambda^2\theta^2\sigma^2\right)\end{aligned}$$

Hence, the optimal θ is given by the argmax of:

$$-\exp\left(-\lambda\left(1+r+\theta\epsilon\right)+\frac{1}{2}\lambda^2\theta^2\sigma^2\right)+\beta\mathbb{E}_\epsilon\left[1_{\tilde{\epsilon}>0}M(\theta)+1_{\tilde{\epsilon}\leq 0}(1-M(\theta))\right]$$

Let's differentiate the above equation. We get the first order condition for an ϵ-type asset manager:

$$\begin{aligned}&-\lambda^2\sigma^2\left(\theta-\frac{\epsilon}{\lambda\sigma^2}\right)\exp\left(-\lambda(1+r)-\lambda\theta\epsilon+\frac{1}{2}\lambda^2\theta^2\sigma^2\right)\\&+\beta\mathbb{E}_\epsilon\left[1_{\tilde{\epsilon}>0}-1_{\tilde{\epsilon}\leq 0}\right]m(\theta)=0\end{aligned}$$

But,

$$\mathbb{P}_\epsilon(\tilde{\epsilon}>0)-\mathbb{P}_\epsilon(\tilde{\epsilon}\leq 0)=2\left[\mathbb{P}_\epsilon(\tilde{\epsilon}>0)-\frac{1}{2}\right]=2\left[\mathbb{P}\left(\mathcal{N}(0,1)>-\frac{\epsilon}{\sigma}\right)-\frac{1}{2}\right]=C(\epsilon)$$

Hence we get the result. □

If we now use the fact that the solution $\epsilon\mapsto\theta(\epsilon)$ transports distribution f toward distribution m, we see that the problem, once resolved, can be written in a differential way:

Proposition 14 (Differential equation for $\epsilon\mapsto\theta(\epsilon)$). *Let's consider the function $\epsilon\mapsto\theta(\epsilon)$ that gives the optimal θ for each type. If θ is C^1 then it verifies the following differential equation:*

$$-\lambda^2\sigma^2\left(\theta-\frac{\epsilon}{\lambda\sigma^2}\right)\exp\left(-\lambda(1+r)-\lambda\theta\epsilon+\frac{1}{2}\lambda^2\theta^2\sigma^2\right)\frac{d\theta}{d\epsilon}+\beta f(\epsilon)C(\epsilon)=0\quad(*)$$

Moreover, θ must verify $\theta(0)=0$.

Proof. To go from the distribution f of the types to the distribution m of the θ's, we need a coherence equation that is simply:

$$m(\theta)\theta'(\epsilon)=f(\epsilon)$$

Now, if we take the different first order conditions FOC_ϵ and multiply by $\theta'(\epsilon)$ we get the ODE we wanted to obtain.

Now, because $C(0)=0$, the equation (FOC_0) is simply

$$-\lambda^2\sigma^2\theta\exp\left(-\lambda(1+r)+\frac{1}{2}\lambda^2\theta^2\sigma^2\right)=0$$

and the unique solution of this equation is $\theta=0$. □

If we return to the Markowitz problem ($\beta = 0$), we see that the solution is simply given by $\epsilon \mapsto \theta_0(\epsilon) = \frac{\epsilon}{\lambda\sigma^2}$. Our problem with $\beta > 0$ is therefore written:

$$\theta'(\epsilon) = \frac{\beta C(\epsilon) f(\epsilon)}{\lambda^2\sigma^2 \exp\left(-\lambda(1 + r + \theta(\epsilon)\epsilon) + \frac{1}{2}\lambda^2\sigma^2\theta(\epsilon)^2\right)} \frac{1}{\theta(\epsilon) - \theta_0(\epsilon)}, \quad \lim_{\epsilon \to 0} \theta(\epsilon) = 0$$

This is not a usual Cauchy problem since the condition in 0 is meaningful only at the limit. However, we should point out that the solution will be odd and that we can therefore restrict ourselves to $\epsilon > 0$. Also, $\theta(\epsilon)$ must be increasing, which implies $\theta(\epsilon) > \theta_0(\epsilon)$ and hence greater risk-taking in our model than in the Markowitz model.

Now we can completely solve the problem and we get:

Proposition 15 (Existence and Uniqueness). *There exists a unique function θ that verifies the equation $(*)$ with the two additional constraints:*

- $\theta(\epsilon) > \theta_0(\epsilon) = \frac{\epsilon}{\lambda\sigma^2}$
- $\lim_{\epsilon \to 0} \theta(\epsilon) = 0$

Proof. Let's start with the proof of the uniqueness.

Let's consider a solution θ of the problem and let's introduce the function z defined by:

$$z(\epsilon) = \frac{1}{2}\theta(\epsilon)^2 - \theta_0(\epsilon)\theta(\epsilon)$$

If we want to invert this equation and get θ as a function of z then we get:

$$\theta(\epsilon) = \theta_0(\epsilon) \pm \sqrt{\theta_0(\epsilon)^2 + 2z(\epsilon)}$$

but since $\theta(\epsilon) > \theta_0(\epsilon)$ we clearly can invert the equation and get:

$$\theta(\epsilon) = \theta_0(\epsilon) + \sqrt{\theta_0(\epsilon)^2 + 2z(\epsilon)} := \Theta(\epsilon, z(\epsilon))$$

Now, if we differentiate the equation that defines z we have:

$$z'(\epsilon) = \theta'(\epsilon)\theta(\epsilon) - \theta_0(\epsilon)\theta'(\epsilon) - \frac{1}{\lambda\sigma^2}\theta(\epsilon) = \theta'(\epsilon)\left(\theta(\epsilon) - \theta_0(\epsilon)\right) - \frac{1}{\lambda\sigma^2}\theta(\epsilon)$$

$$\Rightarrow z'(\epsilon) = \frac{\beta C(\epsilon) f(\epsilon)}{\lambda^2\sigma^2 \exp\left(-\lambda(1 + r + \theta(\epsilon)\epsilon) + \frac{1}{2}\lambda^2\sigma^2\theta(\epsilon)^2\right)} - \frac{1}{\lambda\sigma^2}\theta(\epsilon)$$

$$\Rightarrow z'(\epsilon) = \frac{\beta C(\epsilon) f(\epsilon)}{\lambda^2\sigma^2 \exp\left(-\lambda(1 + r + \Theta(\epsilon, z(\epsilon))\epsilon) + \frac{1}{2}\lambda^2\sigma^2\Theta(\epsilon, z(\epsilon))^2\right)} - \frac{1}{\lambda\sigma^2}\Theta(\epsilon, z(\epsilon))$$

From Cauchy-Peano we know that there is a unique solution z of this equation that verifies $z(0) = 0$. This solution is defined in a neighborhood V of 0. From this we can prove that we can locally define, in the neighborhood V, θ by $\theta(\epsilon) = \theta_0(\epsilon) + \sqrt{\theta_0(\epsilon)^2 + 2z(\epsilon)}$. This solution is unique because the differential equation that defines z has a unique solution for $\epsilon > 0$ (because of monotonicity). Hence, since there is no problem outside of 0 (*i.e.* the Cauchy-Lipschitz theorem can be applied directly) the uniqueness is proved.

Now, we want to prove that there exists a solution on the whole domain. For that let's consider the following ODE:

$$z'(\epsilon) = \frac{\beta C(\epsilon) f(\epsilon)}{\lambda^2\sigma^2 \exp\left(-\lambda(1 + r + \Theta(\epsilon, z(\epsilon))\epsilon) + \frac{1}{2}\lambda^2\sigma^2\Theta(\epsilon, z(\epsilon))^2\right)} - \frac{1}{\lambda\sigma^2}\Theta(\epsilon, z(\epsilon))$$

We know that there is a local solution z (defined on a neighborhood V of 0) satisfying this equation with $z(0) = 0$.

If we define θ_{loc} on V (or more exactly on an open subset of V that contains 0, because it is not a priori defined on V) as:

$$\theta_{loc}(\epsilon) = \theta_0(\epsilon) + \sqrt{\theta_0(\epsilon)^2 + 2z(\epsilon)}$$

then, we have a local solution of the equation $(*)$ that satisfies the two additional conditions. Let's consider now $\hat{\epsilon}$ in V. We can apply the Cauchy Lipschitz theorem to the equation $(*)$ with the Cauchy condition $\theta(\hat{\epsilon}) = \theta_{loc}(\hat{\epsilon})$ on the domain $\{(\epsilon, \theta)/\epsilon > 0, \theta > \theta_0(\epsilon)\}$ and consider θ the maximal solution of the problem. This maximal solution clearly satisfies $\lim_{\epsilon\to 0}\theta(\epsilon) = 0$. We want to show that there is in fact no upper bound for the maximal domain.

Suppose there is such an upper bound $\bar{\epsilon}$. Since θ is increasing, we have either:

$$\lim_{\epsilon\to\bar{\epsilon}}\theta(\epsilon) = +\infty$$

or

$$\lim_{\epsilon\to\bar{\epsilon}}\theta(\epsilon) = \theta_0(\bar{\epsilon})$$

We are going to show that these two cases are impossible.

Suppose first that $\lim_{\epsilon\to\bar{\epsilon}}\theta(\epsilon) = +\infty$. Then, we can suppose there exists an interval $(\underline{\epsilon}, \bar{\epsilon})$ such that $\forall\epsilon \in (\underline{\epsilon}, \bar{\epsilon}), \theta(\epsilon) > \theta_0(\epsilon) + 1$. Hence, on $(\underline{\epsilon}, \bar{\epsilon})$ we have:

$$\theta'(\epsilon) \leq \frac{\beta C(\epsilon) f(\epsilon)}{\lambda^2\sigma^2 \exp\left(-\lambda(1 + r + \theta(\epsilon)\epsilon) + \frac{1}{2}\lambda^2\sigma^2\theta(\epsilon)^2\right)}$$

$$\Rightarrow \theta'(\epsilon) \leq \frac{\beta C(\epsilon) f(\epsilon)}{\lambda^2\sigma^2} \exp\left(\lambda(1 + r) + \lambda\theta(\epsilon)\epsilon - \frac{1}{2}\lambda^2\sigma^2\theta(\epsilon)^2\right)$$

But $\lambda\theta(\epsilon)\epsilon - \frac{1}{2}\lambda^2\sigma^2\theta(\epsilon)^2 \leq \frac{\epsilon^2}{2\sigma^2}$ so that:

$$\forall \epsilon \in (\underline{\epsilon}, \overline{\epsilon}), \theta'(\epsilon) \leq \frac{\beta C(\epsilon) f(\epsilon)}{\lambda^2\sigma^2} \exp\left(\lambda(1+r) + \frac{\epsilon^2}{2\sigma^2}\right)$$

Hence,

$$\forall \epsilon \in (\underline{\epsilon}, \overline{\epsilon}), \theta(\epsilon) \leq \theta(\underline{\epsilon}) + \int_{\underline{\epsilon}}^{\epsilon} \frac{\beta C(\xi) f(\xi)}{\lambda^2\sigma^2} \exp\left(\lambda(1+r) + \frac{\xi^2}{2\sigma^2}\right) d\xi$$

This implies that we cannot have $\lim_{\epsilon \to \overline{\epsilon}} \theta(\epsilon) = +\infty$.

Now, let's consider the remaining possibility that is $\lim_{\epsilon \to \overline{\epsilon}} \theta(\epsilon) = \theta_0(\overline{\epsilon})$. The intuitive reason why this case is also impossible is that the slope when θ crosses the line associated to the solution θ_0 should be infinite and this cannot happen. To see that more precisely let's consider the following ODE:

$$\epsilon'(\theta) = \frac{\lambda^2\sigma^2 \exp\left(-\lambda(1 + r + \theta\epsilon(\theta)) + \frac{1}{2}\lambda^2\sigma^2\theta^2\right)}{\beta C(\epsilon(\theta)) f(\epsilon(\theta))} (\theta - \theta_0(\epsilon(\theta)))$$

Let's apply the Cauchy–Lipschitz theorem to the above equation on the domain $(\mathbb{R}^{+*})^2$ with the Cauchy condition $\epsilon(\theta_0(\overline{\epsilon})) = \overline{\epsilon}$. We have a local solution defined on a small interval $[\theta_0(\overline{\epsilon}) - \eta, \theta_0(\overline{\epsilon}) + \eta]$ and this solution exhibits a local minimum at $\theta_0(\overline{\epsilon})$. However, we can build another solution of the above Cauchy problem since the inverse of the maximal solution θ satisfies the equation and can be prolonged to satisfy the Cauchy condition. Therefore, because of the local minimum, the two solutions are different and this is absurd.

The conclusion is that the maximal interval has no upper bound.

Now, by symmetry the solution is defined on $\mathbb{R}$. □

One thing remains to be done. In fact, if we have found a function $\theta(\epsilon)$ that verifies the differential equation and hence a distribution m coherent with the first order condition, we still need to check that the second order condition is verified to be sure that we characterized a maximum of the optimization criterion. This is the purpose of the following proposition:

Proposition 16 (Second order condition). *Let's introduce*

$$\Gamma(\epsilon, \theta) = -\lambda^2\sigma^2 \left(\theta - \frac{\epsilon}{\lambda\sigma^2}\right) \exp\left(-\lambda(1+r) - \lambda\theta\epsilon + \frac{1}{2}\lambda^2\theta^2\sigma^2\right) + \beta m(\theta) C(\epsilon)$$

Let's consider the unique function $\theta(\epsilon)$, given by the preceding proposition, that satisfies $\forall \epsilon, \Gamma(\epsilon, \theta(\epsilon)) = 0$ and the conditions of the above proposition.
We have:

$$\partial_\theta \Gamma(\epsilon, \theta(\epsilon)) < 0$$

Proof. First, let's differentiate the first order condition $\Gamma(\epsilon, \theta(\epsilon)) = 0$ with respect to ϵ. We get:

$$\partial_\epsilon \Gamma(\epsilon, \theta(\epsilon)) + \theta'(\epsilon)\partial_\theta \Gamma(\epsilon, \theta(\epsilon)) = 0$$

Thus, the sign of $\partial_\theta \Gamma(\epsilon, \theta(\epsilon))$ is the sign of $-\partial_\epsilon \Gamma(\epsilon, \theta(\epsilon))$ and we need to prove that $\partial_\epsilon \Gamma(\epsilon, \theta(\epsilon)) > 0$.

But:

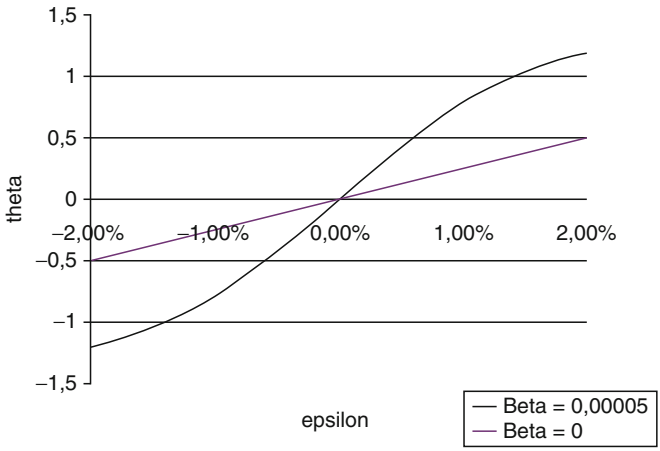
$$\partial_\epsilon \Gamma(\epsilon, \theta) = \lambda \exp\left(-\lambda(1+r) - \lambda\theta\epsilon + \frac{1}{2}\lambda^2\theta^2\sigma^2\right)\left(1 + \lambda^2\sigma^2\theta\left(\theta - \frac{\epsilon}{\lambda\sigma^2}\right)\right) + \beta m(\theta) C'(\epsilon)$$

This expression is positive for $\theta = \theta(\epsilon)$ since $\theta(\epsilon) \geq \frac{\epsilon}{\lambda\sigma^2}$ □

6.4 Example

Now that existence and uniqueness have been proved, we can try to compute numerically a solution. To know the shape of the curve, it's indeed important to compute the function $\theta(\epsilon)$ for an example and to compare it to the linear function $\theta_0(\epsilon)$ we usually obtain in the non-competitive case. This is what we are doing now.

Let us consider the following case. We put $r = 2\%$, $\sigma = 20\%$ and $\lambda = 1$. We put $s = 1\%$ the standard deviation associated to f and we make β small: $\beta = 5 \times 10^{-5}$. Numerically, we obtain the following result:



The conclusion is that the introduction of the mean field m overturns the Markowitz model. The Markowitz model indeed supposes that each agent reasons

as if he were alone and that's not true in practice. Surprisingly perhaps, even a small influence of competition (β) completely changes the shape of the solution and induces asset managers to take risker positions, both bullish and bearish.

7 Mean Field Games Model of Growth and Pareto Distribution of Salaries

7.1 Introduction to the Model Mechanisms

We shall construct an economic growth model based on human capital using the mean field games approach. The idea is to consider a large number of people who will endeavor to increase their human capital in order to increase their salary. Increasing one's human capital usually has two effects: it leads to increased competence and therefore salary, and also, *ceteris paribus*, a reduction in the number of people one is in competition with and, as a result, an increased salary. To take advantage of these two effects, there is obviously a cost. However, this cost is not the same for each individual since it is a priori easier for someone with poor qualifications to resort to training than for an agent whose human capital is close to what economists call the technological frontier (see [2]).

We consider a large number of agents, each having human capital that we term q. This human capital is distributed in the population according to a distribution function we term m (the associated cumulative distribution function is F and $\overline{F} = 1 - F$ is the tail function).

Let us now define the salary function. If we take a Cobb–Douglas production function (see [22]), it is clear[13] that the salary can be written in the form:

$$w(q, m(t,q)) = \begin{cases} C\frac{q^{\alpha}}{m(t,q)^{\beta}}, & \text{if } q \text{ is in the support of } m(t,\cdot) \\ 0 & \text{otherwise} \end{cases}$$

If we suppose that m is a distribution function that decreases with q, we find the two effects mentioned above.

The costs of increasing human capital must also be made explicit, and we express them as follows:

$$H\left(\frac{dq}{dt}, \overline{F}(t,q)\right) = \frac{E}{\varphi}\frac{\left(\frac{dq}{dt}\right)^{\varphi}}{\overline{F}(t,q)^{\delta}}, \quad \forall q \text{ in the support of } m(t,\cdot)$$

[13] In general, if we consider two production factors x_1 and x_2, taking a Cobb–Douglas production function means that the production y is of the form $y = Ax_1^{\alpha_1}x_2^{\alpha_2}$. Hence if x_1 is labor, the wage simply is $\frac{dy}{dx_1}$.

Here, E is a constant that indicates the inefficiency of the human capital production mechanism and δ and φ are two constants. This functional form means that the cost depends on the growth intensity a of human capital ($dq_t = a(t, q_t)dt$) but also on the proximity to the technological frontier, because of the tail function $\overline{F}$.

The parameters α, β, δ and φ are positive and, to successfully do our calculations, we shall suppose that $\alpha + \beta = \varphi$, $\beta = \delta$ and $\varphi > 1$, thus leaving two degrees of freedom.

7.2 The Optimization Problem and the Associated PDEs

Let us now move on to the problem of optimization of agents. We assume that they will maximize their wealth over time, which is coherent if one is situated in a stylized world without liquidity constraint.

The agents' problem is the following maximization:

$$Max_{(q_s),q_0=q} \int_0^\infty \left[w(q_s, m(s, q_s)) - H(a(s, q_s), \overline{F}(s, q_s))\right] e^{-rs} ds$$

To solve this problem, we must first specify the initial distribution of human capital. If we take a Pareto distribution for human capital, i.e. a distribution of the form:

$$m(0, q) = k \frac{1}{q^{k+1}} 1_{q \geq 1}$$

we see that the initial distribution of salaries is also a Pareto distribution, which conforms to reality, at least for distribution tails. We therefore opt for this specification.

The optimization problem can be solved, since it is deterministic, by using the classic Euler–Lagrange tools. However, as we shall later introduce uncertainty, we prefer to solve it with mean field games tools.

To do this, we introduce the Bellman function J:

$$J(t, q) = Max_{(q_s),q_t=q} \int_t^\infty \left[w(q_s, m(s, q_s)) - H(a(s, q_s), \bar{F}(s, q_s))\right] e^{-r(s-t)} ds$$

The mean field games PDEs that concern J and m are then written in the following form:

$$(HJB) \qquad w(q, m(t, q)) + \partial_t J + Max_a \left(a \partial_q J - H(a, \bar{F}(t, q))\right) - rJ = 0$$

$$(Kolmogorov) \qquad \partial_t m(t, q) + \partial_q (a(t, q) m(t, q)) = 0$$

where $a(t, q) = ArgMax_a \left(a \partial_q J - H(a, \bar{F}(t, q))\right)$ is the optimal control.

By using the specific forms chosen, we obtain:

$$C\frac{q^\alpha}{m(t,q)^\beta}+\frac{\varphi-1}{\varphi}\frac{1}{E^{\frac{1}{\varphi-1}}}\overline{F}(t,q)^{\frac{\beta}{\varphi-1}}(\partial_q J)^{\frac{\varphi}{\varphi-1}}+\partial_t J-rJ=0$$

$$\partial_t m(t,q)+\partial_q\left(\left(\frac{\overline{F}(t,q)^\beta}{E}\partial_q J(t,q)\right)^{\frac{1}{\varphi-1}}m(t,q)\right)=0$$

and the optimal control is:

$$a(t,q)=\left(\frac{\overline{F}(t,q)^\beta}{E}\partial_q J(t,q)\right)^{\frac{1}{\varphi-1}}$$

7.3 Solution

We can give explicit solutions[14,15]:

Proposition 17 (Resolution of the PDEs). *If $\varphi(\varphi-1)<\beta k$, there is a unique triple (J,m,γ) that satisfies both the PDEs and the additional equation on the optimal control function: $a(t,q)=\gamma q$.*

Solutions are of the following form:

$$m(t,q)=k\frac{\exp(\gamma k t)}{q^{k+1}}1_{q\geq\exp(\gamma t)}$$

$$J(t,q)=B\exp(-\beta k\gamma t)q^{\beta k+\varphi}1_{q\geq\exp(\gamma t)}$$

where γ and B are related by $\gamma=\left(\frac{B}{E}(\beta k+\varphi)\right)^{\frac{1}{\varphi-1}}$

Proof. First of all, the additional condition is equivalent to a constant growth rate for q_t and therefore, we obtain the Pareto distribution $m(t,\cdot)$ stated above.
Therefore, we have the following equation for $\partial_q J(t,q)$ if $q\geq\exp(\gamma t)$:

$$\partial_q J(t,q)=E(\gamma q)^{\varphi-1}\overline{F}(t,q)^{-\beta}=E(\gamma q)^{\varphi-1}e^{-\beta k\gamma t}q^{\beta k}$$

Hence (the constant being zero),

$$J(t,q)=\frac{E}{\beta k+\varphi}\gamma^{\varphi-1}e^{-\beta k\gamma t}q^{\beta k+\varphi}$$

[14] There are some additional restrictions about the parameters for the integral in the criterion to be defined at equilibrium (see [22]).

[15] As always, this solution of the PDEs does not automatically induce a solution of the control problem and a verification theorem still need to be written.

If we plug this expression into the Hamilton–Jacobi–Bellman equation we get:

$$\frac{C}{k^\beta} q^{\beta k+\varphi} e^{-\beta k\gamma t} + \frac{\varphi-1}{\varphi} E\gamma^\varphi q^{\beta k+\varphi} e^{-\beta k\gamma t}$$
$$-\beta k\gamma \frac{E}{\beta k+\varphi}\gamma^{\varphi-1} e^{-\beta k\gamma t} q^{\beta k+\varphi} - r\frac{E}{\beta k+\varphi}\gamma^{\varphi-1} e^{-\beta k\gamma t} q^{\beta k+\varphi} = 0$$

From this we get:

$$\frac{C}{k^\beta} + \frac{\varphi-1}{\varphi} E\gamma^\varphi - \beta k \frac{E}{\beta k+\varphi}\gamma^\varphi - r\frac{E}{\beta k+\varphi}\gamma^{\varphi-1} = 0$$
$$\frac{C}{k^\beta} + \frac{(\varphi-1)\varphi - k\beta}{\varphi(k\beta+\varphi)} E\gamma^\varphi - r\frac{E}{\beta k+\varphi}\gamma^{\varphi-1} = 0$$

Since $\varphi(\varphi-1) < \beta k$, γ is unique. □

Even though we cannot prove that there is uniqueness, this solution is very interesting since γ corresponds in fact to the rate of growth of human capital, which is the same for everyone. Furthermore, we see that the solution m is always Pareto-type, which is remarkable since in practice salary distribution tails are indeed Pareto-type (see [7, 8, 48, 49] for more details on the analysis of wealth distribution).

7.4 Underlying Mechanisms

The fact that there is a regular growth path merits spending a few moments on the underlying economic mechanism. To begin with, the basic reason why people change their human capital is due to two effects. First, there is a pure wage effect since, *ceteris paribus*, wage increases with human capital. However, this effect cannot explain by itself the continuous improvement of human capital at a constant growth rate. The effect needed to ensure a convincing explanation is an *escape competition* effect.[16] A given individual taken at random in the population is threaten by people who have less human capital than he has (say $\tilde{q}$). Indeed, if part of those people were to improve their human capital so that they end up with a human capital $\tilde{q}$ they will compete with our individual on the labor market, reducing her wage. This effect is the origin of continuous growth in our model. We have here a continuum of agents and therefore, for any given individual, there is always a threat.[17] We think therefore that the Schumpeterian effect which basically assumes that people will not improve their human capital if the gains are too small is reduced to nothing because there is always a potential competitor and that's why a Darwinian effect

[16] See [4–6] for the link between growth and competition.

[17] In practice everybody thinks there are people less gifted than he is...

(competition effect) dominates. Let's indeed highlight how tough is the threat effect. Each agent knows that every one is threaten by every one, and that fear will induce behaviors that will make the frightening event happen and be more important.

This model shows that the growth process is not only due to those who innovate, that is to say "researchers" near the technological frontier, but is in fact a process that involves the whole population and is fostered by those who are far from the technological frontier and threaten the leaders by improving their human capital. The process revealed is therefore very mean field games, if we can put it like that, since it brings into play an overall social dimension.

7.5 A Stochastic Generalization

Let us now move on to a more stochastic model.

We suppose that $dq_t = a(t, q_t)dt + \sigma q_t dW_t$ where W s a Brownian common to all the agents. If therefore we put q^m as the minimum human capital (this is in fact a new state variable that evolves according to $dq_t^m = a(t, q_t^m)dt + \sigma q_t^m dW_t$ where a is here the optimal control), we see that the Bellman function can be written $J = J(t, q, q^m)$ and the PDEs are:

$$Max_a\, C\frac{q^\alpha}{m(t,q)^\beta} - \frac{E}{\varphi}\frac{a^\varphi}{\overline{F}(t,q)^\beta} - rJ$$
$$+\partial_t J + a\partial_q J + \frac{\sigma^2}{2}q^2\partial^2_{qq}J + a'\partial_{q^m}J + \frac{\sigma^2}{2}{q^m}^2\partial^2_{q^m q^m}J + \sigma^2 q q^m \partial^2_{qq^m}J = 0$$

where a' is nothing but $a(t, q_t^m)$, exogenous in the optimization.

The optimal control is given by:

$$a(t,q) = \left(\frac{\overline{F}(t,q)^\beta}{E}\partial_q J(t,q)\right)^{\frac{1}{\varphi-1}}$$

Lemma 11. *If $a(t, q) = \gamma q$, then the probability distribution function of the q's is $m(t, q) = k\frac{(q_t^m)^k}{q^{k+1}}1_{q\geq q_t^m}$.*

Proof. Assuming $a(t, q) = \gamma q$ we get:

$$q_t = q_0 \exp\left(\left(\gamma - \frac{\sigma^2}{2}\right)t + \sigma W_t\right) = q_0 q_t^m$$

$$\Rightarrow m(t,q) = k\frac{\exp\left(k(\gamma - \frac{\sigma^2}{2})t + \sigma k W_t\right)}{q^{k+1}}1_{q\geq \exp((\gamma-\frac{\sigma^2}{2})t+\sigma W_t)}$$

$$= k\frac{(q_t^m)^k}{q^{k+1}}1_{q\geq q_t^m}$$

□

Proposition 18 (Resolution of the PDEs[18]). *If* $\varphi(\varphi-1) < \beta k$ *and* $r > \frac{\sigma^2}{2}\varphi(\varphi-1)$*, then, there is a unique growth rate* γ *compatible with the problem and* J *is of the form:*

$$J(q,q^m) = Bq^{\beta k+\varphi}(q^m)^{-\beta k}1_{q\geq q^m}$$

where γ *and* B *are related by* $\gamma = \left(\frac{B}{E}(\beta k+\varphi)\right)^{\frac{1}{\varphi-1}}$

Moreover, γ *is given by:*

$$\frac{\varphi(\varphi-1)-\beta k}{\varphi}\gamma^\varphi = \left(r-\varphi(\varphi-1)\frac{\sigma^2}{2}\right)\gamma^{\varphi-1} - \frac{C(\varphi+\beta k)}{Ek^\beta}$$

Proof. First, if $a(t,q) = \gamma q$ then,

$$\partial_q J(t,q,q^m) = E(\gamma q)^{\varphi-1}\overline{F}(t,q)^{-\beta} = E\gamma^{\varphi-1}q^{\beta k+\varphi-1}(q^m_t)^{-\beta k}$$

From this we deduce that the solution is of the stated form with $B = \frac{E}{\beta k+\varphi}\gamma^{\varphi-1}$.

If we want to find B or γ we need to plug the expression for J in the Hamilton–Jacobi–Bellman equation. This gives:

$$q^{\beta k+\varphi-1}(q^m)^{-\beta k}\left[\frac{C}{k^\beta} - \frac{E}{\varphi}\gamma^\varphi - rB + \gamma(\beta k+\varphi)B - \beta k\gamma B\right.$$
$$\left. + \frac{\sigma^2}{2}B\left((\beta k+\varphi)(\beta k+\varphi-1) + (-\beta k)(-\beta k-1) + 2(\beta k+\varphi)(-\beta k)\right)\right] = 0$$
$$\frac{C}{k^\beta} - \frac{E}{\varphi}\gamma^\varphi + \gamma\varphi B - \left(r-\varphi(\varphi-1)\frac{\sigma^2}{2}\right)B = 0$$
$$\frac{C(\beta k+\varphi)}{Ek^\beta} - \frac{\beta k+\varphi}{\varphi}\gamma^\varphi + \varphi\gamma^\varphi - \left(r-\varphi(\varphi-1)\frac{\sigma^2}{2}\right)\gamma^{\varphi-1} = 0$$
$$\frac{\varphi(\varphi-1)-\beta k}{\varphi}\gamma^\varphi = \left(r-\varphi(\varphi-1)\frac{\sigma^2}{2}\right)\gamma^{\varphi-1} - \frac{C(\varphi+\beta k)}{Ek^\beta}$$

As before, it's clear that, given our hypotheses, this equation has a unique solution. □

One consequence of these solutions is that growth is greater in the presence of a risk factor, even though this risk is common to everyone.

8 Mathematical Perspectives

The examples above clearly show that many kinds of nonlinear problems arise in the context of mean field games models. For most of them, these nonlinear problems are new systems of coupled nonlinear equations which, in the case the state of the

[18] For the "transversality" condition, see [22].

agents is described by continuous variables and the time variable is continuous, are partial differential equations. In all situations, the main novelty of these systems is the mixed "forward-backward" nature of the equations composing these systems. In general, no classical mathematical theory could be involved to tackle them. Furthermore, in the "partial differential" case for instance, the scope of the necessary new mathematical theory is quite large since many classical Partial Differential Equations (such as Hamilton–Jacobi–Bellman equations, Nonlinear heat or porous media equations, kinetic equations such as Vlasov or Boltzmann equations, compressible Euler equations of Fluid Mechanics, general semilinear elliptic equations, Hartree equations in Quantum Mechanics, optimal transportation problems, ...) are in fact particular cases of mean field games systems! This is to be expected since all these models arise in Mechanics and Physics model in a "mean field" fashion where the mean field sums up the collective or average behaviour of a large number of interacting particles (which can be seen as agents without any possibility of choosing their actions!)

Both the novelty of the mean field games models and the "range" of problems explain why numerous (and delicate) mathematical issues are being raised by mean field game theory.

To conclude, we set a brief (and thus far from exhaustive) list of issues for which some mathematical understanding is available (although a lot of open questions remain):

- *Justification of the derivation of mean field games models from N-player Nash equilibria*:
  A general analytical framework has been developed by J.-M. Lasry and P.-L. Lions ([34–37] and [38]) that allows to derive rigorously the mean field games equations from N-player Nash equilibria. This framework is of independent mathematical interest and has many other applications (limits of equations when the dimension goes to infinity, interacting particle systems, large deviations for stochastic partial differential equations, ...)
- *Expansion in N as the number of players N goes to infinity*:
  Such an expansion has been rigorously established for a large class of examples of mean field games models (at least in situations where the uniqueness of solutions holds for the limit mean field games system).
- *Existence and regularity results*:
  For large classes of models, the existence and regularity of solutions is now understood.
- *Uniqueness results*:
  Two uniqueness regimes have been identified: the case of a small horizon and the case of a "monotone" coupling. In addition, non-uniqueness examples are available that show that there does not seem to be any other general uniqueness regime.
- *Stability questions*:
  Of course, closely related to uniqueness is the issue of the stability of solutions which is indeed, true in the uniqueness regimes. It is worth pointing out that there are many notions of stability (small perturbations of data, horizon going to infinity, ...) which are all of interest.

– *Interpretation of mean field games models as control problems*:
  For a substantial class of mean field games models, it is possible to show that the mean field games system corresponds to a global optimal control problem of a certain partial differential equation. Roughly speaking, the system is then described as the coupling of the equations governing the state of a system and its dual state.
– *Numerical Approaches*:
  Various numerical methods or approaches have been proposed such as direct discretizations (finite elements) of the systems, discretization of the associated control problem (when there is one, see above), various iteration strategies, or the addition of an extra time variable ("relaxation time").
– *Limiting situations*:
  One relevant class of limiting situations corresponds to what could be called a planning problem. Instead of prescribing the initial state ("density") of the agents population and the terminal "cost-reward" profile for each agent as it is the case for "classical" mean field games models, one prescribes the state of the population (agents) both initially and at the end of the time interval (in other words, at both ends). In that case, the unknown terminal "cost-reward" function can be thought as the incentive scheme for each player which will lead to the desired final state of the population. Most of the preceding mathematical results can now be extended to that "limiting" class of models.

Acknowledgements The research results presented here could not have been made without the financial support of the Chair "Finance and Sustainable Development". Hence, the authors express their gratitude to the Chair and all the associated partners.

Also, the "Conseil Français de l'Energie" deserves a special acknowledgment for its specific financial support on the mean field games application to exhaustible resources.

References

1. Abramowitz, M., Stegun, I.: Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover, New York (1964)
2. Acemoglu, D., Aghion, P., Zilibotti, F.: Distance to frontier, selection, and economic growth. J. Eur. Econ. Assoc. **4**(1) (2006)
3. Achdou, Y., Capuzzo-Dolcetta, I.: Mean field games: Numerical methods (To appear)
4. Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P.: Competition and innovation: An inverted-u relationship. Q. J. Econ. **120**(2) (2005)
5. Aghion, P., Harris, C., Howitt, P., Vickers, J.: Competition, imitation and growth with step-by-step innovation. Rev. Econ. Stud. **68**(3), 467–492 (2001)
6. Aghion, P., Howitt, P. (To appear)
7. Atkinson, A.: Top incomes in the united kingdom over the 20th century. J. R. Stat. Soc. **168**, 325-343 (2005)
8. Atkinson, A., Piketty, T.: Top Incomes over the Twentieth Century: A Contrast between European and English-Speaking Countries. Oxford University Press, Oxford (2007)
9. Aumann, R.: Markets with a continuum of traders. Econometrica **32**(1/2), 1–17 (1964)
10. Bardi, M., Capuzzo-Dolcetta, I.: Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman equations. Birkhäuser, Boston (1997)

11. Bensoussan, A., Frehse, J.: Nonlinear elliptic systems in stochastic game theory. J. Reine Angew. Math. **350**, 23–67 (1984)
12. Bensoussan, A., Frehse, J.: Ergodic bellman systems for stochastic games in arbitrary dimension. Proc. Roy. Soc. A. **449**, 65–77 (1995)
13. Carmona, G.: Nash equilibria of games with a continuum of players. FEUNL Working paper series no. 466 (2004)
14. Chakrabarti, B.K., Chatterjee, A., Yarlagadda, S.: Econophysics of Wealth Distributions. Springer (2005)
15. Cordier, S., Pareschi, L., Toscani, G.: On a kinetic model for a simple market economy. J. Stat. Phys. **120**(1/2), 253–277 (2005)
16. Cross, R., Grinfeld, M., Lamba, H.: A mean-field model of investor behaviour. Journal of Physics: Conference Series 55 (2006)
17. David, H.A., Nagaraja, H.N.: Order Statistics (Third edition). Wiley, New Jersey (2003)
18. Ekeland, I.: Le pétrole sera-t-il bradé? Pour la Science (2007)
19. Farkas, I., Helbing, D., Vicsek, T.: Mexican wave in an excitable medium. Nature **419**, 131–132 (2002)
20. Giraud, P.N., Guéant, O., Lasry, J.M., Lions, P.L.: A mean field game model of oil production in presence of alternative energy producers (To appear)
21. Gomes, D., Mohr, J., Sousa, R.: Discrete time, finite state space mean field games. Journal de Mathématiques Pures et Appliqués **93**, 308–328 (2010)
22. Guéant, O.: Mean field games and applications to economics. Ph.D. thesis, Université Paris-Dauphine (2009)
23. Guéant, O.: A reference case for mean field games. Journal de Mathématiques Pures et Appliquées **92**, 276–294 (2009)
24. Guéant, O.: A reference case for mean field games. Cahier de la Chaire Finance et Développement Durable **10** (2009)
25. Guéant, O., Lasry, J.M., Lions, P.L.: Mean field games and oil production. In: Lasry, J.M., Lautier, D., Fessler, D. (eds.) The Economics of Sustainable Development. Economica (2010)
26. Guesnerie, R.: An exploration of the eductive justifications of the rational-expectations hypothesis. Am. Econ. Rev. **82**(5), 1254-1278 (1992)
27. Horst, U.: Ergodic fluctuations in a stock market model with interacting agents: the mean field case. Discussion paper No. 106, Sonderforschungbereich 373, Humboldt Universität, Berlin (1999)
28. Hotelling, H.: The economics of exhaustible resources. J. Polit. Econ. **39**(2), 137–175 (1931)
29. Kalai, E.: Private information in large games. Discussion paper 1312. Center for Mathematical Studies in Economics and Management Science. Northwestern University (2000)
30. Kalai, E.: Ex-post stability in large games. Discussion paper 1351. Center for Mathematical Studies in Economics and Management Science. Northwestern University (2001)
31. Kalai, E.: Large robust games. Econometrica **72**(6) (2004)
32. Khan, A., Sun, Y.: Non-cooperative games with many players (2002)
33. Lachapelle, A., Salomon, J., Turinici, G.: Computation of mean field equilibria in economics. Mathematical Models and Methods in Applied Sciences no. 1, 1–22 (2010)
34. Lasry, J.M., Lions, P.L.: Jeux à champ moyen i. le cas stationnaire. C. R. Acad. Sci. Paris **343**(9), 619–625 (2006)
35. Lasry, J.M., Lions, P.L.: Jeux à champ moyen ii. horizon fini et contrôle optimal. C. R. Acad. Sci. Paris **343**(10), 679–684 (2006)
36. Lasry, J.M., Lions, P.L.: Mean field games. Jpn. J. Math. **2**(1), 229–260 (2007)
37. Lasry, J.M., Lions, P.L.: Mean field games. Cahiers de la Chaire Finance et Développement Durable (2) (2007)
38. Lions, P.L.: Théorie des jeux à champs moyen et applications. Cours au Collège de France, http://www.College-de-france.fr/default/EN/all/equ_der/cours_et_seminaire s.htm (2007–2008)
39. Lions, P.L.: Mathematical Topics in Fluid Mechanics. Oxford Science Publications, Clarendon Press, Oxford (Vol. 1 (1996); Vol. 2 (1998))

40. Lucas, R., Sargent, T.: Rational expectations and econometric practice. University of Minnesota Press, Minnesota (1981)
41. Markowitz, H.: Portfolio selection. J. Finance **7**(1), 77–91 (1952)
42. Muth, J.F.: Rational expectations and the theory of price movements. Econometrica **29**(3), 315–335 (1961)
43. Nash, J.: The bargaining problem. Econometrica **18**, 155–162 (1950)
44. Nash, J.: Equilibrium points in n-person games. Proc. Natl. Acad. Sci. USA **36**, 48–49 (1950)
45. Nash, J.: Non-cooperative games. Ann. Math. **54**, 286–295 (1951)
46. Nash, J.: Two-person cooperative games. Econometrica **21**, 128–140 (1953)
47. Neumann, J.V., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (1944)
48. Piketty, T.: Income inequality in france, 1901-1998. J. Polit. Econ. **111**(5), 1004–1042 (2003)
49. Piketty, T., Saez, E.: Income inequality in the united states, 1913-1998. Q. J. Econ. **118**(1), 1–39 (2003)
50. Touzi, N.: Stochastic control and application to finance. Scuola Normale Superiore, Pisa. Special Research Semester on Financial Mathematics (2002)

The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices

David Hobson

Abstract This set of lecture notes is concerned with the following pair of ideas and concepts:

1. The Skorokhod Embedding problem (SEP) is, given a stochastic process $X = (X_t)_{t\geq 0}$ and a measure μ on the state space of X, to find a stopping time τ such that the stopped process X_τ has law μ. Most often we take the process X to be Brownian motion, and μ to be a centred probability measure.
2. The standard approach for the pricing of financial options is to postulate a model and then to calculate the price of a contingent claim as the suitably discounted, risk-neutral expectation of the payoff under that model. In practice we can observe traded option prices, but know little or nothing about the model. Hence the question arises, if we know vanilla option prices, what can we infer about the underlying model?

If we know a single call price, then we can calibrate the volatility of the Black–Scholes model (but if we know the prices of more than one call then together they will typically be inconsistent with the Black–Scholes model). At the other extreme, if we know the prices of call options for all strikes and maturities, then we can find a unique martingale diffusion consistent with those prices.

If we know call prices of all strikes for a single maturity, then we know the marginal distribution of the asset price, but there may be many martingales with the same marginal at a single fixed time. Any martingale with the given marginal is a candidate price process. On the other hand, after a time change it becomes a Brownian motion with a given distribution at a random time. Hence there is a 1–1 correspondence between candidate price processes which are consistent with observed prices, and solutions of the Skorokhod embedding problem.

These notes are about this correspondence, and the idea that extremal solutions of the Skorokhod embedding problem lead to robust, model independent prices and hedges for exotic options.

