SWITCHING TYPE VALUATION AND DESIGN PROBLEMS IN GENERAL OTC CONTRACTS WITH CVA, COLLATERAL AND FUNDING ISSUE.

Sapienza University
PhD program in Economics and Finance.

Giovanni Mottola PhD
Start by doing what’s necessary; then do what’s possible;
and suddenly you are doing the impossible.

(San Francesco)
1. Introduction and literature review. ........................................... 6
   1.1 Research objectives and contributions. .................................. 6
   1.2 Review of the literature. ..................................................... 9

2. Model framework and assumptions. ........................................ 13
   2.1 Introduction ........................................................................ 13
   2.2 Model framework ................................................................. 15
   2.3 Price process and CVA of a contract without CSA ................. 16
   2.4 Price process and CVA in presence of CSA ......................... 21
   2.5 Contingent CSA, price process and CVA .............................. 25
   2.6 Funding with CVA and collateral. ......................................... 27
   2.7 Counterparty objective and the fundamental trade-off .......... 28

3. Stochastic switching control problem formulation. ....................... 34
   3.1 Introduction ........................................................................ 34
   3.2 Modeling dynamics and controls of the problem .................. 34
   3.3 Modeling the objective functional ....................................... 39
   3.4 Problem recursion and alternative reformulation ................. 43

4. Analytical approach to solution of optimal switching control problems 46
   4.1 Introduction ........................................................................ 46
   4.2 Dynamic programming principle and HJB equations in switching control
       problems: main issues .......................................................... 48
   4.3 Viscosity solution and variational inequalities formulation .......... 52
   4.4 Analysis and properties of the cost functions ....................... 55
   4.5 Value function behavior and switching strategy conditions ....... 58

5. Snell envelope approach, RBSDE and connections with QVI ............ 65
   5.1 Introduction ........................................................................ 65
   5.2 Switching control problem solution via Snell envelope: theory and assumptions 67
   5.3 Verification theorem, existence and uniqueness of the solution..... 70
   5.4 Solution connection with Reflected Backward SDEs and variational inequalities 73
   5.5 The optimal switching solution recast as an iterative optimal stopping procedure 77
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. Numerical solution approach and implementation</td>
<td>79</td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>79</td>
</tr>
<tr>
<td>6.2 Numerical solution: the basic underlying idea.</td>
<td>80</td>
</tr>
<tr>
<td>6.3 Numerical procedure definition and analysis</td>
<td>82</td>
</tr>
<tr>
<td>6.4 Numerical implementation and examples.</td>
<td>91</td>
</tr>
<tr>
<td>6.5 Numerical issues and further developments.</td>
<td>101</td>
</tr>
<tr>
<td>7. Generalization to stochastic games and the price-hedge problem</td>
<td>102</td>
</tr>
<tr>
<td>7.1 Introduction</td>
<td>102</td>
</tr>
<tr>
<td>7.2 On defaultable stochastic game option of switching type.</td>
<td>107</td>
</tr>
<tr>
<td>7.3 Game solution in a ”special case” and further analysis.</td>
<td>115</td>
</tr>
<tr>
<td>7.4 Price and hedging issue for a general contract with contingent collateralization</td>
<td>121</td>
</tr>
<tr>
<td>8. Applications of switching control model in finance</td>
<td>127</td>
</tr>
<tr>
<td>8.1 Some interesting cases.</td>
<td>127</td>
</tr>
<tr>
<td>8.2 Optimal swap closure/stopping time.</td>
<td>128</td>
</tr>
<tr>
<td>9. Conclusions</td>
<td>132</td>
</tr>
<tr>
<td>10. References</td>
<td>133</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS.

Special acknowledgements go to professor Enrico Biffis who has pushed me to build the present work, to professor Rita D’Ecclesia for her help and advises in building this work and throughout all the PhD years and to professor Fausto Gozzi for the fruitful discussions about some important issues of the work.

I want to remind also the important contribution through their speech and seminars held during the Imperial PhD visiting period of prof. Davis, prof. Houghston, prof. Brigo, dr. Zheng, dr. Barnett and dr. Cesari. Special thanks, goes also to my dear friends dr. Andrea Santacroce and dr. Romano Trabalzini, for their involvements at various levels in some very useful and special (almost for me!) discussions and talks.
1. INTRODUCTION AND LITERATURE REVIEW.

Eureka!
(Archimede)

1.1 Research objectives and contributions.

The present research work explores and develops some issues related to the valuation and risk management of general OTC contracts in presence of counterparty risk, namely credit value adjustment (CVA), collateralization and funding.

The counterparty risk valuation and impact in pricing and risk management has gained great importance and relevance in everyday financial markets. Particularly in the aftermath of the last crisis and Lehman Brothers default, this risk has been regularly traded by the financial institutions on their over the counter positions via the charge of the well-known “Credit Value Adjustment” (CVA, we refer to the chapter two for its formal definition), which was one of the major causes of the relevant losses suffered during the crisis due to the deterioration of the counterparty credit/rating condition.

Almost at the same time, the OTC deals, especially credit derivatives and interest rate swap, have seen a large increment of the collateral provision and collateralization mechanisms defined within the contract. These are typically defined in a specific annex known as "credit support annex" (CSA)\(^1\).

The reason for this is that setting a collateralization mechanism allows the counterparties to reduce or even null the CVA (this depends on the type of collateralization, full or partial and on margining), that is the integral of the expected weighted exposures on counterparty default risk taken over the life of the underlying contract or equivalently the cost for the setting of a dynamic hedging strategy to minimize this risk.

To remark the relevance of this issue, also the Basel Committee has intervened on counterparty risk defining new provisions in relation to the new Basel three accord. The main relevant points of the intervention are the following (see Albanese et al. (2011))

1. the need to make explicit the CVA as a capital charge for all the old contracts;

2. the introduction of an additional capital charge for the volatility of the CVA (note that until five times the CVA charge can be imposed for the vol risk) and to cover the unhedged default risk;

\(^1\) We refer for its formal definition to the second chapter.
3. the introduction of a central counterparty clearing house that operates as "guarantor" of the transactions with the obligation to set only full collateralised procedure with the counterparties that transfer to it their credit/default risk.

So counterparty risk represents a crucial element to assess and analyze. At the moment, the recent literature has dealt with the modeling of CVA dynamics and collateral mechanisms within specific derivatives products and at portfolio level. In particular, collateralisation and counterparty risk mitigation mechanisms, these typically follow some fixed schemes (full or partial) with no reference to possible optimal design problems of these collateralization mechanisms at contract level (where the contract is a defaultable claim\(^2\)); hence, considering the introduction of collateral and the funding issues in the picture has naturally lead us to think to the risk management problem of a defaultable claim in terms of a stochastic control problem of switching type.

In this work, we think to a possible real situation where two defaultable counterparties want to define a CSA in which both of them need the flexibility and the possibility to activate the collateralization during the life of the underlying claim/contract. We refer specifically to a contingent risk mitigation mechanism that allows the counterparty to switch from zero to full/perfect collateralization (or even partial) and switch back whenever she wants until maturity \(T\) paying some switching costs and taking into account the running costs that emerge over time.

The running costs that we model and consider in the analysis of this problem are - by one side - those related to CVA namely counterparty risk hedging costs and - by the other side - the collateral and funding/liquidity costs that emerge when collateralization is active.

We can summarize the characteristics and the basic idea underlying the problem (that we show to admit a natural formulation as a stochastic (multiple) impulse/switching control problem) through the so defined contingent CSA scheme shown below (Fig. 1.1), in which - by considering also the funding issue in the picture - is present a third party, an external funder assumed for convenience default free \((\lambda = 0)\).

In order to give a brief but faithful account of the work, we highlight by chapter the main logical steps and contributions of our research.

The second chapter is central having connections with all the other chapters (except the last one in which some applications are presented); here we set the underlying framework which is a reduced form one, we state the definitions of the objects (mainly the processes that enter functionals, dynamics and controls) used in the setting of the problem and impose the working assumptions under which the problem is tackled.

---

\(^2\) A claim in which counterparty risk affects:
- only one part (unilateral counterparty risk valuation) that now turn to be relevant if one considers the view of a central clearing;
- both the parties (bilateral counterparty risk valuation);
- more than two parties, that can happen if one introduces in the picture the funding issue that is the funder that can be considered not "risk-free" or for products like CDS, CDS basket or CDO's.
1. Introduction and literature review.

In **chapter three** we give a working formulation (not as general as possible) of our problem as a *stochastic switching control* one, which is one of the main original contribution of the work.

The **chapter four** and **five** build on the formulation of the third chapter in order to tackle the main issue of problem solution existence and uniqueness. In particular, we show the difficulties that come from tackling this problem analytically⁵ (chapter four) that lead us to solve it (in chapter five) through suitable *stochastic methods* (mainly the nonlinear *Snell envelope* and *reflected backward SDE* representation). The main contributions here are the problem analysis in chapter four, the *verification theorem* and the existence and uniqueness of the solution in chapter five (which are derived on the basis of the works of El Karoui et al. (1997) and Djehiche et al. (2008)).

In the **sixth chapter** we build on the results of the fifth one in order to define the algorithm and the numerical solution approach which is based on an *iterative optimal stopping procedure* combined with the *Longstaff-Schwartz method*. The algorithm step definition and implementation in the case of a defaultable interest rate swap are our main original contribution of the chapter.

In **chapter seven** we generalize the analysis allowing the strategic interaction between the parties of the contingent CSA, relaxing two of the assumptions made in the second chapter. In particular, we show that the new problem has a natural formulation as a generalized *stochastic differential game of switching type*. The study of the solution and of the equilibrium (Nash) for this game is a very hard problem in general. Anyway, we provide an analysis of the general problem, a definition of equilibrium for this type of game and a solution to the problem in a special ”*symmetric*” case.

---

⁵ We show in particular that a *dynamic programming principle* and a *viscosity solution approach* would be the key to prove the existence of the solution but the theorem should be build and this is out of the scope of the work.
We also provide a solution approach through - *hedging strategy decomposition* and *reflected BSDE* - for the *pricing and hedging* problem (in general, without referring the problem to a particular claim/payoff) of a defaultable claim with contingent CSA of switching type. These are our main original contribution here.

In the last chapter (chapter eight) we conclude the work presenting some interesting applications of our model and focussing on the problem determination of the optimal time to unwind a swap in presence of asymmetries between three agents.

The existence of the optimal switching strategy for our control problem represents an important result compared to the current use of fixed type (full or partial) collateralisation mechanisms, and this is highlighted also from the results of our numerical solution.

This suggests also the greater efficiency, in terms of cost minimization, at systemic/market level of the contingent/switching CSA respect to full collateralized OTC positions with central clearings in the OTC markets. Of course to show this rigorously further and also different type of studies at market level are necessary.

The relevance of our problem solution can be more striking if one extends the view from the single contract to a portfolio of positions, in the sense that it could be desirable (coherently with the preferences of the risk manager) to minimize the (bilateral) CVA, together with the funding and collateral costs, of a set of position hold with an other counterparty. But also at a single contract level, for example in case of OTC deals between financial institutions and sovereigns, a full collateralization can be expensive in terms of liquidity, cost of funding and opportunity costs (given the greater ”dimension” of the deal). That is why can be convenient and more efficient to set a contingent counterparty risk mitigation mechanism of switching type.

So, we conclude by remarking that the problem - although mainly theoretic - can have practical important applications - as we also show in chapter eight - that are actual and relevant, contributing to spread the research in the field.

### 1.2 Review of the literature.

Although the recent history of counterparty risk research field, the body of literature has become rapidly very large. As regards our research problem, given its characteristic highlighted above, it spreads also through the fields of stochastic control problems and stochastic differential games.

So, in order to give a possibly clear account of the references to our research work, we can divide the literature review in three parts as follows:

1. **Counterparty risk references.** For what concerns the counterparty risk literature, the recent monographic works of Gregory (2010) and Cesari (2009) provide a detailed analysis of the problem involved both from the pricing/hedging and risk management point of view. In particular, they tackle the problem related to CVA modeling and hedging highlighting the possible approaches based on dynamic and static hedging strategies in presence of collateral
mechanisms too, focusing especially on operational and computational aspects (at portfolio level in particular) and showing examples and applications to different type of derivative products. Also Bielecki, Cepey et al. (2011) have recently faced the CVA hedging issue in a Markov copula framework proposing a dynamic hedging strategy based on an abstract contract, called "rolling CDS", written on the default event of the counterparty derived via the mean-variance hedging approach - given the market incompleteness and the impossibility to replicate the CVA.

For what concerns the importance and the role of collateral in credit risk mitigation and CVA reduction, other than the already mentioned works of Cesari (2010), Gregory (2009), are worth of mention the works of Brigo et al (2011) that model and determine the effects of collateralization on CVA and on valuation of defaultable claims; Cossin e Aparicio (2001), working under a structural framework, for first propose to use the collateral as a suitable control instrument of the counterparty risk and using it to set a stochastic impulse control problem (with controlled diffusion dynamic) solved via variational inequalities numerical methods.

Johannes and Sundaresan (2003) instead, analyze the role of collateral in the determination of the market swap rates, highlighting its relevance and impact. Their observations are extended by Fujii and Takahashi (2010, 2011) considering the collateral role in term structure modeling and - adopting a Duffie and Huang framework and using the Gauteaux functional derivative - they are able to show the relevant impact that collateral mechanism defined in both unilateral and bilateral CSA can have on CVA (underlining the relative issues of the cost of funding, the collateral currency choice problem and the model risk especially in presence of exotic products).

Other recent works on collateral issue are those of Pieterbarg (2010) who introduces it in the classic Black and Scholes framework; Bielecki et al. (2011) that show, in a Markov copula framework, collateralization and CVA impact through the spread value adjustment in the case of CDS products; Brigo et al. (2011) that generalize the CVA valuation in presence of CSA including the funding issue and Crepey (2011) that for first rigorously tackles the price-hedge problem for a general contract in presence of CSA and funding.

b) Switching control models and numerical solution references. Given the characteristics of our problem shown before, we review also the literature and already known results in optimal switching control theory on which we build our results. The reference literature is that of impulse control problems, also known as sequential starting and stopping problem especially in the economic literature related to optimal investment decision and real option valuation.

Well known papers in these contexts are those of Brennan and Schwartz (1985), the first to propose a two-modes switching model for the life cycle of an investment in the natural resource industry. Other relevant references are those of Dixit (1989) that studies starting/stopping control problems in relation to real/investment options; Brekke and Øksendal (1991 and 1994) that face the switching control problem solution analytically and provide a
1. Introduction and literature review.

first verification theorem for a fairly general stochastic switching control model with markovian/Ito diffusion as dynamic; Duckworth and Zervos (2000, 2001) and Zervos (2003) use the framework of generalized impulse control to tackle generalizations of Brennan-Schwartz model, in the case where the decision to start and stop the production process is done over an infinite time horizon and the market price process of the underlying commodity \( X \) is a diffusion process.

Therefore, Hamadene and Hdhiri (2006) extend the model considering a jump diffusion process while Pham and Vath (2007) derive a closed solution form of viscosity type - using the smooth fit principle - in a two regime switching model with one-dimensional diffusion process and a smooth profit function (with linear growth). Then Djehiche, Hamadene and Popier (2008) generalize it showing the existence of a solution to a general multiple switching problem with finite horizon and general adapted stochastic process using the Snell envelope and reflected BSDE techniques. It is worth of mention also the paper of El Karoui, Pardoux et al. (1995) which is the main reference for most of the theorems and their generalizations related to reflected BSDE, Snell envelope and the deep connection with the viscosity solution of system of PDE with obstacles (also called variational inequalities).

As regards the numerical solution approach, we refer mainly to the works of Longstaff-Schwartz (2001) and to Carmona and Ludkovski (2006) (further details can be found in chapter six).

In relation to finance and in particular to counterparty risk modeling in presence of collateral and mitigation mechanisms, except the already mentioned paper of Cossin and Aparicio (2001), as far as we know, no other control problems as the one that we develop in this work are known.

c) Stochastic differential games references. As regards this part, the body of literature related to stochastic differential games is very wide. In relation to our problem, for what concerns the analytical solution approach, we refer in particular to the central works of Bensoussan and Friedman (1977) in relation to the solution of non-zero-sum stochastic differential games (SDG) with stopping \( \{\tau_1, \tau_2\} \) as controls and Fleming and Souganidis (1989) in relation to zero-sum SDG that approach the solution and determine the game equilibrium through dynamic programming principle and viscosity solution. As regards the probabilistic solution approach, the main references are the work of Cvitanic and Karatzas (1996) that for first establishes the connection between the solution of zero-sum Dynkin game and that of doubly reflected BSDE (other than its analytical solution), El Karoui and Hamadene (2003) that generalize the existence and uniqueness results for zero and non-zero-sum game with ”risk sensitive” controls and Hamadene and Zhang (2008) that use Snell envelope technique to show the existence of a Nash equilibrium for non-zero-sum Dynkin game in a non-markovian framework and Hamadene and Zhang (2010) that tackle the solution of a general switching control problem via general system of BSDE.

These studies have found fruitful application in finance and we refer mainly to the american game option problem as defined in Kifer (2000) (also known as israeli option).
So, let us conclude by remarking that much details on the reference are generally post in the introductive section of the chapters or within the text where we need to refer or use some known results or techniques.
2. MODEL FRAMEWORK AND ASSUMPTIONS.

Although to penetrate into the intimate mysteries of nature and thence to learn the true causes of phenomena is not allowed to us, nevertheless it can happen that a certain fictive hypothesis may suffice for explaining many phenomena.

L. Euler

2.1 Introduction

The introduction of a contingent/optional collateral mechanism inside the CSA deed assumed to be designed by two counterparties of a given defaultable claim has lead us to study the solution of a stochastic control problem of switching type in order to risk manage this type of generalized contract. For the related price and hedge issue, we will see in chapter seven that also the pricing functional equation for this type of problem is strictly related to the solution of a stochastic control problem.

As already mentioned in the introductive chapter, the model analysis will be focussed on the multiple switching type mechanism that include the stopping time and the related *american* or *bermudan* type of problems as a special case.

The complexity of the problem that we want to tackle depends on following modeling choices and factors:

1) the choice of the functional objective for the party and of the switching costs functions;

2) the underlying claim of the contract that can be characterized by a vanilla or exotic payoff, by the number of counterparties/default intensities and stochastic factor (and the related volatilities and correlation);

3) the CSA/collateral cashflows (that can be in cash or securities and present thresholds, MTA, netting, close out amounts, rehypotecation clause, break clause) and the funding cashflows.

In the following section we will come back on each of these points and we will impose some working assumptions in order to make the framework tractable. This approach can be extended to the case \( n > 1 \) that is of a portfolio of \( n \) position/contract in order to set a model for the optimal management of the counterparty risk integrated with collateral and funding issues, but we leave...
this investigation for further research. Hence, we remain in the case \( n = 1 \) and we recall the basic underlying idea and motivation of the analysis: the full/partial collateralization cannot always an optimal/efficient choice for the party of a general deal, because she can ask for more flexibility in the setting of the mitigation mechanism. This flexibility can be ensured by a switching type collateralization, but this impose the problem of how to design it in order to get economic sense? Which optimality criterion can be defined for the counterparty to ensure the existence of an optimal switching strategy for the problem? How to price and hedge this special type of contingent claim?

In order to try to answer to these questions, we start by assuming that two counterparty, one called \( A \) and the other one \( B \) both risky (defaultable), has defined in relation to an underlying claim a risk mitigation mechanism that allow to party \( A \) to switch from “zero” to ”partial” or ”full collateralization” at any time \( t \) during the life of the underlying contract, fixed as \( T \).

It’s worth noting that this case in which both parties are risky (bilateral CVA) would require a bilateral switching mechanism, but this would necessitate other assumptions and a different analysis that we leave for the generalization part.

So, in order to keep the model as general as possible and tied to market real practices but also allowing tractability - keeping the dimensionality of the problem low - we assume that both the name has the same default risk, say \( \lambda^A = \lambda^B \). This implies the symmetry of the (bilateral) CVA between the party and the switching decision too will be ”symmetrically optimal” in the sense that if for \( A \) is optimal to switch to full collateral in \( t = \tau \in [0, T] \), also the other party agrees with this decision in the sense that counterparty \( B \) won’t switch back immediately after \( A \). In other words, no strategic interaction between the parties of the deal enter the picture. Formally, this symmetry is described by the following relations

\[
CV_A - DV_A = CV_B - DV_B \implies BCVA_A = BCVA_B
\]

where the last equality has to be considered in absolute value terms and the \( DVA \) is the well-known ”debt value adjustment”\(^1\). This symmetry relation will be more clear when we formulate the objective functional of the agents, that being a quadratic function allows us to use it to get other results too.

Under different assumptions, if the default intensities of the party would be different, there could be conflicts in the choice of collateralization strategy that should take in consideration the counterparty strategy too. As already mentioned, this is a generalization of our problem that could be better formalized through a generalization of the stochastic differential game notion but we don’t need to introduce it here thanks to the hypothesis of ”symmetry” that allows to tackle the problem from the perspective of just one part (because for the other part it would be the same).

Given these assumptions, we start with the formulation of the framework and defining all the

\(^1\) It incorporates the value of the gains that come from the own default of the party.
ingredients of our model for the stochastic switching control problem.