D. Hobson (✉)
Department of Statistics, Zeeman Building, University of Warwick,
Coventry CV4 7AL, UK
e-mail: D.Hobson@warwick.ac.uk

A.R. Carmona et al. (eds.), *Paris-Princeton Lectures on Mathematical Finance 2010*,
Lecture Notes in Mathematics 2003, DOI 10.1007/978-3-642-14660-2_4,
© Springer-Verlag Berlin Heidelberg 2011

1 Motivation

Consider the problem of pricing and hedging a one-touch digital option on an underlying. The standard approach to such a problem is to postulate a stochastic model for the underlying, perhaps the Osborne–Samuelson–Black–Scholes (exponential Brownian motion) model, and to price the option as the discounted expectation under the risk-neutral measure. In perfect frictionless markets this approach is justified by the theory of replication, and allows either counter-party in a transaction to eliminate the market risk.

However, the success of a replicating strategy is predicated on the fundamental truth of the model on which it is based. Although market risk (the known unknown) is eliminated, model risk (the unknown unknown) remains. The classical hedging strategies remove hedging risk, but leave agents exposed to Knightian uncertainty.

If the problem at issue is to price and hedge the exotic digital option then it seems reasonable to assume that simpler, vanilla options (such as call options) would also be traded. Then, at the very least, the volatility in the exponential Brownian motion can be calibrated with reference to the price of a traded call. However, this leads to a potential inconsistency, since many calls may be traded, each with their distinct volatilities. Ideally we should use a model which calibrates perfectly to the full spectrum of traded calls. However, in principle there are many such models, and associated with each model which is consistent with the market prices of (liquidly) traded options, there may be a different price for the exotic. Instead, one might attempt to characterise the class of models which are consistent with the market prices of options. This is a very challenging problem, and a less ambitious target is to characterise the extremal elements of this set, and especially those models for which the price of the exotic is maximised or minimised.

Suppose the one-touch digital option is written on a forward price $(S_t)_{0\le t\le T}$, and that the payoff is given by $\tilde{F} = I_A$ where I is the indicator function and $A = \{S_t \ge B, \text{ for some } t \in [0,T]\}$. Here the payoff is made at time T, we take 0 to be the current time and we assume that the barrier B is above the initial price $B > S_0$, and that S is right continuous. If we write H_B for the first time the underlying reaches the barrier then we have $H_B = \inf\{u > 0 : S_u \ge B\}$, and $\tilde{F} = I_{\{H_B \le T\}}$.

The key observation is contained in the following inequality which is valid for any $K < B$:

$$I_{\{H_B\le T\}} \le \frac{(S_T - K)^+}{B-K} + \frac{(S_{H_B} - S_T)}{B-K} I_{\{H_B\le T\}}. \tag{1}$$

Note that (1) is a path-wise inequality. The left-hand-side is the payoff of the option. The two terms on the right-hand-side have simple financial interpretations as the payoff from $1/(B-K)$ call options with strike K (with maturity taken to match that of the exotic), and the gains from trade from a forward transaction in S, struck the first time, if ever, that the underlying crosses the barrier. Provided the call with strike K is traded, and provided it is possible to invest forward in S, then

the right-hand-side of (1) describes a super-replicating strategy for the one-touch-digital; furthermore by no-arbitrage it determines an upper bound on the price of the one-touch-digital given by $C(K)/(B-K)$, where $C(k)$ is the traded price of a call option on the forward strike with k (By definition, the forward transaction is costless).

Since the strike K in (1) is arbitrary we can optimise over K. If we write the arbitrage-free price of the one-touch digital as $\mathcal{P}(\tilde{F})$ then

$$\mathcal{P}(\tilde{F}) \leq \inf \frac{C(K)}{B-K} \tag{2}$$

where the infimum is taken over traded strikes $K < B$.

In deriving this bound the only assumption that has been used is that it is possible to generate constant multiples of the payoff $(S_{H_B} - S_T)I_{\{H_B \leq T\}}$ at zero cost, which we interpret as an investment in the forward market. Note that we have imposed no probabilistic structure, we have not written down a stochastic model $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ nor have we postulated the existence of a martingale measure $\mathbb{Q}$. Instead, the bound is based on the ability to hedge using simple strategies in the forward market, and to take a static initial positions in vanilla securities at their prevailing time-0 market price. In this sense the resulting bound is model-independent. The explicit modelling of the dynamics of S has been replaced by the requirement that the model is consistent with observed call prices, (which places implicit conditions on the stochastic properties of the underlying price process) and δ-hedging is replaced by super-hedging with a portfolio of call options.

The quantity $\overline{D} = \inf_{K<B} \frac{C(K)}{B-K}$ gives a model-independent bound on the price of a digital option, in the sense that if we introduce any model for the underlying for which call prices are given by the pricing function $C(k)$, then the discounted expected payoff of the option under (any) martingale measure is less than $\overline{D}$.

The question immediately arises: is this bound best possible? We will show in Sect. 2.7 below, by demonstrating a model which matches call prices and for which the price of the one-touch digital is equal to the bound, that this is indeed the case, and the bound is tight. (The model we give may not be realistic, but unless it can be ruled out, the bound $\overline{D}$ cannot be refined.)

Return to the characterisation problem of describing the models which are consistent with observed option prices. It turns out that knowledge of the prices of puts and calls is equivalent to knowledge of the marginal distribution of the underlying asset under the measure used by the market for pricing. Further, if we assume temporarily that the market uses a pricing measure under which the forward price is a martingale, then the problem is to find a martingale with a given law at time T. However, any martingale is a time-change of Brownian motion. Hence the characterisation problem is reduced to the following: find stopping times such that the stopped Brownian motion has the given law. This is the classical Skorokhod embedding problem (SEP), first introduced (and solved) by Skorokhod [54].

The idea of relating the range of no-arbitrage prices of exotics to the prices of vanilla puts and calls via solutions of the Skorokhod embedding problem was first developed by Hobson [31] in the context of lookback options. The example above of

a one-touch barrier option, which allows perhaps for the simplest exposition of the theory, is from Brown et al. [9]. Recent applications include Cox and Oblój [18, 19] to two-sided barrier options, Hobson and Neuberger [34] to forward-starting straddles and Dupire [24] and Carr and Lee [12] to volatility options, but, in principle, the method applies to a wide range of path-dependent derivatives. The connection between model independent option price bounds and the Skorokhod embedding problem can be exploited in both directions. Existing solutions to the SEP (see Oblój [43] for a comprehensive survey) can be used to give bounds on the prices of financial options; conversely questions about the range of no-arbitrage prices for path-dependent options motivates the search for further solutions of the SEP.

2 Model Independent Option Pricing

2.1 The Classical Approach

The standard approach in derivative pricing is to begin by postulating a model (or a parametric family of models) for the price process $(P_t)_{t\geq 0}$, supported on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}))$. The price of a contingent claim with payoff $F_T = F(P_t; 0 \leq t \leq T)$, payable at T, is calculated as a discounted expectation:

$$\mathbb{E}^{\mathbb{Q}}[e^{-rT}F_T] = \mathbb{E}^{\mathbb{Q}}[e^{-rT}F(P_t; 0 \leq t \leq T)].$$

Here $\mathbb{Q}$ is an equivalent (local) martingale measure. In a complete market, the rationale for this pricing rule is supported by the replication (martingale representation)

$$e^{-rT}F_T = \mathbb{E}^{\mathbb{Q}}[e^{-rT}F_T] + \int_0^T \theta_t(dP_t - rP_t dt), \qquad \mathbb{Q} \text{ a.s.}$$

(Note that since $\mathbb{Q}$ is equivalent to $\mathbb{P}$, any statement which holds $\mathbb{Q}$ almost surely, also holds $\mathbb{P}$ almost surely). The quantity θ_t determines the investment or hedge in the risky asset.

For example, in the exponential Brownian motion parametric family of models we have that the asset price P follows

$$\frac{dP}{P} = \sigma dW + \gamma dt$$

and then for a call option with strike K, maturity T and payoff $F_T = (P_T - K)^+$ we have the price $\mathcal{C}(K, T; P_0; r, \gamma, \sigma)$, as given by the Black–Scholes formula.

Here the quantities K and T are characteristics of the derivative, P_0 is the observed initial value of the asset, r is assumed known (or observed from bond prices), the drift γ is irrelevant, since we price under the risk-neutral measure, and the volatility σ is estimated from historical data.

When the derivative has a payoff which is convex in the price process P_t (for example, a call option) then the model price $\mathcal{C} = \mathcal{C}(\sigma)$ is increasing in σ. Hence in practice, the volatility is selected via a calibration process in which the price of a liquidly traded option $\mathcal{C}^{\text{traded}}$ is used to determine the implied volatility σ_I as the solution to $\mathcal{C}^{\text{traded}} = \mathcal{C}(\sigma_I)$. In this sense the real purpose of the pricing function $\mathcal{C}(K,T)$ is as an extrapolation device. See Figlewski [26] for a discussion of other choices of extrapolation functions.

2.2 Model Misspecification and Incomplete Markets

In a complete market the price $\mathbb{E}^{\mathbb{Q}}[e^{-rT}F_T]$ and the hedge (or replication strategy) θ_t are uniquely determined, and replication is perfect provided that the model provides an exact description of reality.

Clearly, this last provision never holds. However, we have the following reassuring result, see Bergman et al. [6], El Karoui et al. [25], Hobson [33] and Janson and Tysk [37].

Theorem 1. *Suppose that the goal is to price and hedge a claim on P_T with convex payoff function $F_T = F(P_T)$. Suppose that the claim is priced and hedged under a diffusion model $dP_t = P_t\hat{\sigma}(P_t)dB_t + rP_tdt$ and that the model is sufficiently regular that the solution, denoted $\hat{P}$, is such that $(e^{-rt}\hat{P}_t)_{t\geq 0}$, is a true martingale. The model is Markovian, so that we may write the replicating hedge of the claim F as $\hat{\theta}_t = \hat{\theta}(P_t,t)$.*

Now suppose that the real-world dynamics are such that $dP_t = P_t\sigma_t dB_t + rP_tdt$ and that $\sigma_t \leq \hat{\sigma}(P_t)$. Then

$$\hat{\mathbb{E}}^{\mathbb{Q}}[e^{-rT}F_T] + \int_0^T \hat{\theta}(P_t,t)(dP_t - rP_tdt) \geq e^{-rT}F(P_T), \qquad \mathbb{Q}\ a.s.,$$

where $\hat{\mathbb{E}}^{\mathbb{Q}}$ denotes the fact that expectation has been calculated under a martingale measure for the model with volatility $\hat{\sigma}$.

The content of the theorem is that if an agent uses a model which overestimates volatility, then provided she uses both the model price and the model hedge, then she will still succeed in super-replicating the option on a path-wise basis. We will extend this idea of path-wise super-replication so that it does not rely on an assumption that the true volatility is bounded by a model volatility, but rather so that super-replication holds whatever the behaviour of the underlying.

In general, in incomplete markets there are several alternative criteria which can be used for option pricing. One alternative is to select arbitrarily (or otherwise) a martingale measure, perhaps by fixing the market prices of risk of those Brownian motions orthogonal to those driving the traded assets to be zero. Alternatively, utility indifference pricing incorporates the agents' attitudes to risk to find a price at which an option seller will be prepared to enter into a contract for sale of a derivative. The extreme version of utility indifference pricing is super-replication, whereby the agent is not willing to accept any risk.

These approaches are all conditional on an assumption that the dynamics of P follow a given model, or family of models. Often, (e.g. in the case of utility indifference pricing) the investor needs knowledge of the drift parameter γ (which is notoriously difficult to estimate) as well as the volatility σ.

2.3 A Reverse Approach: Recovering a Model from Prices

The starting point of our philosophy is to take option prices (at least those of liquidly traded options) as exogenously given by the market, and to use those prices to learn about the stochastic properties of the underlying.

Since market prices are determined under the measure used by the market for pricing, the conclusions we draw are already phrased in a fashion whereby they can be used for pricing other derivatives.

Lemma 1 (Breeden and Litzenberger [8]). *Fix $T \in (0, \infty)$. Suppose that call prices with maturity T are known for every $K \in (0, \infty)$. Then assuming call prices are calculated as the discounted expected payoff under a model $\mathbb{Q}$, so that*

$$\mathcal{C}(K,T) = \mathbb{E}^{\mathbb{Q}}[e^{-rT}(P_T - K)^+],$$

we have

$$\mathbb{Q}(P_T > K) = e^{rT} \left| \frac{\partial}{\partial K} \mathcal{C}(K,T) \right|, \tag{3}$$

and, provided $\mathcal{C}$ is twice-differentiable in K

$$\mathbb{Q}(P_T \in dK) = e^{rT} \frac{\partial^2}{\partial K^2} \mathcal{C}(K,T) \tag{4}$$

If the law of P_T under $\mathbb{Q}$ has atoms, then $\mathbb{Q}(P_T > K)$ is given by the right derivative in (3), and $\mathbb{Q}(P_T \geq K)$ by the left derivative. In this case (4) must be understood in a distributional sense. However, for the present we assume that there are no atoms and $\mathcal{C}$ is smooth (Fig. 1).

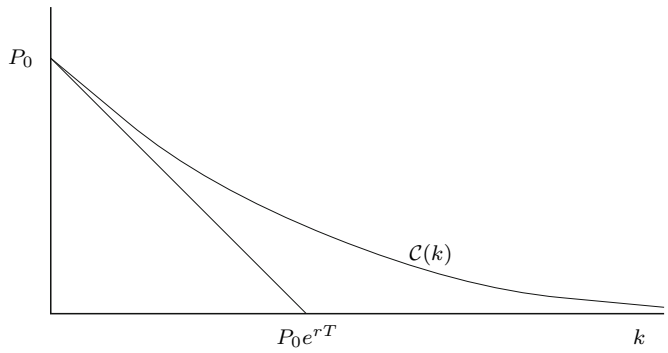


Fig. 1 No arbitrage considerations force that the call price curve $\mathcal{C}$ is a decreasing convex function of k (for maturity T fixed) with $\mathcal{C}(0) = P_0$ and $\mathcal{C}(k) \geq (P_0 - Ke^{-rT})^+$

It follows from the Lemma that call prices, as represented by the set of decreasing convex functions are in one-to-one correspondence with the marginal laws of P_T under the pricing measure. The above result applies for individual maturities. If we also have information about the dependence of option prices on maturity we can recover a price process.

Theorem 2 (Krylov [39], Gyöngy [28], Dupire [23]). *Suppose call prices are known for every $K \in (0, \infty)$ and every $T \in (0, \overline{T})$. Assuming that $\mathcal{C}(T, K)$ is sufficiently differentiable, there exists a unique diffusion of the form*

$$dP_t = P_t \sigma(P_t, t) dB_t + r P_t dt$$

such that

$$\mathbb{E}[e^{-rT}(P_T - K)^+] = \mathcal{C}(K, T)$$

In particular, $\sigma(p, t)$ solves

$$\frac{1}{2} K^2 \sigma(K, T)^2 \mathcal{C}_{KK} - rK\mathcal{C}_K - \mathcal{C}_T = 0.$$

The idea in Krylov [39] is to take any process $(Y_t)_{t \geq 0}$ and to show that there exists a diffusion process X_t with the same marginal distributions. Gyöngy refined the analysis and derived the relationship $\sigma(k, t)^2 = \mathbb{E}[(dY)_t^2 | Y_t = k]$. Dupire has a slightly different starting point in that rather than assuming the existence of a process Y he begins with the call price surface $\mathcal{C}(K, T)$.

The Dupire construction gives a unique Markovian martingale which agrees with the doubly infinite family of call option prices. In principle, it is then possible to give the prices of any other derivatives. However, the diffusion assumption is a strong assumption (and is inconsistent with stochastic volatility, for example), and although the Dupire diffusion calibrates perfectly to the initial structure of call prices, there is no guarantee that it will match derivative prices at any later time. (We consider in Sect. 6.3, the existence of other martingales which also match the marginals implicit in the call price function.)

Both the Breeden and Litzenberger [8] and Dupire [23] results give us information about how the marginal distributions of the asset (under the pricing measure) can be determined from call prices. More generally, given any set of option prices we can ask:

- Is there a model consistent with those prices? (i.e. can we find $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$ such that $\mathbb{E}^{\mathbb{Q}}[e^{-rT} F_T] = f$ for each claim F with associated traded price f)
- If such a model exists is it unique?
- If there is no such model, is there an arbitrage?

2.4 Notation

It is very convenient to remove the impact of interest rates by switching to a notation in which prices are expressed in discounted terms. Define $S_t = e^{-rt}P_t$, then $(S_t)_{t\geq 0}$ is the forward price, and is a martingale under a pricing measure. Furthermore, we write $C(k)$ for the time-zero price of an option on the forward S with strike k (and maturity T):

$$C(k) = \mathbb{E}^{\mathbb{Q}}[(S_T - k)^+] = \mathbb{E}^{\mathbb{Q}}[e^{-rT}(P_T - ke^{rT})^+] = \mathcal{C}(ke^{rT})$$

where $\mathcal{C}$ denotes the price of a call option on P_T. Then $\mathbb{Q}(S_T > k) = |C'(k+)|$.

When written without a superscript C denotes the traded set of options prices. However, under a martingale model for which $S_T \sim \mu$, we also have that call prices are given by

$$C_\mu(k) = \mathbb{E}[(X - k)]^+, \qquad X \sim \mu;$$

and we use C with a subscript to denote call prices which correspond to S_T having a given law. Later we use P_μ to denote put prices under law μ:

$$P_\mu(k) = \mathbb{E}[(k - X)]^+, \qquad X \sim \mu.$$

In the financial context, μ will be the law of a positive random variable, often with mean 1. However, the quantities C_μ and P_μ are well defined for any μ with first moment, and we will frequently consider the case where μ is the law of a centred random variable.

2.5 The Pricing Problem

The goal of the next two sections is to develop a formal framework within which we can attempt to determine model-independent bounds and hedging strategies. Our set-up is an extension of the framework developed by Föllmer and Schied [27] and especially Cox and Obłój [18, 19].

Let Ω^S describe the space of possible paths for the asset price process. A natural candidate is to let Ω^S be the space of continuous non-negative trajectories with a given initial value S_0. The advantage of using this choice is that we can then define the natural filtration associated with the coordinate process, and then first hitting times will be stopping times with respect to this filtration. Note that we do not want to specify the probability triple (and especially the probability measure) exogenously, but rather to construct models (and measures) which are consistent with observed prices. See Cox and Obłój [19] for further discussion on this issue.

Suppose we are given a family Θ of admissible trading strategies $(\theta_t)_{0\leq t\leq T}$ with associated terminal gains from trade processes $G^\theta = \int_0^T \theta_t dS_t$. The space

$G^{\Theta} = \{G^{\theta}; \theta \in \Theta\}$ represents a space of costless, forward transactions. Implicitly, we are assuming zero transactions costs for trades on the forward market, or at least that such costs are negligible on strategies that we class as admissible.

Suppose we are given a family $\{F_\alpha, \alpha \in A\}$ of traded payoffs (random variables), and suppose that V is a vector space with basis $F_\alpha, \alpha \in A$.

Suppose we are also given a map (pricing functional), $\mathcal{P} : V \oplus G^{\Theta} \to \mathbb{R}$ such that $\mathcal{P}$ is both linear, and monotone:

$$\begin{aligned} \mathcal{P}(v + G^{\theta}) &= \mathcal{P}(v); \\ \mathcal{P}(\lambda_1 F_1 + \lambda_2 F_2) &= \lambda_1 \mathcal{P}(F_1) + \lambda_2 \mathcal{P}(F_2); \\ F_1 \geq F_2 &\Rightarrow \mathcal{P}(F_1) \geq \mathcal{P}(F_2). \end{aligned}$$

Note that the monotonicity plays the role of a no-arbitrage condition. The idea is that $\mathcal{P}$ represents the prices at which the traded payoffs can be bought in the market. In the analysis we will assume that these payoffs can be bought and sold at the same price, but once we attempt to value a hedging portfolio of traded payoffs we can use the market bid or ask prices as appropriate.

Now suppose we add a family of payoffs $\{F_{\tilde{\alpha}}; \tilde{\alpha} \in \tilde{A}\}$, and let $\tilde{V}$ be the vector space with basis $\{F_\alpha; \alpha \in A\} \cup \{F_{\tilde{\alpha}}; \tilde{\alpha} \in \tilde{A}\}$. The problem is to decide when $\mathcal{P}$ can be extended to a linear function from $\tilde{V} \oplus G^{\theta}$ to $\mathbb{R}$.

Definition 1. The family $\{F_\alpha; \alpha \in A\}$ is the set of vanilla securities.
The family $\{F_{\tilde{\alpha}}; \tilde{\alpha} \in \tilde{A}\}$ is the set of exotic securities.

Equivalently the set $\{F_\alpha; \alpha \in A\}$ is the set of payoffs whose prices are precisely determined by the market, and the set $\{F_{\tilde{\alpha}}; \tilde{\alpha} \in \tilde{A}\}$ is the set of payoffs for which we wish to determine a price, relative to the prices of the vanilla securities.

We have not yet been specific about the set of admissible trading strategies Θ, and the definition of Θ may be chosen according to the context. We certainly want to exclude strategies which allow for arbitrage. The main example is when Θ is the linear space generated by $\theta^\tau = (\theta_t^\tau)_{0 \leq t \leq T}$ for stopping times τ, where $\theta_t^\tau = I_{\{t \in (\tau, T]\}}$. These are the gains from trade obtained from simple combinations of forward contracts. However, in Sect. 5.1 we will want to allow for more general strategies.

2.6 The Hedging Problem

Definition 2. We say Π^V is a static vanilla portfolio if

$$\Pi^V = \sum_{i=1}^{n} \pi_{\alpha_i} F_{\alpha_i} \qquad n < \infty, \alpha_1, \ldots \alpha_n \in A$$

We say Π is a semi-static portfolio if Π is the sum of a static vanilla portfolio and an admissible trading portfolio.

Definition 3. Π is a semi-static superhedging portfolio for $\hat{F}$ if Π is a semi-static portfolio and $\Pi \geq \hat{F}$ almost surely.

Lemma 2. *If $\mathcal{P}(F_\alpha) = f_\alpha$ for $\alpha \in A$ and if $\Pi = \sum_{i=1}^n \pi_{\alpha_i} F_{\alpha_i} + G^\theta$ is a semi-static superhedging portfolio for $\hat{F}$, then for any extension $\mathcal{P}$ from V to $\tilde{V}$ we must have*

$$\mathcal{P}(\hat{F}) \leq \sum_{i=1}^n \pi_{\alpha_i} f_{\alpha_i}$$

Proof. If $\mathcal{P}$ can be extended to the linear space spanned by $\{F_\alpha : \alpha \in A\} \cup \hat{F}$ then

$$\mathcal{P}(\hat{F}) \leq \mathcal{P}(\Pi) = \sum_{i=1}^n \pi_{\alpha_i} f_{\alpha_i}$$

by monotonicity and linearity. □

Remark 1. In the set-up of Sect. 2.3 an implicit assumption is that option prices are both bid and ask prices and that there are zero transaction costs. This is a necessary assumption in order to identify a precise correspondence between prices and probabilities under the risk-neutral measure for use in the pricing problem. However, once we have identified a candidate model-independent super-replicating strategy involving vanilla options, in the hedging problem we can use market bid or ask prices (depending on whether the hedge involves a short or long position) as appropriate. In this way the theory extends to situations in which there are bid-ask spreads, even on liquidly traded securities. Note, moreover, that the position in vanilla securities is static over time, so these transaction costs are incurred once only, and unlike classical model-dependent δ-hedging, the semi-static hedging strategy remains feasible in a model with transaction costs.

By restricting the class of admissible trading strategies in the forward, we can also allow for transaction costs on the forward. For example, in the next section we could restrict the class of admissible trading portfolios to those which involve at most a single forward transaction, and therefore incur transaction costs at most once.

2.7 An Example: A Digital Option

We return to the example first discussed in the opening section; namely given the prices of call options what are the possible prices of a one-touch digital option, a security which makes a unit payment at T if ever $S_t \geq B$ for some $t \leq T$.

We assume that forward contracts with maturity T are traded at zero cost, and that

$$\Theta = \left\{ \theta : \theta_t = \sum_{i=1}^p \gamma_i I_{\{\tau_i < t \leq T\}} \right\} \qquad p \in \mathbb{N}, \gamma_i \in \mathbb{R}$$

where $(\tau_i)_{i\leq p}$ are stopping times. In fact the only property we need for Θ is that $\Theta \supseteq \{cI_{\{H_B<t\leq T\}}; c \in \mathbb{R}\}$.

We assume that the vanilla options are precisely the set of call options with maturity T, parameterised by the strike K, and that a continuum of such calls are traded (one for each possible strike), together with the bond which pays a unit value at time T. In the notation of this section, $\{F_\alpha; \alpha \in \mathcal{A}\} = \{F_K; K \in \mathbb{R}_+\} = \{(S_T - K)^+; K \in \mathbb{R}_+\}$, and for $K \in (0, \infty)$ the payoff $F_K = (S_T - K)^+$ has price $f_K = C(K)$. Note that to preclude arbitrage in the vanilla options we must have that C is a decreasing convex function, and that by our assumption that S is a forward price $C(0) = S_0$, $C'(S_0) \geq -1$ and $C(K) \geq (S_0 - K)^+$. We also make the very natural assumption that $\lim_{K\uparrow\infty} C(K) = 0$. (As observed by Davis and Hobson [20] this does not follow from no arbitrage alone, but rather from a martingale assumption, and can be violated if the asset price allows a bubble, see e.g. Cox and Hobson [16]. Davis and Hobson [20] introduce the notion of a *weak arbitrage* and Cox and Oblój [19] introduce the parallel notion of a *weak free lunch with vanishing risk*, which they use to describe some of the consequences of the failure of this assumption.)

Now we consider the contract $\tilde{F} = I_{\{H_B\leq T\}}$ with $B \geq S_0$. For any $K < B$ we have (recall (1))

$$\tilde{F} \leq \frac{1}{(B-K)} F_K + G^\theta$$

where F_K is the payoff of a vanilla option $(S_T - K)^+$, and $\theta_t = -I_{\{H_B<t\leq T\}}/(B - K)$ so that $G^\theta = (S_{H_B} - S_T)/(B - K)$. There are two cases to be checked, namely when the left-hand-side is zero or one. If the barrier is not reached, then the left-hand-side is null and there is only one non-zero term on the right-hand-side, which is non-negative. If the barrier is reached, then the second term on the right-hand-side of (1) is present and moreover $S_{H_B} \geq B$ (with equality if S is continuous). Taking the value $(S_T - K)$ rather than the positive part $(S_T - K)^+$ again only makes the right-hand-side smaller, and leads to equality in (1) on $(H_B \leq T)$.

It follows immediately that $\mathcal{P}(\tilde{F}) \leq C(K)/(B - K)$, and since K is arbitrary we conclude $\mathcal{P}(\tilde{F}) \leq \inf_{K<B} C(K)/(B - K) = C(K^*)/(B - K^*) =: \overline{D}$, where $K^* = \arg\inf_{K<B}(C(K)/(B - K))$.

Now we want to argue that this bound is best possible in the given framework. To do this we exhibit a model under which the fair prices (i.e. expected values) of the vanilla securities are equal to the bound $\overline{D}$. Since we are interested in pricing issues, it makes sense to work directly in a martingale measure, so that there is no distinction between $\mathbb{P}$ and $\mathbb{Q}$.

Write μ for the law of S_T. Let $b(K) = \mathbb{E}[S_T|S_T \geq K] = \int_{y\geq k} y\mu(dy)/\int_{y\geq k} \mu(dy)$; b is called the barycentre function. By definition it is left-continuous, increasing and $b(K) \geq K \vee \mathbb{E}[S_T]$, at least for K smaller than the upper bound on the support of μ, where it is defined. Furthermore, if μ has a density then $b(K)$ is continuous.

Given B there exists $\hat{K} = \hat{K}(B)$ such that $b(\hat{K}+) \geq B \geq b(\hat{K})$ and then we can find a set A with

$$(S_T > \hat{K}) \subseteq A \subseteq (S_T \geq \hat{K})$$

for which $B = \int_A y\mu(dy)/\int_A \mu(dy)$, or in other words

$$0 = \int_A (y - B)\mu(dy) = \int_A (y - \hat{K})\mu(dy) + (\hat{K} - B)\mu(A).$$

It follows that $\mu(A) = C(\hat{K})/(B - \hat{K})$ so that

$$\mathbb{P}(S_T > \hat{K}) \leq \frac{C(\hat{K})}{B - \hat{K}} \leq \mathbb{P}(S_T \geq \hat{K}).$$

Let $a = \int_{A^c} y\mu(dy)/\int_{A^c} \mu(dy)$, and let $S_{T/2}$ take the values B or a with probabilities $(x - a)/(B - a)$ and $(B - x)/(B - a)$ respectively.

On the set $S_{T/2} = B$, let S_T have the law μ restricted to A. Similarly, on $S_{T/2} = a$, let S_T have the law μ restricted to A^c.

Suppose S_t is constant except for a pair of jumps at $T/2$ and T. See Fig. 2. Then, by construction, $\mathbb{E}[S_T; A] = B\mathbb{P}(A)$ and $\mathbb{E}[S_T; A^c] = a\mathbb{P}(A^c)$ so that S_t is a martingale. Moreover, if prices $\mathcal{P}$ are given by expectation with respect to $\mathbb{P}$ then

$$\mathcal{P}(\tilde{\mathcal{F}}) = \mathbb{P}(H_B \leq T) = \mathbb{P}(S_{T/2} = B) = \mu(A) = \frac{C(\hat{K})}{B - \hat{K}}.$$

It remains to show that $\hat{K}$ is a minimiser of $C(K)/(B-K)$. This can be deduced from the fact that the bound $\mathcal{P}(\tilde{F}) \leq C(K)/(B - K)$ holds for any martingale model with the correct marginals. We have

$$\mathbb{P}(H_B \leq T) = \mu(A) = \frac{C(\hat{K})}{(B - \hat{K})} \geq \inf_{K < B} \frac{C(K)}{(B - K)} \geq \sup \tilde{\mathbb{P}}(H_B \leq T) \geq \mathbb{P}(H_B \leq T)$$

where the supremum is taken over martingale models, denoted by $\tilde{\mathbb{P}}$, with the correct marginal laws at time T. Hence there is equality throughout and $\hat{K}$ is a minimiser (Fig. 3).

The above analysis exhibits a pure-jump model for which $S_T \sim \mu$ (and the price of the barrier option is equal to the robust upper bound $C(K^*)/(B - K^*)$). Although potentially the simplest model, it is not unique. Indeed, we can construct a continuous price process model with the same properties.

Let $(W_t^0)_{t\geq 0}$ be Brownian motion, started at zero, and for $\alpha < 0 < \beta$, let $H^0_{\alpha,\beta} = \inf\{u : W_u^0 \notin (\alpha, \beta)\}$. Then $W^0_{H^0_{\alpha,\beta}} \in \{\alpha, \beta\}$ and $\mathbb{P}(W^0_{H^0_{\alpha,\beta}} = \alpha) = \beta/(\beta - \alpha) =: p$, and we have embedded a two-point distribution in Brownian motion. Now set $X_u = W^0_{u/(1-u)\wedge H^0_{\alpha,\beta}}$; then X is a martingale with $X_1 \sim p\delta_\alpha + (1 - p)\delta_\beta$.

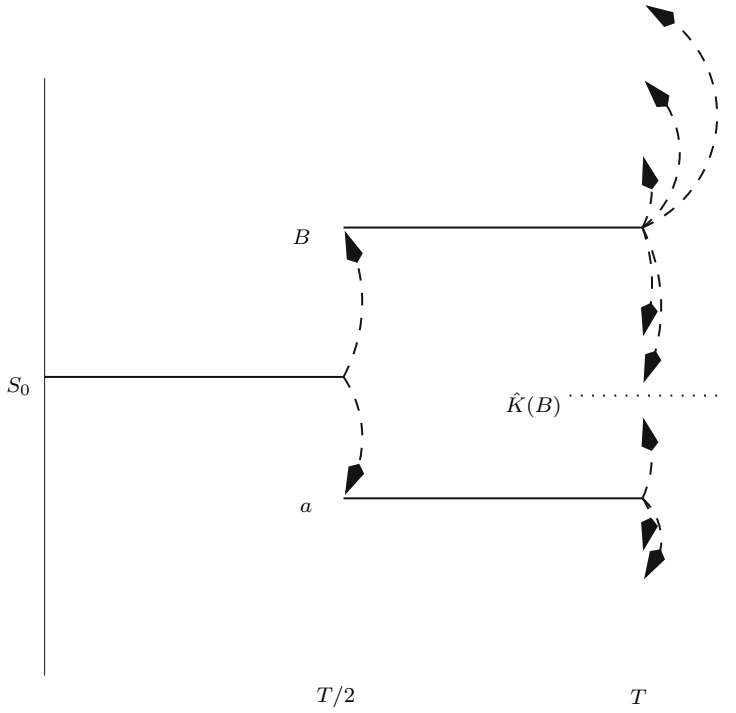


Fig. 2 A jump model which matches options prices, and for which the maximum value of the digital option is attained. The price process is constant except for jumps at $T/2$ and T. Those paths which are at B (respectively a) at $T/2$ end up above (below) $\hat{K}(B)$ after the jump at time T

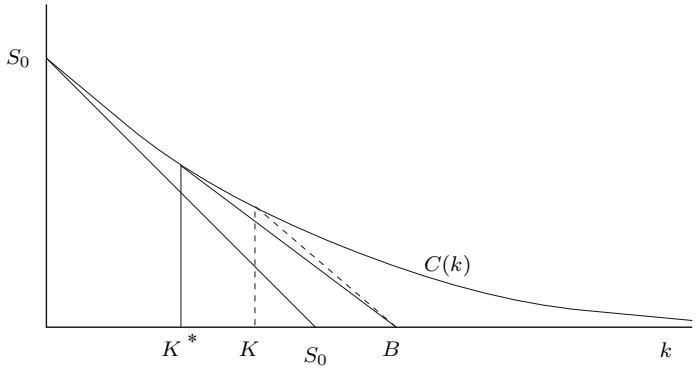


Fig. 3 The minimiser $K^* = K^*(B)$ is chosen such that the tangent to C at K^* crosses the x-axis at B

With this in mind, define $H^s_{a,B} = \inf\{u : W^s_u \notin (a, B)\}$ where $W^s_t = S_0 + W^0_t$, and $S_0 = s$. By construction $S_0 + W^0_{H^s_{a,B}} \in \{a, B\}$. If we set

$$S_t = S_0 + W^0_{2t/(T-2t)\wedge H^0_{a-s,B-s}} = W^s_{2t/(T-2t)\wedge H^s_{a,B}}$$

then $(S_t)_{0\le t\le T/2}$ is a continuous martingale which takes values in $\{a, B\}$ at time $T/2$. To complete the construction (i.e. to construct the process S_t over $(T/2, T]$) we need to discuss how to embed any centred distribution in Brownian motion (and not just a two point distribution). This is the Skorokhod embedding problem.

3 The Skorokhod Embedding Problem

3.1 The Classical Skorokhod Embedding Problem

Suppose we are given a stochastic process $X = (X_t)_{t\ge 0}$ on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F})_{t\ge 0}, \mathbb{P}, (X_t)_{t\ge 0})$. Suppose X has state space I.

The Skorokhod embedding problem (SEP) for (X, μ) is, given μ on I, to find (where possible) a stopping time τ such that the law of the stopped process is μ. We write $\tau \in \bar{\mathcal{T}}(X, \mu)$. The classical version of the problem takes $X \equiv W$, a Brownian motion on $\mathbb{R}$, started at 0, and μ a centred probability measure, in which case we suppress W from the notation and write $\tau \in \bar{\mathcal{T}}(\mu)$.

Oblój [43] provides a comprehensive survey of the literature of solutions to the Skorokhod embedding problem. He lists more than twenty solutions, some of which we will discuss below.

In the sequel the key properties of Brownian motion that we will use are the fact that Brownian motion leaves any interval in finite time almost surely, (so $H_a := \inf\{u : W_u = a\} < \infty$ a.s.); the martingale property; the consequent fact that for $a < x < b$, $\mathbb{P}^x(H_a < H_b) = (x-a)/(b-a)$; and the fact that Brownian motion has no intervals of constancy. Only occasionally will we use the speed of Brownian motion, but then we have that $\mathbb{E}^x[H_a \wedge H_b] = (b-x)(x-a)$.

We shall often decompose Brownian motion into its constituent excursions away from zero, which are parameterised by the local time L_t. See Rogers [49] for an introduction to Brownian excursion theory. If we write J_t for the maximum process $J_t = \sup_{0\le s\le t}\{W_s\}$ (we reserve the notation M_t for martingales, and S_t for forward prices) then by Lévy's Theorem $(L_t, |W_t|) \sim (J_t, J_t - W_t)$, and there is an equivalent decomposition of a Brownian path into the excursions below the maximum. The local time is parameterised in such a way that the rate of excursions with maximum (in modulus) greater than a is a^{-1}, for any positive a.

3.2 Doob's Solution

Our first solution is attributed by Rogers and Williams [50] to Doob, and is not intended as a genuine solution to the SEP, but rather to show that without some refinements on the class of solutions to the SEP, the problem is trivial.

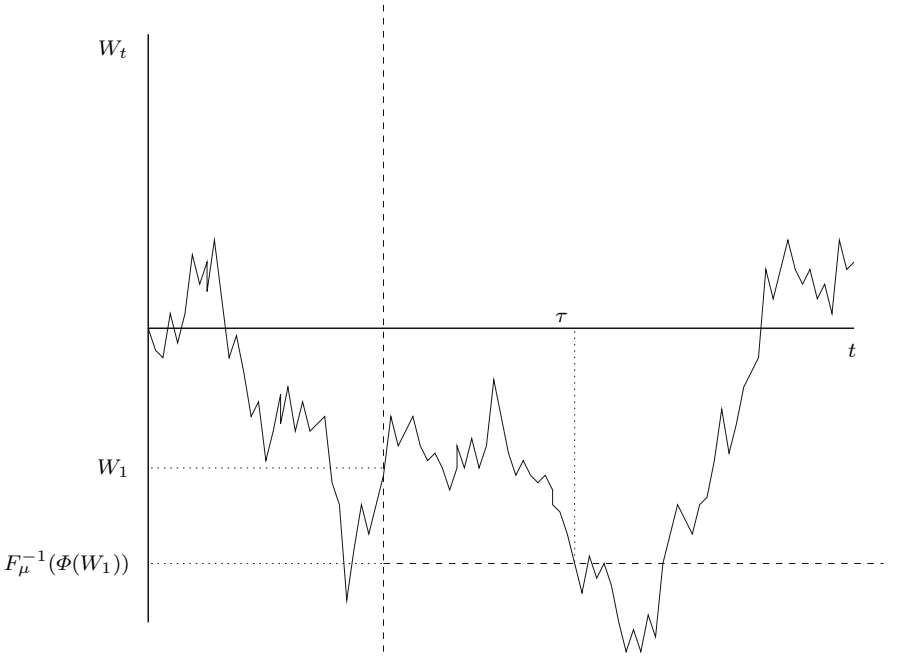


Fig. 4 The Doob construction. Run the Brownian motion until time 1. Conditional on the value at time 1, construct a new target level and run the Brownian motion until it first reaches this level. For a carefully chosen rule for selecting the target level, this construction embeds μ

Given μ, define F_μ to the distribution function $F_\mu(x) = \mu((-\infty, x]) = \mathbb{P}(X \leq x | X \sim \mu)$. Let Φ be the distribution function of a standard normal distribution, and let $Z = F_\mu^{-1}(\Phi(W_1))$. Then Z has law μ:

$$\mathbb{P}(Z \leq x) = \mathbb{P}(F_\mu^{-1}(\Phi(W_1)) \leq x) = \mathbb{P}(W_1 \leq \Phi^{-1}(F_\mu(x))) = F_\mu(x).$$

Now let $\tau = \inf\{u \geq 1 : W_u = F_\mu^{-1}(\Phi(W_1))\}$. Then $W_\tau \sim F_\mu^{-1}(\Phi(W_1)) \sim \mu$. See Fig. 4.

Note that in general $F_\mu^{-1}(\Phi(x)) \neq x$ and thus, conditional on $W_1 = x$, the expected value of τ is infinite. Then, unless μ has a standard normal distribution, $\mathbb{E}[\tau] = \infty$. In particular, $(W_{t\wedge\tau})_{t\geq 0}$ is not uniformly integrable.

3.3 Hall's Solution

Let $c = \int_0^\infty x\mu(dx)$. By the centring property we must also have $c = \int_{-\infty}^0 |x|\mu(dx)$. Choose a pair of random variables $U \in (-\infty, 0)$, $V \in [0, \infty)$ with joint law ρ where

$$\rho(du, dv) = \frac{(|u| + v)}{c}\mu(du)\mu(dv)$$

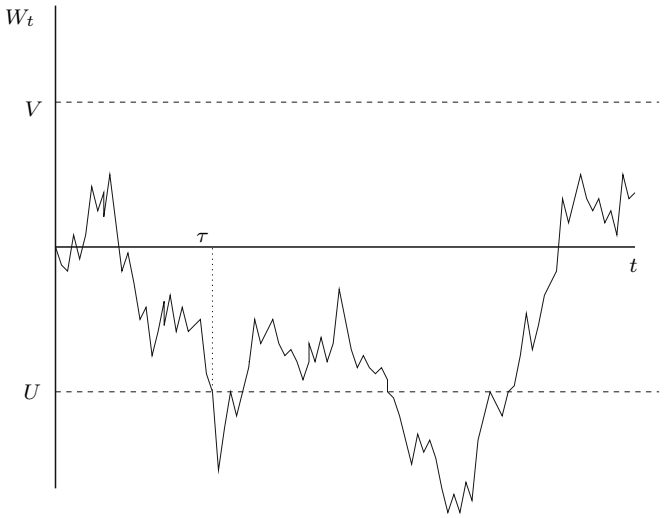


Fig. 5 The Hall construction. Given random variables $U < 0 < V$, we stop the Brownian motion the first time it leaves the interval $[U, V]$. For the appropriate joint distribution of (U, V), this leads to an embedding of μ

(For example, if μ is a uniform distribution on $(-1, 1)$ then $c = 1/4$ and $\rho(du, dv) = (|u| + v)$.)

Then, conditional on $(U = u, V = v)$, let $\tau_{u,v} = \inf\{t \geq 0 : W_t \notin (u, v)\}$. We claim that $\tau = \tau_{U,V}$ embeds μ. See Fig. 5.

We have, for $u < 0$,

$$\begin{aligned}\mathbb{P}(W_\tau \in du) &= \int_{v\in[0,\infty)} \mathbb{P}(U \in du, V \in dv)\mathbb{P}(W_{\tau_{u,v}} = u | U \in du, V \in dv) \\ &= \int_{v=0}^{\infty} \rho(du, dv)\frac{v}{|u| + v} = \mu(du)\left(\int_{v=0}^{\infty} \mu(dv)\frac{v}{c}\right) = \mu(du).\end{aligned}$$

A similar calculation applies for $v > 0$, and the Hall embedding is a solution of the SEP for (W, μ).

For this construction we can also calculate $\mathbb{E}[\tau]$. Again conditioning on the exogenous random variables U and V,

$$\begin{aligned}\mathbb{E}[\tau] &= \mathbb{E}[\mathbb{E}[\tau_{U,V}|U, V]] \\ &= \int_{-\infty}^{0}\int_{0}^{\infty} dv|u|v\rho(du, dv) \\ &= \int_{-\infty}^{0} \mu(du)\int_{0}^{\infty} \mu(dv)\frac{|u|^2 v + |u|v^2}{c}\end{aligned}$$

$$= \int_{-\infty}^{0} u^2 \mu(du) \int_{0}^{\infty} \mu(dv) \frac{v}{c} + \int_{-\infty}^{0} \frac{|u|}{c} \mu(du) \int_{0}^{\infty} \mu(dv) v^2$$
$$= \int_{-\infty}^{\infty} u^2 \mu(du),$$

so that at least in the case where μ has finite variance, $\mathbb{E}[\tau]$ is equal to that variance. As we shall see below this is best possible. In this sense, the Hall solution is optimal; however, it has the undesirable feature of requiring independent randomisation.

3.4 Properties of Good Solutions

Briefly we return to the general setting of a stochastic process X on I, and let τ and σ be stopping times.

Definition 4. The stopping time τ is *minimal* if $\sigma \leq \tau$ and $X_\sigma \sim X_\tau$ implies $\sigma = \tau$ almost surely.

Minimality is a natural concept for good solutions of the SEP in a general context, but now we return to the Brownian setting.

Theorem 3 (Monroe [42]).

1. *Minimal stopping times exist; given τ there exists a minimal stopping time σ with $W_\sigma \sim W_\tau$.*
2. *Hitting times are minimal; all stopping times with finite expectation are minimal.*
3. *Suppose τ is such that $\mathbb{E}[W_\tau] = 0$. Then τ is minimal if and only if $W_{t\wedge\tau}$ is uniformly integrable.*

The equivalence between uniformly integrable stopping times and minimal stopping times in the centred case allows us to translate an analytic condition which is useful for concluding that $\mathbb{E}[W_{\sigma\wedge\tau}] = 0$ into a path-wise condition describing good stopping times, where the adjective 'good' might better be expressed as 'not obviously inefficient'.

We prove some of Theorem 3.

Proof (that stopping times with finite expectation are minimal.).
If $\mathbb{E}[\tau] < \infty$ then $\mathbb{E}[W_\tau] = 0$ and $\mathbb{E}[W_\tau^2] = \mathbb{E}[\tau]$. If $\sigma \leq \tau$ and $W_\sigma \sim W_\tau$, then

$$\mathbb{E}[\sigma] = \mathbb{E}[W_\sigma^2] = \mathbb{E}[W_\tau^2] = \mathbb{E}[\tau]$$

and $\tau = \sigma$ almost surely. □

Corollary 1. *If μ has a second moment and $\tau \in \bar{\mathcal{T}}(\mu)$ then either $\mathbb{E}[\tau] = \int x^2 \mu(dx)$ or $\mathbb{E}[\tau] = \infty$.*

Proof (that if $W_{t\wedge\tau}$ is uniformly integrable, then τ is minimal.). Suppose that $\sigma \le \tau$ and $W_\sigma \sim W_\tau$. Then for all a,

$$\mathbb{E}[W_\tau - a; W_\tau \ge a] = \mathbb{E}[W_\sigma - a; W_\sigma \ge a] = \mathbb{E}[W_\tau - a; W_\sigma \ge a]$$

where we use the identity in law, and the martingale and uniformly integrability properties. Now $\mathbb{E}[W_\tau - a; A]$ is maximised over sets $A \subset \Omega$ by choices A for which, modulo null sets $(W_\tau > a) \subseteq A \subseteq (W_\tau \ge a)$. But $\mathbb{E}[W_\tau - a; W_\sigma \ge a]$ attains this maximum and hence modulo null sets, $(W_\tau > a) \subseteq (W_\sigma \ge a) \subseteq (W_\tau \ge a)$. Hence $W_\tau = W_\sigma$ almost surely.

Now let η be any other stopping time with $\sigma \le \eta \le \tau$. Then

$$W_\eta = \mathbb{E}[W_\tau|\mathcal{F}_\eta] = \mathbb{E}[W_\sigma|\mathcal{F}_\eta] = W_\sigma = W_\tau, \qquad \text{almost surely.}$$

Hence W is constant on (σ, τ), and since W has no intervals on constancy we must have $\sigma = \tau$ as required. □

In the centred Brownian case we write $\tau \in \mathcal{T}(\mu)$ if $\tau \in \bar{\mathcal{T}}(\mu)$ and τ is minimal.

Corollary 2. *If μ is centred and has support contained in an interval I and if $\tau \in \mathcal{T}(\mu)$ then τ is less than or equal to the first exit time from I.*

3.5 The Azéma–Yor Solution

The Azéma–Yor [2] solution is based on a stopping rule which is the first entrance of the joint process (W_t, J_t) into a domain $\mathcal{D}^J_\mu$. Let β be an increasing (but not necessarily continuous, nor strictly increasing) process such that $\beta(j) \le j$. (If j_0 is such that $\beta(j_0) = j_0$ then we insist that $\beta(j) = j$ for all $j \ge j_0$.) Let b be the left-continuous inverse to β so that $b = \beta^{-1}$, and let $\tau_\beta = \inf\{t > 0 : W_t \le \beta(J_t)\}$. See Fig. 6.

Our approach is based on Rogers [49]. By the representation of the Brownian path in terms of excursions down from the maximum we have that the probability that J_τ is greater than j is equal to the probability that no excursion down from the maximum gets below $\beta(J_\cdot)$ before J reaches j. Let n denote Itô excursion measure and let A_z be the set where the excursion down from the maximum when the maximum equals z gets below $\beta(z)$: $A_z = \{\max_{0<u<\xi_z}(e_z(u)) > z - \beta(z)\}$. Here ξ_z is the lifetime of the excursion down from the maximum which occurs when the maximum is at z, and $(e_z(u))_{0<u<\xi_z}$ is an excursion away from 0 of Brownian motion. Then

$$\mathbb{P}(J_\tau \ge j) = \mathbb{P}(n\,(\cup_{z\le j} A_z) = 0) = \exp\left(-\int_0^j \frac{dz}{z - \beta(z)}\right)$$

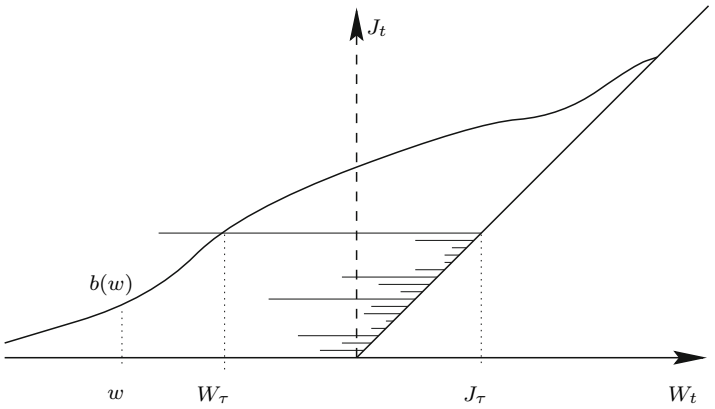


Fig. 6 The Azéma–Yor embedding. The figure shows a plot of the excursions of W_t down from the maximum, represented by the *horizontal lines* in (W_t, J_t) space. The Azéma–Yor stopping time is the first time that the Brownian path crosses the line given by the barycentre b

where we use the fact that the rate of excursions whose maximum modulus exceeds a is a^{-1}, and the fact that for a Poisson random variable Z of mean λ, $\mathbb{P}(Z = 0) = e^{-\lambda}$.

Having described the construction it remains to describe how to choose β so that τ is a solution of the Skorokhod embedding problem for (W, μ). Note that the law of J_τ has a continuous density, except perhaps for an atom at j_0. Moreover, for j such that $\beta(j) < j$,

$$(W_\tau > \beta(j)) \subseteq (J_\tau \geq j) \subseteq (W_\tau \geq \beta(j)), \tag{5}$$

and, at least where b is continuous, we have

$$\mathbb{P}(W_\tau \geq w) = \exp\left(- \int_0^{b(w)} \frac{dz}{z - \beta(z)} \right), \tag{6}$$

and, if b is differentiable,

$$-\frac{\mathbb{P}(W_\tau \in dw)}{\mathbb{P}(W_\tau \geq w)} = -\frac{b'(w)}{b(w) - w}.$$

If $W_\tau \sim \mu$ then $\mathbb{P}(W_\tau \geq w) = \mu([w, \infty)) =: \bar{\mu}(w)$, and if w is a continuity point then $(b(w) - w)\mu(dw) = b'(w)\bar{\mu}(w)$. It follows that

$$\begin{aligned} \frac{d}{dw}\left(\int_w^\infty z\mu(dz) \right) &= -w\mu(dw) = b'(w)\bar{\mu}(w) - b(w)\mu(dw) \\ &= \frac{d}{dw}\left(b(w) \int_w^\infty \mu(dz) \right). \end{aligned}$$

Integrating, and using the fact that at w equal to the lower limit on the support of μ we have $\int_w^\infty z\mu(dz) = 0$ we find that b is equal to the barycentre function:

$$b(w) = \frac{\int_w^\infty z\mu(dz)}{\int_w^\infty \mu(dz)} = \mathbb{E}^{X\sim\mu}[X|X \geq w]. \tag{7}$$

Example 1. If $\mu \sim U[-1,1]$ then $b(w) = (w+1)/2$ (for $-1 \leq w \leq 1$) and $\beta(j) = 2j-1$ for $0 \leq j \leq j_0 = 1$.

If E_1 is an exponential random variable rate 1 and μ is the centred distribution given by $\mu \sim (E_1 - 1)$ then $b(w) = w+1$ (for $w > -1$) and $\beta(j) = j-1$ for $j < j_0 = \infty$.

If $\mu = (p\delta_1 + p\delta_{-1} + (1-2p)\delta_0)$ with $p < 1/2$ then for $w \leq -1$, $b(w) = 0$; for $-1 < w \leq 0$, $b(w) = p/(1-p)$ and for $0 < w < 1$, $b(w) = 1$. Then β is also piecewise constant (for $j \leq 1$): for $0 \leq j < p/(1-p)$ we have $\beta(j) = -1$; for $p/(1-p) \leq j < 1$ we have $\beta(j) = 0$, and $\beta(j) = j$ for $j \geq 1$.

We have that the Azéma–Yor construction gives a solution of the SEP for (W, μ). The question arises, is it minimal? By Monroe's result (Theorem 3), it is equivalent to ask if $W_{t\wedge\tau}$ is uniformly integrable, and for the Azéma–Yor stopping rule it is possible to verify by calculation that this is the case.

By a Theorem of Azéma, Gundy and Yor [4] a (necessary and) sufficient condition for uniform integrability is that $\lim_{x\uparrow\infty} \mathbb{P}(\sup_{t\leq\tau} |W_t| \geq x) \to 0$. For the Azéma–Yor stopping time,

$$\begin{aligned}\mathbb{P}(\sup_{t\leq\tau}|W_t| \geq x) &= \mathbb{P}(H_{-x} < H_{b(-x)}) \\ &\quad + \mathbb{P}(H_{b(-x)} < H_{-x})\mathbb{P}\left(\sup_{t\leq\tau} W_t \geq x \,\middle|\, H_{b(-x)} < H_{-x}\right) \\ &= \frac{b(-x)}{x+b(-x)} + \frac{x}{x+b(-x)}\frac{\mathbb{P}(W_\tau \geq \beta(x))}{\mathbb{P}(W_\tau \geq -x)}.\end{aligned}$$

Now $b(-x) \to 0$ and $\bar\mu(-x) \to 1$ so that

$$\begin{aligned}\lim_{x\uparrow\infty} x\mathbb{P}(\sup_{t\leq\tau}|W_t| \geq x) &= \lim_{x\uparrow\infty} x\bar\mu(\beta(x)) = \lim_{y\uparrow\infty} b(y)\bar\mu(y) \\ &= \lim_{y\uparrow\infty} \mathbb{E}[Y|Y \geq y]\mathbb{P}(Y \geq y) = \lim_{y\uparrow\infty} \mathbb{E}[Y; Y \geq y] = 0.\end{aligned}$$

Thus, if we are given μ and we define the barycentre function and its inverse, and the Azéma–Yor stopping time relative to this inverse, then $W_{t\wedge\tau}$ is uniformly integrable. However, if we consider a general increasing function β (for example, $\beta(j) = 1 - 1/j$), then there is no reason for the corresponding stopping rule to be associated with a uniformly integrable stopped process $W_{t\wedge\tau}$.

Note that from (7) we have

$$(b(w) - w)\mathbb{P}(W_\tau \geq w) = \int_w^\infty (z-w)\mu(dz) = \mathbb{E}[(W_\tau - w)^+].$$

Furthermore, by Doob's submartingale inequality, for any uniformly integrable stopping time, and any j,

$$\begin{aligned} 0 = \mathbb{E}[W_\tau - j; J_\tau \geq j] &= \mathbb{E}[W_\tau - w; J_\tau \geq j] + (w-j)\mathbb{P}(J_\tau \geq j) \\ &\leq \mathbb{E}[W_\tau - w, W_\tau \geq w] + (w-j)\mathbb{P}(J_\tau \geq j). \end{aligned} \tag{8}$$

Then

$$\mathbb{P}(J_\tau \geq j) \leq \inf_{w<j} \frac{\mathbb{E}[(W_\tau - w)^+]}{j-w} = \inf_{w<j} \frac{C_\mu(w)}{j-w}. \tag{9}$$

This is suggestive of the model independent bound $\inf_{K<B} C(K)/(B-K)$ from (2).

There is equality in (8) if and only if

$$(W_\tau > w) \subseteq (J_\tau \geq j) \subseteq (W_\tau \geq w). \tag{10}$$

Comparing (10) with (5) we see that there is equality in (8) for the Azéma–Yor construction provided $w = \beta(j)$. We have the following result:

Proposition 1. *Amongst all minimal solutions of the SEP for (W, μ) the Azéma–Yor stopping rule maximises $\mathbb{P}(J_\tau \geq j)$ simultaneously for all j.*

It is crucial in the proposition that we restrict attention to uniformly integrable stopping times. If we consider all embeddings then the problem is degenerate, and for each $j > 0$ there is an embedding for which $\mathbb{P}(J_\tau \geq j) = 1$. To see this run the Brownian motion until the first return to zero after the first hit on j, and then use a favourite embedding thereafter.

A corollary of the proposition is that there exists a model for which the forward price is continuous and the price of a one-touch digital option is equal to the upper bound $C(K^*)/(B-K^*)$. Moreover, the bound is attained by a process in which the sets where the maximum is large correspond exactly to the sets where the terminal value is large.

3.6 Solutions of the SEP and Candidate Price Processes

Let $(M_t)_{t\geq 0}$ be a continuous martingale, null at 0, such that $M_T \sim \mu$. Then by the Dambis–Dubins–Schwarz Theorem (see e.g. Revuz and Yor [48, Theorem V.1.6]), $M_t = W_{\langle M \rangle_t}$ so that $\langle M \rangle_T$ is a solution of the SEP for (W, μ).

Conversely, if $\tau \in \mathcal{T}(\mu)$ then

$$M_t = W_{t/(T-t)\wedge\tau}$$

is a martingale with $M_T \sim \mu$.

Further, if μ is an integrable but non-centred measure (with mean m_μ), then we can construct a centred measure μ^0 by a parallel shift $\mu^0((-\infty, x]) = \mu((-\infty, x + m_\mu])$. If M_T^0 and W_τ^0 embed μ^0, then $m_\mu + M_T^0$ and $m_\mu + W_\tau^0$ embed μ.

Putting these two ideas together, there is a one-to-one correspondence between minimal solutions of the Skorokhod Embedding Problem, and candidate martingale price processes whose marginal distribution at T is consistent with call prices. If μ is supported on $\mathbb{R}^+$, then by Corollary 2, $(M_t)_{0\le t\le T}$ is non-negative.

3.7 Azéma Martingales, the Azéma–Yor Embedding and Super-Replication

Suppose we are given an increasing (differentiable) function $F(j)$, and consider the problem of maximising $\mathbb{E}[F(J_\tau)]$ over minimal embeddings of μ. For ease of exposition suppose that μ is such that the barycentre b and its inverse β are continuous.

Given a positive function h define $H(j, w) = \int_0^j h(r)(r-w)dr$. The Azéma martingale is $H(J_t, W_t) = \int_0^{J_t} h(r)(r - W_t)dr$. It is easy to see from Itô's formula that

$$dH(J_t, W_t) = -\left(\int_0^{J_t} h(r)dr\right) dW_t + h(J_t)(J_t - W_t)dJ_t.$$

The final term in this expression is zero since $J_t = W_t$ whenever the finite variation process J_t is not constant, and hence $H(J_t, W_t)$ is a martingale.

We aim to design H such that

$$\arg\max_j\{F(j) - H(j, w)\} = b(w).$$

This requires the choice $h(j) = F'(j)/(j - \beta(j))$ and then

$$H(j, w) = \int_0^j F'(r)\frac{(r - w)}{r - \beta(r)}dr = F(j) - F(0) + \int_0^j F'(r)\frac{(\beta(r) - w)}{r - \beta(r)}dr,$$

so that

$$F(j) - H(j, w) = F(0) + \int_0^j F'(r)\frac{(w - \beta(r))}{r - \beta(r)}dr.$$

We want to maximise this expression. Note that the integrand is positive for $\beta(r) < w$ (i.e. $r < b(w)$) and negative for $\beta(r) > w$ (i.e. $r > b(w)$). Hence the expression is maximised at $j = b(w)$ and

$$F(j) - H(j, w) \le F(0) + \int_0^{b(w)} F'(r)\frac{(w - \beta(r))}{r - \beta(r)}dr \equiv G(w)$$

where the above is taken as the definition of G. In particular, $F(j) \leq G(w) + H(j,w)$ (uniformly in $j > 0, w \in \mathbb{R}$) and applying this result at (J_τ, W_τ) we obtain

$$F(J_\tau) \leq G(W_\tau) + H(J_\tau, W_\tau),$$

on a path-by-path basis, with equality for the Azéma–Yor embedding.

Taking expectations, then provided τ is such that $H(J_{t\wedge\tau}, W_{t\wedge\tau})$ is uniformly integrable, it follows that $\mathbb{E}[F(J_\tau)] \leq \mathbb{E}[G(W_\tau)]$. If we further restrict attention to minimal stopping times which are solutions of the SEP for (W,μ) then

$$\sup_{\tau\in\mathcal{T}(\mu)} \mathbb{E}[F(J_\tau)] \leq \int_{\mathbb{R}} G(x)\mu(dx).$$

Although we have derived this result for differentiable F and regular μ it is clear that it can be extended to arbitrary (increasing) functions. For example, for the increasing function $F(j) = I_{\{j\geq B\}}$ we have

$$\sup_{\tau\in\mathcal{T}(\mu)} \mathbb{P}[J_\tau \geq B] \leq \mathbb{E}\left[\frac{W_\tau - \beta(B)}{B-\beta(B)} I_{\{W>\beta(B)\}}\right].$$

This yields another proof of Proposition 1.

3.8 The Perkins Solution

The Azéma–Yor stopping time is based on a stopping rule which compares the value of the Brownian motion to a function of the running maximum. The stopping time has the property that it maximises the law of the maximum amongst the class of uniformly integrable embeddings.

The Perkins [47] embedding has the property that it minimises the law of the maximum of the stopped Brownian motion, amongst all embeddings (There is no need to impose a minimality criterion here; if τ is not minimal then there exists $\sigma \leq \tau$ with $W_\sigma = W_\tau$ and then $J_\sigma \leq J_\tau$). In fact the Perkins embedding has the property that it simultaneously minimises the law of the maximum J_τ and maximises the law of the minimum $I_\tau = \inf\{W_s : 0 \leq s \leq \tau\}$.

Our interpretation of the Perkins construction is based on the following path-wise inequality, which plays an analogous role to (1):

$$I_{\{H_B\leq\tau\}} \geq I_{\{W_\tau\geq B\}} + \frac{(W_\tau - B)^+}{(B-K)} - \frac{(K-W_\tau)^+}{(B-K)} + \frac{W_{H_B} - W_\tau}{(B-K)} I_{\{H_B\leq\tau\}} \quad (11)$$

There is equality in (11) if on $H_B > \tau$ we have $K \leq W_\tau < B$ and on $H_B \leq \tau$ we have either $W_\tau \geq B$ or $W_\tau \leq K$.

Under the assumption that τ is minimal, taking expectations in (11) we find

$$\mathbb{P}(J_\tau \geq B) \geq \mathbb{P}(W_\tau \geq B) + \frac{1}{B-K}\left(C_\mu(B) - P_\mu(K)\right) \tag{12}$$

where $P_\mu(K) = \mathbb{E}^{X\sim\mu}[(K - X_T)^+]$. Since (12) holds for any K we can find a best bound of this type by taking the supremum over $K < B$, and then we find that K is chosen to maximise

$$\frac{C_\mu(B) - P_\mu(K)}{B-K}. \tag{13}$$

The final bound is

$$\mathbb{P}(J_\tau \geq B) \geq |C'_\mu(B-)| + \sup_{K<B} \frac{1}{B-K}\left(C_\mu(B) - P_\mu(K)\right). \tag{14}$$

If we can find a model such that the bound is attained, then it will follow that (14) is a best bound.

Fix B and suppose $X \sim \mu$. Then there exists $\hat{K}$ and A with $(X < \hat{K}) \cup (X \geq B) \subseteq A \subseteq (X \leq \hat{K}) \cup (X \geq B)$ such that

$$0 = \int_A (k-B)\mu(dk) = \int_{k\geq B} (k-B)\mu(dk) + \int_{A\setminus(X\geq B)} (k-B)\mu(dk). \tag{15}$$

Suppose τ_1 is such that $W_{\tau_1} \in \{\hat{K}, B\}$. On $W_{\tau_1} = B$ we embed the mass μ restricted to A; on $W_{\tau_1} = \hat{K}$ we embed the mass μ restricted to A^c. Let τ_P be this second embedding, then τ_P embeds μ. By construction, $(J_{\tau_P} \geq B) \equiv (W_{\tau_1} = B)$, and from (15), $C_\mu(B) = (B-\hat{K})(\mu(A) - \mu([B,\infty))) + P_\mu(\hat{K})$. Then,

$$\begin{aligned}
\mathbb{P}(J_{\tau_P} \geq B) \;=\; \mu(A) &= \mu([B,\infty)) + \frac{C_\mu(B) - P_\mu(\hat{K})}{(B-\hat{K})} \\
&\leq \mu([B,\infty)) + \sup_{K<B} \frac{C_\mu(B) - P_\mu(K)}{(B-K)} \\
&\leq \inf_{\tau\in\mathcal{T}(\mu)} \mathbb{P}(J_\tau \geq B) \\
&\leq \mathbb{P}(J_{\tau_P} \geq B)
\end{aligned}$$

Hence, there is equality throughout, and $\hat{K}$ maximises (13).

The above construction gives a solution of the Skorokhod embedding which minimises $\mathbb{P}(J_\tau \geq B)$ for a fixed B. The Perkins embedding has the property that it minimises $\mathbb{P}(J_\tau \geq B)$ over all embeddings simultaneously for all $B > 0$. (It also minimises $\mathbb{P}(I_\tau \leq C)$ for all $C < 0$, where $I_t = \inf\{W_u; u \leq t\}$.) Define

$$\alpha(j) = \arg\max_{\alpha<0} \frac{C_\mu(j) - P_\mu(\alpha)}{j-\alpha} \qquad \gamma(i) = \arg\min_{\gamma>0} \frac{P_\mu(i) - C_\mu(\gamma)}{\gamma - i}$$

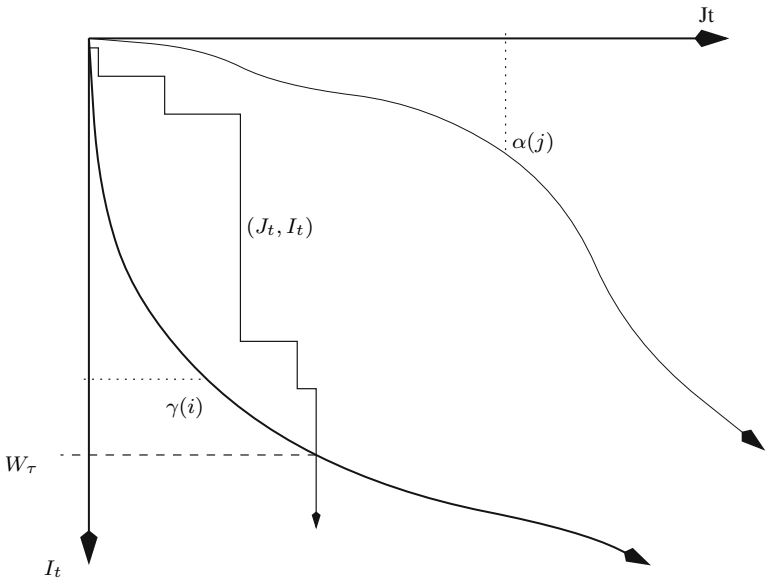


Fig. 7 The Perkins embedding. Given monotonic $\alpha(j)$ and $\gamma(i)$, the stopping rule $\tau = \inf\{u{:}W_u \notin \{\alpha(J_u), \gamma(I_u)\}$ is such that when W is stopped it is either at a new maximum, or at a new minimum

Let $\tau = \inf\{U : W_u \notin (\alpha(J_u), \gamma(I_u))\}$ Then $W_\tau \sim \mu$ and $\mathbb{P}(J_\tau \geq B)$ attains the lower bound from (14) uniformly in B. See Fig. 7.

3.9 Financial Interpretation of the Perkins Construction

Suppose that S_t is a forward price, and that calls are traded with maturity T on S. Suppose that such calls are traded with a continuum of strikes K with prices $C(K)$ which are consistent with no-arbitrage. This is equivalent to assuming we know the marginal distribution of S_T.

One candidate model for which option prices are given by C is when the price process $(S_t)_{0\leq t<T}$ is constant and then there is a jump at T, taken such that $\mathbb{E}[(S_T - K)^+] = C(K)$. For this price process, for $B > S_0$, $\mathbb{P}(J_T^S \geq B) = \mathbb{P}(S_T \geq B)$, where $J_t^S = \sup_{0\leq u\leq t} S_u$. This is the lowest possible value for $\mathbb{P}(J_T^S \geq B)$ given the law of S_T.

Now suppose that the price process S_t is assumed to be continuous. The effect of this assumption is to ensure that if we buy S the first time that the price gets to level B or above, then the price paid is B.

Consider the one-touch digital option with barrier $B > S_0$ and maturity T. In particular, consider the payoff $\tilde{F} = I_{\{\sup_{0\leq t\leq T} S_t \geq B\}}$, and consider the problem of finding a highest, model-independent, lower bound on the price of this option.

In the setting of the general theory we have $F_K = (S_T - K)^+$, with $f_K = C(K)$, and $\Theta \supseteq \{cI_{(H_B,T]}\}$. Then, modulo translation to non-centred random variables, the analysis of the previous section applies and

$$\mathcal{P}(\tilde{F}) \geq |C'(B-)| + \sup_{K<B} \frac{C(B) - P(K)}{(B-K)},$$

with equality for any model for which $(S_T < \hat{K}(B)) \subseteq (H_B \leq T) \setminus (S_T \geq B) \subseteq (S_T \leq \hat{K}(B))$ where $\hat{K}(B) = \arg\max_{K<B}(C(B) - P(K))/(B-K)$.

3.10 Barrier Options

As another example of a model-independent bound for an exotic option consider a knock-in barrier call option on the forward price S_t with maturity T, strike K and barrier $B > \max\{S_0, K\}$. The option has payoff $(S_T - K)^+ I_{\{H_B \leq T\}}$, paid at T.

If we place ourselves in the setting of Sect. 2.7 then we have the fundamental inequality, for $k \in [K, B)$;

$$(S_T - K)^+ I_{\{H_B \leq T\}} \leq \frac{B-K}{B-k}(S_T - k)^+ + \frac{k-K}{B-k}(S_{H_B} - S_T) I_{\{H_B \leq T\}}.$$

The first term is a semi-static hedging strategy in calls, and the second is the result of a forward investment in the underlying. This yields the price bound on the barrier option

$$\mathcal{P}((S_T - K)^+ I_{\{H_B \leq T\}}) \leq (B-K) \inf_{k \in [K,B)} \frac{C(k)}{B-k} \tag{16}$$

Recall the definition of the barycentre function b in (7) and its inverse β. To emphasize the dependence of these quantities on the law μ representing the marginal distribution of S_T we use a subscript and write b_μ and β_μ. Then, for $K \geq \beta_\mu(B)$ we find that the infimum in (16) is attained at $k = K$ and the upper bound on the knock-in barrier option is equal to the price of the call without the knock-in feature. However, for $K < \beta_\mu(B)$ the upper bound is strictly less than the call, and is equal to $(B-K)C(\beta_\mu(B))/(B-\beta_\mu(B))$. It is possible to construct a model based on the Azéma–Yor embedding and a time-change to show that this price can be attained.

For the lower bound, if there are no assumptions on the underlying then we have the trivial inequality

$$(S_T - K)^+ I_{\{H_B \leq T\}} \geq (S_T - K)^+ I_{\{S_T \geq B\}}$$

with associated price bound

$$\mathcal{P}((S_T - K)^+ I_{\{H_B \leq T\}}) \geq C(B) + (B-K)|C'(B-)|.$$

This price bound is attained for a jump model with a single jump at time T.

If we are prepared to assume that the underlying price process is continuous then we can refine the lower bound (This assumption has the effect that $S_{H_B} \equiv B$, so that if we go short the forward the first time that the underlying equals or exceeds the barrier, then the gains from trade are $(B - S_T)$). For $k \in [K, B)$ we have the inequality

$$(S_T - K)^+ I_{\{H_B \le T\}} \ge (S_T - K)^+ - \frac{B-K}{B-k}\left\{(S_T - k)^+ - (S_T - B)^+\right\} + (B-K) I_{\{S_T \ge B\}} + \frac{k-K}{B-k}(S_{T\wedge H_B} - B)$$

(To see this, split into two cases, $(H_B \le T)$ and $(H_B > T)$, and observe that the right-hand-side is piecewise linear, so that it is sufficient to check that the inequality holds where there are changes in value or kinks, i.e. at K, k and B.) Then

$$\mathcal{P}((S_T - K)^+ I_{\{H_B \le T\}}) \ge C(K) + (B-K)|C'(B-)| - \inf_{k \in [K,B)} \left\{ \frac{B-K}{B-k}(C(k) - C(B)) + \frac{(k-K)(B-S_0)}{(B-k)} \right\}$$

This bound can be attained by a time-change of the Perkins solution to the Skorokhod embedding problem.

See Brown et al. [10] for a further discussion of barrier options, including knock-out barriers, and knock-in and knock-out puts, and Cox and Oblój [18, 19] for a discussion of two-sided barriers.

3.11 Potential Theory and the Chacon–Walsh Embedding

To date we have supposed that the initial law of W is trivial (and then by a parallel shift we may as well assume that $W_0 \sim \delta_0$). Now we suppose that $W_0 \sim \mu_0$. Given centred probability measures μ_0 and μ_1, we ask when does there exist a stopping time τ such that

$$W_0 \sim \mu_0, W_\tau \sim \mu_1 \text{ and } W_{t\wedge\tau} \text{ is uniformly integrable?} \tag{17}$$

Clearly, if we can decide when it is possible to solve (17), and if we can find constructions where they exist, then we can solve the iterated problem: given $(\mu_i)_{0\le i\le n}$ find an increasing sequence of stopping times $0 = \tau_0 \le \tau_1 \le \ldots \le \tau_n$ (where possible) such that

$$W_{\tau_i} \sim \mu_i, \qquad W_{t\wedge\tau_n} \text{ is uniformly integrable.} \tag{18}$$

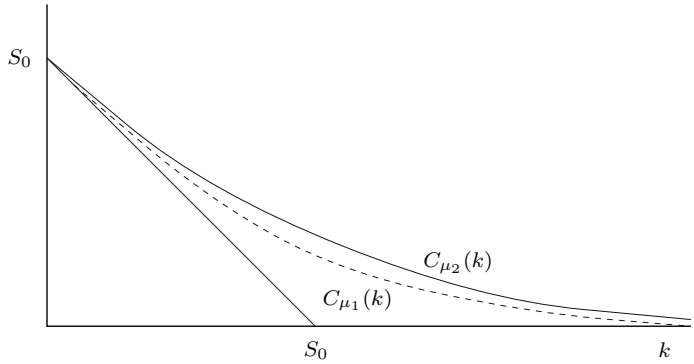


Fig. 8 Given a pair of call price functions satisfying the no-arbitrage conditions, there exists a model for the forward price which is consistent with those prices provided the call prices are increasing in maturity

Theorem 4 (Röst [52]). *A necessary and sufficient condition for the existence of a solution to* (17) *is* $C_{\mu_0}(K) \le C_{\mu_1}(K)$ *for all* K *(Fig. 8).*

Proof. Necessity follows from an application of the conditional Jensen's inequality. Since $(w-K)^+$ is convex as a function of w,

$$C_{\mu_1}(K) = \mathbb{E}[(W_\tau - K)^+] = \mathbb{E}[\mathbb{E}[(W_\tau - K)^+|\mathcal{F}_0]] \ge \mathbb{E}[(W_0 - K)^+] = C_{\mu_0}(K).$$

Sufficiency follows from one of the constructions given below. □

Put-call parity is a consequence of the identity $(y-k)^+ - (y-k)^- = y-k$. Given also $(y-k)^+ + (y-k)^- = |y-k|$ we have $2(y-k)^+ = |y-k| + (y-k)$. Hence $C_{\mu_0}(K) \le C_{\mu_1}(K)$ is equivalent to $U_{\mu_0}(K) \ge U_{\mu_1}(K)$ where $U_\eta(K) = -\mathbb{E}^{X\sim\eta}|X-K|$ is the potential (or minus the price of a straddle). Then the fact that C is a convex function, with $C_\eta(k) \ge k^-$ implies that U_η is concave and $U_\eta(x) \le -|x|$. Kinks in the derivative U'_η correspond to atoms of η; conversely straight segments of U_η correspond to intervals where η places no mass.

The Chacon–Walsh solution of the SEP is based on potentials. We return to the case where $W_0 = 0$ and μ is a centred probability measure, but it is immediate from the inductive nature of the construction that provided the potentials are decreasing, then the same construction will work as an embedding for any initial law.

Let $U_0(x) = U_{\delta_0}(x) = -|x|$, and $\tau_0 = 0$.

Fix $a_1 < 0 < b_1$. Set $\tau_1 = \inf\{t > \tau_0 : W_t \notin (a_1, b_1)\}$, $\mu_1 \sim \mathcal{L}(W_{\tau_1})$ and $U_1 \equiv U_{\mu_1}$. Then $U_1(x) = -|x|$ for $x \le a$ and $x \ge b$, and $U_1(x) = a_1 - (b_1 + a_1)(x-a_1)/(b_1-a_1)$ for $x \in (a_1, b_1)$. See Fig. 9.

We now proceed inductively. Fix $a_n < b_n$. Set $\tau_n = \inf\{t > \tau_{n-1} : W_t \notin (a_n, b_n)\}$, $\mu_n \sim \mathcal{L}(W_{\tau_n})$, $U_n \equiv U_{\mu_n}$. Then U_n is a continuous concave function with $U_n(x) = U_{n-1}(x)$ outside (a_n, b_n) and U_n linear on (a_n, b_n).

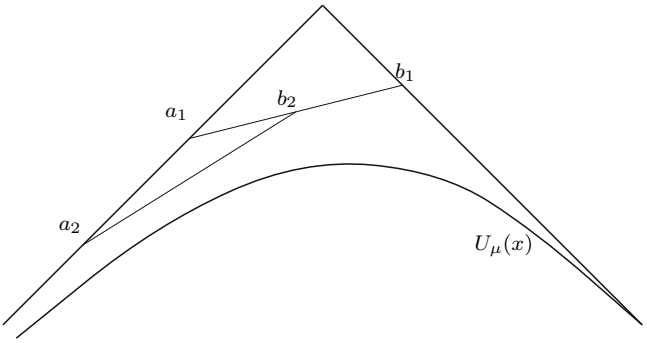


Fig. 9 Potentials in the Chacon–Walsh picture

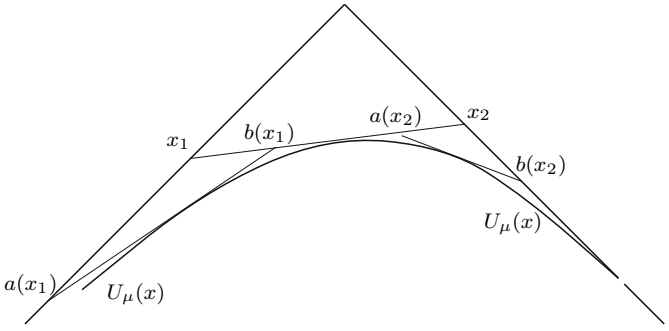


Fig. 10 The Dubins construction in the Chacon–Walsh picture. Shown are the potentials at the first and second stages of the iteration, together with the starting and terminal potentials

The idea is to choose the points (a_n, b_n) appropriately in such a way that $U_n \downarrow U_\mu$. Then the increasing sequence of stopping times τ_n converges (to τ say) and $W_\tau \sim \mu$.

There is a large amount of freedom in the choice of the sequence (a_n, b_n). A natural choice is given by the Dubins [22] solution to the SEP. In the Dubins construction the sequence (a_n, b_n) is defined as follows.

Let $\mathcal{A}_0 = \{0\}$. Given U_μ let $(a(0), b(0))$ be the x-coordinates of the points where the tangent to U_μ at x crosses $U_{\delta_0}(x) = -|x|$. (If there is an atom of μ at zero, then U_μ has a kink there, and there are many candidate tangents to U_μ. We are free to choose any of these; for definiteness we can choose the tangent with slope equal to the right derivative of U_μ at x_0.) Set $\mathcal{A}_1 = \{a(0), b(0)\}$ and let μ_1 be the law with corresponding potential $U_1(x) = \min\{U_0(x), U_\mu(0) + xU'_\mu(0+)\}$. See Fig. 10.

Given U_μ, $\mathcal{A}_n$ and potential U_n, for each $x \in \mathcal{A}_n$ let $(a(x), b(x))$ be the x-coordinates of the points where the tangent to U_μ (with slope $U'_\mu(x+)$) intersects U_n. Note that if $x_i < x_j \in \mathcal{A}_n$ then $a(x_i) \le b(x_i) \le a(x_j) \le b(x_j)$. Then let $\mathcal{A}_{n+1} = \{a(x_i), b(x_i); x_i \in \mathcal{A}_n\}$, and μ_{n+1} the law with corresponding potential $U_{n+1} = \min_{x_i \in \mathcal{A}_n}\{U_n(x), U_n(x_i) + (x - x_i)U'_n(x_i+)\}$.

The Azéma–Yor embedding [2] can also be phrased in terms of a Chacon–Walsh potential picture. Rather than considering a sequence of intervals, consider a continuum of intervals, parameterised by j, such that $b(j) = j$ and $a(j)$ is the x-coordinate of the point where the tangent to U_μ passes through the point $(j, -j)$. (Note that $a(j)$ is not necessarily unique, but it is increasing, so we can make it unique by making the function a right continuous.) We have

$$\tau_j = \inf\{t \geq \tau_{j-}; W_t \notin (a(j), j)\}$$

A few lines of calculus show that $a(j)$ is the inverse barycentre function.

3.12 The Skorokhod Solution in the Potential Picture

Skorokhod's solution [54] of the Skorokhod embedding problem can also be represented in the Chacon–Walsh picture. The construction is similar to the Hall solution, and the idea is to choose a single exogenous random variable L, to use that random variable to define an interval, and to set τ to be the first time that Brownian motion leaves that interval.

Let $c = \int_{(0,\infty)} x\mu(dx)$, so that $U_\mu(0) = -2c$. For $l \in (0, 2c)$, let $V_+(l)$ be the x-coordinate of the point in $\mathbb{R}^+$ such that the tangent to U_μ at $V_+(l)$ crosses the y-axis at $y = -l$. (If U_μ has linear sections then there may be an interval of such $V_+(l)$, in which case we take the largest. With this definition the increasing function $V_+(l)$ is right-continuous.) Similarly, define the decreasing function $V_-(l)$ to be the x-coordinate of the point in $\mathbb{R}^-$ such that the tangent to U_μ at $V_+(l)$ crosses the y-axis at $y = -l$. See Fig. 11.

If μ has a continuous distribution function then we take L to have distribution given by $\mathbb{P}(L \geq l) = \mu((V_-(l), V_+(l))$, for $0 \leq l \leq 2c$. More generally, to allow for atoms we take $\mathbb{P}(L \geq l)$ to equal one half the difference in slopes

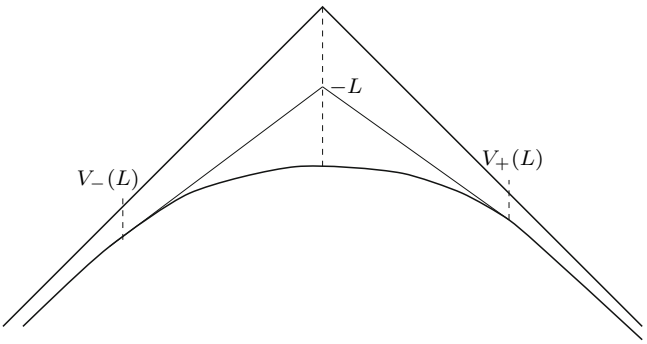


Fig. 11 The Skorokhod embedding in the potential picture. The points $V_\pm(l)$ are determined by the (unique) tangents to U_μ which pass through $(0, -l)$

between the tangents joining $(V_-(l), U_\mu(V_-(l)))$ with $(0, -l)$ and $(0, -l)$ with $(V_+(l), U_\mu(V_+(l)))$. Finally we set $\tau_l = \inf\{u : W_u \notin (V_-(l), V_+(l))\}$, and then $W_{\tau_L} \sim \mu$.

Rather than prove this result, we give an example. Suppose μ is the law of a uniform random variable on $(-1, 1)$. Then $U_\mu(x) = -(x^2 + 1)/2$, $c = 1/4$, and $V_\pm(l) = \pm\sqrt{1 - 2l}$. Then, by symmetry, for $w \in (0, 1)$,

$$\mathbb{P}(W_{\tau_L} \le w) - \mathbb{P}(W_{\tau_L} \ge 0) = \frac{1}{2}\mathbb{P}(V_+(L) \le w) = \frac{1}{2}\mathbb{P}(L \ge (1 - w^2)/2) = \frac{w}{2}$$

as required.