2. Model framework and assumptions.

To begin the analysis is convenient to describe the framework in which we work and to give some useful definitions of the processes and variables involved\(^2\). The framework is the typical one of the reduced-form models literature: we are in continuous time and we have a probability space described by the triple \((\Omega, \mathcal{G}, \mathbb{Q})\) in which lie two strictly positive random time \(\tau_i\) for \(i \in \{A, B\}\), which represent the default times of the counterparties considered in our model. In addition, we define the default process \(H_i^t = 1_{\{\tau_i \leq t\}}\) and the relative filtration \(\mathbb{H}^i\) generated by \(H_i^t\) for any \(t \in \mathbb{R}^+\). This implies that the full filtration of the model is given by \(\mathbb{G} = \mathbb{F} \vee \mathbb{H}^A \vee \mathbb{H}^B\) where \(\mathbb{F}\) is the (risk-free) market filtration usually generated by a Brownian motion \(W\) (or a vector \(\hat{W}\)) adapted to it, under the real measure \(\mathbb{Q}\). In the rest of the work we are used to denote with \(\mathbb{G}_t = \sigma(\mathbb{F}_t \vee \mathbb{H}^A \vee \mathbb{H}^B)\) that is the sigma algebra of the enlarged market filtration. In addition, we remember that all the processes we consider, in particular \(H_i^t\), are c\`adl\`ag semimartingales \(\mathbb{G}\) adapted and \(\tau_i\) are \(\mathbb{G}\) stopping times (which means to be in a Skorohod topology).

For convenience, let us define the first default time of the counterparties as \(\tau = \tau_A \wedge \tau_B\) which also represent the ending/extinction time of the underlying contract, with the corresponding indicator process \(H_t = 1_{\{\tau \leq t\}}\). This framework can be easily generalized to the case of claim with a third reference name also defaultable, like in CDS contracts, or with multiple defaultable reference names, like in credit basket type claims.

For what concerns the underlying market model it is assumed arbitrage-free, meaning that it admits a spot martingale measure \(\mathbb{Q}^*\) (not necessarily unique) equivalent to \(\mathbb{Q}\). A spot martingale measure is associated with the choice of the savings account \(B_t\) (so that \(B^{-1}\) as discount factor) as a numeraire that, as usual, is given by \(\mathcal{F}_t\)-predictable process

\[
B_t = \exp \int_0^t r_s ds, \quad \forall t \in \mathbb{R}^+ \tag{2.1}
\]

where the short-term \(r\) is assumed to follow an \(\mathbb{F}\)-progressively measurable stochastic process (whatever it is the choice of the term structure model for itself).

We then define the Az\`ema supermartingale \(G_t = \mathbb{Q}^*(\tau > t|\mathcal{F}_t)\) with \(G_0 = 1\) and \(G_t > 0\) \(\forall t \in \mathbb{R}^+\) as the survival process of the default time \(\tau\) with respect to the filtration \(\mathbb{F}\). It is important in the valuation formula when one switch from the full filtration \(\mathcal{G}_t\) to the market one \(\mathcal{F}_t\) that we recall in general here for any \(\mathbb{Q}^*\)-integrable and \(\mathcal{F}_T\)-measurable random variable \(Y\)

\[
\mathbb{E}^{\mathbb{Q}^*}(1_{\{T<\tau\}} Y|\mathcal{G}_t) = 1_{\{\tau>t\}} G_t^{-1} \mathbb{E}^{\mathbb{Q}^*}(G_T Y|\mathcal{F}_t) \tag{2.2}
\]

It’s worth of noting that the process \(G\) being a bounded \(\mathbb{G}\)-supermartingale it admits a unique Doob-Meyer decomposition \(G = \mu - \nu\), where \(\mu\) is the martingale part and \(\nu\) is a predictable increasing process. In particular, \(\nu\) is assumed to be absolutely continuous with respect to the

\(^2\) In particular, we follow, the paper of Bielecki, Jeanblanc, Rutkowski.
Lebesgue measure, so that \( dv_t = v_t dt \) for some \( \mathbb{F} \)-progressively measurable, non-negative process \( v \). So that, we can define now the default intensity \( \lambda \) as the \( \mathbb{F} \)-progressively measurable process that is set as \( \lambda_t = G_t^{-1} v_t \) so that \( dG_t = d\mu_t - \lambda_t G_t dt \) and the cumulative default intensity is defined as follows
\[
\Lambda_t = \int_0^t G_u^{-1} \nu_u = \int_0^t \lambda_u du,
\]
(2.3)
The default intensities \( \lambda_t \) are generally extracted from the market quoted credit spreads and on this base they can be used in a predefined stochastic model for their simulation.

For convenience, we assume that the \textit{immersion property} holds in our framework, so that every càdlàg \( \mathcal{G} \)-adapted (square-integrable) process is also \( \mathcal{F} \)-adapted. In particular, the processes and random variable that we encounter in the following sections will be in general \( \mathcal{G} \)-adapted or predictable and often \( \mathcal{G}_\tau \)-measurable, but because the default time \( \tau \) is inaccessible in the reduced framework, one needs to work with \( \mathcal{F} \)-adapted or predictable process called \textit{pre-default value} processes, denoted with \( X_- \) that is the left limiting process of \( X \). The following lemma is a classic results\(^3\) that allows this change of filtration.

\textbf{Lemma 2.2.1.} Set \( J := 1 - H = 1_{\{t \leq \tau\}} \). For any \( \mathcal{G} \)-adapted, respectively \( \mathcal{G} \)-predictable process \( X \) over \([0, T]\), there exists a unique \( \mathcal{F} \)-adapted, respectively \( \mathcal{F} \)-predictable, process \( \tilde{X} \) over \([0, T]\), called the \textit{pre-default value process} of \( X \), such that \( JX = J\tilde{X} \), respectively \( J_-X = J_-\tilde{X} \).

\textbf{Remarks 2.2.1.} 1) In particular, in our framework we assume also true that, \( \mathcal{F} \)-local martingales stopped at \( \tau \) are also \( \mathcal{G} \)-local martingale, which is a just a modification of the immersion hypothesis.
2) Therefore is worth to remind that \( \mathcal{F} \)-adapted càdlàg process cannot jump at \( \tau \) so one has that \( \Delta X = 0 \) almost surely, for every \( \mathcal{F} \)-adapted càdlàg process \( X \).

These are technical conditions necessary to make more workable the framework and the model that being characterized by defaultable processes, whose value at default is inaccessible, allow to set off jump behavior and all the problems related to consider in the model also close out clause and the contract replacement costs (for which we refer for example to Brigo, Capponi et al (2011)). Anyway, this will be clear in the next section.

\section{Price process and CVA of a contract without CSA.}

Recalled the framework, we start giving some general definitions of the price process (or NPV) and CVA for a defaultable claim for which no CSA that is no collateralization and other mitigation mechanism has been defined by the parties within the relative contract.

First of all, by a defaultable claim signed between two risky counterparty maturing at time \( T \) we mean the quadruple \( (X; A; Z; \tau) \), where \( X \) is an \( \mathcal{F}_T \)-measurable random variable, \( A = (A_t)_{t \in [0, T]} \)

\(^3\) See Bielecki, Crepey et al. (2009).
2. Model framework and assumptions.

is an $\mathcal{F}$-adapted, continuous process of finite variation with $A_0 = 0$, $Z = (Z_t)_{t \in [0,T]}$ is an $\mathcal{F}$-predictable process, and $\tau = \tau_A \land \tau_B$ is the first default time of one of the counterparties (assumed completely inaccessible in the reduced form framework). Generalizing, one should also include in $\tau$ definition the possibility of simultaneous default of the counterparties, namely the event $\tau^* = \{\tau_A = \tau_B\}$, with the relative default indicator $H^*_i = 1_{\{\tau_A = \tau_B\}}$. In the following, we keep the more parsimonious setting because the simultaneous default is not relevant in our symmetric analysis. So, let us star with the definition of the clean dividend $D^{rf}_t$ and clean price process $S^{rf}_t$ which can be seen as $\mathcal{F}$-adapted processes for a fictitious contract free of counterparty risk and funding risk. These contracts are also called somewhere else in the literature as exchange traded contracts defaultable free.

**Definition 2.3.1 (Clean dividend and price process).** The clean dividend process $D^{rf}_t$ of a counterparty default free (exchange traded) contract is the $\mathcal{F}$-adapted process described by the final payoff $X$, the cashflows $A$ and $\tau = \tau^i = \infty$ as follows

$$D^{rf}_t = X \mathbb{1}_{[T,\infty]}(t) + \sum_{i \in \{A,B\}} \left( \int_{[t,T]} dA^i_u \right) t \in [0,T]. \quad (2.4)$$

The clean price process $S^{rf}_t$ would be simply represented by the integral over time of the dividend process under the relative pricing measure, that is

$$S^{rf}_t = B_t \mathbb{E}^Q \left( \int_{[t,T]} B^{-1}_u dD^{rf}_u | \mathcal{F}_t \right) t \in [0,T]. \quad (2.5)$$

Similarly, the following definitions for dividend and price process are valid in the case of bilateral counterparty risk, in which we need to consider the default times, the recovery process and the close out rules that are set in the so called CSA. Here, no CSA is present, so that no collateralization mechanisms is activate, but we will introduce it in the next section.

**Definition 2.3.2 (Bilateral risky dividend and price process).** The dividend process $D_t$ of a defaultable claim with bilateral counterparty risk $(X; A; Z; \tau)$, is defined as the total cash flows of the claim until maturity $T$ that is formally

$$D_t = X \mathbb{1}_{\{T < \tau\}} \mathbb{1}_{[T,\infty]}(t) + \sum_{i \in \{A,B\}} \left( \int_{[t,T]} (1 - H^i_u) dA^i_u + \int_{[t,T]} Z_u dH^i_u \right) t \in [0,T] \quad (2.6)$$

for $i \in \{A,B\}$.

Similarly, the (ex-dividend) price process $S_t$ of a defaultable claim with bilateral counterparty risk maturing in $T$ is defined as the integral of the discounted dividend flow under the risk neutral measure $Q^*$, namely the NPV of the claim, that is formally

$$NPV_t = S_t = B_t \mathbb{E}^{Q^*} \left( \int_{[t,T]} B^{-1}_u dD_u | \mathcal{G}_t \right) t \in [0,T]. \quad (2.7)$$
The financial interpretation of the dividend process $D$ is as follows: $X$ is the promised payoff at maturity, $A$ represents the process of promised dividends (in case of not defaulting $(1 - H_t = 1_{\{\tau > t\}})$ that are all the cashflows of the underlying contract, while the process $Z$, termed the recovery process, specifies the recovery payoff at default.

We remark that $Z$ is usually specified in the literature through a recovery rate $R^i_c$, namely a fraction of the contract value that is recovered after the default of the counterparty. Because of the difficulty to model this stochastic quantity (which depends on a series of factors), usually it is set as orthogonal to the other stochastic factor involved, constant or a deterministic quantity $F_{\tau}$-adapted or predictable. In more general terms, it can be also modeled as a process given its dependence on the value of the underlying contract at default and in this case it would be $G_\tau$-adapted\(^4\). In our framework, we have defined $Z$ as an $F$-predictable process, in fact thanks to lemma 2.2.1, the process has no jumps at $\tau$ so that $Z_{\tau} = Z_{\tau-}$. In particular, we assume (as it is commonly done in literature) that at default of one of the counterparties, the other party recover a fraction of the clean value of the contract right before the default time (if his exposure at default is positive, minus the whole exposure if is negative and the counterparty default) so that no replacement costs are involved.

Given these definitions for the price and dividend process, we are able to state the following general definition for the bilateral CVA (at a given time $t \in \mathbb{R}^+$)

$$BCVA_t = S^f_{\tau} - S_t \quad \forall \ t \in [0, T].$$

as the difference between the counterparty risk free and risky price process/NPV of the contract. To better understand the ingredients of this stochastic object and in order to make it more explicit, let us state the following proposition.

**Proposition 2.3.3 (Bilateral CVA).** The bilateral CVA process of a defaultable claim with bilateral counterparty risk $(X; A; Z; \tau)$ maturing in $T$ satisfies the following relation\(^5\)

$$BCVA_t = B_t \mathbb{E}_{Q^*}\left[1_{\{t < \tau_B \leq T\}} B_{\tau}^{-1} (1 - R_c^B)(S^c_{\tau})^{-} | G_t \right] +$$

$$- B_t \mathbb{E}_{Q^*}\left[1_{\{t < \tau_A \leq T\}} B_{\tau}^{-1} (1 - R_c^A)(S^c_{\tau})^{+} | G_t \right]$$

(2.8)

for every $t \in [0, T]$, where $R_c^i$ for $i \in \{A, B\}$ is the counterparty recovery rate (process).

**Proof.** The proposition proof follows the same simplified lines of the proposition 2.9 of Bielecki et al. (2011). From definition 2.3.2, we explicit the recovery process which is typically defined as a deterministic fraction of the exposure at default in case the NPV (or mark to market) of the contract is positive for the non defaulted counterparty minus the whole value of the negative exposure ($NPV < 0$) when the other party default. Considering the exposures value as the

\(^4\)For example, the recovery process could be defined as a fraction of some function of the NPV of the contract $R_c(t) = g(NPV_c, t)$.

\(^5\)The formulation is seen from the point of view of $B$. Being symmetrical between the party, just the signs change.
2. Model framework and assumptions.

risk-free price process right before default (pre-default values) \( \tau = \tau^- \), we have formally

\[ Z_t = R_t^i(S_t^f)^+ - (S_t^f)^- \quad \forall \ t = \tau \in [0, T] \]

for \( i \in \{ A, B \} \). So, we can write

\[
\int_{[t,T]} Z_u dH_u^i = \int_{[t,T]} (R_t^i(S_t^f)^+ - (S_t^f)^-)(1 - H_u^-)dH_u^i = (R_t^i(S_t^f)^+ - (S_t^f)^-)\mathbb{1}_{\{t<\tau \leq T\}}
\]

and setting for convenience \( X = 0 \) and imposing for the promised dividends

\[
\int_{[t,T]} dA_u^i = \int_{[t,T]} (1 - H_u^i)dA_u^i = \int_{[t,T]} k_u\mathbb{1}_{\{\tau > u\}}du
\]

where \( k \) has to be intended as the net dividend flow between \( A \) and \( B \), then we can substitute these relations in the integral of counterparty risky dividend process \( D_t \), getting (from the point of view of counterparty \( B \))

\[
\int_{[t,T]} B_t^{-1}dD_u = B_t^{-1}(S_t^f)^+ - R_t^B(S_t^f)^-\mathbb{1}_{\{t<\tau \leq T\}} + B_t^{-1}(R_t^A(S_t^f)^+ - (S_t^f)^-)\mathbb{1}_{\{t<\tau \leq T\}} = + \int_{[t,T]} B_t^{-1}k_u\mathbb{1}_{\{\tau > u\}}du
\]

Now, since

\[ 1 = \mathbb{1}_{\{\tau \leq t\}} + \mathbb{1}_{\{\tau > T\}} + \mathbb{1}_{\{t<\tau \leq A \leq T\}} + \mathbb{1}_{\{t<\tau \leq B \leq T\}} \]

and being \((R_t^i(S_t^f)^+ - (S_t^f)^-) = S_t^f + (1 - R_t^i)(S_t^{f^-})\), by recognizing in the last term the risk free price process definition \( S_t^f \) and recalling the defaultable one for \( S_t \) we get

\[
S_t = B_tE^{Q^*}\left( \mathbb{1}_{\{\tau > T\}} \int_{[t,T]} B_u^{-1}dD_u^f | \mathcal{G}_t \right)
\]

\[
+ B_tE^{Q^*}\left( \mathbb{1}_{\{t<\tau \leq A \leq T\}} + \mathbb{1}_{\{t<\tau \leq B \leq T\}} \right)E^{Q^*}\left( \int_{[\tau,T]} B_u^{-1}dD_u^f | \mathcal{G}_t \right) | \mathcal{G}_t
\]

\[
- B_tE^{Q^*}(B_t^{-1}(1 - R_t^A)(S_t^f)^+\mathbb{1}_{\{t<\tau \leq A \leq T\}} | \mathcal{G}_t)
\]

\[
+ B_tE^{Q^*}(B_t^{-1}(1 - R_t^B)(S_t^f)^-\mathbb{1}_{\{t<\tau \leq B \leq T\}} | \mathcal{G}_t) \quad (2.10)
\]

from which we obtain in the end the proof of the proposition

\[
S_t = S_t^f - B_tE^{Q^*}(B_t^{-1}(1 - R_t^A)(S_t^f)^+\mathbb{1}_{\{t<\tau \leq A \leq T\}} | \mathcal{G}_t)
\]

\[
+ B_tE^{Q^*}(B_t^{-1}(1 - R_t^B)(S_t^f)^-\mathbb{1}_{\{t<\tau \leq B \leq T\}} | \mathcal{G}_t) \circ . \quad (2.11)
\]

Therefore, from the last proposition we recall the proper definition of the single components of the BCVA, namely the unilateral (CVA) and the DVA.
Definition 2.3.4 (CVA and DVA). The Unilateral Credit Value Adjustment is defined as follows

$$CV_A_t = B_t E_Q^* \left[ 1_{\{t<\tau_A \leq T\}} B_{\tau}^{-1} (1 - R_c^A)(S_{\tau}^{f})^+ | \mathcal{G}_t \right] t \in [0, T],$$

while the Debt Value Adjustment is formally expressed as follows

$$DV_A_t = B_t E_Q^* \left[ 1_{\{t<\tau_B \leq T\}} B_{\tau}^{-1} (1 - R_c^B)(S_{\tau}^{f})^- | \mathcal{G}_t \right] t \in [0, T]$$

so that (from last proposition)

$$BCVA_t = CV_A_t - DVA_t, t \in [0, T].$$

Remarks 2.3.1. From the last statements, the bilateral CVA can be easily interpreted as a sum of the expected costs/revenues - to which we refer to as exposures (positive or negative) - coming from default of the counterparty or from its own defaults $\tau^i$. In fact, we remark that BCVA is usually expressed also in terms of credit exposures defined as the potential loss that may be suffered by one of the counterparty due to the other party’s default. Different measure are used for CVA computation depending on payoff type and driving factors underlying the pricing problem (one can refers to Gregory (2011) or Cesari et al. 2010), but the most popular ones are Potential Future Exposure (PFE), Expected Positive Exposure (EPE) and Expected Negative Exposure (ENE). In particular we recall here the definition of EPE and ENE.

Definition 2.3.5 (EPE and ENE). The Expected Positive and Negative exposures of a contract with bilateral default risk are defined as follows

$$EPE_t = E_Q^* \left[ (1 - R_c^A)(S_{\tau}^{f})^+ | s = \tau_A \right], s \geq t \in [0, T],$$

$$ENE_t = E_Q^* \left[ (1 - R_c^B)(S_{\tau}^{f})^- | s = \tau_B \right], s \geq t \in [0, T].$$

To conclude, we remark (see for example Bielecki, Brigo et al. (2011)) that in case of independence between discount factor and default intensities the CVA process can be represented in terms of EPE and ENE as follows

$$BCVA_t = B_t E_Q^* \left( 1_{\{t < \tau_A \leq T\}} B_{\tau}^{-1} (1 - R_c^A)(S_{\tau}^{f})^+ \right) - B_t E_Q^* \left( 1_{\{t < \tau_B \leq T\}} B_{\tau}^{-1} (1 - R_c^B)(S_{\tau}^{f})^- \right)
= B_t \int_{[t, T]} B_{\tau}^{-1} E_Q^* \left( (1 - R_c^A)(S_{\tau}^{f})^+ | \tau_A = s, \tau_A \leq \tau_B \right) Q^s(\tau_A \in ds, s \leq \tau_B)
- B_t \int_{[t, T]} B_{\tau}^{-1} E_Q^* \left( (1 - R_c^B)(S_{\tau}^{f})^- | \tau_B = s, \tau_B \leq \tau_A \right) Q^s(\tau_B \in ds, s \leq \tau_A)
= B_t \int_{[t, T]} B_{\tau}^{-1} EPE_t^s Q^s(\tau_A \in ds, s \leq \tau_B) - B_t \int_{[t, T]} B_{\tau}^{-1} ENE_t^s Q^s(\tau_B \in ds, s \leq \tau_A)$$

6 This orthogonality hypothesis will be useful in the implementation and numerical part.
for \( t \in [0, T] \) where from the framework section we know that \( Q^*(\tau_i \in ds, s \leq \tau_i) = \lambda_i := G^{-1}(t)Q^*(t = \tau_i \in dt) \), namely the counterparty default intensities.

### 2.4 Price process and CVA in presence of CSA.

The CSA can be described, in general terms, as an articulated part of the OTC contracts devoted to mitigate the counterparty risk through the definition of a collateral agreement between the parties over the life of the underlying claim\(^7\). It has been largely used also to reduce the capital requirements imposed by the Basel committee on portfolio exposures and to give a more competitive price to the CVA charges.

As already mentioned, this type of agreement usually contains a lot of factors, clauses and events that are complex to model. To get an idea of this, we recall some of the main components of a CSA and sending to the relative literature for more details:

- **Close out clause:** it defines the value of the claim that is due or received in case of default of the counterparty or of its own default;
- **Break clause:** it’s an option to close/end the contract without any costs;
- **Netting clause:** this clause is largely used if the counterparty has multiple positions with the other party, to net their values in case of default and then proceed to the close out settlement;
- **Additional termination event:** option to settle the claim if a predefined event, usually related to variable like credit ratings or balance-sheet indicator (like Net-assets), is verified;
- **Threshold:** is a predefined amount that indicates the value of the claim above which the collateral is called by one of the parties;
- **Minimum transfer amount** (MTA): is the minimum amount above which the collateral can be called and transferred between the parties (it is usually set together with the threshold);
- **Reuse and rehypotecation clause:** are clauses that allows the parties to use the collateral for other transaction especially if it is not cash collateral. Otherwise the CSA specifies that the collateral has to be post in a segregated account and remunerated to a specific rate (usually the overnight);
- **Other parameters:** the CSA usually defines also the 1) base currency in which collateral is post; 2) the type, one way or two way; 3) the securities that are eligible to be used as collateral and 4) the margin/call frequency.

\(^7\) Now, thanks to diffusion of central clearings to mitigate the counterparty risk, this agreement has been considering, as a third party involved, the central clearing as a guarantee of the transaction.
To simplify things, we assume in our model that no break clause and additional termination event are set in the CSA between the parties, because these clauses would impose the dependence of the optimal switching/control strategy on other factors not strictly related to CVA and increase the dimension and the complexity of the problem (without gaining in terms of explanatory power). Therefore, for convenience we assume that the collateral is post in a segregated account so that no reuse and rehypotecation are allowed and the collateral can only be made up by cash (or highly liquid securities).

With these things in mind, we focus the analysis on the most common type of collateralization, partial collateralization and full/perfect collateralization, on which we base the formalization of the contingent CSA of switching type that we want to model.

a) **CSA with partial collateralization.** This is one of the most common counterparty risk mitigation mechanism in which the collateral can be called by both the parties (namely, a two way CSA), at the margin dates specified within the CSA, if the fair value $S_{rf}^t$ of the contract, that is its clean price (NPV), is greater/minor of specified thresholds $\Gamma_A$, $\Gamma_B$ plus/minus the minimum transfer amount MTA. We note that the collateral process/account here does not depend on the counterparty risky price, because this would make it recursive and more complicate. In the following we state a possible definition of the collateral process from the point of view of counterparty $A^8$.

**Definition 2.4.1 (Collateral account/process).** Let us define for the bilateral CSA of a contract the positive/negative threshold with $\Gamma_i$ for $i = \{A, B\}$ and the positive minimum transfer amount with MTA. The collateral process $\text{Coll}_t : [0, T] \to \mathbb{R}$ is a stochastic $\mathcal{F}_t$-adapted process$^9$ defined as follows

$$\text{Coll}_t = \mathbb{1}_{\{S_{rf}^t > \Gamma_B + MTA\}}(S_{rf}^t - \Gamma_B) + \mathbb{1}_{\{S_{rf}^t < \Gamma_A - MTA\}}(S_{rf}^t - \Gamma_A), \quad (2.18)$$
on the time set $\{t < \tau\}$, and

$$\text{Coll}_t = \mathbb{1}_{\{S_{rf}^\tau > \Gamma_B + MTA\}}(S_{rf}^\tau - \Gamma_B) + \mathbb{1}_{\{S_{rf}^\tau < \Gamma_A - MTA\}}(S_{rf}^\tau - \Gamma_A), \quad (2.19)$$
on the set $\{\tau \leq t < \tau + \Delta\}$.

Considerations similar to the ones relatives to the recovery rate, are valid here for the threshold $\Gamma_i$ and MTA, that we have chosen to set as constant parameters but in general they can be modeled as deterministic or process/functions of the exposures or NPV.

Now, given the presence of the collateral, is important to define the CSA close-out cashflows at counterparty default time that enter in the dividend process $D_t$ and consequently in the price process $S_t$ (not clean) and CVA of the defaultable claim. It is worth of mention that the CSA

---

8 We also underline that the collateral process in general can depend on the clean price process in a path dependent way especially if the collateral is not cash based and in presence of rehypo end reuse, cases that we have left out of the analysis here.

9 In the literature, depending on the CSA provisions and modeling choice, the process is also considered $\mathcal{F}_t$-predictable or in general adapted to $\mathcal{G}_t$. 
close-out cashflows are in general $G_\tau$-adapted, but as we know from the framework hypothesis, to define them we will take always the clean price of the contract right before $\tau = \tau^-$, so they are actually $G$-predictable.

In order to define it formally, we need to distinguish the possible close out events and the related cashflows:

**Case** $\{ t = \tau = \tau_B \} \implies$ Here the counterparty defaults so after the collateral transfer takes place, if the uncollateralized mark to market is negative, that is if $S^r_t - Coll_t < 0$, $A$ closes out the position by paying the defaulting counterparty the uncollateralized mark to market. If it is positive, $A$ closes out the position and receives a fraction $R^B_c$ of the uncollateralized mark to market from the counterparty. So the close-out payment is defined as

$$\bar{R}^B = R^B_c (S^r_t - Coll_t)^+ - (S^r_t - Coll_t)^-$$

**Case** $\{ t = \tau = \tau_A \} \implies$ Here $A$ defaults, so after the collateral transfer takes place, if the uncollateralized mark to market is positive, that is if $S^r_t - Coll_t > 0$, the counterparty closes out the position by paying the defaulting counterparty the uncollateralized mark to market. If it is negative, the counterparty closes out the position and receives a fraction $R^A_c$ of the uncollateralized mark to market from $A$. Hence we get the close-out payment

$$\bar{R}^A = (S^r_t - Coll_t)^+ - R^A_c (S^r_t - Coll_t)^-$$

**Case** $\{ t = \tau = \tau_A = \tau_B \} \implies$ In case that both the parties default simultaneously if the uncollateralized mark to market is negative, the counterparty receives a fraction $R^B_c$ of the uncollateralized mark to market; however, if the uncollateralized mark to market is positive, the investor receives a fraction $R^A_c$ of the uncollateralized mark to market, so that the close out cashflow will be

$$\bar{R}^{A,B} = R^B_c (S^r_t - Coll_t)^+ - R^A_c (S^r_t - Coll_t)^- = -(S^r_t - Coll_t)$$

So, putting all the things together, we can state the following definition of CSA close-out cashflows (from the point of view of $A$).

**Definition 2.4.2 (CSA close-out cashflows).** Let us define the recovery rate $R^i_c$ for $i \in \{ A, B \}$ a $G$-predictable process set as a positive constant and the uncollateralized mark to market as $\theta_t = S^r_t - Coll_t$ (for $t = \tau$). The CSA close-out cashflows is the real valued left limited $G_{\tau^-}$-measurable process defined as follows

$$C^{f}_{t}^{CSA} = Coll_t + 1_{(t=\tau_B)}(R^B_c \theta^+_t - \theta^-_t) - 1_{(t=\tau_A)}(R^A_c \theta^-_t - \theta^+_t) - 1_{(t=\tau_A=\tau_B)}\theta_t.$$
Now, substituting all the CSA close out in the defaultable dividend, say $D_t^{CSA}$, we get

$$\int_{[t,T]} B_u^{-1} dD_u^{CSA} = B_T^{-1} X1_{[T,\infty]}(t) + \int_{[t,T]} B_u^{-1} \mathbb{1}_{(\tau>u)} dA_u^{t}$$

$$+ B_\tau^{-1} Coll_\tau + B_\tau^{-1} R^A \mathbb{1}_{\{t<\tau=t_A\leq T\}}$$

$$+ B_\tau^{-1} R^B \mathbb{1}_{\{t<\tau=t_B\leq T\}} - B_\tau^{-1} R^{AB} \mathbb{1}_{\{t<\tau=t_A=t_B\leq T\}}.$$

Expanding this relation in terms of $S^f_t$ and $D^f_t$, taking the expectation and using arguments similar to the proof of proposition 2.3.3, one can obtain the generalized price process in presence of CSA with partial collateralization, say $S^{CSA}_t$, that is

$$S^{CSA}_t = B_t \mathbb{E}^Q \left( \int_{[t,T]} B_u^{-1} dD^f_u | G_t \right)$$

$$+ B_t \mathbb{E}^Q \left( \mathbb{1}_{\{t<\tau=t_B\leq T\}} B_\tau^{-1} (1-R^B)(S^f_\tau - Coll_\tau)^- | G_t \right)$$

$$- B_t \mathbb{E}^Q \left( \mathbb{1}_{\{t<\tau=t_A\leq T\}} B_\tau^{-1} (1-R^A)(S^f_\tau - Coll_\tau)^+ | G_t \right).$$

From the last relations follows the following proposition on the bilateral CVA in presence of CSA.

**Proposition 2.4.3 (CVA (bilateral) with CSA).** The bilateral CVA for a defaultable claim maturing in $T$ and mitigated by a partial collateralization satisfies the following relation

$$BCVA_t^{CSA} = B_t \mathbb{E}^Q \left[ \mathbb{1}_{\{t<\tau=t_B\leq T\}} B_\tau^{-1} (1-R^B)(S^f_\tau - Coll_\tau)^- | G_t \right]$$

$$- B_t \mathbb{E}^Q \left[ \mathbb{1}_{\{t<\tau=t_A\leq T\}} B_\tau^{-1} (1-R^A)(S^f_\tau - Coll_\tau)^+ | G_t \right] \quad (2.20)$$

\forall t \in [0, T], where the first term is the (unilateral) CVA and the second the DVA (collateralised).

**Proof.** The proof follows the same line of the proposition 2.3.3.

b) **CSA with perfect/full collateralization.** This type of collateralization can be easily understood as the limit case of the partial collateralization when the delta between the margin dates, say $\Delta t_m = t_m - t_{m-1}$, tends to zero and no thresholds and minimum transfer amount are defined in the related CSA. So, we get in this case that the collateral process, say $Coll_t^{Perf}$, is always equal to the mark to market, namely to the (default free) price process $S^f_t$ of the underlying claim, that is formally (by taking the definition 2.4.1)

$$Coll_t^{Perf} = \mathbb{1}_{\{S^f_t > 0\}}(S^f_t - 0) + \mathbb{1}_{\{S^f_t < 0\}}(S^f_t - 0) = S^f_t \ \forall t \in [0, T], \text{ on } \{t < \tau\}. \quad (2.21)$$

and

$$Coll_t^{Perf} = S^f_{\tau^-} \ \forall t \in [0, T], \text{ on } \{\tau \leq t < \tau + \delta t\} \quad (2.22)$$

Then, by plugging the last relation in the definition 2.4.2 of CSA cashflows, one easily gets

$$C_f^{tCSA}_t = Coll_t^{Perf} \ \forall t \in [0, T]. \quad (2.23)$$
Actually, the CSA close out cashflows won’t get sense anymore because the counterparty risk is zeroed with perfect collateral and, in fact, the CVA defined in 2.4.3 become

\[ BCV A_t^{\text{Coll}_{\text{Perf}}} = B_t E_{Q^*} \left[ 1_{\{t<\tau\leq T\}}(0) | G_t \right] - B_t E_{Q^*} \left[ 1_{\{t<\tau\leq T\}}(0) | G_t \right] = 0 \quad \forall t \in [0, T] \quad (2.24) \]

that implies

\[ BCV A_t^{\text{Coll}_{\text{Perf}}} = S_t^{rf} - S_t = 0 \implies S_t = S_t^{rf} \quad \forall t \in [0, T] \quad (2.25) \]

namely the equality between the defaultable price and risk free price of the claim. Under this collateralization rule, the proper discount curve for pricing the deal is the collateral curve (typically expressed by the risk free rate plus a credit spread or the Eonia if the collateral is cash based). This theoretical type of mitigation mechanism, is actually largely used in practice setting a margin frequency usually daily or weekly. In fact, the new Basel three requirements impose this type of collateralization to the market central clearing. It is clear that this mitigation mechanism can be quite expensive in terms of liquidity and funding costs. So, as we need to consider all the relevant factors that impact the optimal switching strategy for our control problem, we are lead to introduce also the funding problem in the picture. Before doing this, we give here some definitions in order to formulate the contingent CSA of switching type that we have assumed to be designed by the parties of the contract.

### 2.5 Contingent CSA, price process and CVA

**Contingent CSA of switching type**. The contingent mechanism of switching type that we assume to be set between the parties in our model, can be seen as a mix of zero and perfect collateralization and it is natural to model it (as will be clearer from the next section) introducing the switching indicators \( z_j \) and times \( \tau_j \) for \( j = 1, \ldots, M \), that affects - being the agent controls - the system dynamic and consequently the processes involved, in particular the collateral process. Let us consider just two possible switching regimes (but it can be easily generalized also to the case of partial collateralization) so that the indicator will take only two values, say \( z_j = \{0, 1\} \):

- when \( z_j = 1 \) the collateral is null, no collateralization is active, that implies a full BCVA (with no CSA);
- while for \( z_j = 0 \) we have a CSA with perfect/full collateralization (which zeroes the BCVA).

So, as done before we give the main definitions for collateral and CVA process which now depend also on these switching control \( \{\tau_j, z_j\}_{j=1}^m \) variables that needs to be determined optimally. How, it will be the object of the following section.

**Definition 2.5.1 (Contingent Collateral process)**. The contingent collateral \( \text{Coll}^C_t \) is the \( \mathcal{F}_t \)-adapted process defined for any time \( t \in [0, T] \) and for every switching time \( \tau_j \in [0, T] \) and
\[ j = 1, \ldots, M, \text{ switching indicator } z_j \text{ and default time } \tau \text{ (defined above as } \min\{\tau_A, \tau_B\} \text{), as follows} \]

\[
\text{Coll}_t^C = S_t^{rf} \mathbb{1}_{\{z_j=0\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}} + 0 \mathbb{1}_{\{z_j=1\}} \mathbb{1}_{\{t_j \leq t \leq t_{j+1}\}} \text{ on } \{t < \tau\} \tag{2.26}
\]
on the set \{t < \tau\}, and

\[
\text{Coll}_t^C = S_t^{rf} \mathbb{1}_{\{z_j=0\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}} + 0 \mathbb{1}_{\{z_j=1\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}} \tag{2.27}
\]
on the set \{\tau \leq t < \tau + \Delta t\} \text{ where we recall that } S_t = \text{Coll}_t^{Perf} \text{ from point b) results.}

Using Definition 2.5.1 and the results stated for close out amounts and bilateral CVA, we get easily the following results. Firstly, setting \( D_t^C \) and \( S_t^C \) as the dividend and price process in presence of the contingent CSA just defined, we can set

\[
\begin{align*}
D_t^C &= D_t^{rf} \mathbb{1}_{\{z_j=0\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}} + D_t \mathbb{1}_{\{z_j=1\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}} \\
S_t^C &= S_t^{rf} \mathbb{1}_{\{z_j=0\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}} + S_t \mathbb{1}_{\{z_j=1\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}}
\end{align*}
\]

for \( t \in [0, T \wedge \tau] \). From these relations it is easy to recover the general formulation for the bilateral CVA process in the contingent case as follows

\[
\begin{align*}
\text{BCVA}_t^C &= S_t^{rf} - S_t^C \\
&= S_t^{rf} - (S_t^{rf} \mathbb{1}_{\{z_j=0\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}} + S_t \mathbb{1}_{\{z_j=1\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}}) \\
&= 0 \mathbb{1}_{\{z_j=0\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}} + \text{BCVA}_t \mathbb{1}_{\{z_j=1\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}}
\end{align*}
\]

so that the following definition is well posed.

**Definition 2.5.2 (BCVA with contingent CSA).** The bilateral CVA of a contract with contingent CSA of switching type, \( \text{BCVA}_t^C \) is the \( \mathcal{G}_t \)-adapted process defined for any time \( t \in [0, T] \), for every switching time \( \tau_j \in [0, T] \) and \( j = 1, \ldots, M \), switching indicator \( z_j \) and default time \( \tau \) (defined above), as follows

\[
\text{BCVA}_t^C = \text{BCVA}_t \mathbb{1}_{\{z_j=1\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}} + 0 \mathbb{1}_{\{z_j=0\}} \mathbb{1}_{\{t_j \leq t < t_{j+1}\}}, \text{ for } t \in [0, T \wedge \tau], \tag{2.28}
\]

where the expression for \( \text{BCVA} \) is known from Proposition 2.3.3.

We end by noting that the CSA closeout cashflows are different from zero only when collateralization is active, that is

\[
Cf_t^{CSA} = \text{Coll}_t^C \forall t = \tau \in [0, T].
\]

**Remarks 2.5.1.** We underline that all these processes are in general càdlàg semi-martingales and is difficult to deal with their dynamics which are also recursive and non linear being affected by agent switching controls. In the next section we will see from the set up of the stochastic control the working assumptions needed to ease the problem analysis.
2. Model framework and assumptions.

2.6 Funding with CVA and collateral.

The funding issue and how to model it in the whole issue of pricing and hedging of derivatives and defaultable claims in particular, is one of the hottest topic in the literature at the moment. The main difficulty that derives from the inclusion of the cost of funding in the valuation, especially of path dependent products, is that the problem becomes recursive and highly nonlinear. The already mentioned paper of Crepey for first has shown a rigorous approach to deal with this problem based on the backward SDE (BSDE) solution which is also coherent with the least square monte carlo procedure based on forward simulation and backward induction. We can anticipate that the same recursive problem affect our switching-type control problem because the value of the optimal strategy depends on future switching decisions, but this will be more clear later.

Here we do not give a complete analysis of the funding/investing issue (one can refer for this also to Brigo et al. (2011)) but we focus on the study of the funding cost precess and its impact on CVA and collateral. The funding/investing problem is strictly related to the cash/liquidity management, a function that is usually covered internally by the treasury function or, externally, by the market or a third part, called the funder. There are different modeling approach, in our specific case we focus on a deal specific/bottom-up\(^{10}\) approach assuming that the funder is a risk-free third party.

We assume the existence of a cash account held by the funder, say \(C_{r}^{\text{Fund}} \geq 0\), that can be positive or negative depending on the funding (> 0) or investing (< 0) strategy that counterparty use in relation to the price and hedge of the specific deal which is also collateralised in our case. Hence, including the funding in the analysis, this has impact on:

- the cashflows related to the hedging portfolio/strategy of the specific deal, especially if the hedge cannot be perfect as in the case of defaultable claims so that the trader needs to fund or invest the cash surplus deriving from his hedging portfolio;

- the cashflows related to CVA hedging; this is usually included in the hedging portfolio of the whole claim;

- the cashflows related to the collateral/CSA defined between the parties; for example, when the exposure trespass a specified threshold and the counterparty has to post/receive the collateral it incurs in funding/opportunity costs.

Formally we have that, for a collateralized (defaultable) claim inclusive of funding/investing costs or cashflows, whose value in \(t = 0\) is replicated by an hedging portfolio \(\Pi_{t}^{\text{Hedge}}\), the funding account \(C_{t}^{\text{Fund}}\) will be

\[
P_{t}(0, C_{t}^{\text{Fund}}) = \Pi_{t}^{\text{Hedge}} + \text{Coll}_{t} + C_{t}^{\text{Fund}} \\
C_{t}^{\text{Fund}} = P_{t}(0, C_{t}^{\text{Fund}}) - \text{Coll}_{t} - \Pi_{t}^{\text{Hedge}} \forall t
\]

From the last equations is evident that the funding cashflows (or account) will be always activated when the deal is collateralized and hedged; moreover we mark the recursive nature of the above

\(^{10}\) The other approach largely used in the literature is called of the "large homogeneous pool".
relation given that the value of the claim in \( t \) depends on the funding strategies after \( t \), but these funding strategies depend themselves on the claim value at the same times after \( t \).

For what concerns our problem, in relation to CVA, whatever it is the case of zero, partial or full collateralization, we eliminate a lot of complications by making the assumption that the CVA value is charged to the counterparty but no hedging portfolio is set, so that the funding costs are null \( C^{Fund} = 0 \) (namely the funding account is not active)\(^{11}\).

As regards the funding issue in relation to the collateralization, under the assumptions of segregation (no rehypo) and collateral made up by cash, we distinguish the following cases:

1) If the counterparty has to post collateral in the margin account, she sustains a funding cost, applied by the external funder, represented by the borrowing rate \( r^\text{borr}_t = r_t + s_t \) that is the risk free rate plus a credit spread \( s \) (that is usually different from the other party). By the other side, the counterparty receives by the funder the remuneration on the collateral post, that is usually defined in the CSA as a risk free rate plus some basis points, so that we can approximate it at the risk free rate, that is \( r_t + bp_t \cong r_t \).

Clearly, we remark that the existence of the above defined rate implies the existence of the following funding assets

\[
\begin{align*}
dB^\text{borr}_t &= (r_t + s_t)B^\text{borr}_t dt \\
\bar{dB}^\text{rem}_t &= (r_t + bp_t)\bar{B}^\text{rem}_t dt
\end{align*}
\]

2) Instead, let’s consider the counterparty that call the collateral: as above the collateral is remunerated at the rate given (as the remuneration for the two parties can be different) by \( \bar{B}^\text{rem} \), but she cannot use or invest the collateral amount (that is segregated), so she sustains an opportunity cost, that can be represented by the rate \( r^\text{opp}_t = r_t + \pi_t \), where \( \pi \) is a premium over the risk free rate.

Hence, we assume the existence of the following assets too

\[
\begin{align*}
\bar{dB}^\text{opp}_t &= (r_t + \pi_t)\bar{B}^\text{opp}_t dt \\
\bar{dB}^\text{rem}_t &= (r_t + \bar{bp}_t)\bar{B}^\text{rem}_t dt
\end{align*}
\]

The objective of the last definitions will be clearer the next paragraph in which we can pass to define the counterparty objective and to go deeper in the analysis of stochastic switching control problem that we want to tackle.