Remark 2. The original Skorokhod construction assumes the existence of a non-negative exogenous random variable Z, and then sets τ to be the first time the Brownian motion leaves the interval $(\rho(Z), Z)$ for a non-positive decreasing function ρ. From our representation of the Skorokhod solution in the potential picture we see $Z \sim V_+(L)$ and $\rho(z) = V_-(V_+^{-1}(z))$.

3.13 The Bass Solution

Whereas our philosophy has been to construct solutions of the Skorokhod embedding problem for Brownian motion with the aim of generating models for the martingale price process via a time-change, Bass [5] constructs a (time-inhomogeneous) martingale diffusion such that $X_1 \sim \mu$, and then uses this to construct a solution of the SEP for (W, μ). As such the Bass construction is interesting in the finance context because it gives directly (a family) of diffusion processes with marginal law μ at time 1.

Suppose that on $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0 \le t \le 1}, \mathbb{P})$, the diffusion process Z solves the stochastic differential equation $dZ_t = \eta(Z_t, t)dB_t + \gamma(Z_t, t)dt$ with $Z_0 = z$, and is such that the law of Z_1 is known and given by the atom-free centred distribution function F_Z. Suppose further that we can calculate the time-1 law of Z_t from any starting point $(Z_t, t)_{0 \le t < 1}$. (The simplest case is when Z is itself a Brownian motion, but other constructions could be based on a Brownian bridge, a Bessel process or an Ornstein–Uhlenbeck process.) Given μ, (and associated distribution function F_μ), fix $g(z) = F_\mu^{-1}(F_Z(z))$. Then g is increasing and $g(Z_1) \sim \mu$.

Now set $X_t = \mathbb{E}[g(Z_1)|\mathcal{F}_t]$. Then $X_t = h(Z_t, t)$, for a function h which is increasing in its first argument, and has inverse $Z_t = H(X_t, t)$ say. Then $X_1 \sim \mu$ and

$$dX_t = h'(Z_t, t)\eta(Z_t, t)dB_t = a(X_t, t)dB_t$$

where $a(x, t) = h'(H(x, t), t)\eta(H(x, t), t)$. Note that $X_0 = \mathbb{E}[g(Z_1)] = 0$.

Given a Brownian motion W and the function a, define the increasing additive functional Γ and its inverse A via

$$\Gamma_0 = 0, \qquad \frac{d\Gamma_u}{du} = \frac{1}{a(W_u, \Gamma_u)^2}, \qquad A_t = \inf\{u : \Gamma_u > t\}.$$

Set $X_t = W_{A_t}$. Then, with $u = A_t$,

$$\frac{dA_t}{dt} = \frac{1}{d\Gamma_u/du} = a(W_u, \Gamma_u)^2 = a(X_t, t)^2,$$

and for some Brownian motion $\hat{W}_t$,

$$dX_t = \sqrt{dA_t/dt}\ \ d\hat{W}_t = a(X_t, t)d\hat{W}_t.$$

Finally, since $X_1 \sim \mu$, we have $W_{A_1} \sim \mu$ and hence $\tau = A_1$ is a solution of the Skorokhod embedding problem for (W, μ).

3.14 A General Schematic for Solutions

Solutions of the Skorokhod embedding problem can be classified, very approximately, into three main types.

The first type (such as Dubins [22] and also Azéma–Yor [2]) fall under the general heading of potential theoretic constructions, and have a pictorial representation in the Chacon–Walsh [13] picture. It is also possible to place the original embedding due to Skorokhod in this category.

For the second type of embedding the aim is to solve the SEP subject to an additional optimality criterion. In particular, given some functional $H(W_t; 0 \le t \le \tau)$, the aim is to solve the SEP for (W, μ) such that H is maximised, in the sense of first order stochastic dominance, or sometimes in the weaker sense of maximised in expectation. As we have seen, the Azéma–Yor [2] and Perkins [47] embeddings fall into this category, but so do the Root/Röst [51, 53] constructions (based on the additive functional t), the Vallois [55, 56] embedding (based on the local time) and the Obłój–Yor [44] solutions (based on the age of the excursion), amongst others. We will discuss the Root and Röst solutions in detail below, because of the direct financial interpretations.

Finally, there are miscellaneous embeddings which fall into neither of the above categories. These include the Hall [29] and Bass [5] solutions described above, and also the Bertoin–Le Jan [7] solution.

4 Other Applications of the Skorokhod Embedding Problem

4.1 The Central Limit Theorem

These notes are primarily concerned with the connection between Skorokhod embeddings and derivative pricing and hedging. However, there is one other important application which deserves comment, namely the use of Skorokhod embeddings to prove the Central Limit Theorem.

Let $X_1, X_2, \dots$ be a sequence of independent identically distributed random variables, with law μ such that μ has mean zero and variance 1. Let $S_n = \sum_{k \le n} X_k$.

Let $B^{(i)}$ be Brownian motions, and let $\tau^{(i)}$ be any of the minimal embeddings of μ in $B^{(i)}$ described to date. The times $\tau^{(i)}$ are independent identically distributed random variables with mean 1, and $B^{(i)}_{\tau^{(i)}}$ are independent. The sequence $(X_i)_{i\ge 1} = (B^{(i)}_{\tau^{(i)}})_{i\ge 1}$ has the properties listed in the previous paragraph.

Let $T_0 = 0$ and $T_j = \sum_{i \le j} \tau^{(i)}$. Now define a Brownian motion $(W_t)_{t\ge 0}$ by

$$W_t = \sum_{i \le j} B^{(i)}_{\tau^{(i)}} + B^{(j+1)}_t \qquad T_j \le t < T_{j+1} \tag{19}$$

Then $W_{T_n} = S_n$, and by Brownian scaling, $S_n/\sqrt{n} = W_{T_n}/\sqrt{n} \sim W_{T_n/n}$.

The idea is that, since $T_n/n \to 1$ and since at fixed times Brownian motion has Gaussian marginals, then in the limit $S_n/\sqrt{n}$ will inherit a standard normal distribution.

Theorem 5. *With the notation and hypotheses of this section $S_n/\sqrt{n}$ converges in law to a standard normal distribution.*

Proof. Let W be the Brownian motion defined in (19). Let $Z^{(n)}_t = W_{nt}/\sqrt{n}$. Then, for each n, $Z^{(n)}$ is a Brownian motion. By the above construction we have

$$\frac{S_n}{\sqrt{n}} = \frac{W_{T_n}}{\sqrt{n}} = Z^{(n)}_{T_n/n} \sim Z_{T_n/n},$$

for a Brownian motion Z. Fix $\epsilon > 0$ and choose $\delta > 0$ so that

$$\mathbb{P}\left(\sup_{t:|1-t|<\delta} |Z_t - Z_1| > \epsilon \right) < \frac{\epsilon}{2}$$

By the strong law of large numbers $T_n/n \to 1$ almost surely, so that we can choose N large, such that for all $n \ge N$,

$$\mathbb{P}(|T_n/n - 1| > \delta) < \frac{\epsilon}{2}$$

Combining these last two displayed equations we conclude that

$$\mathbb{P}(|W_{T_n}/\sqrt{n} - W_n/\sqrt{n}| > \epsilon) = \mathbb{P}(|Z_{T_n/n} - Z_1| > \epsilon) < \epsilon$$

Then $\mathbb{P}(W_n/\sqrt{n} < x - \epsilon) - \epsilon \le \mathbb{P}(S_n/\sqrt{n} < x) \le \mathbb{P}(W_n/\sqrt{n} < x + \epsilon) + \epsilon$ and the result follows.

4.2 Non-Centred Target Laws

There is a very easy way to embed integrable, but non-centred, target laws in Brownian motion started at zero – simply run Brownian motion until it hits the mean, and thereafter use a favourite embedding for a centred distribution. Such an embedding is minimal, see Cox and Hobson [17]. Note however that whereas in the centred case if μ has support in an interval $I = [a, b]$ then for a minimal embedding $\tau \leq H_a \wedge H_b$, in the non-centred case this is no longer true. If the mean of the target distribution is positive then it is no longer possible to have an embedding for which $\tau \leq H_z$ almost surely, for any $z < 0$. (If it were, then $(W_{t\wedge\tau})_{t\geq 0}$ would be bounded below and Fatou's lemma would imply that the law of W_τ had negative mean.)

4.3 Embeddings in Time-Homogeneous Diffusions

Suppose that X is a time-homogeneous diffusion with state space $I \subseteq \mathbb{R}$ with absorbing or inaccessible endpoints, and consider the SEP for (X, μ). The two fundamental questions are: when is it possible to embed μ, and if it is possible, how can it be done?

In fact the problem can easily be reduced to the case of Brownian motion. The relevant insight is due to Azéma and Yor [3], see also Pedersen and Peskir [46] and Cox and Hobson [15].

Suppose that X solves $dX = a(X)dB + b(X)dt$. If s solves $a(x)^2 s''(x)/2 + b(x)s'(x) = 0$, then s is a scale function for X and $Y = s(X)$ is a local martingale with state space $s(I)$. We have $dY = s'(s^{-1}(Y))a(s^{-1}(Y))dB$. We can choose s such that $s(x_0) = 0$ and then by the Dambis–Dubins–Schwarz Theorem $Y_t = W_{\Gamma_t}$ for some Brownian motion W, where $\Gamma_t = \langle Y \rangle_t$.

Let ν be the image of μ under s. If we can solve the problem for (W, ν) then we can construct a solution τ for (Y, ν) via time-change, and then τ is a solution of the SEP for (X, μ). Thus, if $W_\tau \sim \mu$ and if A is the inverse to Γ, then $Y_{A_\tau} = W_\tau \sim \nu$, and $X_{A_\tau} \sim s(Y_{A_\tau}) \sim \mu$.

Note that there is no reason to expect that ν is centred as a measure on $\mathbb{R}$. By the remarks in the previous section we can construct embeddings for non-centred target distributions provided we do not impose any restrictions that τ is less than the first exit time from some interval. However, for the purposes of embedding in non-martingale diffusions we must have that τ is smaller than the first exit time from $s(I)$. If we set $m_\nu = \int_{\mathbb{R}} x\nu(dx)$, then we find that the answer to the question about whether ν can be embedded in Y depends on the joint properties of $s(I)$ and m_ν:
If $s(I) = \mathbb{R}$, then there is a solution for the SEP for any ν,
If $s(I) = (-a, \infty)$ then there is a solution for the SEP for ν if and only if $m_\nu \leq 0$,
If $s(I) = (-\infty, b)$ then there is a solution for the SEP for ν if and only if $m_\nu \geq 0$,
If $s(I) = (-a, b)$ then there is a solution for the SEP for ν if and only if $m_\nu = 0$.

5 The Root and Röst Solutions and Options on Volatility

5.1 The Root and Röst Embeddings

Return to the classical setting of minimal solutions of the SEP for (W, μ) where μ is a centred probability measure. Suppose μ has finite variance. Then there are embeddings τ such that $\mathbb{E}[\tau] = \int_{\mathbb{R}} x^2 \mu(dx)$, (and as described in Corollary 1 above, for any embedding either $\mathbb{E}[\tau] = \int_{\mathbb{R}} x^2 \mu(dx)$, or τ has infinite expectation). In particular, it does not make sense to attempt to find the solution of the SEP for which $\mathbb{E}[\tau]$ is minimised; any minimal stopping time will satisfy this criterion. The task of maximising $\mathbb{E}[\tau]$ over minimal stopping times is equally trivial.

However, the problem of searching for the embedding which minimises (over all stopping times) the variance of τ, or alternatively the embedding which maximises (over uniformly integrable stopping times) the variance of τ leads to a unique solution.

Since we are looking for an embedding which optimises a functional of t evaluated at the stopping time, it is plausible that the stopping rule should depend on the joint distribution of W_t and the additive functional t.

Given $b : \mathbb{R} \to \mathbb{R}_+$, define $\mathcal{R} \equiv \mathcal{R}_b = \{(t, x); t \geq b(x)\}$. Let $\tau_b = \inf\{u : (u, W_u) \in \mathcal{R}_b\}$. We will describe the function b (and sometimes the region $\mathcal{R}_b$) as a barrier (Fig. 12).

Theorem 6 (Root [51], Röst [53]).

1. *Given a centred probability measure μ there exists b such that τ_b is a solution of the SEP for (W, μ); moreover τ_b is minimal.*

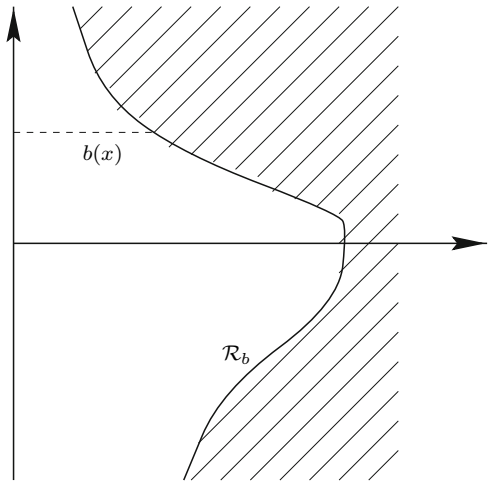


Fig. 12 The Root barrier solution. Given $b : \mathbb{R} \to \mathbb{R}_+$, or equivalently $\mathcal{R} \equiv \mathcal{R}_b = \{(t, x); t \geq b(x)\}$, let $\tau_b = \inf\{u : (u, W_u) \in \mathcal{R}_b\}$ be the time of first entry by (u, W_u) into the shaded region. The idea is to choose b such that $W_\tau \sim \mu$

2. *τ_b minimises the variance of τ amongst all embeddings of the law of W_{τ_b}. More generally τ_b minimises $\mathbb{E}\left[\int_{t\wedge\tau}^{\tau} h(W_s)ds\right]$ for all non-negative functions h and times t.*

The embedding property of the solution is due to Root [51], and henceforth we will label this solution as the Root solution. The optimality properties were conjectured by Kiefer [38] and proved by Röst [53]. Taking $h = 1$ we find that the Root solution minimises the *residual expectation* $\mathbb{E}[(\tau - t)^+] \equiv \mathbb{E}[(\tau - \tau \wedge t)]$ for all t. Since any convex function can be decomposed into a positive combination of such functions (and since $\mathbb{E}[\tau]$ is invariant across minimal embeddings) it follows that the Root barrier solution minimises $\mathbb{E}[f(\tau)]$ for any convex f.

Example 2. 1. Suppose $\mu \sim N(0,1)$. Then the barrier is given by $b(x) = 1$, $\mathcal{R}_b = \{(t,x); t \geq 1\}$ and $\tau_b = 1$.
2. Suppose $\mu \sim (\delta_{-1} + \delta_{+1})/2$. Then $\mathcal{R}_b = \{(x,t) : x \leq -1, x \geq 1\}$, and $\tau_b = H_{\pm 1}$.
3. Suppose $\mu \sim p\delta_{-1} + p\delta_{+1} + (1-2p)\delta_0$ for $0 \leq p \leq 1/2$. Then $\mathcal{R}_b = \{(x,t) : x \leq -1, x \geq 1\} \cup \{(0,t); t \geq t_0(p)\}$. The quantity $t_0(p)$ does not have a simple expression, but it is clear that $t_0(0) = 0$, $t_0(1/2) = \infty$ and that $t_0(p)$ is strictly increasing in p. Hence the existence of a barrier which embeds μ is guaranteed.

There is also a converse to the Root embedding, which is based on a reverse barrier (Fig. 13). The embedding is attributed to Röst by Meilijson [41], see also Oblój [43].

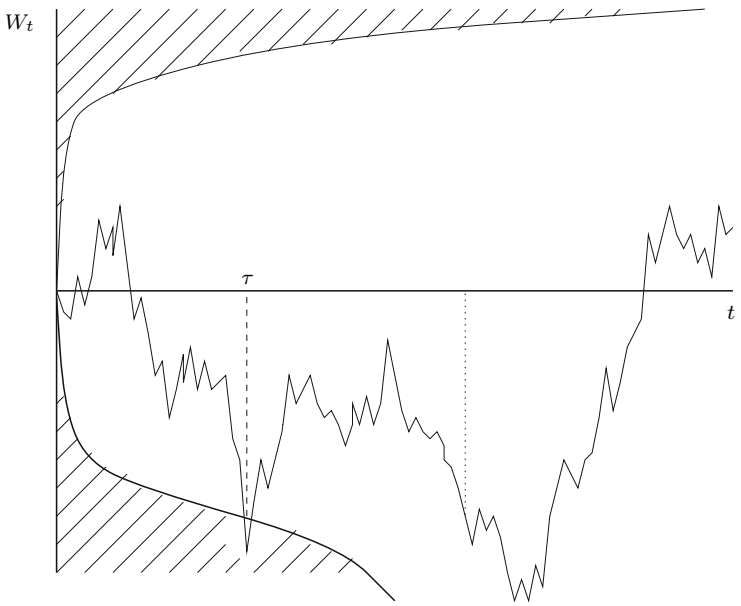


Fig. 13 The Röst reversed barrier. The embedding τ is the first time Brownian motion W enters the shaded region

We assume that μ has no atom at zero. The construction can be extended to include an atom at 0 by independent randomisation.

Theorem 7 (Röst). *Given $b : \mathbb{R} \to \mathbb{R}_+$, define $\bar{\mathcal{R}} \equiv \bar{\mathcal{R}}_b = \{(t,x); t \le b(x)\}$. Let $\bar{\tau}_b = \inf\{u : (u, W_u) \in \bar{\mathcal{R}}_b\}$. Then,*

1. *Given a probability measure μ with no atom at 0, there exists b such that $\bar{\tau}_b$ is a solution of the SEP for (W, μ); moreover $\bar{\tau}_b$ is minimal;*
2. *$\bar{\tau}_b$ maximises the variance of τ amongst all uniformly integrable embeddings of the law of $W_{\bar{\tau}_b}$. More generally $\bar{\tau}_b$ maximises $\mathbb{E}[\int_{t\wedge\tau}^{\tau} h(W_s)ds]$ for all non-negative functions h and times t.*

Remark 3. In fact both the Root and Röst embeddings can be extended to any continuous Markov process $(X_t)_{t\ge 0}$, and the optimality properties still hold. This will be clear from the fact that the proofs we sketch below do not use any properties of Brownian motion beyond the Markov property and continuity of paths.

Sketch of proof of Theorem 6. *Suppose μ consists of a finite collection of atoms at locations $\mathcal{X} = \{x_0 < x_1 < \ldots < x_N < x_{N+1}\}$, with associated probabilities $\{p_0, p_1, \ldots, p_N, p_{N+1}\}$. Suppose also that μ is centred, so that $\sum_{i=0}^{N+1} x_i p_i = 0$.*

We search for an embedding for (W, μ) of barrier form. Since the embedding is minimal, by Corollary 2 we must have that the process is stopped before it leaves the interval $[x_0, x_{N+1}]$. Furthermore, the process should only stop at the points $x_i \in \mathcal{X}$, hence the barrier $\mathcal{R}$ must be of the form $\mathcal{R} = \{(x,b); x_i \in \mathcal{X}, b \ge b_i\}$. Moreover, we must have $b_0 = 0 = b_{N+1}$.

Let $\gamma = \{\gamma_1, \ldots, \gamma_N\}$ be a vector in $\mathbb{R}_+^N$, and augment it with the values $\gamma_0 = 0 = \gamma_{N+1}$. Let $\tau_\gamma = \inf\{u : B_u = x_i, \quad u \ge \gamma_i; 0 \le i \le N+1\}$. Set $\Gamma_\mu = \{\gamma : \mathbb{P}(B_{\tau_\gamma} = x_i) \le p_i; 1 \le i \le N\}$. Note that if $\gamma \in \Gamma_\mu$ we must have $\mathbb{P}(B_{\tau_\gamma} = x_0) + \mathbb{P}(B_{\tau_\gamma} = x_{N+1}) \ge p_0 + p_{N+1}$ to compensate. Then $\gamma \in \Gamma_\mu$ if the associated stopping rule never embeds more than allowed under μ at any point in the interior of the support of μ. The excess mass is embedded at the endpoints.

We claim that if $\hat{\gamma}$ and $\tilde{\gamma}$ are elements of Γ_ν then so is $\underline{\gamma}$ where $\underline{\gamma}_i = \hat{\gamma}_i \wedge \tilde{\gamma}_i$. To verify the claim, fix $i \in \{1, \ldots, N\}$. Without loss of generality we may assume that $\hat{\gamma}_i \le \tilde{\gamma}_i$. Then $\tau_{\underline{\gamma}} = \tau_{\hat{\gamma}}$ on $B_{\tau_{\underline{\gamma}}} = x_i$, and $\tau_{\underline{\gamma}} \le \tau_{\hat{\gamma}}$ otherwise. See Fig. 14. Thus $(\omega : B_{\tau_{\underline{\gamma}}} = x_i) \subseteq (\omega : B_{\tau_{\hat{\gamma}}} = x_i)$ and $\mathbb{P}(B_{\tau_{\underline{\gamma}}} = x_i) \le p_i$. Since i was arbitrary, the claim is proved.

It follows from the claim that Γ_μ has a minimal element. Moreover, this minimal element must embed μ; if not then for some i we must have $\mathbb{P}(B_{\tau_\gamma} = x_i) < p_i$, and for this i we can reduce the value of γ slightly without violating the condition $\mathbb{P}(B_{\tau_\gamma} = x_i) \le p_i$ and only making the values of $(\mathbb{P}(B_{\tau_\gamma} = x_j))_{j\ne i}$ smaller.

It remains to extend from atomic distributions μ_n to general μ. Root [51] achieves this by mapping $[0,\infty] \times [-\infty, \infty]$ to $[0,1] \times [-1,1]$ and putting a distance metric on barriers expressed as sets in the new space. Given the images of the barriers

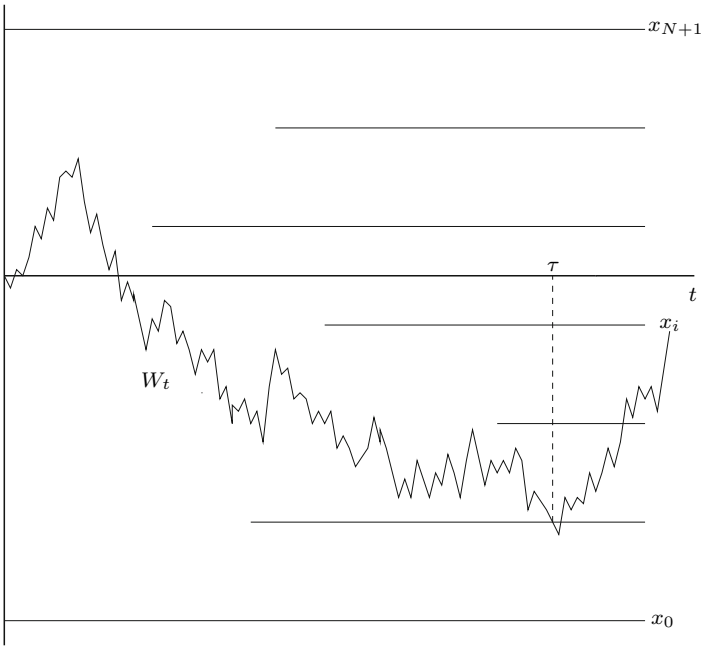


Fig. 14 The embedding property of the Root barrier for atomic measures μ on a finite number of points. The stopping time τ_γ is the first time the Brownian path is at one of the x_i at a time later than γ_i. If γ is such the probability of stopping at x_i is less than p_i, then making γ_i smaller, will increase the probability that $B_{\tau_\gamma} = x_i$, but can only decrease the probability of stopping at any other point

associated with μ_n, there must be a convergent subsequence in the compact space $[0,1]\times[-1,1]$. This must take the form of a barrier, and mapping back to the original space we obtain the solution of the SEP for μ. □

Sketch of proof of optimality of the Root construction. Fix x and t and suppose that $\tau \in \mathcal{T}(\mu)$. We show that the Root barrier stopping time maximises $\mathbb{E}[L^x_{t\wedge\tau}]$ (simultaneously for all x and t) amongst all minimal solutions of the Skorokhod embedding problem for μ.

If so, then the Root stopping time maximises $\mathbb{E}[\int_0^{t\wedge\tau} h(W_s)ds]$ for all positive h, and since $\mathbb{E}[\int_0^{\tau} h(W_s)ds]$ is constant across minimal embeddings of μ, it also minimises $\mathbb{E}[\int_{t\wedge\tau}^{\tau} h(W_s)ds]$.

Let τ be any embedding of μ and suppose that there are paths with $(W_\tau \in dx, \tau < t)$ and $(W_\sigma \in dx, t \le \sigma < \tau)$. Then we can take the terminal mass of W_τ from those paths which pass through (σ, x) and embed that mass by extending the paths from (τ, x) which are currently stopped at $(W_\tau \in dx, \tau < t)$. See Fig. 15. This modification of the stopping rule must increase the value of the local time $L^x_{t\wedge\tau}$.

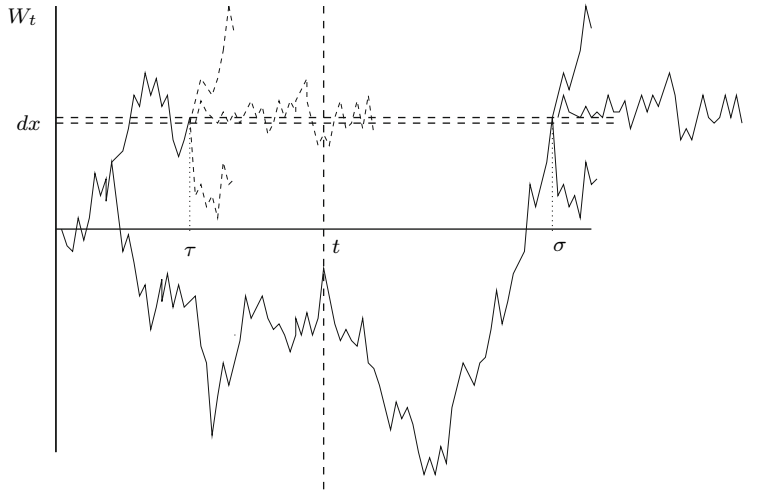


Fig. 15 Optimality of the Root construction. If for some (x,t) there are paths which stop at (x,τ) with $\tau < t$, and there are other paths which continue from (x,σ) with $t < \sigma < \tau$, then by shifting the continuations of paths from (x,σ) so that they are continuations from (x,t) with $t < \tau$, we can only increase the time spent at x before time t. We stop some paths which pass through (x,σ) early, and extend other paths at (x,τ) to compensate

Hence, if τ is optimal, in the sense that it maximises $\mathbb{E}[L^x_{t\wedge\tau}]$ amongst minimal solutions of the Skorokhod embedding problem for μ, then for any pair (x,t) we have that either the stopping rule never involves stopping paths at x before t, or, no paths cross x after t.

Let $\mathcal{R}_t = \{x : \text{no paths cross } x \text{ after } t\}$. Then $\mathcal{R}_t$ is increasing in t and $\mathcal{R} = \cup_{t\geq 0}\mathcal{R}_t$ defines a barrier.

□

The justification for the embedding property and optimality of the Röst reverse barrier is similar. Suppose τ is such that there are paths with $(W_\sigma \in dx, \sigma < \tau < t)$ and $(W_\tau \in dx, \tau > t)$. Then we can take (some of) the terminal mass of W_τ from those paths which aren't stopped at σ (but do stop by time t) and embed that mass by continuing paths from (x,τ) which are currently stopped at $(W_\tau \in dx, \tau < t)$. This modification of the stopping rule must decrease the value of the local time $L^x_{t\wedge\tau}$.

Remark 4. Note that the proofs of the embedding property and of the optimality of the Root and Röst constructions do not rely on any properties of Brownian motion, beyond the Markov property and the continuity of paths. We will exploit this fact when we discuss applications to the pricing of volatility derivatives below.

Example 3. Suppose, that for $\alpha, \beta > 0$

$$\mu(dx) = \frac{\alpha\sqrt{\beta}e^{-\beta x^2/(2(x+\alpha))}}{\sqrt{2\pi}(x+\alpha)^{3/2}}\,dx; \qquad x > -\alpha$$

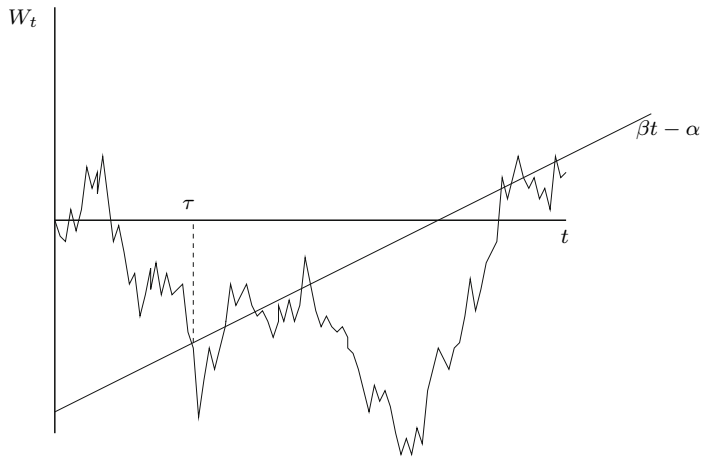


Fig. 16 The law μ is such that Brownian motion stopped the first time it hits the straight line $x = \beta t - \alpha$ has law μ

The associated barrier function is $b(x) = (x + \alpha)/\beta$, which corresponds to the straight line $x = \beta t - \alpha$. See Fig. 16.

Suppose h is the convex function $h(t) = e^{\eta t}$ (with $\eta < \beta^2/2$) and consider the problem of minimising $\mathbb{E}[h(\tau)]$ over stopping times τ which are solutions of the SEP for (W, μ).

Define the constants $\varphi = \beta - \sqrt{\beta^2 - 2\eta}$ (note that $\varphi > 0$) and $\lambda = 2\eta e^{\varphi\alpha}/\varphi^2$, and set $M(t, x) = \lambda - \lambda e^{\varphi x - \varphi^2 t/2}$, so that $M(0, 0) = 0$ and $M = (M(t, W_t))_{t \geq 0}$ is a martingale.

Define $\Gamma(x) = \inf_{t>0}[e^{\eta t} - M(t, x)]$. The infimum is attained at t solving

$$\eta e^{\eta t} = \frac{\lambda\varphi^2}{2} e^{\varphi x - \varphi^2 t/2}$$

(at least for $x > -\alpha$) which reduces to

$$t = \frac{2\varphi(x + \alpha)}{(\varphi^2 + 2\eta)} = \frac{x + \alpha}{\beta} = b(x),$$

where we use $(\varphi - \beta)^2 = \beta^2 - 2\eta$ so that $\varphi^2 + 2\eta = 2\varphi\beta$. It also follows that

$$\Gamma(x) = \left(1 + \frac{2\eta}{\varphi^2}\right) e^{\eta(x+\alpha)/\beta} - \frac{2\eta}{\varphi^2} e^{\varphi\alpha}.$$

Then, by construction,

$$h(t) \geq M(t, x) + \Gamma(x), \qquad \forall t, x,$$

with equality at $t = (x+\alpha)/\beta$. In particular, at $t = \tau$ we have $h(\tau) \geq M(\tau, W_\tau) + \Gamma(W_\tau)$ so that for any τ with $M(t \wedge \tau, W_{t\wedge\tau})$ uniformly integrable,

$$\mathbb{E}[h(\tau)] \geq \mathbb{E}[\Gamma(W_\tau)].$$

Further, for all τ which are solutions to the Skorokhod embedding problem for μ we have

$$\mathbb{E}[e^{\eta\tau}] \geq \int_{\mathbb{R}} \Gamma(x)\mu(dx) = e^{\alpha\varphi}$$

with equality for the Root embedding $\tau = \inf\{u : W_u = \beta u - \alpha\}$.

5.2 Pricing Options on Volatility

Let $(P_t)_{t\geq 0}$ be the price process of an asset, (not necessarily a forward price) and suppose P_t is continuous. Denote the quadratic variation by $\langle \ln P\rangle_t$. The problem is to derive robust, model independent bounds on the prices of call options on volatility, and, following Dupire [24] and Carr and Lee [12] the idea is to use to Root and Röst stopping times, and their optimality properties.

Suppose interest rates are deterministic. (For ease of exposition we will assume that they are constant, and equal to r.) Then $(S_t)_{t\geq 0}$ given by $S_t = e^{-rt}P_t$ is a forward price process, which we may assume to be a martingale under a pricing measure. Note that $\ln S_t = \ln P_t - rt$ so that $\langle \ln P\rangle_t = \langle \ln S\rangle_t$.

Write X_t as shorthand for $\ln S_t$. Then

$$\langle X\rangle_T = \int_0^T \left(\frac{dS_u}{S_u}\right)^2 = \int_0^T \left(\frac{dP_u}{P_u}\right)^2.$$

On the other hand,

$$X_T - X_0 = \int_0^T \frac{dS_u}{S_u} - \frac{1}{2}\int_0^T \left(\frac{dS_u}{S_u}\right)^2,$$

so that

$$\langle X\rangle_T = -2\ln S_T + 2\ln S_0 + \int_0^T \frac{2}{S_u} dS_u.$$

Suppose that call options are traded for all strikes so that $\ln S_T$ is effectively a traded asset. Suppose further that the strategy $\theta_t = 2/S_t$ is an admissible dynamic hedging strategy. Then the integrated squared volatility $\langle X\rangle_T$ has a model free price equal to $2(\ln S_0 - \ln S_T)$.

Now consider call options on $\langle X\rangle_T = \langle \ln P\rangle_T$. In particular, consider the security with payoff $(\langle X\rangle_T - Q)^+$.

Without loss of generality we may assume that $P_0 = S_0 = 1$, so that $X_0 = 0$. The goal is to derive model-independent bounds on the price of $(\langle X\rangle_T - Q)^+$, assuming that the law of S_T is given (by μ say).

Define $M_t = \int_0^t dS_t/S_t$. Then $X_t = M_t - \langle M\rangle_T/2 = M_t - \langle X\rangle_T/2$. We know that the continuous local martingale M can be written as a time-change of Brownian motion: $M_t = W_{\langle X\rangle_t}$ for some Brownian motion W. We also have

$$S_t = e^{X_t} = e^{W_{\langle X\rangle_t} - \langle X\rangle_t/2} = Z_{\langle X\rangle_t}$$

where $(Z_u)_{u\geq 0}$ is the exponential Brownian motion $Z_u = e^{W_u - u/2}$.

Since $S_T \sim \mu$ we have $Z_{\langle \ln S\rangle_T} \sim \mu$, so that $\langle \ln S\rangle_T$ is a solution of the Skorokhod embedding problem for (Z, μ). Conversely, if σ is a solution of the SEP for (Z, μ), and if $S_t = Z_{\sigma\wedge(t/(T-t))}$ then

$$S_T = Z_\sigma \sim \mu \qquad \text{and} \qquad \langle \ln S\rangle_T = \langle X\rangle_T = \sigma.$$

If σ is the Root solution of the SEP for (Z, μ) then we say that $(S_t)_{0\leq t\leq T}$ given by $S_t = Z_{\sigma\wedge t/(T-t)}$ is the Root model. This model is consistent with observed call prices as represented by μ. Then, combining the time-change with Theorem 6 and Remark 3 we have the following theorem.

Theorem 8 (Dupire [24], Carr–Lee [12]). *Suppose call prices satisfy $C(K) = C_\mu(K)$ for some distribution μ with mean S_0. Let σ_R be the Root barrier solution of the Skorokhod embedding problem for (Z, μ).*

Let $((S_t)_{0\leq t\leq T}, \mathbb{P}^R)$ be a model for the forward price for which $S_t = Z_{\sigma_R\wedge(t/(T-t))}$. Let $((S_t)_{0\leq t\leq T}, \tilde{\mathbb{P}})$ be any alternative model for which S_t is a martingale and $S_T \sim \mu$. Then, for any $Q \geq 0$,

$$\mathbb{E}^R[(\langle \ln S\rangle_T - Q)^+] \leq \tilde{\mathbb{E}}[(\langle \ln S\rangle_T - Q)^+]$$

and the price of a call option on quadratic variation for a given set of vanilla call prices is minimised under the Root model.

The Röst (reverse barrier) construction gives an upper bound.

5.3 A Path-Wise Hedging Strategy for Volatility Options

Theorem 8 gives bounds on the prices of call options on volatility, but it does not give an associated super-replicating or sub-replicating strategy. In particular, it does not give a path-wise inequality which could be used to enforce the bounds, in the way that the inequality (1) enforces no-arbitrage bounds on the prices of one-touch digitals. We show below that, subject to solving a variant on the heat equation with a rather strange 'boundary condition', a sub-replicating strategy exists.

Let call prices (for all strikes and maturity T) be such that $S_T \sim \mu$. (By scaling we may assume that $S_0 = 1$ and μ has support $\mathbb{R}^+$ and mean 1.) Suppose that the

Root solution to the Skorokhod embedding problem for (Z, μ) is associated with the barrier b_μ. Consider the problem of pricing an option with payoff $h(\langle \ln S \rangle_T)$ with $h(t)$ convex.

Theorem 9. *Suppose there exists $m(t,z)$ such that $m(0,1) = 0$, $\frac{z^2}{2} m'' + \dot{m} = 0$ and such that*

$$\arg\inf_{t \geq 0} [h(t) - m(t,z)] = b_\mu(z).$$

Suppose further that θ given by

$$\theta_t = m'(\langle \ln S \rangle_t, S_t)$$

is admissible, in the sense that $G^\theta = \int_0^T \theta_t dS_t \in G^\Theta$.

If we define $\Gamma(z) = \inf_{t>0}\{h(t) - m(t,z)\} = h(b_\mu(z)) - m(t, b_\mu(z))$ then

$$h(\langle \ln S \rangle_T) \geq \Gamma(S_T) + G^\theta,$$

path-wise, and thus

$$\mathcal{P}(h\langle \ln S \rangle_T) \geq \int_{\mathbb{R}^+} \Gamma(s)\mu(ds),$$

with equality for the Root model.

Proof. Since $h(t) - m(t,z) \geq \Gamma(z)$ uniformly in t and z we have

$$h(\langle \ln S \rangle_T) \geq \Gamma(S_T) + m(\langle \ln S \rangle_T, S_T).$$

But, by Itô's formula, and using $d\langle \ln S \rangle_t = (dS_t)^2 / S_t^2$,

$$\begin{aligned} m(\langle \ln S \rangle_T, S_T) &= m(0,1) + \int_0^T \left[\dot{m}\, d\langle \ln S \rangle_t + m' dS_t + \frac{S_t^2}{2} m'' \frac{(dS_t)^2}{S_t^2} \right] \\ &= \int_0^T m'(\langle \ln S \rangle_t, S_t) dS_t \\ &= G^\theta. \end{aligned}$$

□

Example 4. Suppose $h(t) = t$ and $m(t,z) = 2\ln z + t$. Then $m(0,1) = 0$ and $\dot{m} + z^2 m''/2 = 0$. Furthermore, $h(t) - m(t,z)$ is independent of t so that, for any μ with Root barrier $b_\mu(z)$,

$$\Gamma(z) = \inf_t \{h(t) - m(t,z)\} = h(b_\mu(z)) - m(b_\mu(z), z) = -2\ln z.$$

Then, provided θ given by $\theta_t = (2/S_t)$ is admissible we have

$$\mathcal{P}(\langle \ln S \rangle_T) = \mathcal{P}(-2\ln S_T) = \int_{\mathbb{R}^+} (-2\ln s)\mu(ds)$$

and the fair price of a security paying $\langle \ln S \rangle_T$ is model independent.

Example 5. In this second example the payoff is non-linear and Theorem 9 gives a model-independent bound, rather than the unique no-arbitrage price. We show how to construct a sub-replicating strategy which enforces the lower bound. The example is based on Example 3.

Suppose that $h(t) = e^{\eta t}$. Suppose further that $S_T \sim \mu$ where

$$\mu(dz) = \frac{\alpha\sqrt{\gamma}}{\sqrt{2\pi}(\alpha + \ln z)^{3/2}} \exp\left(\frac{(\alpha - (2\gamma - 1)\ln z)^2}{8\gamma(\alpha + \ln z)}\right) dz, \qquad z > e^{-\alpha}.$$

Then $b_\mu(z) = (\alpha + \ln z)/\gamma$.

Fix $v = e^{-\alpha}$ and let ψ be the smallest root of $\psi^2 - (1 + 2\gamma)\psi + 2\eta = 0$. (We assume that η is small enough so that this equation has a solution.) Define $\varphi = \psi(\psi - 1)/2$ and $\lambda = \eta\varphi^{-1}v^{-\psi}$.

Now set $m(t, z) = \lambda - \lambda z^\psi e^{-\varphi t}$. If $b(z) = \arg\inf_{t>0}[h(t) - m(t, z)]$ then $b(z)$ solves

$$\eta e^{\eta b(z)} = \varphi\lambda z^\psi e^{-\varphi b(z)}.$$

The choices of constants are such that $b(z) = b_\mu(z)$. Provided θ given by $\theta_t = -\lambda\psi S_t^{\psi-1} e^{-\varphi\langle X\rangle_t}$ is an admissible strategy then, in the notation of Theorem 9, $\Gamma(S_T) + G^\theta$ is a sub-replicating strategy for the volatility option with payoff $h(\langle X\rangle_T)$.

The model-independent lower bound on the price of the option is $\int_{\mathbb{R}^+} \Gamma(s)\mu(ds)$, where $\Gamma(s) = h(b(s)) - m(b(s), s)$.

6 Multiple Stopping Times; Option Price Data for Several Maturities

To date we have assumed that there are a continuum of option prices available, for vanilla European options of a single maturity. But what if options trade with two (or more) maturities? Are the option prices consistent with no arbitrage? Can the call prices from an earlier maturity be used to refine the price bounds? Is it possible to derive bounds for forward starting options?

In terms of the Skorokhod embedding problem, the equivalent problems relate to the existence and properties of solutions when the initial law of the Brownian motion is non-trivial. Recall from Theorem 4 that there exist solutions of the Skorokhod embedding problem for centred initial and terminal laws if and only if $C_\mu(x) \le C_\nu(x)$. This has an immediate Corollary in the finance setting; a necessary and sufficient condition for the absence of arbitrage between a pair of (decreasing, convex) call price functions is that for each fixed strike (in the bond numeraire), call prices are increasing in maturity. (There are extensions of this result to the case where only a finite number of strikes are traded, in which case the answer depends on the ability to extrapolate between the traded strikes to generate convex (in strike) call price functions whose prices are increasing in maturity. See Bühler [11], Davis and Hobson [20] or Cousot [14].)