### 2.7 Counterparty objective and the fundamental trade-off.

Now we have in mind all the relevant elements and variables that the counterparty of a collateralized defaultable claim of contingent type has to take in consideration to define the optimality

\(^{11}\) An alternative simplifying assumption is that the CVA hedging is realized without resorting to funding.
criterion/objective. As already said, the counterparty is assumed to be counterparty risk averse, she wants to minimize the negative impact of CVA, namely the costs of the default of the other party, via the definition of a contingent collateral/CSA, but she wants to do it optimally deciding the optimal times when to switch to a partial or full collateralization. Of course the collateralization implies other running costs, like the funding ones. So, the objective will be to determine over the set of controls represented by the switching time (and indicator) the optimal switching strategy that minimizes over \([0, T]\) the overall running costs related to CVA, collateral and funding and the instantaneous switching costs that we need to model. This kind of problem is highly non linear and recursive, as already highlighted above, and because of in the partial collateralization case the CVA process is complicated by the presence of the CSA close out cash-flows that depends recursively on the controls (the switching times), this makes the analysis more complex. So, for the next sections, we leave aside the partial collateralization and we analyze the problem for the counterparty in the twofold decision case: "switch from zero collateral (full CVA) to full collateralization (zero CVA)" (and viceversa).

We recall here the main working assumption that we use in the analysis of the problem for the following sections:

**Hp 1)** The analysis is done under the "symmetry hypothesis": the parties of the deal has symmetric objectives and the same default intensities \(\lambda_A = \lambda_B\).

**Hp 2)** Both the parties are counterparty risk averse but we assume no strategic interaction for the moment.

**Hp 3)** The analysis is focussed on the full/perfect collateralization case;

**Hp 4)** The CVA cost process is not funded, \(C^{\text{Fund}} = 0\);

**Hp 5)** All the processes considered are intended to be pre-default value processes (as from lemma 2.2.1.)

Some notes on the unilateral counterparty objective. Before passing to the formal analysis and modeling of our problem, let us focus on the unilateral case to make some reasoning useful for the next sections. Firstly, note that the CVA, being a weighted summation of positive terms that is a function \(\varphi(EPE_t, \alpha, Q)\) of \(EPE_t \geq 0\) and positive parameters \(\alpha = (R, B_t)\) scaled by the default probabilities\(^{12}\) (by definition \(\in [0, 1]\)), is always positive \(CVA_t \geq 0\) \(\forall t\).\(^{13}\) So, this term that incorporates the expected loss that \(A\) would incur if \(B\) defaults must be hedged by \(A\) if no collateralization has been set. Because of market incompleteness due to default process, as shown in the literature, the CVA hedging strategy is not perfect, typically in the literature it is a mean-variance hedging type. This implies that the expected costs of the hedging are only approximately equal to the CVA value.

\(^{12}\) This representation is true when the discount factors and default intensities are assumed orthogonal.

\(^{13}\) This is obvious too if we take the definition of the CVA as the difference of the free risk price process which is always greater than the defaultable one.
In theory, the effective objective of counterparty $A$ is the costs minimization, say $C_{\text{Hedge}}(t) \in [t, T]$ related to this hedging strategy that one should determine. For a matter of convenience, because of the approximation underlined above between these expected hedging costs and the CVA value, we are implicitly assuming that the following objectives/functionals are equivalent to solve for counterparty $A$

$$\inf_{\alpha \in \mathbb{R}^k} C_{\text{hedge}}(t; \alpha) = \inf_{\alpha \in \mathbb{R}^k} CVA(t; \alpha) \forall t \leq T.$$  \hspace{1cm} (2.33)

where the hedging costs $C_{\text{hedge}}(t)$ and $CVA(t)$ as $\mathcal{F}_t$-measurable and adapted processes.

This assumptions agrees with the HP 4) given that no hedging portfolio is set to hedge the CVA and so the related funding costs are zeroed. Now, recalling our problem, we have assumed that counterparty $A$ wants to reduce the possible future losses from the CVA component by setting a contingent CSA that allows her to dynamically switch in every $t \leq T$ to full collateral (or partial) and also reversing the switch (passing from full to zero collateral).

At this point, let us make some heuristic reasoning on the switching strategy and on the continuation/switching region of the problem in object. Let us start defining the indicator $z \in \mathbb{Z} := \{0, 1\}$, where $z = 0$ means no collateralization while $z = 1$ means full collateral, in the unilateral case we can state the following proposition.

**Proposition 2.7.1.** If for a given $t \in [0^+, T]$ the collateral is switched to $z = 1$, after the switch if the CVA is not monotonically increasing and the collateral costs are ”large enough”, is optimal to reverse the switch to $z = 0$.

**Proof.** The proof we give is not rigorous, it is based just on a logic reasoning. Here this issue is central because switching from zero to full collateral means set to zero the CVA, that is the counterparty risk is nulled and $S_t = S^f_t$, that is also equal to $\text{Coll}^f_t$ (from section 2.3). But the activation of the full collateralization, (which can be set up with a Clearing House as imposed by the new Basel provisions) is not ”for free”, without any costs, in fact a trade off emerges: as already shown in the paragraph related to the funding issue, when the price process of the claim underlying the contract become negative, counterparty $A$ has to post collateral in a segregated account (given the assumption post above), that is amount of money, remunerated with the free risk rate $r_t \equiv r_t + bp_t$ that needs to be fund (at a rate $r_{\text{borr}}$) and that could be invested at an higher rate of return $r_{\text{opp}}$. So, switching to full collateral determines the emergence of a collateral cost process that can be represented as the sum of fixed costs - that emerge for example from the set up of the account or in general represent operational costs - called ”instantaneous costs” plus funding and opportunity costs that should be computed only when the collateral is post that is when the price process $NPV_t$ becomes negative for counterparty $A$. So, we can calculate this running collateral costs taking the expected negative path of the price process until maturity, that is formally:

$$CC_t = f(t, c, r_{\text{borr}}, r_{\text{opp}}, NPV) = \mathbb{E}_t \left\{ c - \exp(r_t - r_{\text{opp}} - r_{\text{borr}})(T - t) \left[ \int_t^T [NPV_s^-] ds \right] \right\} \hspace{1cm} (2.34)$$

where:
- $c > 0$ is the fixed instantaneous switching cost assumed constant and positive;

- $r_{opp}$ is the opportunity cost set as a deterministic and positive variable but it can be modeled as a stochastic process too;

- $r_{borr}$ is the borrowing/funding cost, set as a given positive variable too;

- $r \equiv r_{rem}$ is the remuneration for the collateral set approximately equal to the risk free rate;

- $[NPV_t^-] = \min[S_t, 0]$ is the expected negative value of the NPV of the claim.

So, the collateral cost process described above gives us the suitable candidate for the cost function of the stochastic control problem that we are going to formalize and highlights the fundamental trade-off between the CVA reduction through setting collateral to $z = 1$ and the relative collateral cost process that emerges. We are now able to explain the sense of the initial statement: in fact if after the switching to $z = 1$ the CVA is not increasing which means that the claim price process have negative paths which lower the CVA (because it takes only the positive path of $S_t$). But along this path the counterparty has to post collateral which increases the cost of collateral process defined above, so depending on the parameters, if these expected costs become greater than CVA, it becomes optimal for counterparty A to switch to $z = 0$ collateral along these paths and time (and eventually switch back to $z = 1$ when the CVA level increase becomes $CVA(t) > CC(t))$.

The following proposition follows easily from the above reasoning.

**Proposition 2.7.2.** If in $t = 0$ the collateral is set to $z = 0$ (or equivalently $z = 1$) and the CVA is greater (smaller) than the collateral/switching costs almost one time $t \in [0^+, T]$ (with $Q\{\omega_t\}$ a.s), then almost one switch to $z = 1$ ($z = 0$) would be optimal.

**Proof:** this is obvious by the assumption 1) relative to counterparty objective and by the hypothesis on CVA and collateral costs $CC$ process. Therefore, this is trivial because otherwise it wouldn’t have economic/financial sense for $A$ the setting of this switching mitigation mechanism.

**Remarks 2.7.1.** We remark that in the ”extreme cases” in which for every $t \in [0^+, T]$ the CVA process remains grater than the collateral costs process, the problem reduces to an optimal stopping/switching time, that is

$$if \ CVA(t) \geq CC(t) \ \forall t \in [0^+, T] \iff \mathcal{C}_{ad} := \{Z, T\} = \{1, \tau_1\}$$

where $\mathcal{C}_{ad}$ is the set of admissible controls, made up (as will be defined later) of all the sequences of switching times $\mathcal{T}$ and indicator $Z$ that becomes a singleton in this case. This condition are intended to be sufficient to get only one switching over the life of the contract.

By the other side, reversing the inequality we have the following condition

$$if \ CVA(t) \leq CC(t) \ \forall t \in [0^+, T] \iff \mathcal{C}_{ad} := \{Z, T\} = \{0, \infty\}$$
that is the switching is never optimal so the stopping time $\tau = \infty$ and $z = 0$. The last condition is economically relevant in the contract design phase because it would give a control for the convenience to set the collateralization (for a given claim). Of course this always depends on the set of parameters of the model and on the underlying contract and payoff too.

**Notes on the bilateral counterparty objective.** Returning to the generalized bilateral case, is worth noting that the observations and derivations done above are not true anymore in this case, because the BCVA level can take positive or negative values, since that in the bilateral case one has to consider the expected incomes coming from its own default (DVA) that must be subtracted to the CVA. For this reason the switching strategy and the optimality criterion is not so clear anymore.

Leaving the setting of the problem as a minimization one, it can be convenient to consider as objective the minimization of the squared level of the BCVA or its variance respect to a given threshold, say $\delta$, in the case of "zero collateral", that becomes a relevant parameter of the problem even in the full collateral case. More specifically, to give sense to our optimal switching problem in the bilateral case that is to minimize the BCVA variance by switching to full collateralization and minimizing the related costs associated to switch (that we need to model), one can:

1. square the BCVA in order to make it always positive and set the threshold $\delta = 0$ to simplify things. Note that in this setting, the switching times would happen as soon as the BCVA become quite different from zero and positive;

2. consider the variance of the BCVA respect to a threshold $\delta$ positive that can be given exogenously, for example it can be a given level of capital or liquidity allocated by the risk management function or by the treasury. An other possibility is to set it, coherently with the definition of variance, as the expected BCVA value calculated at inception in $t = 0$. An other reasonable alternative is to make the threshold endogenous setting it as a control (possibly depending on time) of our minimization problem.

3. differently one can complicate the cost function and the setting by considering, for example, the *Basel 3* constraints defined in the new CVA provisions.

In the case of switching to "full collateral" we model the running costs related to this case in similar way by keeping the same ideas given above, in order to give sense to the minimization problem and the optimality of the switching strategy. The problem would be clearer in the following section where we define formally the running cost functions and the whole bilateral setting. What we should remark here is that the two statement given above in relation to the fundamental trade off that both counterparty has to face in the unilateral CVA case, are not true anymore in the bilateral one because of the different running cost functions that are not linear anymore and the tradeoff is less trivial now.

In particular, if we consider the case 1) in which the threshold $\delta$ is set to zero, the first statement about the optimality of the switching to full collateral will still be true; in fact whatever are the other parameters of the problem, if the BCVA would be greater than the switching costs almost
one time \( t \in [0^+, T] \), the switching will be optimal almost one time during the life of the contract.
If we are in case 2) where \( \delta \neq 0 \), depending on its level and on the other parameters setting of the problem, it could never be optimal the switching to full collateral so invalidating the first statement. About the second statement we need a more careful analysis and will discuss it later when we approach the solution of our switching control problem (chapter four).
3. STOCHASTIC SWITCHING CONTROL PROBLEM FORMULATION.

It is not knowledge, but the act of learning,
not possession but the act of getting there,
which grants the greatest enjoyment.
C.F.Gauss

3.1 Introduction

From the reasoning and definitions of the last chapter is quite clear what kind of stochastic control problem we are facing: we are (counterparty A) in \( t = 0 \) and we want to determine the optimal strategy that minimizes the expected costs deriving from switching the collateral from zero to full or partial and vice versa until the maturity of the underlying defaultable claim. This problem can be seen as a sequential optimal starting and stopping time but is natural to formalize it as a multiple switching control problem with finite horizon, a particular case of the more general class of impulse control models. The main ingredients of this control problem class are the following:

1. Setting of the objective functional of the problem which is generally made up of an objective function, a reward function and a switching/impulse cost function;
2. Setting of the stochastic dynamics that describe the state of the system considered;
3. Setting and modeling of the controls set in which search for the solution of the problem.

We remark that in the next section we do not make any distinction between the CVA and the BCVA, so that the analysis of dynamics and controls for the unilateral and bilateral case we carry on is the same. Anyway, for the objective functional formulation we distinguish the two cases. We remark that everything is intended in the perspective of counterparty A.

3.2 Modeling dynamics and controls of the problem.

From the list set above, the last point is the less problematic because in the switching control problems the set of control has a natural and standard formalization. For the objective functional, we have almost all the elements because of the insights of the latter section, so we model it later. As regards the model for the "system" dynamic, this is more problematic because the natural
choice would be to consider the CVA dynamic given that the CVA enters the objective function of counterparty $A$ (and symmetrically $B$ as discussed above). But it can be shown\footnote{See, for example, Bielecki, Jeanblanc, Rutkowski (2007)} that CVA dynamic is quite complicate and is not an Ito diffusion, so we need to use different solution methods (for general càdlàg processes) that we leave for further research. Here, we make a different choice for the state dynamics which circumvents the problem but at cost of a minor generalization.

For starting, we remember that for our problem we have assumed only counterparty $A$ and $B$ defaultable, which excludes defaultable claims in which more than one default process is present (like CDS or basket). Therefore, we are intended to use the bilateral CVA expression set in the latter section under the hypothesis of interest rate orthogonal to $\tau$, namely the default intensities of counterparty $B$. So in the following, for CVA/BCVA we refer to the formula set in the equation 2.17, that we recall here as follows

$$ BCVA_t = B_t \int_{[t,T]} B_s^{-1} \mathbb{E}((1 - R^A_c) (NPV^{-})_{t} | \tau_A = s, \tau_A \leq \tau_B) \lambda^A_s ds $$

$$ - B_t \int_{[t,T]} B_s^{-1} \mathbb{E}((1 - R^B_c) (NPV^{+})_{t} | \tau_B = s, \tau_B \leq \tau_A) \lambda^B_s ds \quad (3.1) $$

where we have used the default intensity definition for $\lambda^i_t$ and we have made an abuse of notation given the price process definition $S_{t}^{rf} := NPV_t$.

In order to model the system dynamic in our model, let us note that the BCVA, for a given process realization ($\omega \in \Omega$), can be defined as a bounded and $\mathcal{F}_t$-measurable function $f(.)$ of:

- the price process (default free) or NPV, that can be expressed as a function $h(t, X, k)$ (also $\mathcal{F}_t$-measurable ) of time, of some constants ($k$) and a stochastic factor $X$\footnote{This can be in general a stochastic vector.}, that is typically the interest rate process, for example in an interest rate swap;

- of the default process $\tau$ which is determined through the intensity process, say $\lambda$, that for convenience we have assumed equal for both the counterparties $\lambda^A_t = \lambda^B_t$;

- discount factors that presume the existence in the model of a bank account defined by $B_t^{-1} = \exp(- \int_0^t r_s ds)$;

- a vector of constant factors $\alpha = (k, R_c)$ defined in the contract, like the recovery rate $R_C = 1 - LGD_c$ or the ”strike rate” that is included in $k$.

So, formally we have that

$$ CVA(t) = f(t, NPV, \lambda, B_t, \alpha) $$

but being $NPV = h(t, X, k)$ with $k \subset \alpha$ and $B_t = h'(t, X)$, we get that

$$ CVA(t) = f(t, X, \lambda, \alpha). $$

For what concerns the cost function related to collateral that also enters in the objective functional of our problem, we observe that this is a function of other parameters assumed constant for

convenience and again of NPV, so we can set \((X, \lambda)\) as the relevant factors that we need to model the dynamic of our "system", the other elements are just parameters to set\(^3\), say

\[
\alpha' := (\alpha, c, r_{bort}, r_{opp}) = (k, R_c, c, r_{bort}, r_{opp})
\]

It’s worth noting that some of the fixed parameters can enter in the optimization problem and in particular the cost of funding can include a *liquidity spread*, especially in presence of collateral made up of non standard highly liquid securities. This would necessitate a model for it and would generalize and complicate the valuation, but we leave the topic for further research.

Now, we are just left to choose the model for \((X, \lambda)\). Because the typical defaultable claim described by this factors is a swap, \(X\) will be our swap/libor rate that we can model with a single factor stochastic process like the *CIR* or with a two factor model, which is more adapted to capture the movements and correlations of the term structure. Even for the default intensity \(\lambda\) (equal for \(A\) and \(B\)) we can use a stochastic process like CIR or CIR+++, typical choices done in the literature. So, keeping the dimensionality of the problem low, formally we assume that the dynamic of our problem is described by the following system of SDE

\[
\begin{align*}
\frac{dX_t}{X_t} &= \eta(\mu - X_t)dt + \sigma \sqrt{X_t}dW^x_t, X(0) = x_0 \tag{3.2} \\
\frac{d\lambda_t}{\lambda_t} &= \kappa(\gamma - \lambda_t)dt + \nu \sqrt{\lambda_t}dW^\lambda_t, \lambda(0) = \lambda_0 \tag{3.3} \\
\langle X, \lambda \rangle_t &= \rho_{X,\lambda} dt = 0 \quad (\rightarrow \rho_{X,\lambda} = 0), \tag{3.4}
\end{align*}
\]

where both the processes are of CIR type and we have set the correlation parameter of the two Brownian motion \((W^x, W^\lambda)\)\(^4\) to zero without loss of generalization, because as shown, for example, in *Brigo, Pallavicini (2008)*, the effect of correlation between interest rates and default intensities is not much relevant in modeling counterparty risk. The parameters of the vector \([\eta, \mu, \sigma]_x; [\kappa, \gamma, \nu]_\lambda\) are all positive constants\(^5\) describing the mean reversion speed, level and volatility of the relative process and the classical condition for the existence of a strong solution for the given SDE are verified. To keep the origin inaccessible and eliminate possible negative values, is used to impose the following condition on parameter dynamics:

\[
\begin{align*}
2\eta\mu &> \sigma^2 \\
2\kappa\gamma &> \nu^2.
\end{align*}
\]

This formulation can be easily generalized to include jumps in the dynamics if the underlying claim requires it or as already mentioned using the more indicated *two-factor gaussian process* (G2++) or the shifted CIR (CIR++) (see *Brigo, Mercurio (2006)* for details).

For now we keep the easier and more parsimonious version of the dynamic for our problem and, given a smooth function \(\phi(t, X, \lambda)\) of class \(C^3_{x,\lambda}\) with respect to \(X\) and \(\lambda\) and \(C^4_t\) respect to time, and given that our vector dynamic \(Y(t) = [t, X_t, \lambda_t]\) (by the above modeling assumption) is a markovian

\[\text{\footnotesize{Eventually they could be modeled as deterministic or stochastic too.}}\]

\[\text{\footnotesize{Assumed to be already } F_t\text{-adapted and under the objective measure } Q.}\]

\[\text{\footnotesize{Or they can be specified as deterministic function of time.}}\]

Diffusion, it admits the following \textit{second order partial differential operator} called \textit{infinitesimal generator} of the stochastic system:

\[
A\phi(y) = \frac{\partial \phi}{\partial t} + \eta(\mu - x) \frac{\partial \phi}{\partial x} + \kappa(\gamma - \lambda) \frac{\partial \phi}{\partial \lambda} + \frac{\sigma^2}{2} x \frac{\partial^2 \phi}{\partial x^2} + \frac{\nu^2}{2} \lambda \frac{\partial^2 \phi}{\partial \lambda^2}.
\]

We keep this in mind and we remark that the dynamics of the system we are considering is affected by the counterparty choice to switch from zero to full/partial collateral or viceversa, so we need to define the control set for our problem.

In a multiple switching/impulse control type problem (with finite horizon) like this, in which counterparty \(A\) can intervene or "give impulse" to the system every times in \([0, T]\) through a sequence of switching times, say \(\tau_j \in \mathcal{T}\), with \(\mathcal{T} \subset [0, T]\), where we denote the switching set

\(\mathcal{T} := \{\tau_1, \ldots, \tau_M\} = \{\tau_j\}_{j=1}^M\)

with the last switching \(\{\tau_M \leq T\} (M < \infty)\) and the \(\tau_j\) are, by definition of \textit{stopping times}, \(\mathcal{F}_\tau\)-measurable random variable. Therefore, at any of these switching times, we need to define the related set of the \textit{switching/impulse indicator}, which is quite trivial in our case with only two possible switching regime:

\(\mathcal{Z} := \{z_1, \ldots, z_M\} = \{z_j\}_{j=1}^M\)

with the \(z_j\)'s that are \(\mathcal{F}_{\tau_j}\)-measurable switching indicators taking values \(z_j = \{0, 1\} \forall j = 1, \ldots, M\), so we have

\[
\begin{cases}
  z_j = 1 & \Rightarrow \text{"zero collateral" (full CVA)} \\
  z_j = 0 & \Rightarrow \text{"full collateral" (null CVA)}
\end{cases}
\]

Here we are focussing on the simpler case of collateralization leaving the "partial case" aside. So we have that the control set for our problem is formed of all the sequences of indicators and switching times \(\mathcal{C} \in \mathbb{R}_{+}^M\), that is

\(\mathcal{C} = \{\mathcal{T}, \mathcal{Z}\} = \{\tau_j, z_j\}_{j=1}^M\)

with the convention - and without losing generality - that in \(t = 0\) the initial condition of the indicator be \(z_0 = 1\) because usually contracts start with no collateralization at the inception\(^6\).