So, suppose μ and ν are such $C_\mu(x) \le C_\nu(x)$, or equivalently $U_\mu(x) \ge U_\nu(x)$. The aim is to find stopping rules τ which maximise a functional $F(W_t, 0 \le \sigma \le t \le \tau)$ (or in the forward-starting version of the problem $F(W_t, \sigma \le t \le \tau)$) where $W_\sigma \sim \mu$ and $W_\tau \sim \nu$. the question arises: what are the appropriate generalisations of the Azéma–Yor (or Perkins, Root, Röst ...) embeddings to non-zero initial law? In the financial context, new derivatives become relevant, such as the forward-starting straddle option with payoff $|S_{T_2} - S_{T_1}|$. An analysis of this can be considered as a first step towards pricing the discrete time Asian option with payoff $(\sum_{i \le n} S_{T_i} - K)^+$, and thence its continuous time analogue.

6.1 Maximising the Law of the Maximum, with Intermediate Constraints

Consider the problem: find

$$\sup_\tau \mathbb{P}(J_\tau \ge j)$$

where the supremum is taking over uniformly integrable stopping times τ such that there exists $\sigma \le \tau$ for which

$$W_0 \sim \delta_0, \qquad W_\sigma \sim \mu, \qquad W_\tau \sim \nu. \tag{20}$$

This is directly linked to finding candidate models for which $S_{T_1} \sim \mu$ and $S_{T_2} \sim \nu$. It turns out that there are two cases, an easy case, and a hard case.

In the easy case, the barycentres b_μ and b_ν are ordered so that $b_\mu(x) \le b_\nu(x)$ for all x. See Fig. 17. Then the inverse barycentres β_μ and β_ν satisfy $\beta_\nu(j) \le \beta_\mu(j)$.

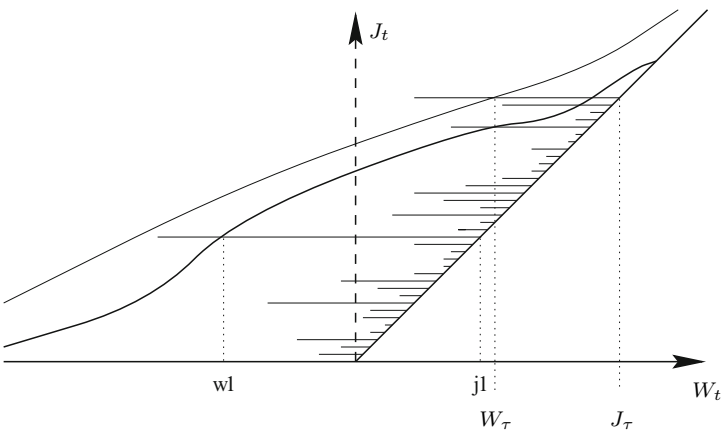


Fig. 17 A pair of barycentres. In this case the barycentres are ordered and it is possible to find σ and τ with $\sigma \le \tau$ where both separately represent the Azéma–Yor construction of the associated measure

In particular, if we define $\sigma = \tau_{b_\mu} = \inf\{u : W_u \le \beta_\mu(J_u)\}$ and $\tau = \tau_{b_\nu} = \inf\{u : W_u \le \beta_\nu(J_u)\}$ then $\sigma \le \tau$ embed μ and ν respectively.

Moreover, since τ is the Azéma–Yor embedding for ν, τ maximises $\mathbb{P}(J_\tau \ge j)$ over all minimal solutions of the Skorokhod embedding problem for (W, ν), and hence must maximise $\mathbb{P}(J_\tau \ge j)$ over the smaller set of embeddings which are subject to the intermediate constraint.

Unfortunately, although $U_\mu(x) \ge U_\nu(x)$ is a necessary condition for $b_\mu(x) \le b_\nu(x),\ \forall x$, it is not sufficient. Suppose for example $\mu \sim (\delta_{-1} + \delta_1)/2$ and $\nu \sim p\delta_{-2} + (1-2p)\delta_0 + p\delta_2$. For $1/4 \le p \le 1/2$ we have $-1 = U_\mu(0) \ge U_\nu(0) = -4p$ and then $U_\mu(x) \ge U_\nu(x)$ for all x. However, for $-1 < x < 0$ we have $b_\mu(x) = 1$ and $b_\nu(x) = 2p/(1-p)$ so that in order to have $b_\mu(x) \le b_\nu(x)$ we must have $p \ge 1/3$. For $p < 1/3$ the barycentres are not ordered.

It follows that the intermediate constraint $W_\sigma \sim \mu$ has an impact on the possible values of $\mathbb{P}(J_\tau \ge j)$. For more on this situation, and the implications for model-independent bounds, see Brown et al. [10].

A similar situation arises if we try to extend the Perkins [47], Root [51], Röst [41] or Vallois [55, 56] constructions to multiple time points. All these constructions utilise a stopping rule based on the first hitting time of the joint process (W_u, A_u) of a domain, where A_t is an additive functional of the Brownian path. If these domains posses a natural ordering, then the construction of optimal stopping times σ, τ satisfying (20) is straightforward. Otherwise the construction of the optimal stopping rule (i.e. the one which maximises A_τ) is much more challenging. See Fig. 18 for an illustration in the Root barrier case.

6.2 *Maximising the Law of the Maximum, with Non-Trivial Initial Law*

Now consider the problem:

$$\sup_\tau \mathbb{P}(J_\tau \ge j)$$

where the supremum is taking over stopping times τ such that

$$W_0 \sim \mu, \qquad W_\tau \sim \nu \tag{21}$$

where μ and ν are centred probability measures with $U_\mu(x) \ge U_\nu(x)$.

An alternative but equivalent problem is to search for pairs of stopping times $\sigma \le \tau$ for which (20) holds, but to only calculate the maximum over the interval (σ, τ). This corresponds to searching for model-independent derivative prices for forward-starting options, e.g. for barrier options where the option is knocked-in only if it is crossed in the period (T_1, T_2).

Several generalisations of solutions of embeddings to the situation with non-zero initial laws are known. Hobson [32] shows how to adapt the potential picture to extend the Azéma–Yor embedding to this case (see Fig. 19) and Hobson

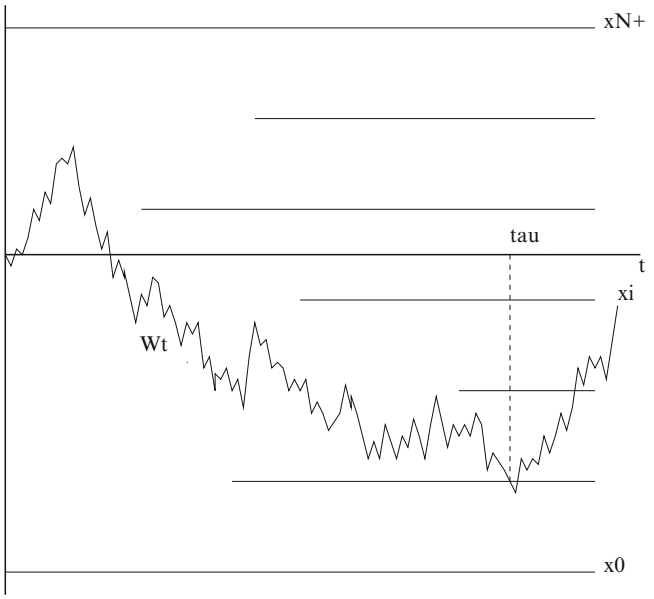


Fig. 18 A pair of Root barriers. Suppose we try to solve the problem (20) of constructing an embedding τ of ν subject to the existence of $\sigma \leq \tau$ for which $W_\sigma \sim \mu$. Let b_1 with associated stopping region $\mathcal{R}_{b_1}$ embed μ; for (different) $\hat{\nu}$ and $\tilde{\nu}$ let $\hat{b}_2$ and $\tilde{b}_2$ be the associated Root barriers. If the barriers are ordered, e.g. $\hat{b}_2 \geq b_1$, then we can define $\sigma = \inf\{u{:}(u, W_u) \in \mathcal{R}_{b_1}\}$ and $\hat{\tau} = \inf\{u : (u, W_u) \in \mathcal{R}_{\hat{b}_2}$ and the problem is solved. But, if the barriers are not ordered (e.g. b_1 and $\tilde{b}_2$), then a more complicated construction is needed

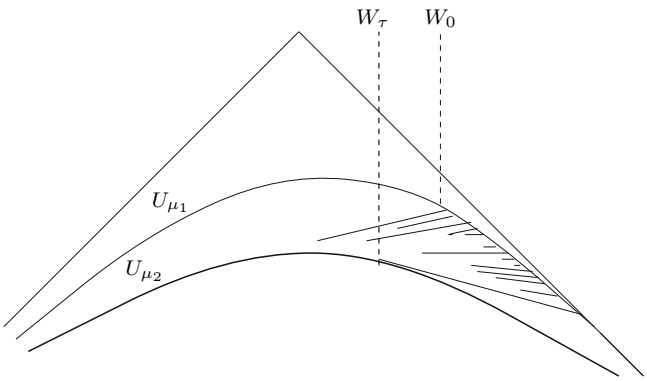


Fig. 19 Potential theory picture of Azéma–Yor story with non-trivial starting law. The idea is to run excursions down from the current maximum along tangents to the potential U_{μ_2} which intersect U_{μ_1} at the current value of the maximum. See Hobson [32]

and Pedersen [35] consider the extension of the Perkins embedding, and give an application to bounds on the prices of barrier options. Indeed, in the case of a Root [51] barrier solution, the same form of construction still holds, and $\tau = \inf\{u{:}(u, W_u) \in \mathcal{R}_b\}$ is an embedding for a suitably chosen barrier b depending on both μ and ν.

6.3 *Maximising the Law of the Maximum, with a Continuum of Marginals*

Suppose we are given the marginal distributions of a martingale X for every t. For there to exist a process with those marginals we must have that the potentials are decreasing in t, (or equivalently the call option prices are increasing in maturity). By the results of Krylov [39], Gyöngy [28] and Dupire [23], if the marginals arose from a continuous process then we can identify a unique diffusion with the same marginals. Now we ask, what other processes might have lead to the same marginals? The Dupire construction gives a canonical process with a given set of marginals, but it is not unique.

In particular, suppose $X_t \sim N(0, t)$, and that X_t is a martingale. Clearly one process with these marginals is Brownian motion, and Brownian motion is the unique continuous Markovian martingale with $N(0, t)$ marginals.

Following Oleszkiewicz [45] we can define a *fake Brownian motion* to be a martingale with marginals $X_t \sim N(0, t)$, which is not Brownian motion. Given Dupire's result, in order to find such a process we need to relax either the assumption of continuity of paths, or the Markov assumption.

It is apparent from the discussion in Sect. 6.2 that several of the construction methods for solutions of the Skorokhod embedding problem for multiple time points extend to the setting of a family of marginals with parameter t, at least under some restrictions on that family. A first example is the Azéma–Yor construction.

Let $\mu_t \sim N(0, t)$, and let $b_t := b_{\mu_t}$. By scaling we have that $b_t(x) = b_1(x\sqrt{t})$, so that the family b_t is increasing in t for each fixed x. In particular, if $\tau_t = \inf\{u : W_u \le \beta_t(J_u)\}$ then τ_t is an increasing family of stopping times such that τ_t embeds μ_t. This is the Madan and Yor [40] fake Brownian motion. The resulting process is a Markovian martingale, for which paths have strictly decreasing continuous parts interleaved with positive jumps.

By extending the methods of Sect. 6.1 from two to a continuum of time-points we can also use the Perkins [47] or Vallois [55, 56] construction to generate other fake Brownian motions. (Note however that the generalisation of the Root [51] construction does not give a fake Brownian motion, since the barriers are simply vertical lines and the resulting stopping times are $\tau_t \equiv t$.)

Interest in the problem of finding fake Brownian motions began with a paper by Hamza and Klebaner [30]. They use a remarkably clever decomposition and recombination of the normal distribution to produce a pure-jump martingale with Gaussian marginals. They also asked the question if there existed continuous

(non-Brownian) martingales with Gaussian marginals. This was answered in the affirmative by Albin [1]. Finally, Oleszkiewicz [45] gave an elegant and simple construction of a continuous fake Brownian motion. In essence this construction makes use of the fact that $R_t \sin \Theta_{A_t}$ is a Brownian motion where $(R_t)_{t\geq 0}$ is a two-dimensional Bessel process, $(\Theta_u)_{u\geq 0}$ is a Brownian motion on the unit circle, and $(A_t)_{t\geq 0}$ is the increasing additive functional such that $A_t = \int_0^t R_s^{-2} ds$. The idea is to write $X_t = \sqrt{t} R \sin(U + W_{\ln t})$ where R has the same distribution as R_1, U is uniform distribution on $[0, 2\pi)$ and W is Brownian motion.

The existence of fake Brownian motions implies that although the Dupire [23] construction gives a model which is consistent with the continuum of traded option prices (assuming such a family exists) – and perhaps gives the *canonical* model consistent with these prices – it is not the unique model with this property. Hence, although the Dupire diffusion might be used to give a guide price for exotic derivatives, these prices are not the unique prices consistent with no-arbitrage.

6.4 Model Independent Bounds on Basket Options

Underlying the discussion in these notes is a philosophy whereby the prices of exotic options are related to the prices of vanilla (traded) options by the construction of super-hedges involving those vanilla options. The (primal) pricing problem is related to a (dual) hedging problem.

The same philosophy can be applied in other contexts, one such being the pricing of basket (exotic) options, given the prices of (vanilla) call options on the individual constituents of the basket. It turns out that the model independent upper bound on the price of a basket option is associated with a model when the assets are co-monotonic, see Dhaene et al. [21] or Hobson et al. [36].

7 Closing Remarks

In a mature market, the prices of liquidly traded vanilla options are not the expectations of the payoff under a stochastic model, but rather they are fixed by supply and demand, or market sentiment. From a derivative pricing perspective the objects of interest become the less liquid exotic derivatives which must be priced relative to the vanilla options.

In the standard case where the vanilla options are precisely the puts and calls, knowledge of vanilla prices is equivalent to knowledge of the marginal distribution of the underlying asset. Then, finding candidate models which fit option prices is equivalent to finding solutions of the Skorokhod embedding problem, and finding the range of no-arbitrage prices for an exotic option is equivalent to finding extremal embeddings which maximise functionals on the Brownian paths.

The associated no-arbitrage bounds may be quite wide. (Conversely, see Brown et al. [9], and Example 5 above, in some special cases the upper and lower bounds coincide, and there is a unique model-independent no-arbitrage price for the exotic option.) However, each bound is associated with a model which achieves the bound, and by investigating features of this model it is possible to determine the characteristics of models which lead to high option payouts. (For example, for the model which attains the upper bound on the price of a lookback option, the overall maximum is an increasing function of the time-T price. Whilst in exponential Brownian motion model the correlation between the final value and the maximum is not perfect, the correlation is quite strong, and therefore for realistic parameter values the model price can be expected to be quite close to the theoretical upper bound.)

The second major advantage of the Skorokhod-embedding-model-independent-bound approach is that since no assumptions are made on the model the resulting strategies must be very simple. Even if an agent believes in a particular stochastic model there may be advantages (especially in the form of lower transaction costs) in following a semi-static hedging strategy involving calls, when compared with a delta-hedging strategy with infinite trading. Cox and Oblój [18] compare the performance of a model independent hedge with the Black Scholes hedge for no-touch options, and conclude that the robust hedge frequently outperforms the classical delta-hedge, even when transaction costs are small.

The optimal strategy involves purchasing an initial portfolio of traded options, but this portfolio is held constant over time, and there is no assumption that elements of the portfolio can be sold at intermediate times. Hence the success of any super-replicating strategy is not contingent upon any model for movements of option prices over time.

Acknowledgements I would like to thank the editors for the invitation to write these lecture notes, and to the audience at the Bachelier seminar in Paris for useful feedback and stimulating questions. Most especially I would like to thank Alex Cox, for his detailed reading of an early draft of these notes, and for his many suggestions for improvements. Responsibility for errors and omissions rests, as always, with the author.

References

1. Albin, J.M.P.: A continuous non-Brownian motion martingale with Brownian motion martingale distributions. Stat. Probab. Lett. **78**(6), 682–686 (2008)
2. Azéma, J., Yor, M.: Une solution simple au problème de Skorokhod, Séminaire de Probabilités, vol. XIII. Lecture Notes in Math., vol. 721, pp. 90–115. Springer, Berlin (1979)
3. Azéma, J., Yor, M.: Une solution simple au problème de Skorokhod, Séminaire de Probabilités, vol. XIII. Lecture Notes in Math., vol. 721, pp. 625–633. Springer, Berlin (1979)
4. Azéma, J., Gundy, R.F., Yor, M.: Sur l'intégrabilité uniforme des martingales continues, Séminaire de Probabilités, vol. XIV. Lecture Notes in Math., vol. 784, pp. 53–61. Springer, Berlin (1980)
5. Bass, R.F.: Skorokhod imbedding via stochastic integrals, Séminaire de Probabilités, vol. XVII. Lecture Notes in Math., vol. 986, pp. 221–224. Springer, Berlin (1983)

6. Bergman, Y.Z., Grundy, B.D., Wiener, Z.: General properties of option prices. J. Finance **51**, 1573–1610 (1996)
7. Bertoin, J., Le Jan, Y.: Representation of measures by balayage from a regular recurrent point. Ann. Probab. **20**(1), 538–548 (1992)
8. Breeden, D.T., Litzenberger, R.H.: Prices of state-contingent claims implicit in options prices. J. Business **51**, 621–651 (1978)
9. Brown, H., Hobson, D., Rogers, L.C.G.: Robust hedging of barrier options. Math. Finance **11**(3), 285–314 (2001)
10. Brown, H., Hobson, D., Rogers, L.C.G.: The maximum maximum of a martingale constrained by an intermediate law. Probab. Theor. Relat. Field. **119**, 558–578 (2001)
11. Bühler, H.: Expensive martingales. Quant. Finance **6**, 207–218 (2006)
12. Carr, P., Lee, R.: Hedging variance options on semi-martingales. Finance Stochast. **14**(2), 179–208 (2010)
13. Chacon, R.V., Walsh, J.B.: One-dimensional potential embedding, Séminaire de Probabilités, vol. X. Lecture Notes in Math., vol. 511, pp. 19–23. Springer, Berlin (1976)
14. Cousot, L.: Conditions on options prices for the absence of arbitrage and exact calibration. J. Bank. Finance **31**(11), 3377–3397 (2007)
15. Cox, A.M.G., Hobson, D.G.: An optimal embedding for diffusions. Stochast. Process. Appl. **111**(1) 17–39 (2004)
16. Cox, A.M.G., Hobson, D.G.: Local martingales, bubbles and options prices. Finance Stochast. **9**, 477–492 (2005)
17. Cox, A.M.G., Hobson, D.G.: Skorokhod embeddings, minimality and non-centred target distributions. Prob. Theor. Relat. Field. **135**, 395–414 (2006)
18. Cox, A.M.G., Oblój, J.: Robust hedging of double touch barrier options (2008, Preprint). arXiv:0808.4012
19. Cox, A.M.G., Oblój, J.: Robust pricing and hedging of double no-touch options. To appear in Finance Stochast. (2010). arXiv:0901.0674
20. Davis, M.H.A., Hobson, D.G.: The range of traded options prices. Math. Finance **17**, 1–14 (2007)
21. Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R., Vyncke, D.: The concept of co-monotonicity in actuarial science and finance: Applications. Insur. Math. Econ. **31**, 133–161 (2002)
22. Dubins, L.E.: On a theorem of Skorohod. Ann. Math. Statist. **39**, 2094–2097 (1968)
23. Dupire, B.: Pricing with a smile. Risk Mag. **7**(1), 18–20 (1994)
24. Dupire, B.: Arbitrage bounds for volatility derivatives. Presentation at PDE and Mathematical Finance KTH, Stockholm (2005)
25. El Karoui, N., Jeanblanc, M., Shreve, S.E.: Robustness of the Black-Scholes formula. Math. Finance **8**, 93–126 (1998)
26. Figlewski, S.: Assessing the incremental value of option pricing theory relative to an informationally passive benchmark. J. Derivatives **10**, 80–96 (2002)
27. Föllmer, H., Schied, A.: Stochastic Finance: An Introduction in Discrete Time, 2nd edn. de Gruyter, Berlin (2002)
28. Gyöngy, I.: Mimicking the one-dimensional marginal distributions of processes having an Itô differential. Prob. Theor. Relat. Field. **71**, 501–516 (1986)
29. Hall, W.: On the Skorokhod embedding theorem, Technical Report 33. Department of Statistics, Stanford University (1968)
30. Hamza, K., Klebaner, F.C.: A family of non-Gaussian martingales with Gaussian marginals. J. Appl. Math. Stochast. Anal. **2007**, 19 (2007); Article Id 92723
31. Hobson, D.G.: Robust hedging of the lookback option. Finance Stochast. **2**(4), 329–347 (1998)
32. Hobson, D.G.: The maximum maximum of a martingale, Séminaire de Probabilités, vol. XXXII. Lecture Notes in Math., vol. 1686, pp. 250–263. Springer, Berlin (1998)
33. Hobson, D.G.: Volatility misspecification, option pricing and super-replication via coupling. Ann. Appl. Probab. **8**, 193–205 (1998)
34. Hobson, D.G., Neuberger, A.: Robust bounds for Forward Start options. To appear in Math. Finance (2010)

35. Hobson, D.G., Pedersen, J.L.: The minimum maximum of a continuous martingale with given initial and terminal laws. Ann. Probab. **30**(2), 978–999 (2002)
36. Hobson, D.G., Lawrance, P., Wang, T.-H.: Static-arbitrage upper bounds for the prices of basket options. Quant. Finance **5**, 329–342 (2005)
37. Janson, S., Tysk, J.: Volatility time and properties of options prices. Ann. Appl. Probab. **13**, 890–913 (2003)
38. Kiefer, J.: Skorokhod imbedding of multivariate RV's and the sample DF. Z. Wahrscheinlichkeitstheorie **24**, 1–35 (1972)
39. Krylov, N.V.: On the relation between differential operators of second order and the solutions of differential equations. In: Krylov, N.V., et al. (eds.) Steklov Seminar 1984, pp. 214–229. Inc. Publications Division, New York (1985)
40. Madan, D.B., Yor, M.: Making Markov martingales meet marginals: With explicit constructions. Bernoulli **8**(4), 509–536 (2002)
41. Meilijson, I.: Skorokhod's problem – embeddings in Brownian motion. Notes from the 3rd Winter School of Probability, Chile (1983)
42. Monroe, I.: On embedding right continuous martingales in Brownian motion. Ann. Math. Statist. **43**, 1293–1311 (1972)
43. Obłój, J.: The Skorokhod embedding problem and its offspring. Probab. Surv. **1**, 321–390 (2004)
44. Obłój, J., Yor, M.: An explicit Skorokhod embedding for the age of Brownian excursions and Azéma martingale. Stochast. Process. Appl. **110**(1), 83–110 (2004)
45. Oleszkiewicz, K.: On fake Brownian motions. Statist. Probab. Lett. **78**, 1251–1254 (2008)
46. Pedersen, J.L., Peskir, G.: The Azéma-Yor embedding in non-singular diffusions. Stochast. Process. Appl. **96**(2), 305–312 (2001)
47. Perkins, E.: The Cereteli-Davis solution to the H^1-embedding problem and an optimal embedding in Brownian motion. In: Seminar on Stochastic Processes, 1985 (Gainesville, Fla., 1985), pp. 172–223. Birkhäuser Boston, MA (1986)
48. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 3rd edn. Springer, Berlin (1999)
49. Rogers, L.C.G.: A guided tour through excursions. Bull. Lond. Math. Soc. **21**(4), 305–341 (1989)
50. Rogers, L.C.G., Williams, D.: Diffusions, Markov Processes and Martingales, vol. 2, Itô Calculus. CUP, Cambridge (2000)
51. Root, D.H.: The existence of certain stopping times on Brownian motion. Ann. Math. Statist. **40**, 715–718 (1969)
52. Rost, H.: The stopping distributions of a Markov Process. Invent. Math. **14**, 1–16 (1971)
53. Rost, H.: Skorokhod stopping times of minimal variance, Séminaire de Probabilités, vol. X, pp. 194–208. Lecture Notes in Math., vol. 511. Springer, Berlin (1976)
54. Skorokhod, A.V.: Studies in the Theory of Random Processes. Addison-Wesley, MA (1965)
55. Vallois, P.: Le problème de Skorokhod sur $\mathbb{R}$: une approche avec le temps local, Seminar on probability, vol. XVII. Lecture Notes in Math. vol. 986, pp. 227–239. Springer, Berlin (1983)
56. Vallois, P.: Quelques inégalités avec le temps local en zero du mouvement Brownien. Stochast. Process. Appl. **41**(1), 117–155 (1992)

Pricing and Hedging in Exponential Lévy Models: Review of Recent Results

Peter Tankov

Abstract These lecture notes cover a major part of the crash course on financial modeling with jump processes given by the author in Bologna on May 21–22, 2009. After a brief introduction, we discuss three aspects of exponential Lévy models: absence of arbitrage, including more recent results on the absence of arbitrage in multidimensional models, properties of implied volatility, and modern approaches to hedging in these models.

Keywords Lévy processes · exponential Lévy models · absence of arbitrage · Esscher transform · implied volatility · smile modeling · quadratic hedging

1 Introduction

Exponential Lévy models generalize the classical Black and Scholes setup by allowing the stock prices to jump while preserving the independence and stationarity of returns. There are ample reasons for introducing jumps in financial modeling. First of all, asset prices do jump, and some risks simply cannot be handled within continuous-path models. Second, the well-documented phenomenon of *implied volatility smile* in option markets shows that the risk-neutral returns are non-gaussian and leptokurtic. While the smile itself can be explained within a model with continuous paths, the fact that it becomes much more pronounced for short maturities is a clear indication of the presence of jumps. In continuous-path models, the law of returns for shorter maturities becomes closer to the Gaussian law, whereas in reality and in models with jumps returns actually become less Gaussian as the horizon becomes shorter. Finally, jump processes correspond to genuinely incomplete markets, whereas all continuous-path models are either complete or "completable" with a small number of additional assets. This fundamental

P. Tankov (✉)
Centre de Mathématiques Appliquées, Ecole Polytechnique, Palaiseau Cedex 91128, France
e-mail: peter.tankov@polytechnique.org

A.R. Carmona et al. (eds.), *Paris-Princeton Lectures on Mathematical Finance 2010*, Lecture Notes in Mathematics 2003, DOI 10.1007/978-3-642-14660-2_5,
© Springer-Verlag Berlin Heidelberg 2011

incompleteness makes it possible to carry out a rigorous analysis of the hedging error and find ways to improve the hedging performance using additional instruments such as liquid European options.

A great advantage of exponential Lévy models is their mathematical tractability, which makes it possible to perform many computations explicitly and to present deep results of modern mathematical finance in a simple manner. This has led to an explosion of the literature on option pricing and hedging in exponential Lévy models in the late 1990s and early 2000s, the literature which now contains hundreds of research papers and several monographs. However, some fundamental aspects such as asymptotic behavior of implied volatility or the computation of hedge ratios have only recently been given a rigorous treatment.

In this survey, after recalling the essential properties of Lévy processes and exponential Lévy models, and introducing the notation used throughout the paper, we discuss, reformulate and extend some recent results on absence of arbitrage, properties of the implied volatility and approaches to option hedging in these models.

For background on exponential Lévy models, the reader may refer to textbooks such as [16, 49] for a more financial perspective or [3, 34] for a more mathematical perspective.

Lévy processes

Lévy processes [3, 7, 48] are stochastic processes with stationary and independent increments. The only Lévy process with continuous trajectories is the Brownian motion with drift; all other representatives of this class admit discontinuities in finite or (countably) infinite number. A general Lévy process can be represented as

$$X_t = \gamma t + B_t + N_t + \lim_{\varepsilon \downarrow 0} M_t^{\varepsilon}, \tag{1}$$

where B is a d-dimensional Brownian motion with covariance matrix A, $\gamma \in \mathbb{R}^d$, N is a compound Poisson process which includes the jumps of X with $|\Delta X_t| > 1$ and M_t^{ε} is a compensated compound Poisson process (compound Poisson minus its expectation) which includes the jumps of X with $\varepsilon < |\Delta X_t| \leq 1$.

The law of a Lévy process is completely identified by its characteristic triplet: the positive definite matrix A, the vector γ and the measure ν on $\mathbb{R}^d$, called the Lévy measure, which determines the intensity of jumps of different sizes: $\nu(A)$ is the expected number of jumps on the time interval $[0, 1]$, whose sizes fall in A. The Lévy measure satisfies the integrability condition

$$\int_{\mathbb{R}^d} 1 \wedge |x|^2 \nu(dx) < \infty,$$

and $\nu(\mathbb{R}^d) < \infty$ if and only if the process has finite jump intensity. The law of X_t at all times t is determined by the triplet, and in particular, the Lévy–Khintchine formula gives the characteristic function:

$$E[e^{iu.X_t}] := \Phi_t(u) = e^{t\psi(u)}, \quad u \in \mathbb{R}^d,$$

$$\text{where} \quad \psi(u) = i\gamma.u + \frac{1}{2}Au.u + \int_{\mathbb{R}^d}(e^{iu.x} - 1 - iu.x1_{|x|\leq 1})\nu(dx).$$

In the above formula and everywhere in this paper, $a.b$ denotes the dot product of vectors a and b.

The jump measure of X, defined by

$$J_X([t_1, t_2] \times A) := \#\{t \in [t_1, t_2] : \Delta X_t \in A\},$$

is a Poisson random measure with intensity $dt \times \nu(dx)$, and using this notation, the representation (1) can be rewritten as

$$X_t = \gamma t + B_t + \int_0^t \int_{|x|>1} xJ_X(ds \times dx) + \lim_{\varepsilon \downarrow 0} \int_0^t \int_{\varepsilon<|x|\leq 1} x\tilde{J}_X(ds \times dx),$$

where $\tilde{J}_X(ds \times dx) := J_X(ds \times dx) - ds \times \nu(dx)$ is the compensated version of J_X.

Exponential Lévy models

The Black–Scholes model

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t$$

can be equivalently rewritten in the exponential form $S_t = S_0 e^{(\mu-\sigma^2/2)t+\sigma W_t}$. This gives us two possibilities to construct an exponential Lévy model starting from a (one-dimensional) Lévy process X: using the stochastic differential equation:

$$\frac{dS_t^{sde}}{S_{t-}^{sde}} = rdt + dX_t, \tag{2}$$

or using the ordinary exponential

$$S_t^{exp} = S_0^{exp} e^{rt+X_t}, \tag{3}$$

where we explicitly included the interest rate r (assumed constant) in the formulas, to simplify notation later on. The subscripts *sde* for stochastic differential equation and *exp* for exponential, used here to emphasize the fact that S^{sde} and S^{exp} are different processes, will be omitted throughout the paper when there is no ambiguity. Sometimes it will be convenient to discount the price processes with the numéraire $B(t,T) = e^{-r(T-t)}$ for some fixed maturity T. In this case $\hat{S}_t := \frac{S_t}{B(t,T)} = e^{r(T-t)}S_t$ and the equations become

$$\frac{d\hat{S}_t}{\hat{S}_{t-}} = dX_t \tag{4}$$

$$\text{or} \quad \hat{S}_t = \hat{S}_0 e^{X_t}, \tag{5}$$

The solution to (4) with initial condition $\hat{S}_0 = 1$ is called the stochastic exponential of X, written explicitly as

$$\mathcal{E}(X)_t = e^{X_t - \frac{1}{2}[X]^c_t} \prod_{s \le t : \Delta X_s \neq 0} (1 + \Delta X_s) e^{-\Delta X_s}.$$

It can become negative if the process X has a big negative jump: $\Delta X_s < -1$ for $s \le t$. However, if X does not have jumps of size less or equal to -1, then its stochastic exponential is positive, and can be represented as the ordinary exponential of another Lévy process [26], which shows that the formulations (4) and (5) are equivalent, and the choice of one or the other is simply a matter of convenience for a particular problem.

Examples of exponential Lévy models

Parametric exponential Lévy models fall into two categories. In the first category, called *jump-diffusion* models, the "normal" evolution of prices is given by a diffusion process, punctuated by jumps at random intervals. Here the jumps represent rare events – crashes and large drawdowns. Such an evolution can be represented by a Lévy process with a nonzero Gaussian component and a jump part with finitely many jumps:

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i, \tag{6}$$

where (Y_i) are i.i.d. and N is a Poisson process.

In the *Merton model* [39], which is the first model of this type, suggested in the literature, jumps in the log-price X are assumed to have a Gaussian distribution: $Y_i \sim N(\mu, \delta^2)$.

In the *Kou model* [32], jump sizes are distributed according to an asymmetric Laplace law with a density of the form

$$\nu_0(dx) = [p\lambda_+ e^{-\lambda_+ x} 1_{x>0} + (1-p)\lambda_- e^{-\lambda_- |x|} 1_{x<0}]dx \tag{7}$$

with $\lambda_+ > 0$, $\lambda_- > 0$ governing the decay of the tails for the distribution of positive and negative jump sizes and $p \in [0, 1]$ representing the probability of an upward jump. The probability distribution of returns in this model has semi-heavy (exponential) tails.

The second category consists of models with an infinite number of jumps in every interval, called *infinite activity* or *infinite intensity* models. In these models, one does not need to introduce a Brownian component since the dynamics of jumps is already rich enough to generate nontrivial small time behavior [13].

The variance gamma process [10, 38] is obtained by time-changing a Brownian motion with a gamma subordinator and has the characteristic exponent of the form:

$$\psi(u) = -\frac{1}{\kappa}\log(1 + \frac{u^2\sigma^2\kappa}{2} - i\theta\kappa u). \tag{8}$$

The density of the Lévy measure of the variance gamma process is given by

$$\nu(x) = \frac{c}{|x|}e^{-\lambda_-|x|}1_{x<0} + \frac{c}{x}e^{-\lambda_+ x}1_{x>0}, \tag{9}$$

where $c = 1/\kappa$, $\lambda_+ = \frac{\sqrt{\theta^2+2\sigma^2/\kappa}}{\sigma^2} - \frac{\theta}{\sigma^2}$ and $\lambda_- = \frac{\sqrt{\theta^2+2\sigma^2/\kappa}}{\sigma^2} + \frac{\theta}{\sigma^2}$.

To define the *tempered stable process*, introduced by Koponen [31] and also known under the name of CGMY model [13], one specifies directly the Lévy density:

$$\nu(x) = \frac{c_-}{|x|^{1+\alpha_-}}e^{-\lambda_-|x|}1_{x<0} + \frac{c_+}{x^{1+\alpha_+}}e^{-\lambda_+ x}1_{x>0} \tag{10}$$

with $\alpha_+ < 2$ and $\alpha_- < 2$.

2 The Esscher Transform and Absence of Arbitrage in Exponential Lévy Models

To find out whether a given exponential Lévy model is suitable for financial modeling, one needs to ensure that it does not contain arbitrage opportunities, a property which, by the fundamental theorem of asset pricing, is guaranteed by the existence of an equivalent martingale measure. The no arbitrage equivalences for exponential Lévy models were studied in [15,28,55] in the one-dimensional unconstrained case and more recently in [30] in the multidimensional case with convex constraints on trading strategies. In this section, we start by reviewing the one-dimensional result, and then provide a multidimensional result (Theorem 2) which is valid in the unconstrained case only but is more explicit than the one in [30] and clarifies the link between the geometric properties of the Lévy measure and the arbitrage opportunities in the model.

In the Black–Scholes model, the unique equivalent martingale measure could be obtained by changing the drift of the Brownian motion. In models with jumps, if the Gaussian component is absent, this is no longer possible, but a much greater variety of equivalent measures can be obtained by altering the distribution of jumps. The following proposition describes the possible measure changes under which a Lévy process remains a Lévy process.

Proposition 1 (see Sato [48], Theorems 33.1 and 33.2). *Let $(X,\mathbb{P})$ be a Lévy process on $\mathbb{R}^d$ with characteristic triplet (A,ν,γ); choose $\eta \in \mathbb{R}^d$ and $\varphi : \mathbb{R}^d \to \mathbb{R}$ with*

$$\int_{\mathbb{R}^d}(e^{\varphi(x)/2} - 1)^2\nu(dx) < \infty. \tag{11}$$

and define

$$U_t := \eta.X^c + \int_0^t \int_{\mathbb{R}^d} (e^{\varphi(x)} - 1)\tilde{J}_X(ds\, dx),$$

where X^c denotes the continuous martingale (Brownian motion) part of X, and $\tilde{J}_X$ is the compensated jump measure of X.

Then $\mathcal{E}(U)_t$ is a positive martingale such that the probability measure $\mathbb{P}'$ defined by

$$\frac{d\mathbb{P}'|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} = \mathcal{E}(U)_t, \tag{12}$$

is equivalent to $\mathbb{P}$ and under $\mathbb{P}'$, X is a Lévy process with characteristic triplet (A, ν', γ') where $\nu' = e^{\varphi}\nu$ and

$$\gamma' = \gamma + \int_{|x|\leq 1} x(\nu' - \nu)(dx) + A\eta. \tag{13}$$

A useful example, which will be the basis of our construction of an equivalent martingale measure is provided by the Esscher transform. Let X be a Lévy process on $\mathbb{R}^d$ with characteristic triplet (A, ν, γ), and let $\theta \in \mathbb{R}^d$ be such that $\int_{|x|>1} e^{\theta.x}\nu(dx) < \infty$. Applying a measure transformation of Proposition 1 with $\eta = \theta$ and $\varphi(x) = \theta.x$, we obtain an equivalent probability under which X is a Lévy process with Lévy measure $\tilde{\nu}(dx) = e^{\theta.x}\nu(dx)$ and third component of the characteristic triplet $\tilde{\gamma} = \gamma + A\theta + \int_{|x|\leq 1} x(e^{\theta.x} - 1)\nu(dx)$. Using Proposition 1, the Radon–Nikodym derivative corresponding to this measure change is found to be

$$\frac{d\mathbb{P}'|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} = \frac{e^{\theta.X_t}}{E[e^{\theta.X_t}]} = \exp(\theta.X_t - \kappa(\theta)t), \tag{14}$$

where $\kappa(\theta) := \ln E[\exp(\theta.X_1)] = \psi(-i\theta)$.

Although the two definitions of an exponential Lévy model, via the ordinary exponential (5) or via the stochastic exponential (4), are equivalent, the set of Lévy processes that lead to arbitrage-free models of the form (5) does not necessarily coincide with the set that yields arbitrage-free models of the form (4). In particular, we shall see that the no-arbitrage conditions for multidimensional stochastic and ordinary exponentials are considerably different. It will be more convenient to find these conditions for models of type (4) first and then deduce the conditions for ordinary exponentials using the transformation $X_t := \ln \mathcal{E}(Y)_t$. In the multidimensional case, this transformation must be applied to each component.

In an exponential Lévy model of type (4), the absence of arbitrage is tantamount to the existence of a probability $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that $\mathcal{E}(X)$ is a $\mathbb{Q}$-martingale. We will see that when this is the case, it is always possible to find a martingale probability $\mathbb{Q} \sim \mathbb{P}$ under which X remains a Lévy process, which means that X itself must be a $\mathbb{Q}$-martingale (cf. Proposition 8.23 in [16]).

We start with the one-dimensional case. In the sequel, $\mathtt{cc}(A)$ denotes the smallest convex cone containing A and $\mathtt{ri}(A)$ denotes the relative interior of the set A, that is, the interior of A in the smallest linear subspace containing A. In particular, $\mathrm{ri}(\{0\}) = \{0\}$.

Theorem 1 (Absence of arbitrage in models based on stochastic exponentials, one-dimensional case). *Let $(X, \mathbb{P})$ be a real-valued Lévy process on $[0, T]$ with characteristic triplet (σ^2, ν, γ). The following statements are equivalent:*

1. *There exists a probability $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that $(X, \mathbb{Q})$ is a Lévy process and a martingale.*
2. *Either $X \equiv 0$ or $(X, \mathbb{P})$ is not a.s. monotone.*
3. *One of the following conditions is satisfied:*

   (i) $\sigma > 0$.
   (ii) $\sigma = 0$ *and* $\int_{|x|\le 1} |x|\nu(dx) = \infty$.
   (iii) $\sigma = 0$, $\int_{|x|\le 1} |x|\nu(dx) < \infty$ *and* $-b \in ri(cc(supp\,\nu))$, *where* $b = \gamma - \int_{|x|\le 1} x\nu(dx)$ *is the drift of* X.

Condition (2) implies that if an exponential Lévy model admits an arbitrage, it can be realized by a buy-and-hold strategy (if X is increasing) or a corresponding short sale (if X is decreasing).

It is easy to see that condition (iii) above is satisfied if and only if $\sigma = 0$, $\int_{|x|\le 1} |x|\nu(dx) < \infty$ and one of the following is true:

- $\nu((-\infty, 0)) > 0$ and $\nu((0, \infty)) > 0$.
- $\nu((-\infty, 0)) > 0$ and $b > 0$.
- $\nu((0, \infty)) > 0$ and $b < 0$.
- The trivial case of a constant process: $\nu = 0$ and $b = 0$.

In other words, when a finite-variation Lévy process has one-sided jumps, it is arbitrage-free if the jumps and the drift point in opposite directions.

Before proceeding with the proof of Theorem 1, we will show that for *one-dimensional* exponential Lévy models of the form (5), the no-arbitrage conditions are actually the same as for stochastic exponentials.

Corollary 1 (Absence of arbitrage in models based on ordinary exponential, one-dimensional case). *Let $(X, \mathbb{P})$ be a real-valued Lévy process on $[0, T]$ with characteristic triplet (σ^2, ν, γ). The following statements are equivalent:*

1. *There exists a probability $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that $(X, \mathbb{Q})$ is a Lévy process and e^X is a martingale.*
2. *Either $X \equiv 0$ or $(X, \mathbb{P})$ is not a.s. monotone.*
3. *One of the following conditions is satisfied:*

   (i) $\sigma > 0$.
   (ii) $\sigma = 0$ *and* $\int_{|x|\le 1} |x|\nu(dx) = \infty$.
   (iii) $\sigma = 0$, $\int_{|x|\le 1} |x|\nu(dx) < \infty$ *and* $-b \in ri(cc(supp\,\nu))$.

Proof. It suffices to show that $\ln \mathcal{E}(X)$ is monotone if and only if X is monotone. From [16, Proposition 8.22] it is easy to see that $\ln \mathcal{E}(X)$ is a finite variation process if and only if X is a finite variation process. In the finite-variation case, the stochastic exponential has a simple form:

$$\mathcal{E}(X)_t = e^{bt} \prod_{s \le t} (1 + \Delta X_s),$$

and it is readily seen that the monotonicity properties of X and $\log \mathcal{E}(X)$ are the same.

Proof (Proof of Theorem 1). We exclude the trivial case $X \equiv 0$ a.s. which clearly does not constitute an arbitrage opportunity (every probability is a martingale measure).

The equivalence $2 \iff 3$ follows from [16, Proposition 3.10].