\textbf{Remarks 3.2.1.} It is worth of mention that, as defined above, the set of switching controls can be made up of maximum \(M\) switching times \(\tau_j\) (and related indicators \(z_j\)) being the counterparty who decides optimally the number of switches that minimize her cost functional. In general one could leave unbounded the number of switched allowed, namely \(M = \infty\) leaving its number determined endogenously by the solution of the stochastic optimization. An other more restrictive possibility that is quite common in CSA contracts is to allow switching only at a given set of

\(^6\) In fact, we could have assumed a contract starting with full collateralization \(z_0 = 0\) at inception in \(t = 0\), but the analysis would have been the same.
3. **Stochastic switching control problem formulation.**

dates \( \tau_j \in [t_1, \ldots, t_n] \subset [0, T] \), typically the margin call dates or even at the coupon exchange dates.

For a matter of notations, we define \( \zeta_j = 1 - z_j \) in order to better distinguish between the two regimes, but the control set is intended to be the same.

At this point, we need to remark - from the definitions of contingent CSA and (bilateral) CVA process, 2.3.4 e 2.3.5 - that the switching controls enter and affects the dynamic of this processes which are described by the system of stochastic differential equations assumed above. In fact, as we know, switching to full collateralization implies \( \text{CVA} = 0 \) (and \( \text{BCVA} = 0 \)), that is \( S_t = S^i_t = \text{Coll}_{perf} \). This means no counterparty risk and so the default intensity dynamic \( d\lambda_t \) won’t be relevant, just \( dX_t \) will be considered until \( z = 0 \), namely until the collateralization will be kept active (\( \zeta = 1 \)). So, formally, we have:

\[
\begin{align*}
\text{if } & \{ z_j = 1 \} \text{ and } \{ \tau_j \leq t < \tau_{j+1} \} \Rightarrow \\
D^C_t &= D_t \\
S^C_t &= S_t \\
\text{BCVA}^C_t &= \text{BCVA}_t \forall t \in [0, T \land \tau]
\end{align*}
\]

so that the relevant dynamic to model the BCVA process in this regime is

\[
\begin{align*}
dX_t &= \eta(\mu - X_t)dt + \sigma \sqrt{X_t}dW^x_t \\
X(0) &= x_0 \\
d\lambda_t &= \nu(\gamma - \lambda_t)dt + \eta \sqrt{\lambda_t}dW^\lambda_t \\
\lambda(0) &= \lambda_0.
\end{align*}
\]

By the other side

\[
\begin{align*}
\text{if } & \{ z_j = 0 \} \text{ and } \{ \tau_j \leq t < \tau_{j+1} \} \Rightarrow \\
D^C_t &= D^f_t \\
S^C_t &= S^f_t = \text{Coll}_{perf} \\
\text{BCVA}^C_t &= 0 \forall t \in [0, T \land \tau]
\end{align*}
\]

so that the relevant dynamic to model the process in this regime will be just

\[
\begin{align*}
dX_t &= \eta(\mu - X_t)dt + \sigma \sqrt{X_t}dW^x_t \\
X(0) &= x_0
\end{align*}
\]

This clearly implies that the infinitesimal generators are actually two, say \( A^z \) and \( A^\zeta \) defined as follows

\[
A^z \phi(y) = \frac{\partial \phi}{\partial t} + \eta(\mu - x) \frac{\partial \phi}{\partial x} + \kappa(\gamma - \lambda) \frac{\partial \phi}{\partial \lambda} + \frac{\sigma^2}{2} x \frac{\partial^2 \phi}{\partial x^2} + \frac{\nu^2}{2} \lambda \frac{\partial^2 \phi}{\partial \lambda^2}. 
\]

\[
(3.6)
\]

\[
\text{if } \{ z_j = 1 \} \text{ and } \{ \tau_j \leq t < \tau_{j+1} \} \text{ and}
\]

\[
A^\zeta \phi(y) = \frac{\partial \phi}{\partial t} + \eta(\mu - x) \frac{\partial \phi}{\partial x} + \frac{\sigma^2}{2} x \frac{\partial^2 \phi}{\partial x^2}
\]

\[
(3.7)
\]
if \( \{ z_j = 0 \} \) and \( \{ \tau_j \leq t < \tau_{j+1} \} \).

To ease the notation, we set our SDE system in vectorial form defining

\[
\begin{align*}
dY_{\text{cad}}(t) &:= \begin{bmatrix}
dt \\
dX_t \\
d\lambda_t \\
dZ_t \\
dY_{\text{cad}}(0) = \\
t = 0 \quad x_0 \\
\lambda_0 \\
Z_0 = 1\end{bmatrix}, \\
\end{align*}
\]

where we are left to define the set of admissible switching control \( \mathcal{C}_{ad} \subseteq \mathcal{C} = \{ T, Z \} \), admissible in the sense that our stochastic (controlled) system described by \( dY_{\text{cad}} \) has a unique solution.

### 3.3 Modeling the objective functional.

The last element left to model to complete the picture of our switching control problem is the objective functional of the agent, namely counteparty \( A \). We start by setting it in the more general bilateral case because it is almost the same as the unilateral one in which only the running cost function relevant in the zero collateralization state is different and we mark this difference below.

A general formulation of the objective functional for a switching control minimization problem can be stated as follows

\[
J(T, Z)(y) = \inf_{u \in \mathcal{C}_{ad}} \mathbb{E}^y \left[ \int_{t=0}^{T} (\exp^{-r(T-s)})F_Z(Y^u(s))ds + G(Y^u(T))1_{\{ T < \infty \}} + \sum_{\tau_j < T} K(Y^u(\tau_j), z_j) \right]
\]

where for notational convenience has been set \( u =: \{ T, Z, T \}^T \) while the function \( F(.) \) represents the agent objective that is in our case the running cost function, \( G(T) \) is the so called "reward function" and \( K(.) \) the instantaneous switching costs of the problem. Reminding the last section, for our problem we have that:

a) for what concerns the function \( G(.) \) or reward function, because our problem is of finite time type and because in \( T \) which is the terminal state/time of our problem we can get or the terminal collateral value which equals (minus)\(^8\) the \( NPV(T) \) if \( z_T = 0 \) (or \( \zeta_T = 1 \))\(^9\), or nothing, because the CVA is zero or almost zero, say \( \epsilon \), for \( t \to T \) if no collateralization is kept active \( z_T = 1 \) (or \( \zeta_T = 1 \)). Formally we set, as we work with the square of the running costs function as we shall see below,

\[
G(Y^u(T)) = \epsilon^2 \times 1_{\{ z_T = 1 \}} + (-NPV(T))^2 1_{\{ z_T = 0 \}}.
\]

\(^7\) So, \( u \) contains \( \mathcal{C}_{ad} \) the set of admissible controls in which the solution must be searched.

\(^8\) The minus is due to the minimization costs setting, so if \( NPVT \) is positive it is not a cost, so we put the minus ahead.

\(^9\) Note that here we are making an abuse of notation setting \( z_T \) instead of \( z_M \) but these are meant to be the same, that is \( M = T \) for the reward function. Therefore, this is one of the easiest way to model this function that can be even complicated.

Actually, because it’s possible in theory that the agent decide to switch just before $T$, to set off this type of behavior\footnote{That is typical of the continuous time framework, when we will tackle the implementation, discretizing this won’t be relevant anymore.} which causes a jump in the value function and to be more general and formal we set $z_{T^-} := \lim_{t \to T} z_t = z_T$ (and similarly for $\zeta_{T^-}$), and setting for convenience $NPV(T^-) = NPV(T)$ and $\epsilon = 0$, we can rewrite the reward function, that is even our terminal/boundary condition, as

$$G(Y^u(t)) = \begin{cases} (-NPV(T))z_{T^-} \Rightarrow & \text{if collateral is active} \\ 0^2z_{T^-} = 0 \Rightarrow & \text{otherwise / no collateral} \end{cases}$$

Here is important to remark that the valuation of the control problem is set in $t$ with $t < \tau \wedge T$ and all the value processes are intended as pre-default values and clearly $\tau = \infty$ in $T$.

b) For what concerns the running cost function $F(.)$, we need to distinguish between the two switching regimes. So if $\{z_j = 0\}$, we know from the latter section that the CVA in the unilateral case and the BCVA in the bilateral one, are the chosen candidate because they are both function of $Y = [t,X,\lambda]$ and some given parameters $\alpha$. So assuming that the counterparty be interested only in the minimization of the CVA/BCVA level or its variance respect to a given threshold $\delta$ (that we have already discussed above), the objective function would be modeled as

$$\begin{align*}
\text{"Unilateral" } F_z(Y^u(t)) &= zCVA(t) \Rightarrow \text{"min CVA level"} \\
F_z(Y^u(t)) &= z[CVA(t) - \delta]^2 \Rightarrow \text{"min CVA variance"}
\end{align*}$$

$$\begin{align*}
\text{"Bilateral" } F_z(Y^u(t)) &= zBCVA(t) \Rightarrow \text{"min BCVA level"} \\
F_z(Y^u(t)) &= z[BCVA(t) - \delta]^2 \Rightarrow \text{"min BCVA variance"}
\end{align*}$$

where $z$ is the control switching indicator that multiply the CVA/BCVA because it is set to $z = 1$ at inception and it can be zeroed when counterparty $A$ decide ”optimally” to switch to full collateral, setting $z = 0$.

So, going for the variance in the bilateral case which is - for technical reasons that we highlight in chapter four - more convenient, we have that

$$\text{if } \{z_j = 1\} \rightarrow F_z(Y^u(s)) = [BCVA(s) - \delta]^2 = [(CVA(s) - DVA(s)) - \delta]^2;$$

while, by the former discussion and reasoning and recalling relation (2.34), we have

$$\text{if } \{z_j = 0\} \rightarrow F_\zeta(Y^u(s)) = NPV(s) - \int_{s}^{T \wedge \tau_j + 1} \beta(s)[NPV(v)]^- ds \in [\tau_j, T],$$

for $j = 1, \ldots, M$, where we have set $\beta(s) = c - \exp(-(r_{borr} + r_{opp})$ and we have considered here the collateral posted or payed at the time of switching that is actually just NPV in $s$ but the
last function is not anymore the true one in the bilateral setting in which we need to redefine the funding/opportunity costs $\beta(s)$. More precisely, recalling our assumptions and definitions of chapter 2), in our bilateral setting the collateralization is full (not partial, minimum transfer amounts and rehypo are considered) so is equal to the NPV process (and we assume just cash collateral or asset highly liquid and tradable), switching to full collateral generate two possible expected costs that we have to take in consideration. Firstly, let us assume the existence of the same asset $B_t^{borr}, B_t^{opp}, B_t^{rem}$ defined in section 2.4) on funding and collateral\(^{11}\).

1. If $\zeta_j = 1$ and $NPV(s) > 0$ ($s \geq \tau_j$), the counterparty has to post the collateral in a segregated account that must be remunerated at the free risk $r$ (by the external funder) but $A$ cannot use or invest that amounts so she has an opportunity costs, say $r_{opp}$ (set as a constant). Because this costs has to be considered until the collateral is active and the NPV is positive, we can define this running cost function and $\beta(s)$ as

$$\int_s^{T \wedge \tau_{j+1}} \beta(v)^+[NPV(v)]^+ dv = \int_s^{T \wedge \tau_{j+1}} \exp(-(r_{opp} - r)v[NPV(v)]^+ dv \ s \in [\tau_j, T].$$

2. By the other side, if $\zeta_j = 1$ and $NPV(s) < 0$ ($s \geq \tau_j$), $A$ has to post the collateral in a segregated account that must be remunerated at the free risk $r$ (by the external funder) but $A$ now has to find the money to do this so she has a funding cost, say $r_{borr}$. Because this costs are alive until the collateral is active and the NPV is negative, we can define this running cost function and $\beta(s)$ in this case as

$$\int_s^{T \wedge \tau_{j+1}} \beta(v)^-[NPV(v)]^- dv = \int_s^{T \wedge \tau_{j+1}} \exp(-(r_{borr} - r)v[NPV(v)]^- dv \ s \in [\tau_j, T].$$

So we can put the things together and being

$$\int_s^{T \wedge \tau_{j+1}} [NPV(v)]^- ds = \int_s^{T \wedge \tau_{j+1}} [NPV(v)]^+ ds + \int_s^{T \wedge \tau_{j+1}} [NPV(v)]^- ds$$

and setting

$$R(s) = \begin{cases} 
-\beta(s)^- = -\exp(-(r_{borr} - r)s & \text{if } \zeta_j = 1 \text{ and } NPV < 0 \\
\beta(s)^+ = \exp(-(r_{opp} - r)s & \text{if } \zeta_j = 1 \text{ and } NPV > 0
\end{cases}$$

and 0 otherwise\(^{12}\), we get that the running switching cost function for the bilateral case is the following

$$F_\zeta(Y^u(s)) = \left( \int_s^{T \wedge \tau_{j+1}} \beta(s)^+[NPV(v)]^+ ds + \int_s^{T \wedge \tau_{j+1}} \beta(s)^-[NPV(v)]^- ds - NPV(s) \right)$$

$$= \left( \int_s^{T \wedge \tau_{j+1}} R(s)[NPV(v)]^- ds - NPV(s) \right) \text{ for } s \in [\tau_j, T], \quad (3.8)$$

\(^{11}\) To ease notation we set to zero the basis spreads $\delta p_t, \delta p^*_t$ over the risk free rate.

\(^{12}\) Note we have set the minus before the exponential in the case in which the $NPV$ is negative in order to obtain a positive value, in fact we’ve considered always positive values as costs according to the minimization objective.
∀τ_j, z_j ∈ \{T, Z\}. Let us note that before the integral we should set a summation over j ≤ M to consider all the possible switches but we set it later in the general objective functional; therefore we have set the minus before the \(NPV(s)\) because if it’s positive the counterparty A receives the collateral which must be subtracted to the expected costs, if it’s negative it must be added, being a sort of instantaneous cost or revenues (depending on the sign). In effects it can be incorporated in the instantaneous cost part of the functional too, but we will see (in chapter 5) that this make the switching boundary stochastic for which the analysis in more involved and the optimal strategy may not exists\(^{13}\).

So, because of our problem is a minimization, given the structure and the definition for \(F_\zeta(Y^u(s))\) it can be positive or negative because of the last term \((NPV(s))\), we square it by keeping the same structure of the other running function related to the zero collateral case, so that we have

\[
i \{z_j = 0\} \rightarrow F(Y^u(s)) = \left( \left( \int_s^{T \land \tau_{j+1}} R(v) |NPV(v)|ds - NPV(s) \right) - \delta \right)^2 \text{ for } s \in [\tau_j, T]
\]

and ∀τ_j, z_j ∈ \{T, Z\}. Here the threshold \(\delta\) is set equal to the running cost function of the other regime switching, but it can be modeled differently.

c) What is left to model is the instantaneous switching costs function. It takes the fixed costs that counterparty has to consider when switches regime. A typical working formulation is the following:

\[
K(Y^u(\tau_j), z_j) = \sum_{j \geq 1} e^{-r\tau_j} c_j(t) 1_{\{\tau_j < T\}}, \quad \forall \tau_j, z_j \in \{T, Z\},
\]

that can be made more explicit as follows

\[
K(Y^u(\tau_j), z_j) = \sum_{j \geq 1} e^{-r\tau_j} \left[ z_j c_{zero} 1_{\{j \geq 2\}} 1_{\{\tau_j < T\}} + \zeta_j c_{full} 1_{\{\tau_j < T\}} \right] \quad \forall \tau_j, z_j \in \{T, Z\}, \quad (3.9)
\]

where we have set

- \(z_j c_{zero} 1_{\{j \geq 2\}}\) is a fixed cost that derives from switching from full to zero collateral \(\Rightarrow z = 1\). This is true only for \(j \geq 2\) because the first switch \((j = 1)\) can be only to full collateral because \(z_0 = 1, z_1 = 0, z_2 = 1\) and so on.

- \(\zeta_j c_{full}\) is the other fixed cost related to switching to full collateral, in fact it is multiplied by \(\zeta\) that is equal to 1 to indicate this switch. The other parameters are all known from the last section or given constants.

\(^{13}\) See Hamadene, Zhang. Switching problem and related system of reflected backward SDEs.

Putting all these pieces together we get our generalized (bilateral) functional

\[ J(T, Z)(y) = \inf_{u \in \{C_{ad}\}} \mathbb{E}^y \left[ \int_0^T \left( \exp^{-rs} \left( \sum_{j \geq 1} \left[ (CVA(s) - DVA(s)) - \delta \right]^2 \mathbb{1}_{\{z_j = 1\}} \right) \right. \right. \]

\[ + \left. \left. \left( \int_u^{T \wedge \tau_j + 1} R(s)[NPV(v)] du - NPV(s) - \delta \right)^2 \mathbb{1}_{\{z_j = 0\}} \right) ds \right] 

\[ + \sum_{j \geq 1} \exp^{-r\tau_j} c_j(t) \mathbb{1}_{\{\tau_j < T\}} \left. \left. \left[ - NPV(T) \xi_T^- \right] \right. \right. \]

\[ \left. \left. \left. - NPV(T) \xi_T^+ \right) \right. \right. \left. \mathbb{1}_{\{z_j = 1\}} \right] \] \quad \forall s, \tau_j \in [t, T \wedge \tau] \quad (3.10) \]

that can be represented in a more compact way, by setting

\[ F_Z(y, u) = \begin{cases} 
\sum_j \left[ (CVA(s) - DVA(s)) - \delta \right]^2 + 0^2 z_T^- & \text{if } \{z = 1\} \\
\sum_j \left( \int_u^{T \wedge \tau_j + 1} R(s)[NPV(v)] du - NPV(s) - \delta \right)^2 + (-NPV(T) \xi_T^-)^2 & \text{if } \{z = 0\} 
\end{cases} \]

\forall s, \tau_j \in [t, T \wedge \tau], \] which allow us to write our control problem as follows

\[ J(T, Z)(y) = \inf_{u \in \{C_{ad}\}} \mathbb{E}^y \left[ \int_0^T \left( \exp^{-rs} F_Z(y, u) \right) ds + \sum_{j \geq 1} \exp^{-r\tau_j} c_j(t) \mathbb{1}_{\{\tau_j < T\}} \right] \mathbb{F}_t \]

subjecting everything to the given dynamics \( dY \) and the admissible control set \( C_{ad} \) it tells us that the objective is to find, if it exists and is unique, the optimal sequence of switching times and indicators \((z_j, \tau_j)^* = u^*\) that minimize the expected costs functional, that is formally to find the value function of the problem \( V(y) \) and \( u^* \in C_{ad} \) such that

\[ V(t, x, \lambda; u) := V(y) = \inf \{ J^u(y); u \in C_{ad} \} = J^{u^*}(y) \quad (3.11) \]

with the terminal condition in \( T \) on the value function that will be

\[ V(T, x; u) := \inf_u \{ (-NPV(T) \xi_T^-)^2, 0^2 z_T^- \} = 0. \quad (3.12) \]

### 3.4 Problem recursion and alternative reformulation

As it is quite evident from a first analysis of the functional objective expression, the problem we are trying to tackle is intrinsically recursive given that the optimal switching decision today depends, path by path, on future switching strategy. In addition, we note that the recursion is partially hidden also inside the running cost functions of both the switching regimes:

\[ \text{Note that we are making an abuse of notation by compacting the reward final condition into the running cost function.} \]

\[ \text{In the unilateral case we would just need to put } DVA = 0 \text{ and } \delta = 0 \text{ and delete the square.} \]

\[ \text{This is the main reason that will lead us in the next section to use the iterative optimal stopping tool together with dynamic programming principle to tackle the solution of the problem.} \]

a) in the "zero collateralization" regime the running cost function of the problem is the following
\[ \{ z_j = 1 \} \rightarrow F_z(Y^u(s)) = [BCVA(s) - \delta]^2 \]
where, from the definition of bilateral CVA above given, as it is an integral over the expected future exposures of the underlying contract, that depend on future switching decision, a recursive backward procedure is necessary to recover paths of this term.

b) Similarly, in the "full collateralization" regime we have a cost function
\[ \{ z_j = 0 \} \rightarrow F_\zeta(Y^u(s)) = \left( \int_s^{T \wedge \tau_j + 1} R(s)[NPV(s)]ds - NPV(s) - \delta \right)^2 \]
here it is easier to see that the integral term takes as terms the expected future NPV whose determination depends on the switching strategy at future times. So also in this case the recursion is present and one must deal with it.

So actually the problem is characterized by a double order of recursion that is analytically very difficult to tackle (as we will show) and also cumbersome from a computational point of view. A possible way to simplify things and to get out from the "curse of recursion" is to kill the recursion almost at one level or part of it. We refer, in particular, to the possibility to kill the recursion in one or both the running cost functions. The most likely candidate is the "full collateral" one, given its dependence on NPV which is more manageable than CVA, especially if the underlying contract is fairly standard and non exotic. The idea - already employed in the counterparty risk literature in relation to the close-out amounts and contract NPV calculation at default via replacement costs referring to exchange traded contracts/transactions - is to "replace" the value or the NPV of the contract that must be computed recursively through times with an exogenous price process that follows the NPV of the contract but it is derived from exchanged traded contracts, say \( NPV^{\text{Exch}}(t) = S^{\text{Exch}}(t) \).

In our case, we need to refer the NPV that enters the cost function, which we remember being a default free process in the full collateralization regime, to a "defaultable free" security regularly and liquidly trade on an authorized market. So, assuming the existence of such security, to simplify things we can substitute in the running cost function (and collateral) in regime \( z_j = 0 \) the exchange traded price process \( NPV^{\text{Exch}}(t) \) whose dynamic is assumed conveniently to be well described by an Ito diffusion as
\[
\begin{align*}
    dNPV^{\text{Exch}}_t & = \chi_t(NPV^{\text{Exch}}_t)dt + \xi_t(NPV^{\text{Exch}}_t)dW^{\text{Exch}}_t \\
    NPV^{\text{Exch}}_0 & = NPV^{\text{Exch}}_0
\end{align*}
\]
for \( W^{\text{Exch}} \) a standard \( \mathbb{Q} \) measurable Wiener process \( \mathcal{F}_t \)-adapted. The drift-diffusion parameters \( (\chi_t, \xi_t) \) respect the usual technical condition in order to ensure existence and uniqueness to the SDE’s solution.

The formulation of the optimization problem remains in substance the same except for the dynamic of the collateral that would be assumed exogenously given. This is the standard approach

to problem recursion but - as it will be clear from the next sections - thanks to some working assumptions that help to simplify computations (in the numerical section) we will not need the process $NPV_t^{Exch}$ to solve the model so we keep in mind the latter formulation of section 3.2 for our optimal switching control model.
4. ANALYTICAL APPROACH TO SOLUTION OF OPTIMAL SWITCHING CONTROL PROBLEMS

We must know, we will know.
D. Hilbert

4.1 Introduction

The switching control problem formulated in the latter chapter (section 2.3), as already noted, can be seen as a special two regime switching control problem with two different (running) cost functions and different dynamics for each regime. This type of problem is well-known in the literature:

Bensoussan and Lions (1982) for first tackled the solution existence for general impulse control problems through variational inequalities approach; Brekke and Oksendal (1994), Pham and Vath (2007) and Duckworth, Zervos (2001) derived a verification theorem and closed form solution in some cases and applications, using the viscosity solution and the variational inequalities approach together with the dynamic programming and the smooth fit principle. In particular, Pham and Vath (2007) derived a closed solution form of viscosity type in two regime switching model in the case of one-dimensional diffusion process and a smooth profit function (with linear growth).

The analytical solution approach to this kind of stochastic problems has had big impulse especially after the development of the notion of viscosity solution introduced in the famous work of Crandall and Lions (1983).

More formally, one can show that given a general stochastic control problem in a markovian framework, namely an objective functional

\[ J(t, x, \alpha) = \mathbb{E} \left[ \int_t^T f(s, X_s, \alpha)ds + g(X_T) \right] \]

for \((t, x) \in [0, T] \times \mathbb{R}^n\), with \(f : [0, T] \times \mathbb{R}^n \times C \to \mathbb{R}\) and \(g : \mathbb{R}^n \to \mathbb{R}\) measurable and a system dynamic driven by a general controlled Ito diffusion,

\[ dX_t = b(X_t, \alpha)dt + \sigma(X_t, \alpha)dW_t \]
\[ X_0 = x \]

\[ x \]

4. Analytical approach to solution of optimal switching control problems

where the control $\alpha = (\alpha)_t$ is a $\mathbb{F}$—progressively measurable process belonging to the control set $\mathcal{C}$ and $X_t$ is an $n$ real valued SDE that satisfies the usual technical conditions on parameters $b(.)$ and $\sigma(.)$ in order to have a unique solution. Under suitable growth and boundedness condition on $f(.)$ and $g(.)$, and set $T$ the stopping time set, one can show that the value function of the problem, namely

$$v(t, x) = \sup_{\alpha \in \mathcal{C}} J(t, x, \alpha)$$

can be found by means of the well known dynamic programming principle, which can be formulated as follows (in the finite horizon case)

$$v(t, x) = \sup_{\alpha \in \mathcal{C}} \sup_{\theta \in T} \mathbb{E} \left[ \int_t^\theta f(s, X_s, \alpha)ds + v(\theta, X_\theta) \right] \quad (t, x) \in [0, T] \times \mathbb{R}^n$$

$$v(t, x) = \sup_{\alpha \in \mathcal{C}} \inf_{\theta \in T} \mathbb{E} \left[ \int_t^\theta f(s, X_s, \alpha)ds + v(\theta, X_\theta) \right] \quad (t, x) \in [0, T] \times \mathbb{R}^n.$$

From the definition of $v(t, x)$ given above and assuming its smoothness in order to apply Itô’s lemma, one can derive the infinitesimal version of the dynamic programming principle that describes the local behavior of the value function when we send the stopping time $\theta \to t$. The result is known as the Hamilton-Jacobi-Bellman equation (HJB), a non linear partial differential equation (typically of second order in stochastic control problems) that takes the following expression

$$-\frac{\partial v}{\partial t}(t, x) - \sup_{\alpha \in \mathcal{C}} \left[ A^\alpha v(t, x) + f(t, x, \alpha) \right] = 0 \quad \forall (t, x) \in [0, T] \times \mathbb{R}^n$$

$$v(T, x) = g(x) \quad \forall (t, x) \in \mathbb{R}^n$$

where $A^\alpha v(t, x)$ is the well known characteristic operator:

$$A^\alpha v(t, x) = b(x, \alpha)D_x v(t, x) + \frac{1}{2} \text{tr}(\sigma(x, \alpha)\sigma(x, \alpha)^\prime D_x^2 v(t, x))$$

$$= \sum_{i=1}^n b_i(x, \alpha) \frac{\partial v}{\partial x_i}(t, x) + \sum_{i,k=1}^n \gamma_{i,k}(x, \alpha) \frac{\partial^2 v}{\partial x_i \partial x_k}(t, x)$$

$\forall \omega \in \Omega \ t \in [0, T]$. This is usually represented also in more compact and general form as follows

$$-\frac{\partial v}{\partial t}(t, x) - H(t, x, D_x v(t, x), D_x^2 v(t, x)) = 0 \quad \forall (t, x) \in [0, T] \times \mathbb{R}^n$$

where for $(t, x, p, M) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{D}_n$

$$H(t, x, p, M) = \sup_{\alpha \in \mathcal{C}} \left[ b(x, \alpha)p + \frac{1}{2} \text{tr}(\sigma^\prime(x, \alpha)M) + f(t, x, \alpha) \right] = 0 \quad \forall (t, x) \in [0, T] \times \mathbb{R}^n$$

is the Hamiltonian operator of the associated control problem.

The central analytical issues here are the proof of the existence (and uniqueness) of a smooth solution for the HJB equation and the verification theorem namely the proof that the candidate solution coincides with the value function of the problem, that is (by Feynman-Kac formula)

$$v(t, x) = \mathbb{E} \left[ \int_t^T f(s, X^*_s, \alpha^*)ds + g(X^*_T) \right]$$
where $X_t^*$ solves the given SDE and $\alpha^*$ is a markovian optimal control. So, the main problem of the dynamic programming approach is the a priori assumption that the value function is at least a $C^{2,1}(x,t)$ class function, which is not necessarily true even in simple cases. In order to tackle this problem a more general definition of solution for the HJB equations, called *viscosity solution*, has been developed which requires only functions locally bounded and allows to study stochastic control problems rigorously and in great generality.

For a full introduction to viscosity solutions theory one can refer to the famous guide of Crandall, Ishii and Lions (1992) or to the Pham (2009) monograph. Here, in this chapter we focus on the analysis of our switching control problem from which we will highlight the following main issues:

a) the impossibility to assume a smooth value function $v$ for the switching nature of the problem;

b) the non linearity of the cost functions $F^i$ (that do not grow linearly and are not lipschitzian) and their recursive nature depending over time on the controls (the switching times $\tau_j$ and indicators $z_j$);

From point a) and the former insights on the viscosity solution approach, surely this is the most suited approach to the analytical solution of this type of switching control problem. In particular, we will derive in section 4.2 and 4.3 the HJB equation for our problem via dynamic programming principle, but for the reason underlined above (at point b) is hard to show here the existence and uniqueness of the solution in viscosity sense. Therefore we do not know any analytical results for similar problems in the literature.

In particular, our PDE problem with interconnected obstacles, being characterized by an optimal switching strategy that is path dependent in the sense that the decision of switch today must take in consideration even the future possible optimal switches, it can be assimilated to a generalized *american/bermudan game problem* for which no analytical solution are available. So, we delay the problem solution to the sequent chapter in which it will result easier to deal with it by means of general probabilistic tools as *Snell envelope* and *backward SDE* which have deep connections with the analytical problem formulation and its viscosity solution. Anyway, in section 4.4 and 4.5 we try to highlight the main problem issues and the the value function behavior, making some simplifying hypothesis and focussing on the analysis of the cost functions and of the switching strategy main drivers.

4.2 Dynamic programming principle and HJB equations in switching control problems: main issues

As already mentioned in the introduction, in this section and the following we derive the analytical formulation for our switching control problem showing, through the *dynamic programming principle*, its analytical representation given by a system of *Hamilton-Jacobi-Bellman PIDE*. In order to make this, given the markovian framework in which has been set our problem stated in equations (3.10-3.11), one can apply the dynamic programming principle that we recall here for switching control problems. The minimal technical conditions needed are the following.
4. Analytical approach to solution of optimal switching control problems

i) The running cost function \( F(y; u) \) for \((y; u) =: (t, x, \lambda, z_j, \tau_j) \in \mathbb{R}^+ \times \mathbb{R}^+ \times [0, T] \times C \) has to be \( \mathcal{F}_t \)-measurable and integrable, namely

\[
\mathbb{E}\left[ \int_t^T |F(Y_s; u)| ds \right] < \infty
\]

it needs to be continuous and to satisfy a linear growth condition (or, also lipschitzian), that is

\[
|F(y, u)| \leq C(1 + |y|) + l(z) \forall (y; u) \in \mathbb{R}^+ \times \mathbb{R}^+ \times [0, T] \times C
\]

for a positive constant \( C \) and a positive function \( l : C \to \mathbb{R}^+ \).

ii) The final reward \( G(y; u) \) has to be bounded and to satisfy a linear or quadratic growth condition.

iii) The diffusion dynamic coefficients has to satisfy the usual condition (typically Lipschitz) in order to ensure solution existence and uniqueness.

So, under this technical conditions we state below the dynamic programming principle for switching control problems. A rigorous proof using the more general holderian condition on \( F(\cdot) \) and one-dimensional diffusion, can be found in Tang and Yong (1993)

**Proposition 4.1 (Dynamic Programming Principle).** For any \((y; u) =: (t, x, \lambda, z_j, \tau_j) \in \mathbb{R}^+ \times \mathbb{R}^+ \times [0, T] \times C \) the value function \( V(\cdot) \) of the problem satisfies

\[
V(y; u) = \inf_{u \in \mathcal{C}_{ad}} \mathbb{E}^y \left[ \int_0^\theta e^{-r(t-s)} F(Y_t; u) dt + e^{-r\theta} V(Y_\theta, u_\theta) + \sum_{j \geq 1} e^{-r\tau_j} c_j 1_{\{\tau_j < T\}} \right] \tag{4.1}
\]

where \( \theta \) is any stopping time that belongs and possibly depending on the controls \( u \in \mathcal{C}_{ad} \).

This tool, together with smooth conditions on \( V(Y_\theta, u_\theta) \), are central to derive the nonlinear PDE, namely the HJB, which characterizes the analytical/variational approach for our problem.

**Hamilton-Jacobi-Bellman equations in our model.**

In order to get the analytical representation for our problem, let’s recall it setting the running cost functions as \( F_Z(\cdot) \) where the switching indicator \( Z := \{z; \zeta\} \), so that we get (from 3.10)

\[
J(T, Z)(y) = \inf_{u \in \mathcal{C}_{ad}} \mathbb{E}^y \left[ \int_0^T (\exp^{-r(T-s)})[F_Z(y, u)] ds + \sum_{j \geq 1} e^{-r\tau_j} c_j 1_{\{\tau_j < T\}} \right] \mathcal{F}_t \tag{4.2}
\]

where\(^2\), to better highlight the technical difficulties involved here, we recall the BCVA relation from (3.1) and making explicit the \( \text{NPV} \) term by considering a plain interest swap, we get the

\(^2\) Note that the final reward \( G(\cdot) \) being a terminal condition has been omitted to simplify notations.
4. Analytical approach to solution of optimal switching control problems

Following running cost functions for each switching regime:

\[ F_z(y, \alpha) := \left[ BCVA(t) - \delta \right]^2 = \left[ (CV A(t) - DV A(t)) - \delta \right]^2 \]

\[ = \left[ \left( \int_{t}^{T \wedge \tau_j} (1 - R_c) \left( \sum_{s = u}^{T} B_u \xi_u (X_u - k) \right)^+ \right) \lambda_u ds + \right. \]

\[ \left. - \int_{s}^{T \wedge \tau_j} (1 - R_c) \left( \sum_{s = u}^{T} B_u \xi_u (X_u - k) \right)^- \lambda_u ds \right] - \delta \right)^2 \quad (4.3) \]

\[ F_\zeta(y', \alpha') := \left[ \left( \int_{s}^{T \wedge \tau_j} R(s) [NPV(u)] ds - NPV(s) \right) - \delta \right]^2 \]

\[ = \left[ \left( \int_{s}^{T \wedge \tau_j} R(s) \sum_{s = u}^{T} B_u \xi_u [X_u - k] ds - \sum_{s = u}^{T} B_s \xi_s [X_s - k] \right) - \delta \right]^2 \quad (4.4) \]

for the switching times \( \{ s \leq \tau_j < T \wedge \tau \} \) with \( s \in [t, T \wedge \tau] \), \( \tau_j \in T \). Note that the default time \( \tau \) is hidden in \( \lambda \), and should be also made explicit in the upper limit of the integral, but we leave it to ease notation. In particular, we remark the setting of \( u = s \in U \subset [t, T] \), where \( U = \{ u, u + 1, \ldots, T \} \) is the subset of the payment schedule that is discrete in this particular case in which the underlying claim is an interest rate swap and we can have the payment time of the claim at time \( s = u \in U \) only after or at the same time of the switching time\(^3 \tau_j \). Therefore, we have the tenor \( \xi_s \), the discount factors \( B_u = \exp^{-\int_{s}^{T} r_s ds} \), \( F_\zeta(\cdot) : \mathbb{R}^+ \times \mathbb{R}^+ \times [0, T] \rightarrow \mathbb{R} \) that depend on both the state variable \((X, \lambda)\) plus a set of parameters \( \alpha \) while \( F_z(\cdot) : \mathbb{R} \times [0, T] \rightarrow \mathbb{R} \) depends only on \( X \) and given parameters. We remark that the expressions for BCVA and \( NPV \),\(^4 \) here referred to the simpler case of the interest rate swap, can be generalized to other contract payoffs (with dividends).

From (4.3-4.4), it is not difficult to see that even in the case of a plain swap the cost functions are very hard to handle. We leave the discussion on their properties for the section 4.4 and for the moment we take as valid the technical conditions in order to apply the proposition 4.1 and we make the central assumption on the smoothness of the value function \( V(\cdot) \) for our problem in order to derive the related \( HJB \) equation.

In particular, from (4.1) we assume the value function \( V(Y_{\theta}, u_{\theta}) \) - already calculated at the generic stopping/switching time \( \theta \) - smooth in order to apply Itô’s lemma and the mean value theorem (details can be found in Pham (2009)).

Then, recalling the definition given in (3.5-3.6) of the \textit{characteristic operator} \( A \phi(y) \) in which we work assuming the existence of a smooth function that become here our value function namely \( V(\cdot) = \phi(\cdot) \in C(t, x, \lambda)^{1,2,2} \), we get the following non linear second order partial differential equation namely the \textit{HJB equation} for our problem

\[ \inf_{u \in \{C_{ad}\}} \left\{ - \frac{\partial \phi}{\partial t} - A(z) \phi(y) + r \phi(y) - F_z \right\} = 0 \implies \]

\(^3\) Note that the upper extreme of the integral is \( T \) but the expected collateral costs should be calculated just until the next switching time, but this presume its knowledge at the moment of valuation and this would suggest a backward induction and recursive procedure for the solution.

\(^4\) In the unilateral case the definitions of the running functions are similar to the bilateral one.
4. Analytical approach to solution of optimal switching control problems

that become explicitly in the ”unilateral case” ⇒

0 = \inf_{u \in \{C_{ad}\}} \left\{ \frac{\partial \phi}{\partial t} + \eta(\mu - x) \frac{\partial \phi}{\partial x} + \kappa(\gamma - \lambda) \frac{\partial \phi}{\partial \lambda} + \frac{\sigma^2}{2} x \frac{\partial^2 \phi}{\partial x^2} + \frac{\nu^2}{2} \lambda \frac{\partial^2 \phi}{\partial \lambda^2} - r \phi(y) \right. \\
+ \left. \sum_j \int_{T \wedge \tau_j} (1 - R_c) \left[ \left( \sum_{s=u}^T B_s \xi_s(x - k) \right)^+ \right] \lambda ds \right\} \quad (4.5)

while in the ”bilateral case” we get ⇒

0 = \inf_{u \in \{C_{ad}\}} \left\{ \frac{\partial \phi}{\partial t} + \eta(\mu - x) \frac{\partial \phi}{\partial x} + \kappa(\gamma - \lambda) \frac{\partial \phi}{\partial \lambda} + \frac{\sigma^2}{2} x \frac{\partial^2 \phi}{\partial x^2} + \frac{\nu^2}{2} \lambda \frac{\partial^2 \phi}{\partial \lambda^2} - r \phi(y) \right. \\
+ \left. \sum_j \left[ \left( \int_{T \wedge \tau_j} (1 - R_c) \left[ \left( \sum_{s=u}^T B_s \xi_s(x - k) \right)^+ \right] \lambda ds + \right. \right. \\
- \left. \left. \int_{T \wedge \tau_j} (1 - R_c) \left[ \left( \sum_{s=u}^T B_s \xi_s(x - k) \right)^- \right] \lambda ds \right) \right] - \delta \right\} \quad (4.6)

After the switching to full collateral, the relevant dynamics and the characteristic operator change together with the relevant cost function, so that the HJB equation to solve becomes

\[ \inf_{u \in \{C_{ad}\}} \left\{ - \frac{\partial \phi}{\partial t} - A^{(\xi)} \phi(y) + r \phi(y) - F_t \right\} = 0 \Rightarrow \]

which is explicitly (here ”unilateral case” = ”bilateral case”) ⇒

0 = \inf_{u \in \{C_{ad}\}} \left\{ \frac{\partial \phi}{\partial t} + \eta(\mu - x) \frac{\partial \phi}{\partial x} + \frac{\sigma^2}{2} x \frac{\partial^2 \phi}{\partial x^2} - r \phi(y) \right. \\
+ \left. \sum_j \left[ \left( \int_{T \wedge \tau_j} R(s) \sum_{u}^T B_u \xi_u(x - k) ds - \sum_{\tau_{j}}^T B_{\tau_j} \xi_{\tau_j}(x - k) \right) - \delta \right]^2 \right\} \quad (4.7)

This system of PIDE (4.6-4.7) has to be coupled with the switching boundary condition in order to complete the analytical representation of our problem, given that one has to consider also the instantaneous switching cost function in the picture. We deal with this issue now and we refer to section 4.3 for the complete formulation of the problem.

Switching costs definition.

As regards the switching costs defined in the instantaneous cost function that are the instantaneous cost associated to switch from zero to full collateralization and viceversa, let us set \( z^c_{\text{zero}} := c^z \) and \( z^c_{\text{full}} := c^\xi \). We also assume them \( F_t \)-predictable, typically deterministic or fixed/constant in \( t \) and of course strictly smaller than the corresponding running cost function \( F(.) > K(.) \). In
4. Analytical approach to solution of optimal switching control problems

particular, one can distinguish in the analysis three cases:

\[ c^z = c^\zeta \]
\[ c^z < c^\zeta \]
\[ c^z > c^\zeta \]

It is clear that the level of the two instantaneous switching costs determines the convenience and the frequency of switching times: in our model, being costs we set them as positive or null and we expect that the different relative level between the switching costs in the two regimes \( c^z \) and \( c^\zeta \) is relevant for the number of switching times and in general for the determination of the optimal switching strategy. In particular we remark that the condition on the process predictability (namely the instantaneous costs are no stochastic) is necessary to ensure the existence of a solution (in viscosity sense) to this type of problems\(^5\).

Switching and continuation region definition.

For what concerns the value function, following Pham (2009), we have that because of the nature of our problem with two switching regime, cost function and dynamics, we can define two value functions \( V^z(y) \) e \( V^\zeta(y) \) (possibly discontinuous) associated to zero and full collateralization. By doing this, is intuitive the definition of the switching region for this problem as

\[ S_z = \left\{ (x, \lambda) > 0 : V^z(y) = V^\zeta(y) + c^z \right\} \]

namely the closed set that indicates when the switching time is optimal that is when the value functions related to the zero collateral equals the one associated to full collateral plus the relative switching costs. Similarly is defined \( S_\zeta \) the switching region when the initial state is \( \zeta \).

By the other side, the continuation region is defined in the same manner but with the strict inequality, that is

\[ C_z = \left\{ (x, \lambda) > 0 : V^z(y) < V^\zeta(y) + c^z \right\} \]

and similarly for \( C_\zeta \).