$3 \Rightarrow 1$. Define a probability $\tilde{\mathbb{P}}$ equivalent to $\mathbb{P}$ by

$$\frac{d\tilde{\mathbb{P}}|_{\mathcal{F}_T}}{d\mathbb{P}|_{\mathcal{F}_T}} = \mathcal{E}\left(\int_0^{\cdot} \int_{\mathbb{R}} (e^{-x^2} - 1) \tilde{J}_X(ds\, dx) \right)_T$$

Under $\tilde{\mathbb{P}}$, X has characteristic triplet $(\sigma^2, \tilde{\nu}, \tilde{\gamma})$ with $\tilde{\nu} = e^{-x^2}\nu$ and $\tilde{\gamma} = \gamma + \int_{|x| \le 1} x(e^{-x^2} - 1)\nu(dx)$. It is easy to see that $E^{\tilde{\mathbb{P}}}[e^{\lambda X_t}] < \infty$ for all $\lambda \in \mathbb{R}$ and all $t > 0$.

Suppose that the convex function $\lambda \mapsto E^{\tilde{\mathbb{P}}}[e^{\lambda X_1}]$ has a finite minimizer λ^*. Then, using the dominated convergence theorem, $E^{\tilde{\mathbb{P}}}[X_1 e^{\lambda^* X_1}] = 0$ which implies that X is a $\mathbb{Q}$-martingale with

$$\frac{d\mathbb{Q}|_{\mathcal{F}_t}}{d\tilde{\mathbb{P}}|_{\mathcal{F}_t}} = \frac{e^{\lambda^* X_t}}{E[e^{\lambda^* X_t}]} \quad \text{(Essher transform)}$$

To show the existence of a finite minimizer λ^*, it is sufficient to prove that $E^{\tilde{\mathbb{P}}}[e^{\lambda X_1}] \to \infty$ as $\lambda \to \infty$, or, equivalently, that the function

$$f(\lambda) = \log E^{\tilde{\mathbb{P}}}[e^{\lambda X_1}] = \frac{\sigma^2}{2}\lambda^2 + \tilde{\gamma}\lambda + \int_{\mathbb{R}} (e^{\lambda x} - 1 - \lambda x 1_{|x| \le 1}) e^{-x^2} \nu(dx).$$

goes to infinity as $\lambda \to \infty$. In case (i), $f''(\lambda) \ge \sigma^2$ which means that $f(\lambda) \to \infty$ as $\lambda \to \infty$. In case (ii),

$$f'(\lambda) = \tilde{\gamma} + \int_{|x|>1} x e^{-x^2} \nu(dx) + \int_{\mathbb{R}} x(e^{\lambda x} - 1) e^{-x^2} \nu(dx),$$

and it is not difficult to check that $\lim_{\lambda \to +\infty} f'(\lambda) = +\infty$ and $\lim_{\lambda \to -\infty} f'(\lambda) = -\infty$ which means that $f(\lambda) \to \infty$ as $\lambda \to \infty$. In case (iii),

$$f'(\lambda) = b + \int_{\mathbb{R}} x e^{\lambda x} e^{-x^2} \nu(dx),$$

and it is easy to see, by examining one by one the different mutually exclusive cases listed after the statement of the theorem, that in each of these cases f' is bounded from below on $\mathbb{R}$ and therefore once again, $f(\lambda) \to \infty$ as $\lambda \to \infty$.

$1 \Rightarrow 2$. It is clear that a process cannot be a martingale under one probability and a.s. monotone under an equivalent probability, unless it is constant.

In the multidimensional case, the no arbitrage conditions for ordinary and stochastic exponentials are different. We start with the simpler case of stochastic exponentials.

Let $(X, \mathbb{P})$ be an $\mathbb{R}^d$-valued Lévy process on $[0, T]$ with characteristic triplet (A, ν, γ). To describe the no-arbitrage conditions, we need to separate the finite and infinite variation components of X. We therefore introduce the linear subspace $\mathcal{L} \subseteq \mathbb{R}^d$ containing all vectors $w \in \mathbb{R}^d$ such that $w.X$ is a finite variation process. From Proposition 3.8 and Theorem 4.1 in [16], it follows that

$$\mathcal{L} = \mathcal{N}(A) \cap \{w \in \mathbb{R}^d : \int_{|x|\leq 1} |w.x|\nu(dx) < \infty\},$$

where $\mathcal{N}(A) := \{w \in \mathbb{R}^d : Aw = 0\}$. Further, denote by $X^{\mathcal{L}}$ the projection of X on $\mathcal{L}$. $X^{\mathcal{L}}$ is a finite variation Lévy process with triplet $(0, \nu^{\mathcal{L}}, \gamma^{\mathcal{L}})$, and we denote its drift by

$$b^{\mathcal{L}} := \gamma^{\mathcal{L}} - \int_{\mathcal{L}\cap\{x:|x|\leq 1\}} x\nu^{\mathcal{L}}(dx).$$

Theorem 2 (Absence of arbitrage in models based on stochastic exponential, multidimensional case). *Let $(X, \mathbb{P})$ be an $\mathbb{R}^d$-valued Lévy process on $[0, T]$ with characteristic triplet (A, ν, γ). The following statements are equivalent:*

1. *There exists a probability $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that $(X, \mathbb{Q})$ is a Lévy process and (X^i) is a $\mathbb{Q}$-martingale for all i.*
2. *For every $w \in \mathbb{R}^d$, the process $w.X$ satisfies one of the equivalent conditions (2) or (3) of Theorem 1.*
3. *$-b^{\mathcal{L}} \in ri(cc(supp\, \nu^{\mathcal{L}}))$.*

Let us comment on the equivalent conditions of the above theorem.

To understand condition (2), assume that for some $w \in \mathbb{R}^d$, the process $w.X$ does not satisfy the equivalent conditions of Theorem 1, meaning that it is either strictly increasing or strictly decreasing. Consider a portfolio where the relative proportions of different assets are kept constant and equal to w_i. The proportions may, of course, change when the underlying assets jump, but it is assumed that they are readjusted to their constant values immediately after the jump. Such a strategy is called a fixed-mix strategy. The discounted value $\hat{V}_t$ of such a portfolio satisfies the SDE $d\hat{V}_t = \hat{V}_{t-} w.dX_t$, and therefore either this portfolio constitutes an arbitrage strategy or

an arbitrage strategy can be obtained by shorting this portfolio. Condition (2) thus implies that a multidimensional exponential Lévy model is arbitrage-free if and only if there are no fixed-mix arbitrage strategies.

The third condition is a concise characterization of arbitrage-free exponential Lévy models in terms of their characteristic triplets. This condition is always satisfied if the process X has no finite-variation components: in this case $\mathcal{L} = \{0\}$ and condition (3) reduces to $0 \in \{0\}$. If the process is of finite variation, this condition reduces to $-b \in \mathrm{ri}(\mathrm{cc}(\mathrm{supp}\,\nu))$, that is, the drift and the finite variation jumps must point in opposite directions.

Proof (Proof of Theorem 2). $1 \Rightarrow 2$ is readily obtained by an application of Theorem 1 to the process $w.X$.

$2 \Rightarrow 1$. By an argument similar to the one in the proof of Theorem 1, we can suppose without loss of generality that for all $\lambda \in \mathbb{R}^d$, $E[e^{\lambda.X_1}] < \infty$. The function $f : \lambda \mapsto E[e^{\lambda.X_1}] < \infty$ is then a proper convex differentiable function on $\mathbb{R}^d$ and if λ^* is a minimizer of this function, $E[X_1^i e^{\lambda^*.X_1}] = 0$ for all $i = 1, \ldots, d$ and we can define an equivalent martingale measure $\mathbb{Q}$ using the Esscher transform

$$\frac{d\mathbb{Q}|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} := \frac{e^{\lambda^*.X_1}}{E[e^{\lambda^*.X_1}]}.$$

Suppose that $w.X$ is a Lévy process satisfying conditions (2) or (3) of Theorem 1. Then it follows from the proof of this theorem that $w.X$ is either constant or $\lim_{\lambda\to\infty} E[e^{\lambda w.X_1}] = \infty$. Hence, the function f is constant along every recession direction, which implies that f attains its minimum (Theorem 27.1 in [43]).

$2 \Rightarrow 3$. Suppose $-b^{\mathcal{L}} \notin \mathrm{ri}(\mathrm{cc}(\mathrm{supp}\,\nu^{\mathcal{L}}))$. Then $-b^{\mathcal{L}}$ can be weakly separated from $\mathrm{cc}(\mathrm{supp}\,\nu^{\mathcal{L}})$ by a hyperplane contained in $\mathcal{L}$, passing through the origin, and which does not contain $-b^{\mathcal{L}}$ or $\mathrm{cc}(\mathrm{supp}\,\nu^{\mathcal{L}})$ completely (Theorems 11.3 and 11.7 in [43]). This means that there exists $w \in \mathcal{L}$ such that

$$b^{\mathcal{L}}.w \geq 0 \quad \text{and} \quad x.w \geq 0, \quad \forall x \in \mathrm{supp}\,\nu^{\mathcal{L}}$$

with either $b^{\mathcal{L}}.w > 0$ or $x.w > 0$ for some $x \in \mathrm{supp}\,\nu^{\mathcal{L}}$. In this case, $\mathrm{ri}(\mathrm{cc}(\mathrm{supp}\,\nu^w))$ is either $\{0\}$ or $(0, \infty)$, where the measure ν^w is defined by $\nu^w(A) := \nu^{\mathcal{L}}(\{x \in \mathcal{L} : w.x \in A\})$. If $b^{\mathcal{L}}.w > 0$, this implies that $-b^{\mathcal{L}}.w \notin \mathrm{ri}(\mathrm{cc}(\mathrm{supp}\,\nu^w))$. If $b^{\mathcal{L}}.w = 0$ then necessarily $x.w > 0$ for some $x \in \mathrm{supp}\,\nu^{\mathcal{L}}$ which means that in this case $\mathrm{ri}(\mathrm{cc}(\mathrm{supp}\,\nu^w)) = (0, \infty)$ and once again $-b^{\mathcal{L}}.w \notin \mathrm{ri}(\mathrm{cc}(\mathrm{supp}\,\nu^w))$. In both cases, we have obtained a contradiction with 2.

$3 \Rightarrow 2$. Assume that $-b^{\mathcal{L}} \in \mathrm{ri}(\mathrm{cc}(\mathrm{supp}\,\nu^{\mathcal{L}}))$ and let $w \in \mathbb{R}^d$. If $w \notin \mathcal{L}$ than $w.X$ has infinite variation and the claim is shown. Assume that $w \in \mathcal{L}$ and let $R^w := \mathrm{ri}(\mathrm{cc}(\mathrm{supp}\,\nu^w))$. R^w can be equal to $\mathbb{R}$, half-axis or a single point $\{0\}$. If $R^w = \mathbb{R}$, there is nothing to prove. In the two other cases, $-b.w \notin R^w$ means that w weakly separates $-b^{\mathcal{L}}$ from $\mathrm{cc}(\mathrm{supp}\,\nu^{\mathcal{L}})$ in such a way that either $b^{\mathcal{L}}.w > 0$ or $x.w > 0$ for some $x \in \mathrm{supp}\,\nu^{\mathcal{L}}$, which is a contradiction with (3).

Case of models based on ordinary exponentials

In multidimensional models of type (5), contrary to the one-dimensional case, the no-arbitrage conditions are not the same as in models of type (4), as the following example illustrates. Let N be a standard Poisson process with intensity λ and define

$$X_t^1 = N_t - \lambda(e-1)t; \qquad S_t^1 = S_0^1 e^{X_t^1}.$$
$$X_t^2 = -N_t - \lambda(e^{-1}-1)t; \qquad S_t^2 = S_0^2 e^{X_t^2}.$$

The linear combination $X^1 + X^2$ is nonconstant and monotone, however the model is arbitrage-free since S^1 and S^2 are easily seen to be martingales.

To check whether a model of type (5) based on an $\mathbb{R}^d$-valued Lévy process X is arbitrage-free, one should construct the equivalent model of type (4) by computing $Y_t^i = \ln \mathcal{E}(X^i)_t$ for $i = 1, \ldots, d$, and then check the conditions of Theorem 2 for the process Y. The following remarks can facilitate this task in some commonly encountered cases:

- The space $\mathcal{L}$ of finite variation components is invariant under the mapping $\ln \mathcal{E}$; therefore, if the process X does not have finite variation components, the model is arbitrage-free.
- If the Lévy measure ν^X of X has full support then the Lévy measure ν^Y of Y satisfies $\mathrm{cc}(\mathrm{supp}\, \nu^Y) = \mathbb{R}^d$, which implies that the model is arbitrage-free.
- If an orthant is contained in the support of ν^X, this orthant will also be contained in $\mathrm{cc}(\mathrm{supp}\, \nu^Y)$.

3 European Options in exp-Lévy Models

Given the results of Sect. 2, in any "reasonable" exponential Lévy model we can assume that there exists a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that the discounted prices of all assets are $\mathbb{Q}$-martingales. In practice, this measure is usually found by calibrating the exponential Lévy model to market quoted prices of European options [5, 17], and the first step in using the model is therefore to obtain fast pricing algorithms for European calls and puts.

Prices of European options in exponential Lévy models can be computed directly from the characteristic function of X which is explicitly known from the Lévy–Khintchine formula. This idea was first introduced to finance by Carr and Madan [10] (for European calls and puts) and later extended and generalized by many authors including [19, 36, 37, 42]. The result given below is a slight generalization of the one in [19], allowing both discontinuous pay-off functions and Lévy processes without a bounded density, such as variance gamma.

We start with a one-dimensional risk-neutral exponential Lévy model in the form (5). Under the risk-neutral probability, the process e^X must therefore be a martingale, a condition which can be expressed in terms of the characteristic triplet of X:

$$\gamma + \frac{A}{2} + \int_{\mathbb{R}} (e^y - 1 - y1_{|y|\leq 1})\nu(dy) = 0.$$

We consider a European option with pay-off $G(S_T) = G(\hat{S}_T)$ at time T and denote by g its log-payoff function: $G(e^x) \equiv g(x)$. As above, we denote by Φ_t the characteristic function of X_t.

Proposition 2. *Suppose that there exists $R \neq 0$ such that*

$$g(x)e^{-Rx} \quad \textit{has finite variation on } \mathbb{R}, \tag{15}$$

$$g(x)e^{-Rx} \in L^1(\mathbb{R}), \tag{16}$$

$$E[e^{RX_{T-t}}] < \infty \quad \textit{and} \quad \int_{\mathbb{R}} \frac{|\Phi_{T-t}(u-iR)|}{1+|u|}du < \infty. \tag{17}$$

Then the price at time t of the European option with pay-off function G satisfies

$$\begin{aligned} P(t, S_t) &:= e^{-r(T-t)}E[G(S_T)|\mathcal{F}_t] \\ &= \frac{e^{-r(T-t)}}{2\pi} \int_{\mathbb{R}} \hat{g}(u + iR)\Phi_{T-t}(-u - iR)\hat{S}_t^{R-iu}du, \end{aligned} \tag{18}$$

where

$$\hat{g}(u) := \int_{\mathbb{R}} e^{iux}g(x)dx.$$

Proof. By integration by parts for Stieltjes integrals,

$$\hat{g}(u + iR) = \int_{\mathbb{R}} g(x)e^{ix(u+iR)}dx = \frac{i}{u + iR} \int_{\mathbb{R}} e^{ix(u+iR)}dg(x). \tag{19}$$

This implies in particular that

$$|\hat{g}(u + iR)| \leq \frac{C}{|u + iR|}, \quad u \in \mathbb{R}. \tag{20}$$

Suppose that $R > 0$ (the case $R < 0$ can be treated in a similar manner) and consider the function

$$f(x) = e^{Rx} \int_x^\infty p(dz),$$

where p denotes the distribution of X_{T-t}. From the assumption (17) it follows

$$\int_{\mathbb{R}} e^{Rx} p(dx) < \infty$$

and therefore $\lim_{x\to\infty} f(x) = 0$. Clearly also $\lim_{x\to-\infty} f(x) = 0$. By integration by parts,

$$\int_{-N}^{N} f(x)dx = \frac{1}{R}\int_{-N}^{N} e^{Rx} p(dx) + \frac{1}{R}(f(N) - f(-N).$$

This shows that $f \in L^1(\mathbb{R})$ and it follows that

$$\int_{\mathbb{R}} e^{-iux} f(x)dx = \frac{\Phi_{T-t}(-u - iR)}{R - iu}.$$

From condition (17) it follows that f can be recovered by Fourier inversion (cf. [46, Theorem 9.11]):

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{iux} \frac{\Phi_{T-t}(-u - iR)}{R - iu} du. \tag{21}$$

Let us now turn to the proof of (18). From (19), (21) and Fubini's theorem,

$$\frac{1}{2\pi}\int_{\mathbb{R}} \hat{g}(u + iR)\Phi_{T-t}(-u - iR)\hat{S}_t^{R-iu} du \tag{22}$$

$$= \frac{1}{2\pi}\int_{\mathbb{R}} dg(x) e^{Rx} \int_{\mathbb{R}} du \frac{e^{iux}\Phi_{T-t}(-u - iR)e^{(R-iu)\log \hat{S}_t}}{R - iu} \tag{23}$$

$$= \int_{\mathbb{R}} dg(x) e^{-R(x-\log \hat{S}_t)} f(x - \log \hat{S}_t) = \int_{\mathbb{R}} dg(x) \int_{x-\log \hat{S}_t}^{\infty} p(dz) \tag{24}$$

$$= \int_{\mathbb{R}} g(x + \log \hat{S}_t) p(dx) = E^Q[G(\hat{S}_T)|\mathcal{F}_t] = E^Q[G(S_T)|\mathcal{F}_t]. \tag{25}$$

Example 1. The *digital* option has pay-off $G(S_T) = 1_{S_T \geq K}$. In this case for all $R > 0$ conditions (15) and (16) are satisfied and

$$\hat{g}(u + iR) = \frac{K^{iu-R}}{R - iu}.$$

Example 2. The European call option has pay-off $G(S_T) = (S_T - K)^+$. Therefore, conditions (15) and (16) are satisfied for all $R > 1$,

$$\hat{g}(u + iR) = \frac{K^{iu+1-R}}{(R - iu)(R - 1 - iu)}.$$

and the price of a call option can be written as an inverse Fourier transform:

$$C(t,S_t) = \frac{e^{-r(T-t)}}{2\pi}\int_{\mathbb{R}} \frac{K^{iu+1-R}\hat{S}_t^{R-iu}\Phi_{T-t}(-u-iR)}{(R-iu)(R-1-iu)}du$$
$$= \frac{S_t}{2\pi}\int_{\mathbb{R}} \frac{e^{k^f(iu+1-R)}\Phi_{T-t}(-u-iR)}{(R-iu)(R-1-iu)}du$$

where k^f is the log forward moneyness defined by $k^f = \ln(K/S_t) - r(T-t)$. This property allows to compute call option prices for many values of k^f in a single computation using the FFT algorithm (see [10] or [16] for details).

4 Implied Volatility

Recall the well-known Black–Scholes formula for call option prices:

$$C^{BS}(t,S_t,T,K,\sigma) = S_tN(d_1) - Ke^{-r(T-t)}N(d_2) \tag{26}$$
$$\text{with} \quad d_{1,2} = \frac{\log(\frac{S_t}{Ke^{-r\tau}}) \pm \tau\sigma^2/2}{\sigma\sqrt{\tau}} \quad \text{and} \quad N(u) \equiv \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-\frac{z^2}{2}}dz,$$

where $\tau = T-t$. If all other parameters are fixed, (26) is an increasing continuous function of σ, mapping $(0,\infty)$ into $((S_t - Ke^{-r\tau})^+, S_t)$. The latter interval is the greatest interval allowed by arbitrage bounds on call option prices. Therefore, given the market price $C_t^*(T,K)$ of a call option, one can always invert (26) and find the value of volatility parameter which, when substituted into the Black–Scholes formula, gives the correct option price:

$$\exists! \quad \Sigma_t(T,K) > 0: \qquad C^{BS}(t,S_t,T,K,\Sigma_t(T,K)) = C_t^*(K,T). \tag{27}$$

This value is called the (Black–Scholes) implied volatility of the option. For fixed (T,K), the implied volatility $\Sigma_t(T,K)$ is in general a stochastic process and, for fixed t, its value depends on the characteristics of the option such as the maturity T and the strike level K: the function $\Sigma_t : (T,K) \to \Sigma_t(T,K)$ is called the *implied volatility surface* at date t (see Fig. 1). Using the log moneyness $k = \log(K/S_t)$ of the option, one can also represent the implied volatility surface as a function of k and time to maturity: $I_t(\tau,k) = \Sigma_t(t+\tau, S_te^k)$. From the independence and stationarity of increments of X, it follows that the definition of implied volatility (27) is equivalent to

$$E[(e^{X_\tau} - e^{k-r\tau})^+] = E[(e^{IW_\tau - \frac{I^2\tau}{2}} - e^{k-r\tau})^+].$$

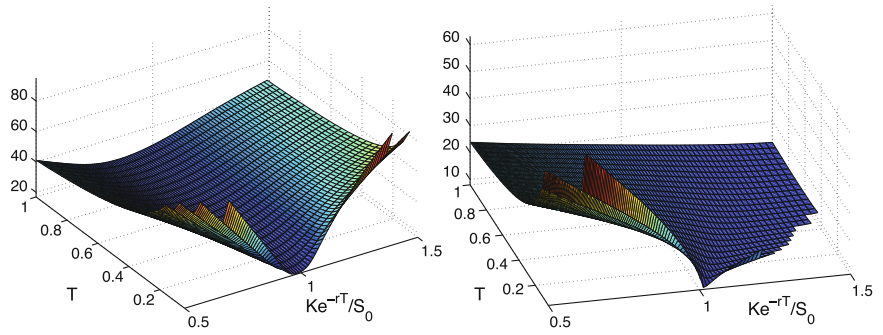


Fig. 1 *Left*: Profile of the implied volatility surface as a function of time to maturity and moneyness for the Merton jump-diffusion model with $\sigma = 15\%$, $\delta = 1$ and $\lambda = 0.1$. *Right*: Implied volatility surface as a function of time to maturity and moneyness for the variance gamma model using parameters taken from [38]. Note the flattening of the skew with maturity

Since each side depends only on (τ, k) and not on t one concludes that in exponential Lévy models, the implied volatility for a given log moneyness k and time to maturity τ does not evolve in time: $I_t(\tau, k) = I_0(\tau, k) := I(\tau, k)$. This property is known as the *floating smile property*.

In exponential Lévy models, the properties of the implied volatility surfaces can be characterized in terms of the asymptotic behavior of the surface for large and small values of strike and maturity. We start with the large and small strike behavior which was first analyzed by Roger Lee [35]; this analysis was subsequently extended and made more precise by Benaim and Friz [23, 24]. Their results, reviewed below, take a particularly simple form in the case of Lévy processes, because the critical exponents do not depend on time. Next, we study the short maturity asymptotics, where it turns out that the behavior of the implied volatility is very different for out of the money (OTM) and at the money (ATM) options. Below, we present some original results for the two cases. Finally, the long-maturity asymptotics were recently studied by Tehranchi [56, 57] and Rogers and Tehranchi [44]. We review their results in the case of Lévy processes, where once again, the formulation is particularly simple and interesting links to the large deviations theory and Cramér's theorem can be made.

Large/small strikes

The limiting slope of time-rescaled implied variance as a function of log-strike turns out to be related to the critical exponents of the moment generating function of the log-price process X, defined by

$$q_t^* = -\inf\{u : E[e^{uX_t}] < \infty\}, \qquad r_t^* = \sup\{u : E[e^{uX_t}] < \infty\}.$$

It is clear that the interval $[-q_t^*, r_t^*]$ is nonempty, because $E[e^{uX_t}] < \infty$ at least for all $u \in [0, 1]$ by the martingale condition.

Proposition 3 (Implied volatility asymptotics at extreme strikes [23]). *Fix $\tau>0$ and suppose that $r_\tau^* \in (0,\infty)$ and $q_\tau^* \in (0,\infty)$ and that the moment generating function blows up in a regularly varying way around its critical exponents (see [23] for a precise definition). Then the implied volatility $I(\tau,k)$ satisfies*

$$\frac{I^2(\tau,-k)\tau}{|k|} \sim \xi(q_\tau^*) \quad \textit{and} \quad \frac{I^2(\tau,k)\tau}{k} \sim \xi(r_\tau^*-1), \quad \textit{as} \quad k \to +\infty,$$

where the function ξ is defined by $\xi(x) = 2 - 4(\sqrt{x^2+x} - x)$.

This proposition extends in a natural way to the case of infinite critical exponents:

$$\frac{I^2(\tau,-k)\tau}{|k|} \xrightarrow{k\to+\infty} 0 \quad \text{if } q_\tau^* = \infty \quad \text{and} \quad \frac{I^2(\tau,k)\tau}{k} \xrightarrow{k\to+\infty} 0 \quad \text{if } p_\tau^* = \infty.$$

This was shown already in the original work of Roger Lee [35].

For Lévy processes, the exponents q^* and r^* do not depend on t and are particularly easy to compute, since the moment generating function is known from the Lévy–Khintchine formula. In particular, the models with exponential tail decay of the Lévy measure such as variance gamma, normal inverse Gaussian and Kou satisfy the necessary conditions for the Proposition 3 and their critical exponents coincide with the inverse decay lengths: $q^* = \lambda_-$ and $r^* = \lambda_+$. Figure 2 shows that the asymptotic linear slope of the implied variance as a function of log strike can be observed for values of k which are not so far from zero, especially for short maturity options.

In Merton model, the tails of the Lévy measure are thinner than exponential and the critical exponents q^* and r^* are infinite. The remark after Proposition 3 then

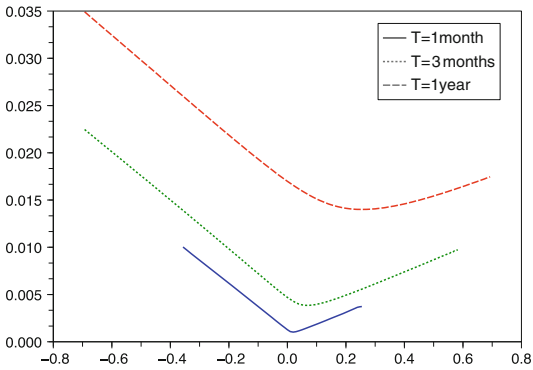


Fig. 2 Smile asymptotics: implied variance multiplied by square root of maturity as function of log-strike in the variance gamma model with parameters taken from [38]: $\sigma = 0.1213$, $\theta = -0.1436$, $\kappa = 0.1686$

only tells us that the limiting slope of the implied variance is zero, but other results in [24] allow to compute the exact asymptotics: for the right tail we have

$$I^2(\tau,k)\tau \underset{k\to\infty}{\sim} \delta \times \frac{k}{2\sqrt{2\log k}}, \quad \text{when } \delta > 0$$

and

$$I^2(\tau,k)\tau \underset{k\to\infty}{\sim} \mu \times \frac{k}{2\log k}, \quad \text{when } \delta = 0,$$

where δ is the standard deviation of the jump size and μ is the mean jump.

Short maturity asymptotics

The short maturity behavior of implied volatility in exponential Lévy models is very different from that observed in stochastic/local volatility models with continuous paths. While in continuous models the implied volatility usually converges to a finite nonzero value as $\tau \to 0$, in models with jumps the implied volatility of out of the money or in the money options blows up. On the other hand, the implied volatility of at-the-money options converges to the volatility of the diffusion component as $\tau \to 0$; in particular it converges to zero for pure jump models. This leads to very pronounced smiles for short maturity options (in agreement with market-quoted smiles). The intuitive explanation of this effect is that in most continuous models, the stock returns at short time scales become close to Gaussian; in particular, the skewness and excess kurtosis converge to zero as $\tau \to 0$. By contrast, in models with jumps, the distribution of stock returns at short time scales shifts further away from the Gaussian law; the skewness and kurtosis explode as $\frac{1}{\sqrt{\tau}}$ and $\frac{1}{\tau}$ respectively.

The short maturity asymptotics of implied volatility smile in exponential Lévy models can be computed by comparing the option price asymptotics in the Black–Scholes model to those in the exponential Lévy model (some results in this direction can be found in Carr and Wu [12] and Roper [45]). To simplify the developments, we suppose that the interest rate is zero. Then the normalized Black–Scholes price satisfies

$$c_{BS}(\tau,k,\sigma) = N(d_1) - e^k N(d_2), \quad d_{1,2} = \frac{-k}{\sigma\sqrt{\tau}} \pm \frac{1}{2}\sigma\sqrt{\tau}.$$

Using the asymptotic expansion of the function N [1], we get, for the ATM options ($k = 0$):

$$c_{BS}(\tau,k,\sigma) \sim \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} \tag{28}$$

and for other options

$$c_{BS}(\tau,k,\sigma) \sim \frac{e^{k/2}}{k^2\sqrt{2\pi}}\sigma^3\tau^{3/2}e^{-\frac{k^2}{2\sigma^2\tau}}, \tag{29}$$

where the notation $f \sim g$ signifies $\frac{f}{g} \to 1$ as $\tau \to 0$.

In every exponential Lévy model satisfying the martingale condition, we have [47]

$$E[(e^{X_\tau} - e^k)^+] \sim \tau \int (e^x - e^k)^+ \nu(dx), \quad \text{for } k > 0 \tag{30}$$

$$E[(e^k - e^{X_\tau})^+] \sim \tau \int (e^k - e^x)^+ \nu(dx), \quad \text{for } k < 0 \tag{31}$$

From these estimates, the following universal result can be deduced: it confirms the numerical observation of smile explosion in exponential Lévy models and gives the exact rate at which this explosion takes place.

Proposition 4 (Short maturity asymptotics: OTM options). *Let X be a Lévy process with Lévy measure ν satisfying $\operatorname{supp}\nu = \mathbb{R}$. Then, for a fixed log moneyness $k \neq 0$, the implied volatility $I(\tau, k)$ in the exponential Lévy model $S_t = S_0 e^{X_t}$ satisfies*

$$\lim_{\tau \to 0} \frac{2I^2(\tau,k)\tau \log \frac{1}{\tau}}{k^2} = 1. \tag{32}$$

Proof. Suppose first that $k > 0$. It is clear that $I(\tau, k)\sqrt{\tau} \to 0$ as $\tau \to 0$ because otherwise the option price would not converge to 0. We then have, from OTM Black–Scholes asymptotics (29):

$$\lim_{\tau \to 0} \frac{c_{BS}(\tau, k, I(\tau,k))}{C_1 I(\tau,k)^3 \tau^{3/2} e^{-\frac{k^2}{2I^2(\tau,k)\tau}}} = 1,$$

where $C_1 > 0$ does not depend on τ. Denote the (normalized) call price in the exponential Lévy model by $c(\tau, k)$. Under the full support hypothesis, $c(\tau, k) \sim C_2\tau$ with $C_2 > 0$ which once again does not depend on τ. By definition of the implied volatility we then have

$$\lim_{\tau \to 0} \frac{C_2\tau}{C_1 I(\tau,k)^3 \tau^{3/2} e^{-\frac{k^2}{2I^2(\tau,k)\tau}}} = 1.$$

Taking the logarithm gives

$$\lim_{\tau \to 0} \{\log(C_2/C_1) + 3\log I(\tau,k) + \frac{1}{2}\log\tau - \frac{k^2}{2I^2(\tau,k)\tau}\} = 0.$$

Now, knowing that $I^2(\tau, k)\tau \to 0$, we can multiply all terms by $I^2(\tau, k)\tau$:

$$\lim_{\tau \to 0} \{I^2\tau \log(C_2/C_1) + \frac{3}{2} I^2\tau \log(I^2\tau) - I^2\tau \log\tau - \frac{k^2}{2}\} = 0.$$

Since the first two terms disappear in the limit, this completes the proof in the case $k > 0$. The case $k < 0$ can be treated in a similar manner using put options.

For ATM options, the situation is completely different, from estimate (28) we will deduce that the implied volatility does not explode but converges to the volatility of the diffusion component.

Proposition 5 (Short maturity asymptotics: ATM options).

1. *Let X be a Lévy process without diffusion component and with Lévy measure ν satisfying $\int_{|x|\leq 1}|x|\nu(dx) < \infty$. Then, the ATM implied volatility $I(\tau,0)$ in the exponential Lévy model $S_t = S_0 e^{X_t}$ falls as $\sqrt{\tau}$ for short maturities:*

$$\lim_{\tau\to 0}\frac{I(\tau,0)}{\sqrt{2\pi\tau}\max\left(\int(e^x-1)^+\nu(dx),\int(1-e^x)^+\nu(dx)\right)} = 1.$$

2. *Let X be a Lévy process with characteristic exponent*

$$\psi(u) = i\gamma u - |u|^\alpha f(u)$$

for $1<\alpha<2$ and some continuous bounded function f satisfying

$$\lim_{u\to+\infty} f(u) = c_+, \quad \lim_{u\to-\infty} f(u) = c_-, \quad 0 < c_1, c_2 < \infty.$$

This includes in particular stable and tempered stable processes with $1<\alpha<2$. Then, the ATM implied volatility $I(\tau,0)$ in the exponential Lévy model $S_t = S_0 e^{X_t}$ falls as $\tau^{1/\alpha-1/2}$ for short maturities:

$$\lim_{\tau\to 0}\frac{I(\tau,0)}{C\tau^{1/\alpha-1/2}\sqrt{2\pi}} = 1.$$

with $C = -\frac{1}{\alpha}\Gamma(-1/\alpha)(c_+^{1/\alpha}+c_-^{1/\alpha})$.

3. *Let X be a Lévy process with a diffusion component with volatility σ and Lévy measure satisfying $\int x^2\nu(dx) < \infty$. Then the ATM implied volatility $I(\tau,0)$ in the exponential Lévy model $S_t = S_0 e^{X_t}$ converges to σ as $\tau\to 0$.*

The short-maturity smile asymptotics are illustrated in Fig. 3: the ATM implied volatility converges to the value of σ and the out of the money and in the money volatilities eventually become very large as τ approaches zero.

Proof.

1. Let b denote the drift of X. Since X is a finite-variation process, the Itô–Tanaka formula applied to the function $(1-e^{X_\tau})^+$ does not yield a local time term, and we obtain

$$\begin{aligned}&E[(1-e^{X_\tau})^+]\\ &\quad = E\left[b\int_0^\tau e^{X_t}1_{X_t\leq 0}dt + \int_0^\tau\int_{\mathbb{R}}\nu(dx)\{(1-e^{X_t+x})^+ - (1-e^{X_t})^+\}\right]dt.\end{aligned}$$

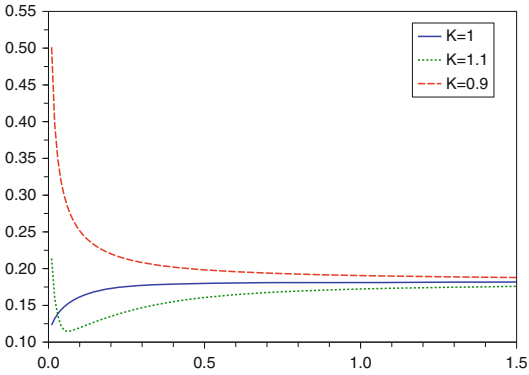


Fig. 3 Smile asymptotics: term structure of implied volatility for different strikes in Merton's model with parameters $\sigma = 0.1$, $\lambda = 2$, $\mu = -0.1$ and $\delta = 0.5$

By L'Hopital's rule,

$$\lim_{\tau\to 0}\frac{1}{\tau}E[(1-e^{X_\tau})^+]$$
$$= b\lim_{\tau\to 0}E[e^{X_\tau}1_{X_\tau\le 0}] + \lim_{\tau\to 0}E\left[\int_{\mathbb{R}}\nu(dx)\{(1-e^{X_\tau+x})^+ - (1-e^{X_\tau})^+\}\right].$$

From Theorem 43.20 in [48], $\frac{X_t}{t} \to b$ almost surely as $t \to 0$. From this we deduce that $\lim_{\tau\to 0} E[e^{X_\tau}1_{X_\tau\le 0}] = 1_{b\le 0}$. Using the dominated convergence for the second term above, we finally obtain

$$\lim_{\tau\to 0}\frac{1}{\tau}E[(1-e^{X_\tau})^+] = b1_{b\le 0} + \int_{\mathbb{R}}\nu(dx)(1-e^x)^+.$$

Since by the martingale condition,

$$b + \int_{\mathbb{R}}(e^x - 1)\nu(dx) = 0,$$

this limit can be rewritten as

$$\lim_{\tau\to 0}\frac{1}{\tau}E[(1-e^{X_\tau})^+] = \max\Big(\int(e^x-1)^+\nu(dx), \int(1-e^x)^+\nu(dx)\Big).$$

Comparing this expression with the Black–Scholes ATM asymptotics, we obtain the desired result.

2. Let p_t denote the density of X_t (which exists and is square integrable under the hypotheses of this part). The ATM call option price is given by

$$c(\tau,0) = \int(e^x-1)^+p_t(x)dx.$$

Let us fix a constant $\beta < -1$ and define

$$\tilde{c}(\tau, 0) = \int (e^x - 1)^+ e^{\beta x} p_t(x) dx.$$

Then it follows from results in [47] that

$$|c(\tau, 0) - \tilde{c}(\tau, 0)| = \int (e^x - 1)^+ (1 - e^{\beta x}) p_t(dx) = O(\tau)$$

as $\tau \to 0$. This means that it is sufficient to study the decay properties of $\tilde{c}$. This function is a scalar product of the square integrable function p_t by the square integrable function $(e^x - 1)^+ e^{\beta x}$. By the Plancherel theorem we then have

$$-\tilde{c}(\tau, 0) = \frac{1}{2\pi} \int \frac{e^{\tau\psi(u)}}{(u - i\beta)(u - i - i\beta)} du = \frac{1}{2\pi} \int \frac{e^{i\tau\gamma u - \tau|u|^\alpha f(u)}}{(u - i\beta)(u - i - i\beta)} du.$$

On the other hand, direct computation shows that

$$\int \frac{e^{i\gamma u\tau} du}{(u - i\beta)(u - i - i\beta)} = O(\tau)$$

as $\tau \to 0$. Then, changing the variable of integration, we obtain

$$-\tilde{c}(\tau, 0) = \frac{\tau^{1/\alpha}}{2\pi} \int \frac{e^{i\gamma z \tau^{1-1/\alpha}} (e^{-|z|^\alpha f(\tau^{-1/\alpha} z)} - 1)}{(z - i\tau^{1/\alpha}\beta)(z - i\tau^{1/\alpha}(1 + \beta))} dz + O(\tau).$$

The dominated convergence theorem then yields

$$-\tau^{-1/\alpha} \tilde{c}(\tau, 0) \to \frac{1}{2\pi} \int_{-\infty}^{0} \frac{e^{-c_-|z|^\alpha} - 1}{z^2} dz + \frac{1}{2\pi} \int_0^{\infty} \frac{e^{-c_+|z|^\alpha} - 1}{z^2} dz$$

as $\tau \to \infty$. This result generalizes the findings of Carr and Wu [12]. Computing the integrals and comparing the result to the Black–Scholes at the money asymptotics, we obtain the final result.

3. Under the conditions of this part, we can write the characteristic exponent of X as $\psi(u) = i\gamma u - f(u)u^2$ for a continuous bounded function f satisfying $\lim_{u\to\infty} f(u) = \frac{\sigma^2}{2}$. Then, exactly as in the previous part, the dominated convergence theorem yields

$$-\frac{\tilde{c}(\tau, 0)}{\sqrt{\tau}} \to \frac{1}{2\pi} \int_{\mathbb{R}} \frac{e^{-\frac{\sigma^2 u^2}{2}} - 1}{u^2} du = \frac{\sigma}{\sqrt{2\pi}},$$

which is equal to the Black–Scholes ATM asymptotics.

Flattening of the skew/smile far from maturity

As the time to maturity τ goes to infinity, the implied volatility $I(\tau, k)$ in an exponential Lévy model converges to a constant value $I(\infty)$ which does not depend on k (see Figs. 1 and 3). As a result, the implied volatility smile flattens for long maturities, a phenomenon which is also observed in the options markets, although with a slower rate. This flattening has been often attributed (e.g., [11]) to the central limit theorem, according to which, for a Lévy process with finite variance, the distribution of increments $(X_\tau - E[X_\tau])/\sqrt{\tau}$ becomes approximately Gaussian as τ goes to infinity. However, contrary to this intuition, the flattening of the smile is not a consequence of the central limit theorem, but, rather, of a "large deviation" principle which governs the tail behavior of the sample average of n i.i.d. random variables. In fact, as observed by Rogers and Tehranchi [44], the implied volatility flattens even in models where log-returns have infinite variance such as the finite moment log-stable process of [11].

To understand this, consider a Lévy process X with $E[X_1] < \infty$. Since in a risk-neutral model $E[e^{X_t}] = 1$, the Jensen inequality implies that $E[X_t] < 0$ for all t. Therefore, by the law of large numbers, $X_t \to -\infty$ almost surely as $t \to \infty$, which means that $e^{X_t} \to 0$ a.s. The exercise of a long-dated call option is thus an event with a very small probability. The probability of such rare events is given by Cramér's theorem, which is the cornerstone of the theory of large deviations, rather than by the CLT.

The normalized price of a call option with log-moneyness k can be written as

$$c(\tau, k) = E(e^{X_\tau} - e^k)^+ = \tilde{\mathbb{P}}[X_\tau \geq k] - e^k \mathbb{P}[X_\tau \geq k],$$

where we introduce the new probability $\tilde{\mathbb{P}}$ via the Esscher transform:

$$\frac{d\tilde{\mathbb{P}}|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} := e^{X_t}.$$

Denote $\alpha = E[X_1]$ and $\tilde{\alpha} = \tilde{E}[X_1]$. An easy computation using Proposition 1 shows that

$$\alpha = -\frac{\sigma^2}{2} - \int_{\mathbb{R}} (e^x - x - 1)\nu(dx) < 0$$
$$\tilde{\alpha} = \frac{\sigma^2}{2} + \int_{\mathbb{R}} (xe^x - e^x + 1)\nu(dx) > 0.$$

To make the probability of a rare event appear, we rewrite the option price as

$$c(\tau, k) = 1 - \tilde{\mathbb{P}}\left[-\frac{X_\tau - \tilde{\alpha}\tau}{\tau} > \tilde{\alpha} - \frac{k}{\tau}\right] - e^k \mathbb{P}\left[\frac{X_\tau - \alpha\tau}{\tau} \geq -\alpha + \frac{k}{\tau}\right].$$

These probabilities can be estimated with the help of the famous Cramér's theorem which gives the exact convergence rate in the law of large numbers.

Theorem 3 (Cramér). *Let $\{X_i\}_{i\geq 1}$ be an i.i.d. sequence of random variables with $E[X_i] = 0$ for all i. Then for all $x \geq 0$*

$$\lim_{n\to\infty} \frac{1}{n} \log P\left[\frac{1}{n}\sum_{i=1}^{n} X_i \geq x\right] = -I(x),$$

where $I(x)$ is the Fenchel transform of the log-Laplace transform of X_1:

$$I(x) = \sup_{\theta}(\theta x - l(\theta)), \qquad l(\theta) = \log E[e^{\theta X_1}].$$

Suppose that the Lévy measure ν is such that[1]

$$\int_{|x|>1} x\nu(d\dot{x}) < \infty \quad \text{and} \quad \int_{|x|>1} xe^{x}\nu(dx) < \infty \tag{33}$$

and define the log-Laplace transforms by

$$\tilde{l}(\theta) := \log \tilde{E}[e^{-\theta(X_1-\tilde{\alpha})}] \quad \text{and} \quad l(\theta) := \log E[e^{\theta(X_1-\alpha)}],$$

and the respective Fenchel transforms by $\tilde{I}$ and I. A direct computation then shows that

$$\tilde{I}(\tilde{\alpha}) = I(-\alpha) = \sup_{\theta}\left\{\frac{\sigma^2}{2}(\theta-\theta^2) - \int_{\mathbb{R}}(e^{\theta x} - \theta e^{x} - 1 + \theta)\nu(dx)\right\},$$

and that the functions $\tilde{I}$ and I are finite and hence, continuous, in the neighborhood of, respectively, $\tilde{\alpha}$ and $-\alpha$. Hence the sup above can be restricted to the interval $\theta \in [0,1]$, since the function being maximized is concave and equal to 0 for $\theta = 0$ and $\theta = 1$. Using Cramér's theorem and the continuity of $\tilde{I}$ and I, we then obtain

$$\lim_{\tau\to\infty} \frac{1}{\tau}\log(1 - c(\tau,k)) = \sup_{\theta\in[0,1]}\left\{\frac{\sigma^2}{2}(\theta-\theta^2) - \int_{\mathbb{R}}(e^{\theta x} - \theta e^{x} - 1 + \theta)\nu(dx)\right\}. \tag{34}$$

Note that this formula is valid for any k, we can even take k to be a function of τ as long as $k = o(\tau)$ as $\tau \to \infty$. Specializing this formula to the Black–Scholes model, where $\nu \equiv 0$ and the sup can be computed explicitly, we get

$$\lim_{\tau\to\infty} \frac{1}{\tau}\log(1 - c_{BS}(\tau,k,\sigma)) = \frac{\sigma^2}{8}.$$

[1] The finite moment log stable process of Carr and Wu [12] satisfies these hypotheses although the variance of the log-price is infinite in this model.

From (34), it follows in particular that the implied volatility satisfies $\tau I^2(\tau,k) \to \infty$ as $\tau \to \infty$ (otherwise the call option price would not converge to 1). Since in the Black–Scholes model the option price depends only on $\tau\sigma^2$ but not on τ or σ separately, we can write

$$\lim_{\tau\to\infty} \frac{1}{\tau I^2(\tau,k)} \log(1 - c_{BS}(\tau I^2(\tau,k), k, 1)) = \frac{1}{8},$$

and combining this with (34), we obtain the final result:

Proposition 6 ([57]). *Let X be a Lévy process with Lévy measure satisfying (33). Then the implied volatility $I(\tau,k)$ in the exponential Lévy model $S_t = S_0 e^{X_t}$ satisfies*

$$\lim_{\tau\to\infty} I^2(\tau,k) = 8 \sup_\theta \left\{ \frac{\sigma^2}{2}(\theta - \theta^2) - \int_{\mathbb{R}} (e^{\theta x} - \theta e^x - 1 + \theta)\nu(dx) \right\}. \quad (35)$$

The exact formula (35) for the limiting long-term implied volatility in an exponential Lévy model is difficult to use in practice: even if for some models such as variance gamma it yields a closed form expression, it is rather cumbersome. However, for small jump sizes, Taylor expansion shows that this expression is not very different from the total variance of the Lévy process:

$$I^2(\infty,k) \approx \sigma^2 + \int x^2 \nu(dx).$$

The smile flattening in exponential Lévy models has thus little to do with the so called aggregational normality of stock returns. One may think that the implied volatility converges to its limiting value faster for Lévy processes to which the central limit theorem applies. However, the results of Rogers and Tehranchi [44] suggest otherwise: they give the following upper bound, valid in exponential Lévy models as soon as $E[|X_t| < \infty]$, for the rate of convergence of the implied volatility skew to zero:

$$\limsup_{\tau\to\infty} \sup_{k_1,k_2\in[-M,M]} \tau \left| \frac{I(\tau,k_2)^2 - I(\tau,k_1)^2}{k_2 - k_1} \right| \le 4, \quad 0 < M < \infty.$$

See also [56] for explicit asymptotics of the derivative $\frac{\partial I(\tau,k)}{\partial k}$ as $\tau \to \infty$.

5 Hedging in Exponential Lévy Models

Exponential Lévy models generally correspond to incomplete markets, making exact replication impossible. Hedging must therefore be interpreted as approximation of the terminal pay-off with an admissible portfolio. The usual practice is to

minimize the expected squared deviation of the hedging portfolio from the contingent claim, an approach known as quadratic hedging. The resulting strategies are often explicitly computable and, more importantly, they are linear, because the hedging portfolios can be interpreted as orthogonal projections of contingent claims onto the closed linear subspace of hedgeable portfolios. To hedge a book of options written on the same underlying, a trader can therefore compute the hedge ratio for every option in the book and then add them up, just like this is typically done with delta hedging. This greatly reduces the computational cost of hedging and is an important advantage of quadratic hedging compared to other, e.g., utility-based approaches.

To define the criterion to be minimized in a mean square sense, two approaches are possible. In the first approach [9, 29, 40], the hedging strategy is supposed to be self-financing, and one minimizes the quadratic hedging error at maturity, that is, the expected squared difference between the terminal value of the hedging portfolio and the option's pay-off:

$$\inf_{V_0,\varphi} E[\,|V_T(\varphi) - H|^2] \quad \text{where} \quad V_T(\varphi) = V_0 + \int_0^T \varphi_t^0 dS_t^0 + \int_0^T \varphi_t dS_t, \quad (36)$$

where S^0 is the risk-free asset. If the interest rate is constant, we can choose the zero-coupon bond with maturity T as the risk-free asset: $S_t^0 = e^{-r(T-t)}$ and after discounting this problem becomes:

$$\inf_{\hat{V}_0,\varphi} E[|V_T(\varphi) - H|^2], \quad \text{where} \quad V_T = \hat{V}_0 + \int_0^T \varphi_t d\hat{S}_t.$$

In the second approach [21, 22, 50, 53], strategies that are not self-financing are allowed, but they are required to replicate the option's pay-off exactly: $V_T(\varphi) = H$. In an incomplete market, this means that the option's seller will have to continuously inject/withdraw money from the hedging portfolio. The cumulative amount of funds injected or withdrawn is called the *cost process*. It is given by

$$C_t(\varphi) = V_t(\varphi) - G_t(\varphi),$$

where

$$V_t(\varphi) = \varphi_t^0 S_t^0 + \varphi_t S_t$$

and G is the *gain process* given by

$$G_t = \int_0^t \varphi_s^0 dS_s^0 + \int_0^t \varphi_s dS_s.$$

The discounted cost process is then given by

$$\hat{C}_t = \varphi_t^0 + \varphi_t \hat{S}_t - \int_0^t \varphi_s d\hat{S}_s.$$

The risk-minimizing strategy, as introduced by Föllmer and Sondermann [22], is a strategy which replicates the option's pay-off, and has the cost process which varies as little as possible, that is, this strategy minimizes, at each date t, the residual cost given by

$$E[(\hat{C}_T - \hat{C}_t)^2 | \mathcal{F}_t]. \tag{37}$$

over all admissible continuations of the strategy from date t onwards. The risk-minimizing strategy always exists in the martingale case (when the discounted stock price is a martingale), but in the general case, it may fail to exist even in the most simple examples [50]. Motivated by this difficulty, Föllmer and Schweizer [21] introduced the notion of *locally risk minimizing strategy*, which corresponds to finding the extremum of (37) with respect to suitably defined small perturbations of the strategy, or, in other words, measuring the riskiness of the cost process locally in time. Local risk minimization is discussed in detail in Sect. 5.2.

The expectations in (37) and (36) are taken with respect to some probability which we have to specify. To begin, let us assume that we have chosen a martingale measure $\mathbb{Q}$ and the expectations in (36) and (37) are taken with respect to $\mathbb{Q}$. In particular, $\hat{S}$ is a martingale under $\mathbb{Q}$. Assume now that $H \in L^2(\Omega, \mathcal{F}, \mathbb{Q})$ and $\hat{S}$ is also square-integrable. If we consider portfolios of the form:

$$\mathcal{S} = \{\varphi \text{ caglad predictable and } E|\int_0^T \varphi_t d\hat{S}_t|^2 < \infty\} \tag{38}$$

then the set $\mathbb{A}$ of attainable pay-offs is a closed linear subspace of $L^2(\Omega, \mathcal{F}, \mathbb{Q})$, and the quadratic hedging problem becomes an orthogonal projection:

$$\inf_{\hat{V}_0, \varphi} E|V_T(\varphi) - H|^2 = \inf_{A \in \mathbb{A}} \|H - A\|^2_{L^2(\mathbb{Q})}. \tag{39}$$

The solution is then given by the well-known Galtchouk–Kunita–Watanabe decomposition [25, 33], which states that any random variable $H \in L^2(\Omega, \mathcal{F}, \mathbb{Q})$ can be represented as

$$H = E[H] + \int_0^T \varphi_t^H d\hat{S}_t + N_T^H, \tag{40}$$

where (N_t^H) is a square integrable martingale orthogonal to $\hat{S}$. The optimal hedging strategy is then given by φ^H and the initial cost of the hedging portfolio is $V_0 = e^{-r(T-t)}E[H]$.

Introducing the martingale $\hat{H}_t := E[H|\mathcal{F}_t]$ generated by H, we have

$$\hat{H}_t = E[H] + \int_0^t \varphi_s^H d\hat{S}_s + N_t^H,$$

and the orthogonality implies

$$\langle \hat{H} - \int_0^{\cdot} \varphi_s^H d\hat{S}_s, \hat{S} \rangle \equiv 0,$$

which means that the optimal hedge ratio may be expressed more explicitly using the predictable covariation of the option price and the stock price:

$$\varphi_t^H = \frac{d\langle \hat{H}, \hat{S} \rangle_t}{d\langle \hat{S}, \hat{S} \rangle_t}. \tag{41}$$

In the martingale setting, optimizing the *global* hedging error (36) we obtain a strategy which is also risk minimizing in the sense of (37). For any strategy φ, we have

$$\begin{aligned} &E[(\hat{C}_T - \hat{C}_t)^2 | \mathcal{F}_t] \\ &\quad = (\hat{H}_t - \hat{V}_t)^2 + E\left[\left(H - \hat{H}_t - \int_t^T \varphi_s d\hat{S}_s \right)^2 \middle| \mathcal{F}_t \right] \\ &\quad = (\hat{H}_t - \hat{V}_t)^2 + E[(N_T - N_t)^2 | \mathcal{F}_t] + E\left[\left(\int_t^T (\varphi_s - \varphi_s^H) d\hat{S}_s \right)^2 \middle| \mathcal{F}_t \right]. \end{aligned}$$

To minimize this expression, we clearly need to take $\varphi = \varphi^H$ and choose φ^0 such that $\hat{V}_t = \hat{H}_t$ for all t. In this case, the discounted cost process is given by

$$\hat{C}_t = \hat{V}_t - \int_0^t \varphi_s^H d\hat{S}_s = E[H] + N_t^H.$$

We shall see in Sect. 5.2 that in the martingale setting, the strategy φ^H which minimizes the terminal hedging error also coincides with the locally risk minimizing strategy of Föllmer and Schweizer [21]. Moreover, it is often easy to compute in terms of option prices. This is no longer true if $\hat{S}$ is not a martingale. However using the *risk-neutral* second moment of the hedging error as a criterion for measuring risk is not very natural: $\mathbb{Q}$ represents a pricing rule and not a statistical description of market events, so the profit and loss (P&L) of a portfolio may have a large variance while its "risk neutral" variance can be small. Nevertheless, to estimate the expected return of a stock, and therefore, to distinguish it from a martingale, one needs historical stock return observations covering an extended period of time, often exceeding the lifetime of the option. Option hedging, on the other hand, is a "local" business, where one tries to cancel out the daily movements of option prices with the daily movements of the underlying and locally, every stock behaves like a martingale. Without contributing to this ongoing argument, we review both approaches in the next two sections.

5.1 Quadratic Hedging in Exponential-Lévy Models Under the Martingale Measure

Although the quadratic hedging problem is "solved" by the Galtchouk–Kunita–Watanabe decomposition, from a practical point of view the problem is of course to compute the risk minimizing hedge φ_t^H. Formulas for φ_t^H with various degrees of explicitness and under various assumptions on the driving process X and on the pay-off G were given in [6,9,20,27] and several other papers. In particular [18] provide the expressions for hedge ratios in the case when the hedging portfolio itself contains options. In the case of European pay-offs and exponential Lévy models, the problem was solved in [29] using Fourier analysis techniques. Their method, reviewed in Sect. 5.2 covers the general case as well as the martingale case. In this section, we provide another Fourier-based result, which is specialized to the martingale setting but works under different regularity assumptions on the pay-off than in [29], which include, for instance, digital options.

Proposition 7 (Quadratic hedge in exponential-Lévy models, martingale case). *Let X be a Lévy process with Lévy measure ν, diffusion coefficient σ, and characteristic function Φ, such that e^X is a martingale and assume:*

i. The log-payoff function satisfies the conditions (15) and (16).
ii. The integrability condition (17) holds for all $t < T$.
iii. The Lévy measure of X satisfies

$$\int_{|x|>1} e^{2(x \vee Rx)} \nu(dx) < \infty. \tag{42}$$

Then the optimal quadratic hedging for a European option with pay-off $G(S_T)$ at date T in an exponential Lévy model $S_t = S_0 e^{rt+X_t}$ amounts to holding a position in the underlying

$$\varphi_t = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{g}(u + iR) \Phi_{T-t}(-u - iR) \hat{S}_{t-}^{R-iu-1} \Upsilon(R - iu) du \tag{43}$$

$$\text{where} \quad \Upsilon(y) = \frac{\kappa(y+1) - \kappa(y) - \kappa(1)}{\kappa(2) - 2\kappa(1)}, \quad \text{and} \quad \kappa(z) := \log E[e^{zX_1}], \tag{44}$$

or, equivalently, $\varphi_t = \varphi(t, S_{t-})$ where:

$$\varphi(t,S) = \frac{\sigma^2 \frac{\partial P}{\partial S}(t,S) + \frac{1}{S} \int \nu(dz)(e^z - 1)[P(t, Se^z) - P(t,S)]}{\sigma^2 + \int (e^z - 1)^2 \nu(dz)} \tag{45}$$

with $P(t,S) = e^{-r(T-t)} E^{\mathbb{Q}}[G(S_T)|S_t = S]$ the option price at date t when the underlying is at the level S.

Remark 1. Condition (42), which is the only assumption imposed in addition to those of Proposition 2, guarantees that both the price process S_t and the option pay-off $G(S_T)$ are square integrable.

Proof. By Itô formula, the discounted stock price dynamics is given by

$$\hat{S}_T = \hat{S}_0 + \int_0^T \hat{S}_t \sigma dW_t + \int_0^T \int_{\mathbb{R}} \hat{S}_t (e^x - 1) \tilde{J}_X(dt \times dx). \tag{46}$$

To prove the proposition using the formula (41), we now need to obtain a similar integral representation for the option's discounted price function $\hat{P}(t, S_t) = e^{r(T-t)} P(t, S_t)$.

Let $t < T$. Applying the Itô formula under the integral sign in (18), we find

$$\begin{aligned} \hat{P}(t, S_t) - \hat{P}(0, S_0) = \frac{1}{2\pi} \int_{\mathbb{R}} du \hat{g}(u + iR) \int_0^t \Phi_{T-s}(-u - iR) \\ (R - iu) \hat{S}_s^{R-iu} \sigma dW_s + \frac{1}{2\pi} \int_{\mathbb{R}} du \hat{g}(u + iR) \\ \int_0^t \Phi_{T-s}(-u - iR) \hat{S}_{s-}^{R-iu} \int_{\mathbb{R}} (e^{(R-iu)z} - 1) \tilde{J}_X(ds \times dz). \end{aligned} \tag{47}$$

Let us first assume that $\sigma > 0$ and study the first term in the right-hand side of (47), which can be written as

$$\int_{\mathbb{R}} \mu(du) \int_0^t H_s^u dW_s$$

where

$$\mu(du) = |\hat{g}(u + iR) \Phi_{T-t}(-u - iR)| du \tag{48}$$

is a finite positive measure on $\mathbb{R}$ and

$$H_s^u = \frac{\sigma \hat{g}(u + iR) \Phi_{T-s}(-u - iR)}{2\pi |\hat{g}(u + iR) \Phi_{T-t}(-u - iR)|} (R - iu) \hat{S}_s^{R-iu}$$

By the Fubini theorem for stochastic integrals (see [41, p. 208]), we can interchange the two integrals in (48) provided that

$$E \int_0^t \mu(du) |H_s^u|^2 ds < \infty \tag{49}$$

Under the assumption (17) it is easy to check that

$$\frac{\Phi_{T-s}(-u - iR)}{|\Phi_{T-t}(-u - iR)|} \leq C$$

for all $s \leq t \leq T$ for some constant $C > 0$ which does not depend on s and t. To prove (49) it is then sufficient to check

$$E \int_0^t \int_{\mathbb{R}} |\hat{g}(u + iR)\Phi_{T-t}(-u - iR)||\hat{S}_s^{2(R-iu)}|^2 (R - iu)^2 du dt < \infty$$

which holds because

$$|\Phi_{T-t}(-u - iR)| \leq Ce^{-(T-t)\frac{\sigma^2 u^2}{2}} \tag{50}$$

Therefore, the first term on the right-hand side of (47) is equal to

$$\int_0^t \tilde{\sigma}_s dW_s, \quad \tilde{\sigma}_s = \frac{\sigma}{2\pi} \int_{\mathbb{R}} du \hat{g}(u + iR)\Phi_{T-s}(-u - iR)(R - iu)\hat{S}_s^{R-iu}. \tag{51}$$

This also shows that $\tilde{\sigma}_s = \sigma S_s \frac{\partial P(s,S_s)}{\partial S}$.

Let us now turn to the second term in the right-hand side of (47). Here we need to apply the Fubini theorem for stochastic integrals with respect to a compensated Poisson random measure [4, Theorem 5] and the applicability condition boils down to

$$E \int_0^t \int_{\mathbb{R}} |\hat{g}(u+iR)\Phi_{T-t}(-u-iR)||\hat{S}_s^{2(R-iu)}|^2 \int_{\mathbb{R}} |e^{(R-iu)z} - 1|^2 \nu(dz) du dt < \infty$$

If $\sigma > 0$, this is once again guaranteed by (50), and when $\sigma = 0$,

$$\int_{\mathbb{R}} |e^{(R-iu)z} - 1|^2 \nu(dz) = \psi(-2iR) - 2\Re\psi(-u - iR).$$

Since, for some $C < \infty$,

$$\begin{aligned} |\Re\psi(-u - iR)\Phi_{T-t}(-u - iR)| &= |\Re\psi(-u - iR)e^{(T-t)\Re\psi(-u-iR)}| \\ &\leq Ce^{\frac{T-t}{2}\Re\psi(-u-iR)}, \end{aligned}$$

the integrability condition is satisfied and we conclude that

$$\hat{P}(t, S_t) - \hat{P}(0, S_0) = \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \int_{\mathbb{R}} \tilde{\gamma}_s(z) \tilde{J}_X(ds \times dz) \tag{52}$$

for all $t < T$ with $\tilde{\sigma}$ as above and

$$\begin{aligned} \tilde{\gamma}_s(z) &= \frac{1}{2\pi} \int_{\mathbb{R}} du \hat{g}(u + iR)\Phi_{T-s}(-u - iR)\hat{S}_{s-}^{R-iu}(e^{(R-iu)z} - 1) \\ &= \hat{P}(s, S_{s-}e^z) - \hat{P}(s, S_{s-}). \end{aligned} \tag{53}$$

The optimal (risk-minimizing) hedge is obtained from formula (41):

$$\hat{\varphi}_t = \frac{\sigma \hat{S}_t \tilde{\sigma}_t + \hat{S}_t \int_{\mathbb{R}} \nu(dz)(e^z - 1)\tilde{\gamma}_t(z)}{\hat{S}_t^2(\sigma^2 + \int_{\mathbb{R}}(e^z - 1)^2 \nu(dz))}.$$

Substituting the expressions for $\tilde{\sigma}$ and $\tilde{\gamma}$ in terms of option prices into the above expression, we obtain (45) directly. On the other hand, the Fourier representations (51) and (53) and an application of Fubini's theorem yield (43).

As a by-product, the martingale representation (52) also yields the expression for the residual risk of a hedging strategy:

$$\begin{aligned} E[\epsilon(\varphi)^2] &= E\left[\int_0^T dt \int_{\mathbb{R}} \nu(dz)\left(\hat{P}(t, S_{t-}e^z) - \hat{P}(t, S_{t-}) - \hat{S}_{t-}\varphi_t(e^z - 1)\right)^2\right] \\ &\quad + E\left[\int_0^T \hat{S}_{t-}^2 \left(\varphi_t - \frac{\partial P}{\partial S}(t, S_{t-})\right)^2 \sigma^2 dt\right]. \end{aligned} \tag{54}$$

This allows us to examine whether there are any cases where the hedging error can be reduced to zero, i.e., where one can achieve a perfect hedge for every option and the market is complete. Hedging error is zero if and only if, for almost all t, there exists $k \in \mathbb{R}$ with:

$$(\sigma S_t \frac{\partial P}{\partial S}, (P(t, S_t e^z) - P(t, S_t))_{z \in \text{supp}\,\nu}) = k(\sigma S_t, (S_t(e^z - 1))_{z \in \text{supp}\,\nu})$$

This is only true in two (trivial) cases:

- The Lévy process X is a Brownian motion with drift: $\nu = 0$ and we retrieve the Black–Scholes delta hedge

$$\varphi_t = \Delta^{BS}(t, S_t) = \frac{\partial P}{\partial S}(t, S_t).$$

- The Lévy process X is a Poisson process with drift: $\sigma = 0$ and there is a single possible jump size: $\nu = \delta_{x_0}(x)$. In this case the hedging error equals

$$E\left[\int_0^T dt \left(\hat{P}(t, S_{t-}e^{x_0}) - \hat{P}(t, S_{t-}) - \hat{S}_{t-}\varphi_t(e^{x_0} - 1)\right)^2\right]$$

so by choosing

$$\varphi_t = \frac{P(t, S_{t-}e^{x_0}) - P(t, S_{t-})}{S_{t-}(e^{x_0} - 1)}$$

we obtain a replication strategy.

In other cases, the market is incomplete (an explicit counter-example may be constructed using power option with pay-off $H_T = (S_T)^\alpha$).

Delta-hedging vs. optimal strategy

We see that the optimal strategy (45) can be represented as a weighted average of the delta hedging $\frac{\partial P}{\partial S}$ and a certain integral involving the sensitivities of the option price to various possible jumps. But how far is the optimal strategy from the pure delta hedging? To answer this question, if option prices are regular (e.g., when $\sigma > 0$) and jumps are small, we can perform a Taylor expansion with respect to the jump size in (45), obtaining

$$\Delta(t,S) = \frac{\partial P}{\partial S} + \frac{S}{2\Sigma^2}\frac{\partial^2 P}{\partial S^2}\int \nu(dz)(e^z-1)^3.$$

where

$$\Sigma^2 = \sigma^2 + \int (e^z-1)^2\nu(dz).$$

Typically in equity markets the jumps are negative and small, therefore $\Delta(t,S) < \frac{\partial P}{\partial S}$ and the optimal strategy represents a small (of the order of third power of jump size) asymmetry correction. This situation is represented in Fig. 4, left graph. On the other hand, for pure-jump processes such as variance gamma, we cannot perform the Taylor expansion, because the second derivative $\frac{\partial^2 P}{\partial S^2}$ may not even exist, and the correction may therefore be quite large (see Fig. 4, right graph).

How big is the hedging error?

To answer this question, we simulated the terminal value of the hedging portfolio and that of the option's payoff over 10,000 trajectories for different strategies and different parameter sets.

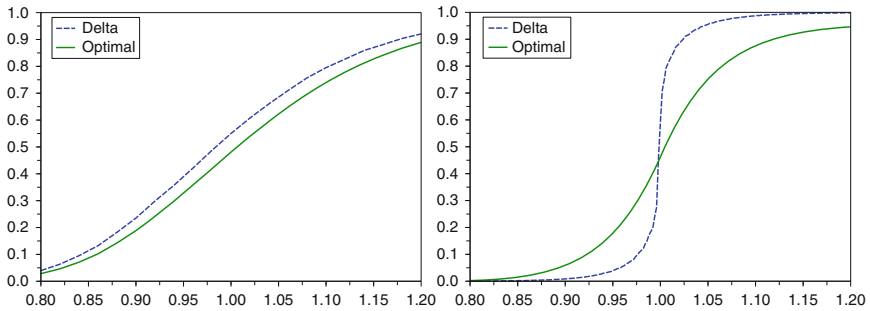


Fig. 4 Hedge ratios for the optimal strategy of Proposition 7 and the delta hedging strategy as function of stock price S. *Left*: hedging with stock in Kou model: the optimal strategy introduces a small asymmetry correction to delta hedging. *Right*: variance gamma model close to maturity (2 days): the optimal strategy is very far from delta hedging

Table 1 Hedging errors for different strategies in Kou model expressed in percentage of the initial stock price. Model parameters were estimated from MSFT time series. The "Black–Scholes" strategy corresponds to delta-hedging in the Black–Scholes model with equivalent volatility

Strategy	Root of mean squared error
Delta hedging	0.0133
Optimal quadratic	0.0133
Delta hedging in Black–Scholes model (error due to discrete hedging)	0.0059
No hedging	0.107

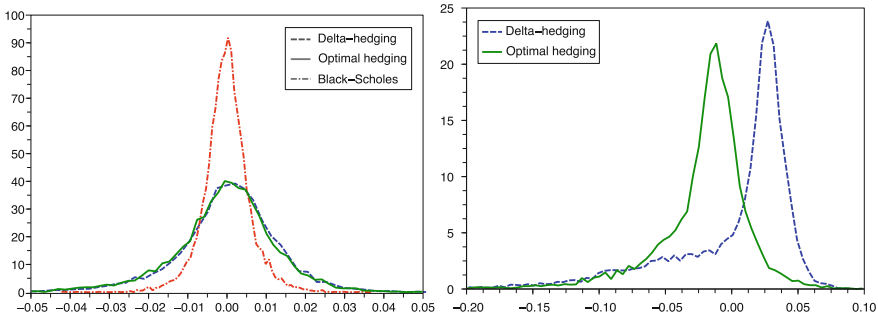


Fig. 5 Histograms of the residual hedging error in Kou model. *Left*: parameters estimated from MSFT time series. This graph shows the residual hedging error in Kou model with the optimal quadratic strategy (*solid line*), in Kou model with the delta-hedging strategy (*dashed line*) and in the Black–Scholes model with the delta-hedging strategy (*dash-dot line*). In the latter case, the error is only due to discrete-time hedging, and this curve was include to assess the magnitude of the discretization error for other tests. *Right*: strong negative jumps

In the first case study, Kou model with parameters estimated from market data (MSFT) during a calm period was used, and the option to hedge was a European put with strike $K = 90\%$ of the spot price and time to maturity $T = 1$ year. The hedging errors are given in Table 1 and the left graph in Fig. 5 shows the P&L histograms. For this parameter set, the optimal strategy is very close to delta hedging, and consequently, the hedging error is the same for delta hedging as for the optimal strategy. On the other hand, this error is very low, it is only twice as big as what we would get in the Black and Scholes model with equivalent volatility (this error in the Black–Scholes model is due to the fact that in the simulations, the portfolio is only rebalanced once a day and not continuously).

In the second case study, Kou model with unfrequent large negative jumps (10%) was used, and we wanted once again to hedge an OTM European put ($K = 90\%$, $T = 1$). The hedging errors are given in Table 2 and the P&L histograms in Fig. 5, right graph. Here we see that first, the optimal strategy has a much better performance than delta-hedging, and second, even this performance may not be sufficient, since the residual error is still of order of 4% of the initial stock price. This means that in this context, the market is "strongly incomplete" and hedging with stock only

Table 2 Hedging errors for different strategies in Kou model expressed in percentage of the initial stock price. A parameter set ensuring the presence of large negative jumps was taken

Strategy	Root of mean squared error
Delta-hedging	0.051
Optimal quadratic	0.041
No hedging	0.156

does not allow to make the risk at terminal date sufficiently small. In this case, to improve the hedging performance, one can include additional liquid assets, such as options on the same underlying, or variance swaps, into the hedging portfolio.

5.2 *Quadratic Hedging in Exponential Lévy Models Under the Historical Measure*

Throughout this section, to simplify notation, we suppose that the interest rate is equal to zero; the formulas for the general case can be obtained by working with discounted claims. Let S be the price process of the underlying, and suppose that it can be written in the form

$$S_t = S_0 + M_t + \int_0^t \alpha_s d\langle M\rangle_s \tag{55}$$

for some square integrable martingale M and some predictable process α. If S is an exponential of a Lévy process X with Lévy measure ν satisfying $\int_{|x|>1} e^{2x}\nu(dx) < \infty$ and diffusion coefficient σ, which can be written as

$$S_t = S_0 + \int_0^t \gamma S_u du + \int_0^t S_u \sigma dW_u + \int_0^t \int_{\mathbb{R}} S_{u-}(e^z - 1)\tilde{J}_X(du \times dz), \tag{56}$$

then the representation (55) holds with

$$\begin{aligned} M_t &= \int_0^t S_u \sigma dW_u + \int_0^t \int_{\mathbb{R}} S_{u-}(e^z - 1)\tilde{J}_X(du \times dz) \\ \langle M\rangle_t &= \int_0^t S_u^2 \left(\sigma^2 + \int_{\mathbb{R}} (e^z - 1)^2 \nu(dz)\right) du \\ \alpha_t &= \frac{\gamma}{S_t\left(\sigma^2 + \int_{\mathbb{R}} (e^z - 1)^2 \nu(dz)\right)} \end{aligned}$$

We then introduce the so-called *mean-variance tradeoff* process

$$K_t := \int_0^t \alpha_s^2 d\langle M\rangle_s.$$

In an exponential Lévy model, the mean-variance tradeoff is deterministic:

$$K_t = \frac{\gamma^2 t}{\sigma^2 + \int_{\mathbb{R}} (e^z - 1)^2 \nu(dz)}.$$

Local risk minimization

The *locally risk minimizing* strategy [21, 51] is a (not necessarily self-financing) trading strategy whose discounted cost process $\hat{C}$ is a martingale orthogonal to M. This strategy is optimal in the sense that we eliminate all the risk associated to the underlying with hedging, and the only part of risk that remains in the cost process is the risk which is orthogonal to the fluctuations of the underlying, and hence, cannot be hedged with it. If the market is complete, then all risk is explained by the underlying and the cost process of a locally minimizing strategy becomes constant, that is, the strategy becomes self-financing. As already mentioned, the locally risk minimizing strategy also has the interpretation of minimizing the residual risk (37) with respect to suitably defined small perturbations of the strategy [51]. Since the cost process is nonconstant, the locally risk minimizing strategy is not a self-financing strategy in general however since C is a martingale with mean zero this strategy is self-financing *on average*.

The locally risk minimizing strategy is closely related to an extension of the Kunita–Watanabe decomposition to semimartingale setting, known as the Föllmer–Schweizer decomposition [21, 50, 53, 54].

Definition 1. Let $H \in L^2(\mathbb{P})$ be a contingent claim. A sum $H = H_0 + \int_0^T \varphi_u^H dS_u + L_T^H$ is called the *Föllmer–Schweizer decomposition* of H if H_0 is $\mathcal{F}_0$-measurable, φ^H is an admissible trading strategy and L^H is a square integrable martingale with $L_0^H = 0$, orthogonal to M.

Given a Föllmer–Schweizer decomposition for the claim H, the locally risk minimizing strategy for H can be constructed by taking $\varphi_t = \varphi_t^H$ for all t, and choosing φ^0 such that the cost process is $C_t = H_0 + L_t^H$ for all t, which amounts to $\varphi_t^0 = H_0 + L_t^H - \varphi_t^H S_t - \int_0^t \varphi_u^H dS_u$.

Relationship with the minimal martingale measure

Define a process Z via $Z := \mathcal{E}(-\int_0^{\cdot} \alpha_s dM_s)$ and assume that Z is a strictly positive square integrable martingale. Then we can define a new measure $\mathbb{Q}^M$ by $\frac{d\mathbb{Q}|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} := Z_t$. By Girsanov–Meyer theorem ([41], Theorem 36 in Chap. 3), we have that (1) $\mathbb{Q}^M$ is a martingale measure, that is, S becomes a martingale under $\mathbb{Q}$ and (2) any square integrable martingale which is orthogonal to M under $\mathbb{P}$ remains a martingale under $\mathbb{Q}$ (although it may no longer be orthogonal to M). This measure is known as the minimal martingale measure [2, 53].

The minimal martingale measure allows to express the Föllmer–Schweizer decomposition in a more explicit form. First, compute the process L^H:

$$L_t^H = E^{\mathbb{Q}^M}[L_T^H|\mathcal{F}_t] = E^{\mathbb{Q}^M}[H|\mathcal{F}_t] - H_0 - \int_0^t \varphi_u^H dS_u.$$

Since $L_0^H = 0$, the initial capital for the Föllmer–Schweizer strategy is $H_0 = E^{\mathbb{Q}^M}[H]$. Let $H_t^M := E^{\mathbb{Q}^M}[H|\mathcal{F}_t]$. The orthogonality condition under $\mathbb{P}$ then yields an analogue of formula (41):

$$\varphi_t^H = \frac{d\langle H^M, S\rangle_t^{\mathbb{P}}}{d\langle S, S\rangle_t^{\mathbb{P}}}.$$

In models with jumps, the minimal martingale measure does not always exist as a probability measure (but may turn out to be a signed measure). In an exponential-Lévy model of the form (56), the density of the minimal martingale measure simplifies to $Z = \mathcal{E}(U)$ with

$$U_t = -\frac{\gamma}{\sigma^2 + \int_{\mathbb{R}}(e^z - 1)^2\nu(dz)}\left\{\sigma W_t + \int_0^t \int_{\mathbb{R}} (e^z - 1)\tilde{J}(ds \times dz)\right\}.$$

By Proposition 1, this yields a probability change if

$$\frac{\gamma(e^x - 1)}{\sigma^2 + \int_{\mathbb{R}}(e^z - 1)^2\nu(dz)} < 1 \quad \forall x \in \operatorname{supp}\nu,$$

which imposes a strong restriction on the drift parameter γ. If this condition is not satisfied, the Föllmer–Schweizer decomposition may still exist, but the interpretation using the minimal martingale measure is no longer valid, and the initial capital may turn out to be negative.

The existence of a Föllmer–Schweizer decomposition has been studied by many authors (see for example [2,53]), and in particular it was shown that the decomposition always exists in the case of exponential Lévy models. For these models, explicit formulas for the coefficients of this decomposition for European options are given in [29]:

Proposition 8 (Föllmer–Schweizer decomposition for European options in exponential Lévy models [29]).

- *Case of exponential pay-offs. Let $z \in \mathbb{C}$ with $S_T^z \in L^2(\mathbb{P})$. Then the contingent claim $H(z) = S_T^z$ admits a Föllmer–Schweizer decomposition with*

$$\varphi(z)_t = \Upsilon(z)e^{\eta(z)(T-t)}S_{t-}^{z-1}$$
$$L(z)_t = e^{\eta(z)(T-t)}S_t^z - e^{\eta(z)T}S_0^z - \int_0^t \varphi(z)_u dS_u,$$

where the coefficients Υ and η are given by

$$\Upsilon(z) = \frac{k(z+1) - k(z) - k(1)}{k(2) - 2k(1)}, \quad \eta(z) = k(z) - k(1)\Upsilon(z),$$

and $k(z) = \log E[e^{zX_1}]$ is the Laplace exponent of X.

- *Case of arbitrary payoffs. Let the option payoff be $H = f(S_T)$ with f of the form*

$$f(s) = \int s^z \Pi(dz)$$

for some finite complex measure Π on a strip $\{z \in \mathbb{C} : R' \leq \Re z \leq R\}$, where $R', R \in \mathbb{R}$ are chosen such that $E[e^{2RX_1}] < \infty$ and $E[e^{2R'X_1}] < \infty$. Then H admits a Föllmer–Schweizer decomposition with coefficients

$$\varphi_t^H = \int \varphi(z)_t \Pi(dz)$$
$$L_t^H = \int L(z)_t \Pi(dz).$$

Example 3. Let N^1 and N^2 be two standard Poisson processes with intensity 1 under $\mathbb{P}$ and suppose that the stock price is given by

$$S_t = \gamma t + 2N_t^1 + N_t^2 - 3t,$$

and that the contingent claim to be hedged is

$$H = 5N_T^1.$$

Define

$$L_t = N_t^1 - 2N_t^2 + t$$

Then L is a $\mathbb{P}$-martingale and

$$[L, S]_t = 2N_t^1 - 2N_t^2$$

which means that L is orthogonal to the martingale part of S under P. It is now easy to check that the Föllmer–Schweizer decomposition for H is given by

$$H = (5 - 2\gamma)T + L_T + 2S_T.$$

The locally risk-minimizing strategy therefore consists in

- Buying 2 units of the risky asset at date $t = 0$ (at zero price) and holding them until maturity.
- Placing $(5 - 2\gamma)T$ at the bank account and dynamically adding/withdrawing money according to the value of L.

The initial cost of this strategy is thus equal to $H_0 = (5 - 2\gamma)T$, which can be both positive and negative (if $\gamma > \frac{5}{2}$), and therefore cannot be interpreted as the price of the claim H. An intuitive explanation is that when the stock returns are very high, one can obtain a terminal pay-off which is (on average) positive even with a negative initial capital.

The minimal martingale measure in this setting is defined by

$$\frac{d\mathbb{Q}^M|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} = Z_t, \quad \frac{dZ_t}{Z_{t-}} = -\frac{\gamma}{5}(2dN_t^1 + dN_t^2 - 3dt).$$

From Proposition 1, we deduce that $\mathbb{Q}^M$ is a probability measure if and only if $\gamma < \frac{5}{2}$, in which case N^1 and N^2 are independent Poisson processes under $\mathbb{Q}^M$, with intensities

$$\lambda_1 = 1 - \frac{2\gamma}{5} \quad \text{and} \quad \lambda_2 = 1 - \frac{\gamma}{5}.$$

Easy calculations show that

- The martingale property of L is preserved under $\mathbb{Q}^M$, and in particular, we can compute

$$H_t^M = E^{Q^M}[H|\mathcal{F}_t] = 5\lambda_1 T + 5(N_t^1 - \lambda_1 t)$$
$$\text{and} \quad \varphi_t^H = \frac{d\langle H^M, S\rangle_t^{\mathbb{P}}}{d\langle S, S\rangle_t^{\mathbb{P}}} = 2.$$

- On the other hand, the orthogonality of S and L is not preserved under $\mathbb{Q}^M$: this would require $[L, S]_t = 2N_t^1 - 2N_t^2$ to be a $\mathbb{Q}^M$-martingale, which holds if and only if $\lambda_1 = \lambda_2$.

Variance-optimal hedging

An alternative approach is to choose a self-financing strategy φ and the initial capital V_0 such as to minimize

$$E^{\mathbb{P}}\Big[(V_0 + G_T(\varphi) - H)^2\Big].$$

under the statistical measure $\mathbb{P}$. This approach, known as mean-variance hedging or variance optimal hedging, is described in many papers including [8,9,14,29,40,54]. The general results concerning existence of optimal strategies are given in [14]. Schweizer [52] studies the case where the mean-variance tradeoff process K is deterministic and shows that in this case, the variance-optimal hedging strategy is also linked to the Föllmer–Schweizer decomposition. Hubalek et al. [29] exploit these results to derive explicit formulas for the hedging strategy in the case of Lévy processes. The following proposition uses the notation of Proposition 8.

Proposition 9 (Mean variance hedging in exponential Lévy models [29]). *Let the contingent claim H be as in the second part of Proposition 8. Then the variance optimal initial capital and the variance optimal hedging strategy are given by*

$$V_0 = H_0$$
$$\varphi_t = \varphi_t^H + \frac{\lambda}{S_{t-}}(H_{t-} - V_0 - G_{t-}(\varphi)), \tag{57}$$

where $\lambda = \frac{\kappa(1)}{\kappa(2)-2\kappa(1)}$ *and*

$$H_t = \int S_t^z e^{\eta(z)(T-t)} \Pi(dz).$$

In the case of exponential Lévy models, and in all models with deterministic mean-variance tradeoff, the variance optimal initial wealth is therefore equal to the initial value of the locally risk minimizing strategy. This allows to interpret the above result as a "stochastic target" approach to hedging, where the locally risk minimizing portfolio H_t plays the role of a "stochastic target" which we would like to follow because it allows to approach the option's pay-off with the least fluctuations. Since the locally risk-minimizing strategy is not self-financing, if we try to follow it with a self-financing strategy, our portfolio may deviate from the locally risk minimizing portfolio upwards or downwards. The strategy (57) measures this deviation at each date and tries to compensate it by investing more or less in the stock, depending on the sign of the expected return (λ is the expected excess return divided by the square of the volatility).

Acknowledgements I would like to thank all the participants of the crash course on financial modeling with jump processes in Bologna on May 21–22, 2009, and especially the organizer, Andrea Pascucci, for stimulating discussions and valuable feedback on my lectures. I am also grateful to the anonymous referee for the careful reading of the manuscript.

This research is supported by the Chair Financial Risks of the Risk Foundation sponsored by Société Générale, the Chair Derivatives of the Future sponsored by the Fédération Bancaire Française, and the Chair Finance and Sustainable Development sponsored by EDF and Calyon.