4.3 Viscosity solution and variational inequalities formulation.

Now, given the HJB equations (4.6-4.7) relative to the two switching regimes and the boundary conditions represented by the above switching/continuation region, by dynamic programming principle and the viscosity solution definition (and the related theorems) one should prove the existence (and uniqueness) of the value functions \( (V^z, V^\zeta) \) for our problem in terms of viscosity

\(^5\) See also the work of Djehiche et al. (2008)
4. Analytical approach to solution of optimal switching control problems

solutions to the following nonlinear system of variational inequalities

\[
\min_{u \in \{C_{ad}\}} \left\{ -\frac{\partial V^z}{\partial t} - A^{(z)}V^z + rV^z - F^z, \ V^z - (V^z + c^z) \right\} = 0 \quad x, \lambda \in D, \ z, \zeta \in Z \quad (4.8)
\]

\[
\min_{u \in \{C_{ad}\}} \left\{ -\frac{\partial V^\zeta}{\partial t} - A^{(\zeta)}V^\zeta + rV^\zeta - F^\zeta, \ V^\zeta - (V^\zeta + c^\zeta) \right\} = 0 \quad x, \lambda \in D, \ z, \zeta \in Z \quad (4.9)
\]

with \(D\) that indicates the domain of the problem and initial conditions, given that \(F_Z(0) = 0, V^z(0) = (-c^z), V^\zeta(0) = (-c^\zeta)\)

and terminal conditions that are

\[
V^z(T, x; u) = 0
\]

\[
V^\zeta(T, x; u) = (-NPV(T)\zeta_T^2).
\]

For a rigorous definition of viscosity solution we refer to definition 5.5 or in particular at the already mentioned work of Pham (2009). As already said, similar switching controls problem has been tackled analytically: under typical condition on the dynamic, assumed to be an Ito diffusion adapted to the Brownian filtration, on the payoff functions assumed linear or Lipschitz and switching costs deterministic, a verification theorem for a general switching control problem (with multiple regimes) has been derived and we reformulate it here for our problem.

**Proposition 4.2 (Verification theorem).** Let \(A^{u,Y}\) denote the infinitesimal generator of the markovian diffusion process \((Y_t)\). Assume the existence of a function \(\phi(t, y, u)\) such that for \(Z = \{z, \zeta\} \in u\) and

\[
\mathcal{D} := \bigcup_{Z} \left\{ (t, y; u) : \phi(t, y; u, z) = \phi(t, y; u, \zeta) + c^\zeta \right\},
\]

\(\phi(.)\) is of class \(C^{1,2,2}([0, T] \times \mathbb{R}^2 \times (\mathbb{R}^+ \times [0, 1]) \cap C^{1,1,1}(\mathcal{D})\) and satisfies the following quasi-variational inequalities (or system of PDEs with obstacles) for each \(z, \zeta \in Z:\)

\[
\begin{align*}
\min \left\{ \phi^z(t, y; u) - (\phi^\zeta(t, y; u) + c^\zeta), \ -\partial_t \phi(t, y; u) - A^{Y,u} \phi(t, y; u) + r \phi(t, y; u) - F^Z(t, y; u) \right\} &= 0, \\
\phi(T, y; u) &= G^Z(T, y; u).
\end{align*}
\]

Then \(\phi(.) = V(.)\) is the optimal value function for our problem.

Is clear that the existence problem of the solution is still there because the verification theorem gives just the sufficient conditions to check once a candidate function is guessed. So, supposed the

\(^6\) See, for example, Brekke and Oksendal (1994) for details on the proof.
Analytical approach to solution of optimal switching control problems

existence of a smooth function that satisfies the above mentioned properties, then this is the value function for the problem. As regards the existence and uniqueness of viscosity solution to this type of stochastic control problem (also in multiple regimes case) some results are known under similar assumptions and that make use of the smooth/continuous-fit principle.

This analytic tool widely used to solve stopping times problems, allows under assumptions on the smoothness of reward function and on regularity of the diffusion process that describe the dynamic, to imply also the smoothness of the value function, which has to be $C^2$ on the continuation region and continuously differentiable ($C^1$) on the switching boundaries (see Pham (2009) or Peskir, Shyraiev (2005).

Unfortunately, the assumptions on which are derived these results are not verified in our problem. In fact, recalling some of the fundamental hypothesis needed to apply the viscosity solution approach, we have

**Hp 1)** The drift and diffusion coefficients for our dynamics that we can express, in general, as $b_Z(.) := b(., Z)$ and $\sigma_Z(.) := \sigma(., Z)$ with $Z = \{z, \zeta\}$ is the index associated to the two regimes and $b, \sigma : D \times Z \to \mathbb{R}^7$, satisfy the Lipschitz condition for some positive constant $C$, that is

$$|b(x, Z) - b(y, Z)| + |\sigma(x, Z) - \sigma(y, Z)| \leq C|x - y| \quad \forall x, y \in D, Z \in Z$$

**Hp 2)** The diffusion coefficient function is assumed to be uniformly elliptic that is

$$\sigma(x, Z) > 0, \quad \forall x \in \text{int}(D), Z \in Z.$$  

**Hp 3)** The running cost functions $F(.,.) : \mathbb{R}^+ \times \mathbb{R}^+ \times [0, T] \to \mathbb{R}$ must be continuous and with linear growth or lipschitzian, that is

$$|F(x, Z) - F(y, Z)| \leq C|x - y| \quad \forall x, y \in D, Z \in Z.$$  

These conditions are necessary to show that the value functions are solutions of the above system of variational inequalities, but - except for the Hp1) and Hp2) that are true in our case because of the assumption done on the dynamic SDEs which have strong and unique solution - the third one (Hp3) is not true, because of our cost functions are recursive integral of non linear (convex) functions so that the HJB equations are actually partial-integro differential equations highly non linear which are hard to solve. Indeed, the existence and uniqueness theorems for this kind of problem are not known in literature and so they should be build. But being a very tough problem to solve analytically, a different probabilistic approach based on the use of Snell Envelope and on results from the theory of backward stochastic differential equation can be employed and we deal with it in the next chapter.

---

7 Where $D$ is the domain in which lie our stochastic problem
4. Analytical approach to solution of optimal switching control problems

4.4 Analysis and properties of the cost functions.

Before turning to the probabilistic view, let us try to highlight some of the properties of the problem that we have just set up trying some heuristic reasoning about the characterization of the optimal strategy and of the value function of our problem. In particular, we proceed by analyzing separately the single running cost functions - as set in equations 4.3 and 4.4 - in the two case of \( \delta = 0 \) and \( \delta > 0 \) and then considering them together.

1) Properties of \( \{ F_z(y, \alpha) \text{ with } \delta = 0 \} \). So we start with the cost function related to the zero collateral regime which is for \( \delta = 0 \)

\[
F_z(y, \alpha) = \left[ BC VA - 0 \right]^2 = \left[ (CV A - DVA) \right]^2
\]

where its single components are real \( \mathbb{F}_t \)-measurable càdlàg process adapted to the filtration generated by the Brownian motion vector \( (W_t^X, W_t^\lambda)_{t \in [0,T]} \) and depend on time \( t \), on the state variables \( (X, \lambda) \) and on a set of given parameters \( \alpha \). More specifically, fixing \( t \) and for given realizations of the state vector \( (X, \lambda) \in [0, T \times \mathbb{R}^+ \times \mathbb{R}^+] \) we can see the CVA = \( \psi^{cva}(t, x, \lambda) \) and DVA = \( \psi^{dva}(t, x, \lambda) \) as non linear functions that present the following properties:

1. \( \psi^{cva}(t, x, \lambda) \) is a non linear function because it takes the integral/summation over time of the expected positive exposure \( EPE_t \) in which we have a max operator , discounted by default intensities \( \lambda > 0 \). So because of the integral takes only positive (or at max zero) values the function is always positive \( \psi^{cva}(.) \in [0, T] \times \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+ \). Similar properties are presented by \( \psi^{dva}(.) \) for which we have the min operator and so the functions takes only negative (or zero) values that is \( \psi^{dva}(.) \in [0, T] \times \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow (-\infty, 0) \).

2. Because of the non linear property highlighted above for CVA and DVA that are càdlàg processes (for \( t \) varying) driven by CIR-type dynamics (see 3.1 and 3.2) for \( (X, \lambda) \) which are continuous markovian diffusion processes part of the squared Bessel processes family, we can deduce the continuity on the respective domains for \( \psi^{cva}(.) \) and \( \psi^{dva}(.) \) for \( t \) varying and for given realizations of the state vector \( (x, \lambda) \).

3. From the last property, being \( x \) and \( \lambda \) realizations of CIR processes which have as transition density a non-central-\( \chi^2 \) and stationary distribution with finite moments of any order\(^9\), then we have that

\[
E[ \sup_{0 \leq t \leq T} |\psi^{cva}(.)|]^p < \infty, \quad (p \geq 1)
\]

\[
E[ \sup_{0 \leq t \leq T} |\psi^{dva}(.)|]^p < \infty, \quad (p \geq 1)
\]

So we can briefly say that \( \psi(\cdot) \in \mathcal{M}_p \) that is they belong to the set of \( p \) integrable and measurable functions (or process for \( t \) varying). Therefore, we can deduce by these

\(^8\) In fact, the whole objective functional is separable in the single running cost functions and the related instantaneous switching costs.

\(^9\) Given that the parameter choice in the dynamic ensure non explosion or absorption in zero of the process
properties that they satisfy almost \textit{quadratic or polynomial growth condition}

\[ |\psi^i(t,y)| \leq C(1 + |y|^\beta), \quad \beta > 1 \quad (t, y) \in [0, T] \times \mathbb{R}^+ \]

4. Fixed \((X, \lambda)\), for \(t \to T\) the CVA and DVA integrals over \(t\) tends to take less terms so these functions are decreasing for \(t\) converging to maturity \(T\).

5. For fixed \(t\) and \(\lambda\), given two realizations \(X(\omega,t)^{10}\) and \(X(\omega',t)\) we have that

\[
\text{if } X(\omega, t) > X(\omega', t) \implies CVA(X(\omega, t)) > CVA(X(\omega', t)),
\]

\[
\text{if } X(\omega, t) > X(\omega', t) \implies DVA(X(\omega, t)) < DVA(X(\omega', t)),
\]

so the CVA is increasing respect to \(X\) (when it contributes to get greater positive exposures). The converse is true for the DVA.

For what concerns the default intensities we have that for greater realizations of \(\lambda\) the CVA and DVA integrals take terms with increasing probability mass so they are both increasing, namely

\[
\text{if } \lambda(\omega, t) > \lambda(\omega', t) \implies CVA(X(\omega, t)) > CVA(X(\omega', t)),
\]

\[
\text{if } \lambda(\omega, t) > \lambda(\omega', t) \implies DVA(X(\omega, t)) > DVA(X(\omega', t)),
\]

Now, given the properties of the single terms of the cost function, we can say that \(\{F_z(y, \alpha)\}\) being the difference of two càdlàg which are in particular continuous processes (squared) is still continuous. Therefore the function is nonlinear sharing the same properties on the moments \(3)\), so that \(F_z(.)\) respects the polynomial or quadratic growth condition. Therefore it satisfy the property \(4)\) respect to time , while it doesn’t matter the sign in \(5)\) and \(6)\) because it’s always positive (or non negative). In fact, the square implies that the objective is to minimize the BCVA significatively different from zero however it takes positive or negative values. Resuming, we have the following proposition.

**Proposition 4.3.** Given realizations of the state variable vector \((t, x, \lambda)\), the running cost function \(\{F_z(y, \alpha)\}\) results to be continuous, non linear, it satisfies the polynomial growth condition and belongs to the set \(\mathcal{M}^p\).

**Proof.** Here everything is quite obvious from the properties stated above. So, we just give a proof of the polynomial growth condition for our running cost function which is a very important condition that we need later for important results related to reflected BSDE’s. The proof clearly depends on the payoff function that one works with for which a typical condition is the square integrability \(F \in \mathcal{M}^2\). So we recall the definition of \(F_z(.)\) in the special case that we’ve considered of a \textit{defaultable interest rate swap} for which we have

\[\text{Actually, when we think to the } X \text{ realization (and } \lambda \text{ too) we have in mind a whole term structure path because this is what is needed to calculate the CVA/DVA.}\]
4. Analytical approach to solution of optimal switching control problems

assumed a unitary notional $N$ (i.e. constant and homogeneous $\forall t \in [0, T]$). Thanks to this assumption, we can derive easily a possible upper bound for the price process of the claim, that is formally

$$UB^{F_\iota} := \int_s^T \sum_{u=s}^T N ds$$

where we have set $M := \sum_{u=s}^T N$ (constant).

This is enough to allow the application of the definition of polynomial growth condition, that is (for $C$ constant, $\beta \geq 1$, $(t, y) \in [0, T] \times \mathbb{R}^+ \times \mathbb{R}^+$)

$$|\Psi(t, y)| \leq C(1 + |y|^\beta) \implies$$

$$|F_\iota(y, \alpha)| \leq C(1 + |y|^\beta) \implies$$

$$|CVA - DVA|^2 \leq C(1 + |y|^\beta) \implies$$

and substituting explicitly (and setting $\gamma := (1 - R_c)$)

$$\left| \int_s^T \gamma [(\sum_{s=u}^T B_s \xi_s(x_s - k))^+|\tau = s] \lambda_s ds - \int_s^T \gamma [(\sum_{s=u}^T B_s \xi_s(x_s - k))^\lambda] \lambda_s ds \right|^2 \leq$$

$$\leq C(1 + |y|^\beta)$$

$$\leq C(1 + |(x, \lambda)|^\beta)$$

where the polynomial growth condition results to be satisfied by substituting the upper bound, namely imposing $C = M(T - s)$ and for $\beta \geq 2$. The proof for $F_\iota(.)$ it can be shown following the same lines $\diamond$.

2) Properties of $\{F_\iota(y, \alpha), \text{with } \delta = 0\}$. For what concerns the running cost function related to the full collateral regime, we recall it here

$$F_\iota(y', \alpha') = \left[ \left( \int_s^{T \land \tau_{j+1}} R(s)[NPV(u)] ds - NPV(s) \right) - 0 \right]^2 = \left[ ECC - NPV(s) \right]^2$$

where for convenience we have defined $ECC$ the first term of the function namely the expected collateral cost process: this is an $\mathbb{F}_s$-measurable and càdlàg process adapted to the filtration of the single Brownian Motion $(W_t^X)_{t \in [0, T]}$, because only the process of $X$ is relevant here. It takes the integral of the realizations of the price process (which is linear in this case and so with polynomial growth) discounted by a deterministic factor and, by construction, the whole term takes only non negative values so $ECC(.) : [0, T] \times \mathbb{R}^+ \to \mathbb{R}^+$. So,
4. Analytical approach to solution of optimal switching control problems

it is continuous for similar reason to property 2), decreasing for \( s \to T \), while respect to \( X \)
realizations, their sign is not relevant.

As regards \( NPV(s) \) it represents the price process calculated in \( s = \tau_j \) so it’s an \( \mathbb{F}_{\tau_j} \)-
measurable process as the latter one which represents the collateral process (because we
have full collateralization) calculated at a switching time \( \tau_j \) (for \( j \in \{1, \ldots, M\} \)). This term
is clearly decreasing in value as \( t \to T \) and can take positive or negative value so that it’s
image is the whole real line.

It is clear that \( F^\zeta(.) \) share the same properties of the above components, and because we
kept the same structure of the other cost function, the difference between the two terms is
always positive because of the square and the objective would be to minimize the expected
collateral costs as soon as they are different from zero (remember we set \( \delta = 0 \)). Concluding,
we resume the property for \( F^\zeta(y, \alpha) \).

**Proposition 4.4.** Given a realization of the state variable vector \((t, x, \lambda)\), the running cost
function \( \{F^\zeta(y, \alpha)\} \) results to be continuous, quadratic, it satisfies the polynomial growth
condition and belongs to the set \( \mathcal{M}^p \).

**Proof.** The proof follows easily from the cost functions properties highlighted above. For
the polynomial condition reasonings similar to proposition 4.3 can be applied.

4.5 Value function behavior and switching strategy conditions.

Now, given the properties of our running cost function we try some reasoning on the optimal
switching strategy and solution useful for the next section. We begin by remarking the fact that
our functions \( F^\zeta(.) \) and \( F^\zeta(.) \) take paths forward in time so the optimal switching strategy depends
on future switches and this suggests a backward recursion to optimally determines the switching
times over the paths and the value function of the problem using the dynamic programming
principle. Here for clarity, we consider a given realization of the state vector \((t, x, \lambda)\)
and take two single periods \( t_1, t_2 \in [0, T] \) and we set the system in the zero collateral regime in \( t_2 \), namely
\( z_{t_2} = 1 \) and for convenience we do not make explicit the conditional expectation of the running
costs that come from wait and not switch - the so called **continuation value** - that is intended as
hidden inside the cost function itself. This term will be made explicit and introduced in the next
chapter when we define the **Snell envelope** tool to approach the model solution.

So let us classify the following outcome in relation to our cost functions:

\[
\text{Case } \{F^\zeta(y, \alpha), \delta = 0\} \Rightarrow A) \{CVA > DVA\} \\
\Rightarrow B) \{CVA < DVA\} \\
\Rightarrow C) \{CVA \equiv DVA\},
\]
4. Analytical approach to solution of optimal switching control problems

Case $\{F_z(y', \alpha'), \delta = 0\} \Rightarrow D \{ECC > NPV(\tau_j)\}$

$\Rightarrow E \{ECC < NPV(\tau_j)\}$

$\Rightarrow F \{ECC \approx NPV(\tau_j)\}$.

Whatever it is the initial condition it is clear that the switching is more likely if one of the cases $\{A, B, D, E\}$ is verified, because the running costs are higher and the objective is to minimize them. Actually, because of the cost functions are both squared, for the decision making the minus sign become positive so what is relevant for the optimality of the switching is the consideration of the difference or the distance between the running cost functions (plus the instantaneous switching costs) between one decision time and the last switching time where an optimal decision has been taken for the Bellman optimality principle. Indeed, if in our case we have in $t_2 \{z_{t_2} = 1\}$ that is zero collateral, we can assume that in $t_2$ the outcome $\{C\}$ for $F_z(\_)$ and $\{D\}$ or $\{E\}$ for $F_\zeta(\_)$ has been verified. Now, in $t_1$ the agent will optimally switch to full collateral $\{\zeta_{t_2} = 1\}$ if she verifies for example the outcome $\{A\}$ or $\{B\}$ for $F_z(\_)$ and $\{F\}$ for $F_\zeta(\_)$, while she won’t switch otherwise.

Obviously here we have considered only some of the possible case for given realizations of the state variable vector. Therefore, as already remarked, we are implicitly considering that the behavior of the solution namely the value functions, that takes the expectation over the paths, strictly follows the behavior of the cost functions.

Anyway, we can conjecture the behavior of the switching strategy considering the simple distance/difference of the cost functions - that become the value functions in the dynamic programming taking the expected infimum over the paths - respect to the instantaneous switching costs. In particular we can state the following propositions.

**Proposition 4.5 (Switching strategy).** Assume as valid the above definition and properties of the running costs $F_z(y, \alpha)$ and $F_\zeta(y', \alpha')$. Define the related instantaneous switching costs $c_z$ and $c_\zeta$ that are assumed to be strictly lower than the related $F(\_)$.

If for a given $t \in [0, T]$ we have set $\{z = 1\}$, as soon as for an other $t' \in [0, T]/\{t\}$ is realized the following condition

\[
i f \ [F_z(t', y, \alpha) - F_\zeta(t', y', \alpha')] \geq c_\zeta \ t' \in [0, T]/\{t\} \implies (4.10)
\]

then it becomes optimal to switch to $\{\zeta = 1\}$ (or $\{z = 0\}$). Otherwise if it’s valid the following

\[
i f \ [F_z(t', y, \alpha) - F_\zeta(t', y', \alpha')] < c_\zeta \ t' \in [0, T]/\{t\} \implies (4.11)
\]

then it’s not optimal to switch to full collateral. Similar conditions are valid when the initial condition is $\{\zeta = 1\}$.

**Proof.** We give just an heuristic proof of the condition (4.10), the remaining can be proved by similar arguments. The proof is quite immediate by recalling the objective functional and

---

11 We recall that this is the discrete time counterpart of the dynamic programming principle.

12 A similar analysis of this type with the difference in the profit functions has been carried in the work on multiple regime switching of Pham and others (2008)
the definition of switching $S$ and continuation region $C$ stated above. Take for simplicity two consecutive discrete times $t_1, t_2 \in [0, T]$ and as condition in $t = t_1 \{z = 1\}$ that implies the following condition on the cost functions

$$F_z(t_1, y, \alpha) < F_\zeta(t_1, y', \alpha') + c_\zeta, \quad t_1 \in [0, T]$$

Then in $t_2$, if the counterparty verifies the opposite condition

$$F_z(t_2, y, \alpha) \geq F_\zeta(t_2, y', \alpha') + c_\zeta, \quad t_2 \in [0, T]$$

then assuming to verify it over the paths and recalling the functional objective one has

$$\min \{F_z(t_2, y, \alpha), F_\zeta(t_2, y', \alpha') + c_\zeta\}$$

and taking the expectation over the paths one can substitute the value function, namely

$$\min \{V^z(y), V^\zeta(y) + c_\zeta\}$$

from which one can recover the definitions of switching and continuation region and hence the optimality of the switching strategy. Also the converse is true $\diamondsuit$.

From the previous proposition we can derive the followings statements that define the conditions under which we get banal switching strategies for our problem that we firstly define here.