References

1. Abramowitz, M., Stegun, I. (eds.): Handbook of Mathematical Functions. Dover, New York (1968)
2. Ansel, J.P., Stricker, C.: Lois de martingale, densités et décomposition de Föllmer-Schweizer. Annales de l'Institut Henri Poincaré **28**, 375–392 (1992)
3. Appelbaum, D.: Lévy Processes and Stochastic Calculus. Cambridge Unviersity Press, Cambridge (2004)
4. Appelbaum, D.: Martingale-valued measures, Ornstein-Uhlenbeck processes with jumps and operator self-decomposability in Hilbert space. In: In Memoriam Paul-André Meyer, Seminaire de Probabilités, vol. XXXIX, pp. 171–196. Springer, Berlin (2006)

5. Belomestny, D., Reiss, M.: Spectral calibration of exponential Lévy models. Finance Stochast. **10**, 449–474 (2006)
6. Benth, F.E., Di Nunno, G., Løkka, A., Øksendal, B., Proske, F.: Explicit representation of the minimal variance portfolio in markets driven by Lévy processes. Math. finance **13**, 55–72 (2003)
7. Bertoin, J.: Lévy Processes. Cambridge University Press, Cambridge (1996)
8. Biagini, F., Guasoni, P.: Mean-variance hedging with random volatility jumps. Stochast. Anal. Appl. **20**, 471–494 (2002)
9. Bouleau, N., Lamberton, D.: Residual risks and hedging strategies in Markovian markets. Stochast. Process. Appl. **33**, 131–150 (1989)
10. Carr, P., Madan, D.: Option valuation using the fast Fourier transform. J. Comput. Finance **2**, 61–73 (1998)
11. Carr, P., Wu, L.: The finite moment logstable process and option pricing. J. Finance **58**, 753–778 (2003)
12. Carr, P., Wu, L.: What type of process underlies options? A simple robust test. J. Finance **58**, 2581–2610 (2003)
13. Carr, P., Geman, H., Madan, D., Yor, M.: The fine structure of asset returns: An empirical investigation. J. Business **75**, 305–332 (2002)
14. Černý, A., Kallsen, J.: On the structure of general mean-variance hedging strategies. Ann. Probab. **35**, 1479–1531 (2007)
15. Cherny, A.S., Shiryaev, A.N.: Change of time and measure for Lévy processes. MaPhySto lecture notes series 2002-13 (2002)
16. Cont, R., Tankov, P.: Financial Modelling with Jump Processes. Chapman & Hall, London (2004)
17. Cont, R., Tankov, P.: Non-parametric calibration of jump-diffusion option pricing models. J. Comput. Finance **7**, 1–49 (2004)
18. Cont, R., Tankov, P., Voltchkova, E.: Hedging with options in models with jumps. In: Stochastic Analysis and Applications – The Abel Symposium 2005. Springer, Berlin (2007)
19. Eberlein, E., Glau, K., Papapantoleon, A.: Analysis of Fourier transform valuation formulas and applications. Appl. Math. Finance **13**, 211–240 (2010)
20. Föllmer, H., Elliott, R.: Orthogonal martingale representation. In: Mayer-Wolf, E., Merzbach, E., Schwartz, A. (eds.) Stochastic Analysis, pp. 139–152. Academic Press, Boston (1991)
21. Föllmer, H., Schweizer, M.: Hedging of contintgent claims under incomplete information. In: Davis, M.H.A., Elliott, R.J. (eds.) Applied Stochastic Analysis, pp. 389–414. Gordon and Breach, NY (1991)
22. Föllmer, H., Sondermann, D.: Hedging of non-redundant contingent claims. In: Hildenbrand, W., Mas-Colell, A. (eds.) Contributions to Mathematical Economics, pp. 205–223. North Holland, Amsterdam (1986)
23. Friz, P., Benaim, S.: Smile asymptotics II: Models with known MGF. J. Appl. Probab. **45**, 16–32 (2008)
24. Friz, P., Benaim, S.: Regular variation and smile asymptotics. Math. Finance **19**, 1–12 (2009)
25. Galtchouk, L.: Représentation des martingales engendrées par un processus à accroissements indépendants (cas des martingales de carré intégrable). Ann. Inst. H. Poincaré Sect. B (N.S.) **12**, 199–211 (1976)
26. Goll, T., Kallsen, J.: Optimal portfolios for logarithmic utility. Stochast. Process. Appl. **89**, 31–48 (2000)
27. Jacod, J., Méléard, S., Protter, P.: Explicit form and robustness of martingale representations. Ann. Probab. **28**, 1747–1780 (2000)
28. Jakubenas, P.: On option pricing in certain incomplete markets. Proceedings of the Steklov Mathematical Institute, 237 (2002)
29. Kallsen, J., Hubalek, F., Krawczyk, L.: Variance-optimal hedging for processes with stationary independent increments. Ann. Appl. Probab. **16**, 853–885 (2006)
30. Kardaras, C.: No-free-lunch equivalences for exponential Lévy models under convex constraints on investment. Math. Finance **19**, 161–187 (2009)

31. Koponen, I.: Analytic approach to the problem of convergence of truncated Lévy flights towards the Gaussian stochastic process. Phys. Rev. E **52**, 1197–1199 (1995)
32. Kou, S.: A jump-diffusion model for option pricing. Manag. Sci. **48**, 1086–1101 (2002)
33. Kunita, H., Watanabe, S.: On square integrable martingales. Nagoya Math. J. **30**, 209–245 (1967)
34. Kyprianou, A.: Introductory Lectures on Fluctuations of Lévy Processes with Applications. Springer, Berlin (2006)
35. Lee, R.: The moment formula for implied volatility at extreme strikes. Math. Finance **14**, 469–480 (2004)
36. Lee, R.W.: Option pricing by transform methods: Extensions, unification and error control. J. Comput. Finance **7**, 51–86 (2004)
37. Lewis, A.: A simple option formula for general jump-diffusion and other exponential Lévy processes. `http://www.optioncity.net` (2001)
38. Madan, D., Carr, P., Chang, E.: The variance gamma process and option pricing. Eur. Finance Rev. **2**, 79–105 (1998)
39. Merton, R.: Option pricing when underlying stock returns are discontinuous. J. Financ. Econ. **3**, 125–144 (1976)
40. Pham, H.: Hedging and optimization problems in continuous financial models. In: Cont, R., Yong, J. (eds.) Mathematical Finance: Theory and Practice, vol. 1 of Series in Contemporary Applied Mathematics, pp. 353–381. Higher Education Press, Beijing (2000)
41. Protter, P.: Stochastic Integration and Differential Equations, 2nd edn. Springer, Berlin (2004)
42. Raible, S.: Lévy Processes in Finance: Theory, Numerics and Empirical Facts, PhD thesis, Freiburg University (2000)
43. Rockafellar, R.T.: Convex Analysis. Princeton University Press, NJ (1970)
44. Rogers, L.C.G., Tehranchi, M.: Can the implied volatility surface move by parallel shifts? Finance Stochast. **14**, 235–248 (2010)
45. Roper, M.: Implied volatility explosions: European calls and implied volatilities close to expiry in exponential Lévy models (2008), arXiv:0809.3305
46. Rudin, W.: Real and Complex Analysis, 3rd edn. McGraw-Hill, NY (1987)
47. Rüschendorf, L., Woerner, J.H.: Expansion of transition distributions of Lévy processes in small time. Bernoulli **8**, 81–96 (2002)
48. Sato, K.: Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press, Cambridge (1999)
49. Schoutens, W.: Lévy Processes in Finance: Pricing Financial Derivatives. Wiley, NY (2003)
50. Schweizer, M.: Risk-minimality and orthogonality of martingales. Stochast. Stochast. Rep. **30**, 123–131 (1990)
51. Schweizer, M.: Option hedging for semimartingales. Stochast. Process. Appl. **37**, 339–363 (1991)
52. Schweizer, M.: Approximating random variables by stochastic integrals. Ann. Probab. **22**, 1536–1575 (1994)
53. Schweizer, M.: On the minimal martingale measure and the Föllmer-Schweizer decomposition. Stochast. Anal. Appl. **13**, 573–599 (1995)
54. Schweizer, M.: A guided tour through quadratic hedging approaches, pp. 538–574. Cambridge University Press, Cambridge (2001)
55. Selivanov, A.V.: On the martingale measures in exponential Lévy models. Theor. Probab. Appl. **49**, 261–274 (2005)
56. Tehranchi, M.R.: Asymptotics of implied volatility far from maturity. J. Appl. Probab. **46**, 629–650 (2009)
57. Tehranchi, M.R.: Implied volatility: Long maturity behavior. In: Encyclopedia of Quantitative Finance. Wiley, NY (2010)

Lecture Notes in Mathematics

For information about earlier volumes
please contact your bookseller or Springer
LNM Online archive: springerlink.com

Vol. 1816: S. Albeverio, W. Schachermayer, M. Talagrand, Lectures on Probability Theory and Statistics. Ecole d'Eté de Probabilités de Saint-Flour XXX-2000. Editor: P. Bernard (2003)
Vol. 1817: E. Koelink, W. Van Assche (Eds.), Orthogonal Polynomials and Special Functions. Leuven 2002 (2003)
Vol. 1818: M. Bildhauer, Convex Variational Problems with Linear, nearly Linear and/or Anisotropic Growth Conditions (2003)
Vol. 1819: D. Masser, Yu. V. Nesterenko, H. P. Schlickewei, W. M. Schmidt, M. Waldschmidt, Diophantine Approximation. Cetraro, Italy 2000. Editors: F. Amoroso, U. Zannier (2003)
Vol. 1820: F. Hiai, H. Kosaki, Means of Hilbert Space Operators (2003)
Vol. 1821: S. Teufel, Adiabatic Perturbation Theory in Quantum Dynamics (2003)
Vol. 1822: S.-N. Chow, R. Conti, R. Johnson, J. Mallet-Paret, R. Nussbaum, Dynamical Systems. Cetraro, Italy 2000. Editors: J. W. Macki, P. Zecca (2003)
Vol. 1823: A. M. Anile, W. Allegretto, C. Ringhofer, Mathematical Problems in Semiconductor Physics. Cetraro, Italy 1998. Editor: A. M. Anile (2003)
Vol. 1824: J. A. Navarro González, J. B. Sancho de Salas, C^∞ – Differentiable Spaces (2003)
Vol. 1825: J. H. Bramble, A. Cohen, W. Dahmen, Multiscale Problems and Methods in Numerical Simulations, Martina Franca, Italy 2001. Editor: C. Canuto (2003)
Vol. 1826: K. Dohmen, Improved Bonferroni Inequalities via Abstract Tubes. Inequalities and Identities of Inclusion-Exclusion Type. VIII, 113 p, 2003.
Vol. 1827: K. M. Pilgrim, Combinations of Complex Dynamical Systems. IX, 118 p, 2003.
Vol. 1828: D. J. Green, Gröbner Bases and the Computation of Group Cohomology. XII, 138 p, 2003.
Vol. 1829: E. Altman, B. Gaujal, A. Hordijk, Discrete-Event Control of Stochastic Networks: Multimodularity and Regularity. XIV, 313 p, 2003.
Vol. 1830: M. I. Gil', Operator Functions and Localization of Spectra. XIV, 256 p, 2003.
Vol. 1831: A. Connes, J. Cuntz, E. Guentner, N. Higson, J. E. Kaminker, Noncommutative Geometry, Martina Franca, Italy 2002. Editors: S. Doplicher, L. Longo (2004)
Vol. 1832: J. Azéma, M. Émery, M. Ledoux, M. Yor (Eds.), Séminaire de Probabilités XXXVII (2003)
Vol. 1833: D.-Q. Jiang, M. Qian, M.-P. Qian, Mathematical Theory of Nonequilibrium Steady States. On the Frontier of Probability and Dynamical Systems. IX, 280 p, 2004.
Vol. 1834: Yo. Yomdin, G. Comte, Tame Geometry with Application in Smooth Analysis. VIII, 186 p, 2004.
Vol. 1835: O.T. Izhboldin, B. Kahn, N.A. Karpenko, A. Vishik, Geometric Methods in the Algebraic Theory of Quadratic Forms. Summer School, Lens, 2000. Editor: J.-P. Tignol (2004)
Vol. 1836: C. Năstăsescu, F. Van Oystaeyen, Methods of Graded Rings. XIII, 304 p, 2004.
Vol. 1837: S. Tavaré, O. Zeitouni, Lectures on Probability Theory and Statistics. Ecole d'Eté de Probabilités de Saint-Flour XXXI-2001. Editor: J. Picard (2004)
Vol. 1838: A.J. Ganesh, N.W. O'Connell, D.J. Wischik, Big Queues. XII, 254 p, 2004.
Vol. 1839: R. Gohm, Noncommutative Stationary Processes. VIII, 170 p, 2004.
Vol. 1840: B. Tsirelson, W. Werner, Lectures on Probability Theory and Statistics. Ecole d'Eté de Probabilités de Saint-Flour XXXII-2002. Editor: J. Picard (2004)
Vol. 1841: W. Reichel, Uniqueness Theorems for Variational Problems by the Method of Transformation Groups (2004)
Vol. 1842: T. Johnsen, A. L. Knutsen, K_3 Projective Models in Scrolls (2004)
Vol. 1843: B. Jefferies, Spectral Properties of Noncommuting Operators (2004)
Vol. 1844: K.F. Siburg, The Principle of Least Action in Geometry and Dynamics (2004)
Vol. 1845: Min Ho Lee, Mixed Automorphic Forms, Torus Bundles, and Jacobi Forms (2004)
Vol. 1846: H. Ammari, H. Kang, Reconstruction of Small Inhomogeneities from Boundary Measurements (2004)
Vol. 1847: T.R. Bielecki, T. Björk, M. Jeanblanc, M. Rutkowski, J.A. Scheinkman, W. Xiong, Paris-Princeton Lectures on Mathematical Finance 2003 (2004)
Vol. 1848: M. Abate, J. E. Fornaess, X. Huang, J. P. Rosay, A. Tumanov, Real Methods in Complex and CR Geometry, Martina Franca, Italy 2002. Editors: D. Zaitsev, G. Zampieri (2004)
Vol. 1849: Martin L. Brown, Heegner Modules and Elliptic Curves (2004)
Vol. 1850: V. D. Milman, G. Schechtman (Eds.), Geometric Aspects of Functional Analysis. Israel Seminar 2002-2003 (2004)
Vol. 1851: O. Catoni, Statistical Learning Theory and Stochastic Optimization (2004)
Vol. 1852: A.S. Kechris, B.D. Miller, Topics in Orbit Equivalence (2004)
Vol. 1853: Ch. Favre, M. Jonsson, The Valuative Tree (2004)
Vol. 1854: O. Saeki, Topology of Singular Fibers of Differential Maps (2004)
Vol. 1855: G. Da Prato, P.C. Kunstmann, I. Lasiecka, A. Lunardi, R. Schnaubelt, L. Weis, Functional Analytic Methods for Evolution Equations. Editors: M. Iannelli, R. Nagel, S. Piazzera (2004)
Vol. 1856: K. Back, T.R. Bielecki, C. Hipp, S. Peng, W. Schachermayer, Stochastic Methods in Finance, Bressanone/Brixen, Italy, 2003. Editors: M. Fritelli, W. Runggaldier (2004)
Vol. 1857: M. Émery, M. Ledoux, M. Yor (Eds.), Séminaire de Probabilités XXXVIII (2005)
Vol. 1858: A.S. Cherny, H.-J. Engelbert, Singular Stochastic Differential Equations (2005)
Vol. 1859: E. Letellier, Fourier Transforms of Invariant Functions on Finite Reductive Lie Algebras (2005)

Vol. 1860: A. Borisyuk, G.B. Ermentrout, A. Friedman, D. Terman, Tutorials in Mathematical Biosciences I. Mathematical Neurosciences (2005)

Vol. 1861: G. Benettin, J. Henrard, S. Kuksin, Hamiltonian Dynamics – Theory and Applications, Cetraro, Italy, 1999. Editor: A. Giorgilli (2005)

Vol. 1862: B. Helffer, F. Nier, Hypoelliptic Estimates and Spectral Theory for Fokker-Planck Operators and Witten Laplacians (2005)

Vol. 1863: H. Führ, Abstract Harmonic Analysis of Continuous Wavelet Transforms (2005)

Vol. 1864: K. Efstathiou, Metamorphoses of Hamiltonian Systems with Symmetries (2005)

Vol. 1865: D. Applebaum, B.V. R. Bhat, J. Kustermans, J. M. Lindsay, Quantum Independent Increment Processes I. From Classical Probability to Quantum Stochastic Calculus. Editors: M. Schürmann, U. Franz (2005)

Vol. 1866: O.E. Barndorff-Nielsen, U. Franz, R. Gohm, B. Kümmerer, S. Thorbjønsen, Quantum Independent Increment Processes II. Structure of Quantum Lévy Processes, Classical Probability, and Physics. Editors: M. Schürmann, U. Franz, (2005)

Vol. 1867: J. Sneyd (Ed.), Tutorials in Mathematical Biosciences II. Mathematical Modeling of Calcium Dynamics and Signal Transduction. (2005)

Vol. 1868: J. Jorgenson, S. Lang, $\mathrm{Pos}_n(\mathrm{R})$ and Eisenstein Series. (2005)

Vol. 1869: A. Dembo, T. Funaki, Lectures on Probability Theory and Statistics. Ecole d'Eté de Probabilités de Saint-Flour XXXIII-2003. Editor: J. Picard (2005)

Vol. 1870: V.I. Gurariy, W. Lusky, Geometry of Müntz Spaces and Related Questions. (2005)

Vol. 1871: P. Constantin, G. Gallavotti, A.V. Kazhikhov, Y. Meyer, S. Ukai, Mathematical Foundation of Turbulent Viscous Flows, Martina Franca, Italy, 2003. Editors: M. Cannone, T. Miyakawa (2006)

Vol. 1872: A. Friedman (Ed.), Tutorials in Mathematical Biosciences III. Cell Cycle, Proliferation, and Cancer (2006)

Vol. 1873: R. Mansuy, M. Yor, Random Times and Enlargements of Filtrations in a Brownian Setting (2006)

Vol. 1874: M. Yor, M. Émery (Eds.), In Memoriam Paul-André Meyer - Séminaire de Probabilités XXXIX (2006)

Vol. 1875: J. Pitman, Combinatorial Stochastic Processes. Ecole d'Eté de Probabilités de Saint-Flour XXXII-2002. Editor: J. Picard (2006)

Vol. 1876: H. Herrlich, Axiom of Choice (2006)

Vol. 1877: J. Steuding, Value Distributions of *L*-Functions (2007)

Vol. 1878: R. Cerf, The Wulff Crystal in Ising and Percolation Models, Ecole d'Eté de Probabilités de Saint-Flour XXXIV-2004. Editor: Jean Picard (2006)

Vol. 1879: G. Slade, The Lace Expansion and its Applications, Ecole d'Eté de Probabilités de Saint-Flour XXXIV-2004. Editor: Jean Picard (2006)

Vol. 1880: S. Attal, A. Joye, C.-A. Pillet, Open Quantum Systems I, The Hamiltonian Approach (2006)

Vol. 1881: S. Attal, A. Joye, C.-A. Pillet, Open Quantum Systems II, The Markovian Approach (2006)

Vol. 1882: S. Attal, A. Joye, C.-A. Pillet, Open Quantum Systems III, Recent Developments (2006)

Vol. 1883: W. Van Assche, F. Marcellàn (Eds.), Orthogonal Polynomials and Special Functions, Computation and Application (2006)

Vol. 1884: N. Hayashi, E.I. Kaikina, P.I. Naumkin, I.A. Shishmarev, Asymptotics for Dissipative Nonlinear Equations (2006)

Vol. 1885: A. Telcs, The Art of Random Walks (2006)

Vol. 1886: S. Takamura, Splitting Deformations of Degenerations of Complex Curves (2006)

Vol. 1887: K. Habermann, L. Habermann, Introduction to Symplectic Dirac Operators (2006)

Vol. 1888: J. van der Hoeven, Transseries and Real Differential Algebra (2006)

Vol. 1889: G. Osipenko, Dynamical Systems, Graphs, and Algorithms (2006)

Vol. 1890: M. Bunge, J. Funk, Singular Coverings of Toposes (2006)

Vol. 1891: J.B. Friedlander, D.R. Heath-Brown, H. Iwaniec, J. Kaczorowski, Analytic Number Theory, Cetraro, Italy, 2002. Editors: A. Perelli, C. Viola (2006)

Vol. 1892: A. Baddeley, I. Bárány, R. Schneider, W. Weil, Stochastic Geometry, Martina Franca, Italy, 2004. Editor: W. Weil (2007)

Vol. 1893: H. Hanßmann, Local and Semi-Local Bifurcations in Hamiltonian Dynamical Systems, Results and Examples (2007)

Vol. 1894: C.W. Groetsch, Stable Approximate Evaluation of Unbounded Operators (2007)

Vol. 1895: L. Molnár, Selected Preserver Problems on Algebraic Structures of Linear Operators and on Function Spaces (2007)

Vol. 1896: P. Massart, Concentration Inequalities and Model Selection, Ecole d'Été de Probabilités de Saint-Flour XXXIII-2003. Editor: J. Picard (2007)

Vol. 1897: R. Doney, Fluctuation Theory for Lévy Processes, Ecole d'Été de Probabilités de Saint-Flour XXXV-2005. Editor: J. Picard (2007)

Vol. 1898: H.R. Beyer, Beyond Partial Differential Equations, On linear and Quasi-Linear Abstract Hyperbolic Evolution Equations (2007)

Vol. 1899: Séminaire de Probabilités XL. Editors: C. Donati-Martin, M. Émery, A. Rouault, C. Stricker (2007)

Vol. 1900: E. Bolthausen, A. Bovier (Eds.), Spin Glasses (2007)

Vol. 1901: O. Wittenberg, Intersections de deux quadriques et pinceaux de courbes de genre 1, Intersections of Two Quadrics and Pencils of Curves of Genus 1 (2007)

Vol. 1902: A. Isaev, Lectures on the Automorphism Groups of Kobayashi-Hyperbolic Manifolds (2007)

Vol. 1903: G. Kresin, V. Maz'ya, Sharp Real-Part Theorems (2007)

Vol. 1904: P. Giesl, Construction of Global Lyapunov Functions Using Radial Basis Functions (2007)

Vol. 1905: C. Prévôt, M. Röckner, A Concise Course on Stochastic Partial Differential Equations (2007)

Vol. 1906: T. Schuster, The Method of Approximate Inverse: Theory and Applications (2007)

Vol. 1907: M. Rasmussen, Attractivity and Bifurcation for Nonautonomous Dynamical Systems (2007)

Vol. 1908: T.J. Lyons, M. Caruana, T. Lévy, Differential Equations Driven by Rough Paths, Ecole d'Été de Probabilités de Saint-Flour XXXIV-2004 (2007)

Vol. 1909: H. Akiyoshi, M. Sakuma, M. Wada, Y. Yamashita, Punctured Torus Groups and 2-Bridge Knot Groups (I) (2007)

Vol. 1910: V.D. Milman, G. Schechtman (Eds.), Geometric Aspects of Functional Analysis. Israel Seminar 2004-2005 (2007)

Vol. 1911: A. Bressan, D. Serre, M. Williams, K. Zumbrun, Hyperbolic Systems of Balance Laws. Cetraro, Italy 2003. Editor: P. Marcati (2007)

Vol. 1912: V. Berinde, Iterative Approximation of Fixed Points (2007)

Vol. 1913: J.E. Marsden, G. Misiołek, J.-P. Ortega, M. Perlmutter, T.S. Ratiu, Hamiltonian Reduction by Stages (2007)

Vol. 1914: G. Kutyniok, Affine Density in Wavelet Analysis (2007)

Vol. 1915: T. Bıyıkoğlu, J. Leydold, P.F. Stadler, Laplacian Eigenvectors of Graphs. Perron-Frobenius and Faber-Krahn Type Theorems (2007)

Vol. 1916: C. Villani, F. Rezakhanlou, Entropy Methods for the Boltzmann Equation. Editors: F. Golse, S. Olla (2008)

Vol. 1917: I. Veselić, Existence and Regularity Properties of the Integrated Density of States of Random Schrödinger (2008)

Vol. 1918: B. Roberts, R. Schmidt, Local Newforms for GSp(4) (2007)

Vol. 1919: R.A. Carmona, I. Ekeland, A. Kohatsu-Higa, J.-M. Lasry, P.-L. Lions, H. Pham, E. Taflin, Paris-Princeton Lectures on Mathematical Finance 2004. Editors: R.A. Carmona, E. Çinlar, I. Ekeland, E. Jouini, J.A. Scheinkman, N. Touzi (2007)

Vol. 1920: S.N. Evans, Probability and Real Trees. Ecole d'Été de Probabilités de Saint-Flour XXXV-2005 (2008)

Vol. 1921: J.P. Tian, Evolution Algebras and their Applications (2008)

Vol. 1922: A. Friedman (Ed.), Tutorials in Mathematical BioSciences IV. Evolution and Ecology (2008)

Vol. 1923: J.P.N. Bishwal, Parameter Estimation in Stochastic Differential Equations (2008)

Vol. 1924: M. Wilson, Littlewood-Paley Theory and Exponential-Square Integrability (2008)

Vol. 1925: M. du Sautoy, L. Woodward, Zeta Functions of Groups and Rings (2008)

Vol. 1926: L. Barreira, V. Claudia, Stability of Nonautonomous Differential Equations (2008)

Vol. 1927: L. Ambrosio, L. Caffarelli, M.G. Crandall, L.C. Evans, N. Fusco, Calculus of Variations and Non-Linear Partial Differential Equations. Cetraro, Italy 2005. Editors: B. Dacorogna, P. Marcellini (2008)

Vol. 1928: J. Jonsson, Simplicial Complexes of Graphs (2008)

Vol. 1929: Y. Mishura, Stochastic Calculus for Fractional Brownian Motion and Related Processes (2008)

Vol. 1930: J.M. Urbano, The Method of Intrinsic Scaling. A Systematic Approach to Regularity for Degenerate and Singular PDEs (2008)

Vol. 1931: M. Cowling, E. Frenkel, M. Kashiwara, A. Valette, D.A. Vogan, Jr., N.R. Wallach, Representation Theory and Complex Analysis. Venice, Italy 2004. Editors: E.C. Tarabusi, A. D'Agnolo, M. Picardello (2008)

Vol. 1932: A.A. Agrachev, A.S. Morse, E.D. Sontag, H.J. Sussmann, V.I. Utkin, Nonlinear and Optimal Control Theory. Cetraro, Italy 2004. Editors: P. Nistri, G. Stefani (2008)

Vol. 1933: M. Petkovic, Point Estimation of Root Finding Methods (2008)

Vol. 1934: C. Donati-Martin, M. Émery, A. Rouault, C. Stricker (Eds.), Séminaire de Probabilités XLI (2008)

Vol. 1935: A. Unterberger, Alternative Pseudodifferential Analysis (2008)

Vol. 1936: P. Magal, S. Ruan (Eds.), Structured Population Models in Biology and Epidemiology (2008)

Vol. 1937: G. Capriz, P. Giovine, P.M. Mariano (Eds.), Mathematical Models of Granular Matter (2008)

Vol. 1938: D. Auroux, F. Catanese, M. Manetti, P. Seidel, B. Siebert, I. Smith, G. Tian, Symplectic 4-Manifolds and Algebraic Surfaces. Cetraro, Italy 2003. Editors: F. Catanese, G. Tian (2008)

Vol. 1939: D. Boffi, F. Brezzi, L. Demkowicz, R.G. Durán, R.S. Falk, M. Fortin, Mixed Finite Elements, Compatibility Conditions, and Applications. Cetraro, Italy 2006. Editors: D. Boffi, L. Gastaldi (2008)

Vol. 1940: J. Banasiak, V. Capasso, M.A.J. Chaplain, M. Lachowicz, J. Miękisz, Multiscale Problems in the Life Sciences. From Microscopic to Macroscopic. Będlewo, Poland 2006. Editors: V. Capasso, M. Lachowicz (2008)

Vol. 1941: S.M.J. Haran, Arithmetical Investigations. Representation Theory, Orthogonal Polynomials, and Quantum Interpolations (2008)

Vol. 1942: S. Albeverio, F. Flandoli, Y.G. Sinai, SPDE in Hydrodynamic. Recent Progress and Prospects. Cetraro, Italy 2005. Editors: G. Da Prato, M. Röckner (2008)

Vol. 1943: L.L. Bonilla (Ed.), Inverse Problems and Imaging. Martina Franca, Italy 2002 (2008)

Vol. 1944: A. Di Bartolo, G. Falcone, P. Plaumann, K. Strambach, Algebraic Groups and Lie Groups with Few Factors (2008)

Vol. 1945: F. Brauer, P. van den Driessche, J. Wu (Eds.), Mathematical Epidemiology (2008)

Vol. 1946: G. Allaire, A. Arnold, P. Degond, T.Y. Hou, Quantum Transport. Modelling, Analysis and Asymptotics. Cetraro, Italy 2006. Editors: N.B. Abdallah, G. Frosali (2008)

Vol. 1947: D. Abramovich, M. Mariño, M. Thaddeus, R. Vakil, Enumerative Invariants in Algebraic Geometry and String Theory. Cetraro, Italy 2005. Editors: K. Behrend, M. Manetti (2008)

Vol. 1948: F. Cao, J-L. Lisani, J-M. Morel, P. Musé, F. Sur, A Theory of Shape Identification (2008)

Vol. 1949: H.G. Feichtinger, B. Helffer, M.P. Lamoureux, N. Lerner, J. Toft, Pseudo-Differential Operators. Quantization and Signals. Cetraro, Italy 2006. Editors: L. Rodino, M.W. Wong (2008)

Vol. 1950: M. Bramson, Stability of Queueing Networks, Ecole d'Eté de Probabilités de Saint-Flour XXXVI-2006 (2008)

Vol. 1951: A. Moltó, J. Orihuela, S. Troyanski, M. Valdivia, A Non Linear Transfer Technique for Renorming (2009)

Vol. 1952: R. Mikhailov, I.B.S. Passi, Lower Central and Dimension Series of Groups (2009)

Vol. 1953: K. Arwini, C.T.J. Dodson, Information Geometry (2008)

Vol. 1954: P. Biane, L. Bouten, F. Cipriani, N. Konno, N. Privault, Q. Xu, Quantum Potential Theory. Editors: U. Franz, M. Schuermann (2008)

Vol. 1955: M. Bernot, V. Caselles, J.-M. Morel, Optimal Transportation Networks (2008)

Vol. 1956: C.H. Chu, Matrix Convolution Operators on Groups (2008)

Vol. 1957: A. Guionnet, On Random Matrices: Macroscopic Asymptotics, Ecole d'Eté de Probabilités de Saint-Flour XXXVI-2006 (2009)

Vol. 1958: M.C. Olsson, Compactifying Moduli Spaces for Abelian Varieties (2008)

Vol. 1959: Y. Nakkajima, A. Shiho, Weight Filtrations on Log Crystalline Cohomologies of Families of Open Smooth Varieties (2008)
Vol. 1960: J. Lipman, M. Hashimoto, Foundations of Grothendieck Duality for Diagrams of Schemes (2009)
Vol. 1961: G. Buttazzo, A. Pratelli, S. Solimini, E. Stepanov, Optimal Urban Networks via Mass Transportation (2009)
Vol. 1962: R. Dalang, D. Khoshnevisan, C. Mueller, D. Nualart, Y. Xiao, A Minicourse on Stochastic Partial Differential Equations (2009)
Vol. 1963: W. Siegert, Local Lyapunov Exponents (2009)
Vol. 1964: W. Roth, Operator-valued Measures and Integrals for Cone-valued Functions and Integrals for Cone-valued Functions (2009)
Vol. 1965: C. Chidume, Geometric Properties of Banach Spaces and Nonlinear Iterations (2009)
Vol. 1966: D. Deng, Y. Han, Harmonic Analysis on Spaces of Homogeneous Type (2009)
Vol. 1967: B. Fresse, Modules over Operads and Functors (2009)
Vol. 1968: R. Weissauer, Endoscopy for GSP(4) and the Cohomology of Siegel Modular Threefolds (2009)
Vol. 1969: B. Roynette, M. Yor, Penalising Brownian Paths (2009)
Vol. 1970: M. Biskup, A. Bovier, F. den Hollander, D. Ioffe, F. Martinelli, K. Netočný, F. Toninelli, Methods of Contemporary Mathematical Statistical Physics. Editor: R. Kotecký (2009)
Vol. 1971: L. Saint-Raymond, Hydrodynamic Limits of the Boltzmann Equation (2009)
Vol. 1972: T. Mochizuki, Donaldson Type Invariants for Algebraic Surfaces (2009)
Vol. 1973: M.A. Berger, L.H. Kauffmann, B. Khesin, H.K. Moffatt, R.L. Ricca, De W. Sumners, Lectures on Topological Fluid Mechanics. Cetraro, Italy 2001. Editor: R.L. Ricca (2009)
Vol. 1974: F. den Hollander, Random Polymers: École d'Été de Probabilités de Saint-Flour XXXVII – 2007 (2009)
Vol. 1975: J.C. Rohde, Cyclic Coverings, Calabi-Yau Manifolds and Complex Multiplication (2009)
Vol. 1976: N. Ginoux, The Dirac Spectrum (2009)
Vol. 1977: M.J. Gursky, E. Lanconelli, A. Malchiodi, G. Tarantello, X.-J. Wang, P.C. Yang, Geometric Analysis and PDEs. Cetraro, Italy 2001. Editors: A. Ambrosetti, S.-Y.A. Chang, A. Malchiodi (2009)
Vol. 1978: M. Qian, J.-S. Xie, S. Zhu, Smooth Ergodic Theory for Endomorphisms (2009)
Vol. 1979: C. Donati-Martin, M. Émery, A. Rouault, C. Stricker (Eds.), Śeminaire de Probablitiés XLII (2009)
Vol. 1980: P. Graczyk, A. Stos (Eds.), Potential Analysis of Stable Processes and its Extensions (2009)
Vol. 1981: M. Chlouveraki, Blocks and Families for Cyclotomic Hecke Algebras (2009)
Vol. 1982: N. Privault, Stochastic Analysis in Discrete and Continuous Settings. With Normal Martingales (2009)
Vol. 1983: H. Ammari (Ed.), Mathematical Modeling in Biomedical Imaging I. Electrical and Ultrasound Tomographies, Anomaly Detection, and Brain Imaging (2009)
Vol. 1984: V. Caselles, P. Monasse, Geometric Description of Images as Topographic Maps (2010)
Vol. 1985: T. Linß, Layer-Adapted Meshes for Reaction-Convection-Diffusion Problems (2010)
Vol. 1986: J.-P. Antoine, C. Trapani, Partial Inner Product Spaces. Theory and Applications (2009)
Vol. 1987: J.-P. Brasselet, J. Seade, T. Suwa, Vector Fields on Singular Varieties (2010)
Vol. 1988: M. Broué, Introduction to Complex Reflection Groups and Their Braid Groups (2010)
Vol. 1989: I.M. Bomze, V. Demyanov, Nonlinear Optimization. Cetraro, Italy 2007. Editors: G. di Pillo, F. Schoen (2010)
Vol. 1990: S. Bouc, Biset Functors for Finite Groups (2010)
Vol. 1991: F. Gazzola, H.-C. Grunau, G. Sweers, Polyharmonic Boundary Value Problems (2010)
Vol. 1992: A. Parmeggiani, Spectral Theory of Non-Commutative Harmonic Oscillators: An Introduction (2010)
Vol. 1993: P. Dodos, Banach Spaces and Descriptive Set Theory: Selected Topics (2010)
Vol. 1994: A. Baricz, Generalized Bessel Functions of the First Kind (2010)
Vol. 1995: A.Y. Khapalov, Controllability of Partial Differential Equations Governed by Multiplicative Controls (2010)
Vol. 1996: T. Lorenz, Mutational Analysis. A Joint Framework for Cauchy Problems *In* and *Beyond* Vector Spaces (2010)
Vol. 1997: M. Banagl, Intersection Spaces, Spatial Homology Truncation, and String Theory (2010)
Vol. 1998: M. Abate, E. Bedford, M. Brunella, T.-C. Dinh, D. Schleicher, N. Sibony, Holomorphic Dynamical Systems. Cetraro, Italy 2008. Editors: G. Gentili, J. Guenot, G. Patrizio (2010)
Vol. 1999: H. Schoutens, The Use of Ultraproducts in Commutative Algebra (2010)
Vol. 2000: H. Yserentant, Regularity and Approximability of Electronic Wave Functions (2010)
Vol. 2001: T. Duquesne, O.E. Barndorff-Nielson, O. Reichmann, J. Bertoin, K.-I. Sato, J. Jacod, C. Schwab, C. Klüppelberg, Lévy Matters I (2010)
Vol. 2002: C. Pötzsche, Geometric Theory of Discrete Nonautonomous Dynamical Systems (2010)
Vol. 2003: A. Cousin, S. Crépey, O. Guéant, D. Hobson, M. Jeanblanc, J.-M. Lasry, J.-P. Laurent, P.-L. Lions, P. Tankov, Paris-Princeton Lectures on Mathematical Finance 2010. Editors: R.A. Carmona, E. Cinlar, I. Ekeland, E. Jouini, J.A. Scheinkman, N. Touzi (2011)

Recent Reprints and New Editions

Vol. 1702: J. Ma, J. Yong, Forward-Backward Stochastic Differential Equations and their Applications. 1999 – Corr. 3rd printing (2007)
Vol. 830: J.A. Green, Polynomial Representations of GL_n, with an Appendix on Schensted Correspondence and Littelmann Paths by K. Erdmann, J.A. Green and M. Schoker 1980 – 2nd corr. and augmented edition (2007)
Vol. 1693: S. Simons, From Hahn-Banach to Monotonicity (Minimax and Monotonicity 1998) – 2nd exp. edition (2008)
Vol. 470: R.E. Bowen, Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms. With a preface by D. Ruelle. Edited by J.-R. Chazottes. 1975 – 2nd rev. edition (2008)
Vol. 523: S.A. Albeverio, R.J. Høegh-Krohn, S. Mazzucchi, Mathematical Theory of Feynman Path Integral. 1976 – 2nd corr. and enlarged edition (2008)
Vol. 1764: A. Cannas da Silva, Lectures on Symplectic Geometry 2001 – Corr. 2nd printing (2008)

LECTURE NOTES IN MATHEMATICS

Springer

Edited by J.-M. Morel, F. Takens, B. Teissier, P.K. Maini

Editorial Policy (for Multi-Author Publications: Summer Schools/Intensive Courses)

1. Lecture Notes aim to report new developments in all areas of mathematics and their applications - quickly, informally and at a high level. Mathematical texts analysing new developments in modelling and numerical simulation are welcome. Manuscripts should be reasonably self-contained and rounded off. Thus they may, and often will, present not only results of the author but also related work by other people. They should provide sufficient motivation, examples and applications. There should also be an introduction making the text comprehensible to a wider audience. This clearly distinguishes Lecture Notes from journal articles or technical reports which normally are very concise. Articles intended for a journal but too long to be accepted by most journals, usually do not have this “lecture notes" character.
2. In general SUMMER SCHOOLS and other similar INTENSIVE COURSES are held to present mathematical topics that are close to the frontiers of recent research to an audience at the beginning or intermediate graduate level, who may want to continue with this area of work, for a thesis or later. This makes demands on the didactic aspects of the presentation. Because the subjects of such schools are advanced, there often exists no textbook, and so ideally, the publication resulting from such a school could be a first approximation to such a textbook. Usually several authors are involved in the writing, so it is not always simple to obtain a unified approach to the presentation.

   For prospective publication in LNM, the resulting manuscript should not be just a collection of course notes, each of which has been developed by an individual author with little or no co-ordination with the others, and with little or no common concept. The subject matter should dictate the structure of the book, and the authorship of each part or chapter should take secondary importance. Of course the choice of authors is crucial to the quality of the material at the school and in the book, and the intention here is not to belittle their impact, but simply to say that the book should be planned to be written by these authors jointly, and not just assembled as a result of what these authors happen to submit.

   This represents considerable preparatory work (as it is imperative to ensure that the authors know these criteria before they invest work on a manuscript), and also considerable editing work afterwards, to get the book into final shape. Still it is the form that holds the most promise of a successful book that will be used by its intended audience, rather than yet another volume of proceedings for the library shelf.
3. Manuscripts should be submitted either to Springer’s mathematics editorial in Heidelberg, or to one of the series editors. Volume editors are expected to arrange for the refereeing, to the usual scientific standards, of the individual contributions. If the resulting reports can be forwarded to us (series editors or Springer) this is very helpful. If no reports are forwarded or if other questions remain unclear in respect of homogeneity etc, the series editors may wish to consult external referees for an overall evaluation of the volume. A final decision to publish can be made only on the basis of the complete manuscript; however a preliminary decision can be based on a pre-final or incomplete manuscript. The strict minimum amount of material that will be considered should include a detailed outline describing the planned contents of each chapter.

   Volume editors and authors should be aware that incomplete or insufficiently close to final manuscripts almost always result in longer evaluation times. They should also be aware that parallel submission of their manuscript to another publisher while under consideration for LNM will in general lead to immediate rejection.

4. Manuscripts should in general be submitted in English. Final manuscripts should contain at least 100 pages of mathematical text and should always include
   - a general table of contents;
   - an informative introduction, with adequate motivation and perhaps some historical remarks: it should be accessible to a reader not intimately familiar with the topic treated;
   - a global subject index: as a rule this is genuinely helpful for the reader.

   Lecture Notes volumes are, as a rule, printed digitally from the authors' files. We strongly recommend that all contributions in a volume be written in LaTeX2e. To ensure best results, authors are asked to use the LaTeX2e style files available from Springer's web-server at

   ftp://ftp.springer.de/pub/tex/latex/svmultt1/ (for summer schools/tutorials).

   Additional technical instructions are available on request from: lnm@springer.com.
5. Careful preparation of the manuscripts will help keep production time short besides ensuring satisfactory appearance of the finished book in print and online. After acceptance of the manuscript authors will be asked to prepare the final LaTeX source files (and also the corresponding dvi-, pdf- or zipped ps-file) together with the final printout made from these files. The LaTeX source files are essential for producing the full-text online version of the book. For the existing online volumes of LNM see: www.springerlink.com/content/110312

   The actual production of a Lecture Notes volume takes approximately 12 weeks.
6. Volume editors receive a total of 50 free copies of their volume to be shared with the authors, but no royalties. They and the authors are entitled to a discount of 33.3% on the price of Springer books purchased for their personal use, if ordering directly from Springer.
7. Commitment to publish is made by letter of intent rather than by signing a formal contract. Springer-Verlag secures the copyright for each volume. Authors are free to reuse material contained in their LNM volumes in later publications: a brief written (or e-mail) request for formal permission is sufficient.

Addresses:

Professor J.-M. Morel, CMLA,
École Normale Supérieure de Cachan,
61 Avenue du Président Wilson,
94235 Cachan Cedex, France
E-mail: Jean-Michel.Morel@cmla.ens-cachan.fr

Professor F. Takens, Mathematisch Instituut,
Rijksuniversiteit Groningen, Postbus 800,
9700 AV Groningen, The Netherlands
E-mail: F.Takens@math.rug.nl

Professor B. Teissier,
Institut Mathématique de Jussieu,
UMR 7586 du CNRS,
Équipe "Géométrie et Dynamique",
175 rue du Chevaleret
75013 Paris, France
E-mail: teissier@math.jussieu.fr

For the "Mathematical Biosciences Subseries" of LNM:

Professor P.K. Maini, Center for Mathematical Biology
Mathematical Institute, 24-29 St Giles,
Oxford OX1 3LP, UK
E-mail: maini@maths.ox.ac.uk

Springer, Mathematics Editorial I, Tiergartenstr. 17,
69121 Heidelberg, Germany,
Tel.: +49 (6221) 487-8259
Fax: +49 (6221) 4876-8259
E-mail: lnm@springer.com