**Definition 4.6 (Banal switching strategies).** Define a banal switching strategy as the one that reduce the optimal switching control to an optimal stopping one that is only one switch is optimal and the one for which the switch is never optimal. Formally we have that

$$"\text{Single optimal switch}" \iff C_{ad} := \{Z, T\} = \{1, \tau_1\}$$

$$"\text{Switch never optimal}" \iff C_{ad} := \{Z, T\} = \{0, \infty\}$$

**Proposition 4.7 (Banal switching strategies conditions).** Assume as valid the same hypothesis of Proposition 4.1. Going forward in time, if for initial condition in $t = 0$ we have $\{z_0 = 1\}$, to get banal switching strategies the followings conditions has to be satisfied:

$$\\text{if } \{F_z(t, y, \alpha) < F_\zeta(t, y', \alpha') + c_\zeta\} \forall t \in [0, T] \iff C_{ad} := \{Z, T\} = \{0, \infty\}$$

$$\\text{if } \{F_z(t, y, \alpha) > F_\zeta(t, y', \alpha') + c_\zeta\} \forall t \in [0, T] \iff C_{ad} := \{Z, T\} = \{1, \tau_1\}$$

The conditions in the case of $\{\zeta_0 = 1\}$ as initial condition are recovered similarly.

**Proof.** Easy by the Definition 4.6 and using similar arguments of the Proposition 4.5.

**Remarks 4.1.** Note that going backward, it must be considered the terminal condition or reward. In particular, depending on the sign of the $NPV(T)$ (that one derive by the simulation
of the paths forward) one can derive the optimal terminal switching condition:
- if $NPV(T) > 0$, because of minus in the condition it is less than zero (the other terminal condition in the zero collateral case) so it’s optimal to start the backward recursion with switching set to $\{\zeta = 1\}$;
- otherwise if $NPV(T) < 0$, then $\{z = 1\}$ would be the optimal choice.

These conditions relates to the existence of non banal or nonsense solution for our general control problem. In fact it would be relevant from a theoretical point of view to define the technical conditions and assumptions under which the banal solutions are excluded or, equivalently, that the solution exists and it is characterized by more than a single switch. But this takes us back right to the problem of proving the existence and uniqueness of solution already highlighted before.

At intuitive level, the existence of a non banal solution should be ensured if our cost functions or equivalently the relative value functions are not too distant path by path on the whole domain and tends to cross each other for different realizations and times. Of course this should depends on the set of parameters and on the payoff of the underlying contract.

In the next sections we will try to solve this problem through a different probabilistic approach proving under some general assumptions the existence and uniqueness of a non banal solution for our switching control problem.

We will also analyze numerically the problem in order to verify the predicted behavior of the solution and to determine the parameters values for which the solution exists and it is of non banal type.

In particular, we can be more explicit on the behavior of the solution marking its dependence on parameters. In fact, under the symmetry assumption $\lambda_A = \lambda_B$ and also $R^A_c = R^B_c$, and the orthogonality hypothesis $X \perp \lambda$ and assuming to use the same paths to estimate the expected exposures $NPV^+$ and $NPV^-$, we can show that these contribute to the CVA and DVA terms (respectively) that is, by definition 2.2.3 and 2.2.4, the BCVA. So, given that $NPV = NPV^+ + NPV^-$, this would give the same path by path contributions to BCVA so that the cost function defined in equation 4.3 can be simplified as follows

$$F_z(y, \alpha) = \left[ \left( \int_s^{T \wedge \tau_j} (1 - R_c) \left( \sum_{s=0}^{T} B_u \xi_u (X_u - k) \right) \lambda_u ds \right)^2 \right]$$ (4.16)

where $\delta = 0$, and recalling the other running cost function defined in (4.4), in which is possible to move $NPV(\tau_j)$ between the instantaneous switching costs, getting

$$F_z(y, \alpha) = \left[ \left( \int_s^{T \wedge \tau_j} R_{Fund}(s) \left( \sum_{s=0}^{T} B_u \xi_u (X_u - k) \right) ds \right)^2 \right]$$ (4.17)

with now $c_{\xi \tau j} = c_{\xi \tau j}^{full} + NPV(\tau_j)$.

From these relations, we can compare the two switching regime cost functions obtaining some conditions on the parameters that would give us useful indications on the solution behavior in particular on the existence of banal solutions and the optimal switching strategy. These results
can be stated in the following proposition\textsuperscript{13}.

**Proposition 4.8. (Optimal switching conditions):** Under the following assumptions:

a) symmetry hypothesis \( \lambda_A = \lambda_B, R^A_c = R^B_c \) (and \( \delta^A = \delta^B = 0 \))

b) orthogonality condition \( X \perp \lambda \)

and taking as valid the hypothesis and the conditions of Proposition 4.5, given as initial regime condition \( \{z = 1\} \), our optimal switching condition can be defined on average over the paths of the cost functions \( F_z(\cdot) \) as follows

\[
\min \left\{ F_z(y, \alpha), F_\zeta(y', \alpha') + c_\zeta \right\}
\] (4.18)

then the following condition depending on problem parameters must hold (on average)

\[
\min \left\{ (1 - R_c)\lambda(t), R_{Fund}(t) + c_\zeta \right\}
\] (4.19)

and similarly for \( \zeta = 1 \).

**Proof.** By simply comparing the cost function equations 4.16 e 4.17, we see that under the HP a) and b) the condition

\[
\left[ \left( \int_{s}^{T \wedge \tau_j} (1 - R_c) \left( \sum_{u \in U} B_u \xi_u (X_u - k) \right) \lambda_u ds \right)^2 \right] \lesssim \left[ \left( \int_{s}^{T \wedge \tau_j} R_{Fund} (s) \left( \sum_{u \in U} B_u \xi_u (X_u - k) \right) ds \right)^2 \right]
\]

being the integral terms - the expected NPV - equal (\( \mathbb{Q} \) a.s) on average over the paths, the inequality will be determined by

\[
[(1 - R_c)\lambda(t)]^2 \lesssim [R_{Fund}(s)]^2
\]

that impact on the processes paths, and by the instantaneous switching cost \( c_z \) (or \( c_\zeta \)). So introducing the switching costs in the inequalities it is easy to verify the validity of Proposition 4.5, that can be expressed also as in equation 4.18 and hence the 4.19 is immediately derived ∙

**Notes on** \( F(\cdot) \) **with** \( \delta > 0 \). For what concerns the more general cost functions in which the threshold \( \delta \) is different from zero, similar considerations and conditions are valid. In particular, one can assume the same \( \delta \) for both the running cost functions \( F_z(\cdot) \) and \( F_\zeta(\cdot) \) or even different, for example in \( F_z(\cdot) \) the \( \delta_z \) could be the minimum level of CVA allowed while in \( F_\zeta(\cdot) \), \( \delta_\zeta \) could be a minimum transfer amount (MTA) above which the collateral can be called that is the switching becomes likely.

\textsuperscript{13} Its relevance will be more clear in the numerical section too.
4. Analytical approach to solution of optimal switching control problems

Fig. 4.1: Example of possible non banal switching strategy for given $F_z(.)$, $F_\zeta(.)$ with $\{c_z = c_\zeta = 0\}$ and $\{z_0 = 1\}$.

Considering the simpler case in which $\{\delta_z = \delta_\zeta > 0\}$, the properties of the running cost functions remain the same, what it’s going to change are just the possible outcomes and the running cost functions but the conditions in the Propositions 4.1 and 4.2 are still valid. More specifically, we have the following running cost functions

$$F_z(y, \alpha) = \left[BCVA - \delta\right]^2 = \left[(CVA - DVA) - \delta\right]^2$$

$$F_\zeta(y', \alpha') = \left[(R_{fund}(\tau_j) \int_{\tau_j}^{T} [NPV(s)] ds - NPV(\tau_j)) - \delta\right]^2 = \left[(ECC - NPV(\tau_j)) - \delta\right]^2$$

with possible outcomes given by

**Case** $\{F_z(y, \alpha), \delta > 0\}$ \Rightarrow $A) \{BCVA > \delta\}$

\Rightarrow $B) \{BCVA < \delta\}$

\Rightarrow $C) \{BCVA \equiv \delta\},$

**Case** $\{F_\zeta(y', \alpha'), \delta > 0\}$ \Rightarrow $D) \{(ECC - NPV(\tau_j)) < \delta\}$

\Rightarrow $E) \{(ECC - NPV(\tau_j)) < \delta\}$

\Rightarrow $F) \{(ECC - NPV(\tau_j)) \equiv \delta\}.$

Similar observations about the optimal switching strategy are valid here because the outcomes and structure of the functions are not changed. What is relevant to note is that the setting of the threshold $\delta$ is central to establish the existence of the optimal switching strategy. In the present case, in fact, the setting of a given threshold which results to be on average over the path lower/higher respect to CVA more than the collateral costs or viceversa, this would imply (intuitively)
a grater probability to get *banal switching strategies*.

In general, the exogenous definition of the threshold for our cost functions can affect the existence of the solution. By the other side, we expect that imposing different thresholds $\delta_z \neq \delta_z$ can help to solve the existence problem of non banal solution but at cost of a more complicated behavior of the switching strategy. A further possible generalization would make the thresholds endogenous as a control of the problem or setting it as depending on time and some other parameter $\delta =: n(t, a)$ or even stochastic.
5. SNELL ENVELOPE APPROACH, RBSDE AND CONNECTIONS WITH QVI

If one has really technically penetrated a subject, things that previously seemed in complete contrast, might be purely mathematical transformations of each other.

J. von Neumann

5.1 Introduction

In this chapter we start to deal with the probabilistic approach to the solution of the optimal switching control based on deep results from the optimal stopping and backward SDE (BSDE) theory which have connections with the (viscosity) solutions of nonlinear PDEs with obstacles namely the system of (quasi)variational inequalities like the one derived in (4.8-4.9).

This elegant approach has its roots in the seminal works of El Karoui et al. (1995) which deals with reflected BSDEs solution and Cvitanic and Karatzas (1996) that generalize it to the doubly reflected case and its application to Dynkin games. Following El Karoui et al. (1995), let us recall that the reflected BSDEs are a class of BSDE whose solution \( Y \) is constrained to stay above a given process, called obstacle which is represented in our case by the switching condition. This is possible introducing an increasing process which pushes the solution upwards, above the obstacle.

More formally, let \( W = (W_t)_{0 \leq t \leq T} \) be a standard \( d \)-dimensional Brownian motion on a standard filtered probability space \((\Omega, \mathcal{F}, \mathbb{F} , \mathbb{Q})\) with \( \mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T} \) the natural filtration of \( W \) and given the pair of terminal condition and generator \((\xi, f)\) satisfying the following conditions\(^1\):

\textit{i)} \( \xi \in L^2 \) namely is square integrable;

\textit{ii)} \( f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R} \) such that \( f(t, 0, 0) \) is \( \mathbb{F} \)-predictable and \( f(t, y, z) \) is progressively measurable for all \( y, z \) and it satisfies a uniform Lipshitz condition in \((y, z)\), namely exists a constant \( K \) such that

\[ |f(t, y_1, z_1) - f(t, y_2, z_2)| \leq K(|y_1 - y_2| + |z_1 - z_2|) \quad \forall \ y_1, y_2, \forall \ z_1, z_2. \]

In addition, let us introduce the continuous process \((L_t)_{0 \leq t \leq T}\) belonging to the set of continuous and square integrable process such that

\[ \mathbb{E} \left[ \sup_{0 \leq t \leq T} |L_t|^2 \right] < \infty \]

\(^1\) For details on BSDE theory we refer in particular to the works of Pham (2009) and Yong, Zhu (1999).
and satisfying $\xi \geq L_T$.

Now, we can state that the solution to the reflected BSDE with terminal condition and generator $(\xi, f)$ and obstacle $L$ is given by the triple $(Y, Z, K) \in \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^+$ of progressively measurable and (square) integrable processes (in particular, $K$ is continuous and increasing and $K_0 = 0$) satisfying

\[
Y_t = \xi + \int_t^T f(s, Y_s, Z_s) + K_T - K_t - \int_t^T Z_s dW_s , \quad 0 \leq t \leq T
\]

\[
Y_t \leq L, \quad 0 \leq t \leq T
\]

\[
0 = \int_0^T (Y_t - L_t) dK.
\]

In their work, El Karoui et al. show the existence of the solution for the above RBSDE in theorem 5.2 through Picard iteration and in section six through penalization method, in theorem 4.1 they prove the uniqueness by comparison principle and in proposition 2.3 and also in section seven they show that the solution admits a probabilistic representation through Snell envelope (see proposition 5.1) relating it to the solution of optimal stopping problems.

This last result, which is valid in general also in a non-markovian framework (when $f$ is a general stochastic process), has deep connection with the analytical representation in a markovian framework where the terminal condition $\xi = g(X_T)$, the generator $f(t, y, x, z)$ and the obstacle process $L = h(X_t)$ become function of an Ito diffusion $X_t$. In this case they show (theorem 8.4) under some technical condition on the function $f(.), g(.)$ and $h(.)$ that the solution of the RBSDE

\[
Y^t,x = v(t, x)
\]

is a deterministic function of $(t, x)$ which is proved to solve in viscosity sense the following variational inequality

\[
\min \left[ -\frac{\partial v}{\partial t} - A^v - f(., v, \sigma D_x v); v - h \right] = 0 \text{ on } [0, T] \times \mathbb{R}^n
\]

\[
v(T, .) = g \text{ on } \mathbb{R}^n
\]

where $A^v$ is the well-known second order differential operator known as the characteristic operator associated to the diffusion $X_t$. Conversely, now the triple $(Y, Z, K)$ solution of the RBSDE it can be proved to be defined as follows

\[
Y_t = v(t, X_t)
\]

\[
Z_t = \sigma'(X_t) D_x v(t, X_t)
\]

\[
K_t = \int_0^t \left( -\frac{\partial v}{\partial t}(s, X_s) - A^v(s, X_s) - f(t, X_s, Y_s, Z_s) \right) ds.
\]

These are the main tools that we need together with the Snell envelope (defined in section 5.2) in order to tackle the solution of our switching control problem.

By doing this task, we refer in particular also to the work of Djehiche et al. (2008) in which the authors prove the existence of the solution for general multiple switching control problem by fairly generalizing the above mentioned results of El Karoui et al. So, by checking all the assumptions, we build on their results in order to show:
5. **Snell envelope approach, RBSDE and connections with QVI**

- the existence (and uniqueness) of the solution to our problem through the Snell envelope technique for optimal stopping;
- the continuity of the value function;
- the connection through reflected BSDE with the analytical problem representation (the system of HJB 4.8 - 4.9) and its viscosity solution set in chapter four.

It is worth of mention also the paper of Hamadene and Zhang (2010) which generalizes the RBSDE formulation in order to deal with switching control problems in which the generator and the obstacles are interconnected and the barrier can be non linear and random (non deterministic as in our case). They are able to show in this case the existence and uniqueness of the solution through a verification theorem, but not the existence of the optimal strategy, even though they provide an approximating optimal strategy $J^\delta$ with the help of some new (and technical) estimates. Let us underline in addition that the results of this section will be fundamental for the numerical part in which - through Monte Carlo simulations - a recursive optimal stopping procedure based on the Snell envelope formulation will be employed. Note, indeed, that the choice of the probabilistic approach has been forced by the recursive nature of our problem and - from the numerical point of view - by the fact that the solution of the HJB system via finite difference is cumbersome in presence of cost functions like our one. Therefore, because of possible generalizations that we have in mind, Monte Carlo based simulation algorithm are a suitable choice.

### 5.2 Switching control problem solution via Snell envelope: theory and assumptions

The objective of this section will be to apply to our problem the probabilistic approach based on the Snell envelope technique. This tool has been employed by Djehiche et al. (2008) to derive the solution of a general switching problem with multiple regimes in finite horizon, which is a generalization of our two regime switching control problem. So, following their work, we try to use the same approach and techniques in order to show the existence of an optimal switching strategy for our problem.

We start by recalling the main definitions and assumptions (recast in a generalized form) needed to prove the next results of the section:

**Def. 1)** The sequence of switching/stopping time $(\tau_j)_{j \geq 1}$ belonging to the set $\mathcal{T}$ over the interval $[0, T]$ (that coincide with the life of the underlying claim) is assumed to be finite $\{j = 1, \ldots, M\}$, increasing $(\tau_j < \tau_{j+1})$ and $\mathbb{F}$-adapted. The same definition is valid for the default time $\tau$ which is always greater than the valuation time $\{t < \tau\}$. In particular, we underline that in the rest of the section the default time is assumed also $\tau \geq \tau_j \ \forall j$, given that working in a pre-default framework, if default happen between two switching times, one takes the pre-default value relative to the last switching. This will be clearer when we pass to the numerical implementation.

**Def. 2)** Similarly, the switching indicator sequence $(Z_j)_{j \geq 1} \in \mathcal{Z}$ is finite and $\mathbb{F}_{\tau_j}$-adapted.
Def. 3) Define $u := ((\tau_j)_{j \geq 1}, (Z_j)_{j \geq 1})$ (for $Z = \{z, \zeta\}$) the sequence of admissible control such that $u \in \mathcal{C}^{ad}$. In particular, note that the set of admissible strategy is finite too, by definition.

Def. 4) Set, for notational convenience, the running cost functions of our problem as follows

$V(y, u) := \mathbb{E}\left[ \int_{t=0}^{T} e^{-r(T-s)} \Psi(t, \mathcal{Y}_s, u_s) ds + \sum_{j \geq 1} e^{-r\tau_j} [z e^z 1_{\{\tau_j < T\}} + \zeta e^\zeta 1_{\{\tau_j < T\}}] \right]$

$= \mathbb{E}\left[ \int_{t=0}^{T} e^{-r(T-s)} \Psi(t, \mathcal{Y}_s, u_s) ds + \sum_{j \geq 1} e^{-r\tau_j} e^Z 1_{\{\tau_j < T\}} \right]$ for $Z = \{z, \zeta\} \in \mathcal{Z}$.

Note that solving this optimal switching problem consists in finding the strategy $u^* := ((\tau_j^*)_{j \geq 1}, (Z_j^*)_{j \geq 1})$ such that

$V(u^*) = \inf_{u \in \mathcal{C}^{ad}} J(y, u)$

that is $V(u^*) \leq V(u)$ \forall $u \in \mathcal{C}^{ad}$.

To keep the same typical notation for the Snell envelope we note that for our minimization problem is equivalent to the maximization of minus the value function, that is

$V(u^*) = \inf_{u \in \mathcal{C}^{ad}} J(y, u) := \sup_{u \in \mathcal{C}^{ad}} (-J(y, u))$.

Note that in the following of the section we denote all the objects without the minus ahead for notational convenience, but they are intended with it (namely negative, or positive where they have a minus ahead).

Furthermore, the following assumptions are needed:

Hp 1) The stochastic factors that drive the dynamic of the system $(X_t)_{0 \leq t \leq T}$ and $(\lambda_t)_{0 \leq t \leq T}$, (that we indicate with the vector $\mathcal{Y}_t$ for brevity) are $\mathbb{R}$-valued processes adapted to the market filtration $\mathbb{F}_t^{x, \lambda} = \sigma\{W_s^{x, \lambda}, s \leq t\}_{t \in [0, T]}$ assumed right-continuous and complete. This assumption is clearly true for the chosen SDEs in our case.

Hp 2) Thanks to the immersion property (section 2.1), all the $\mathbb{G}$-adapted processes $(\nu_t)_{t \in [0, T]}$ that we encounter are also $\mathbb{F}$-adapted. In particular, they are càdlàg, real valued and $\mathbb{Q}$-measurable so that they belong to the set $\mathcal{M}^p = \{E[\sup_{t \leq T} |\nu_t|^p] < \infty\}$. If a given process $(\nu_t)_{t \in [0, T]}$ is also continuous we say that it belong to the set $\mathcal{K}^p = \{E[\int_0^T |\nu_s|^p ds] < \infty\}$. This assumptions are truly valid for our cost functions $F^Z(.) \in \mathcal{M}^p$, by the properties (proposition 4.3, 4.4) stated in the last section. Instead, the switching costs $c^Z_t \in \mathcal{K}^p$, being deterministic and continuous (because of the discounting factor presence).
5. **Snell envelope approach, RBSDE and connections with QVI**

*Hp 3*) The running costs functions need to satisfy at least the *polynomial growth condition*\(^{2}\) and this is true by the stated properties for \(F^Z(\cdot)\). For the switching costs \(c^Z_t\) a technical condition as \(\min\{c^Z_t, c^\zeta_t\} \geq C\) for any \(t \leq T\) and real constant \(C > 0\), is imposed in order to reduce the convenience to switch too many times.

We underline also as working assumption that at maturity the value function of our problem is zero. This is justified by the fact that at maturity all the obligations of the underlying contract are assumed to be paid, so that the terminal condition for our functional (known from chapter three) are valid only until \(T^-\).

With these things in mind, we pass to summarize some of the main results\(^{3}\) on the Snell envelope for optimal stopping that can be seen as a recursive procedure over the stopping time set that produces the optimal stopping/switching strategy and the solution\(^{4}\).

Let \(U = (U_t)_{0 \leq t \leq T}\) be an \(\mathcal{F}\)-adapted and real valued càdlàg process - not necessarily markovian - with square integrable supremum, i.e. \(U \in \mathcal{M}^2\) or uniformly integrable. Then, define for a given stopping time \(\theta\) the set \(\mathcal{T}_\theta = \{\tau \leq T : \mathcal{F} - \text{stopping time} | \theta \leq \tau \text{ a.s.}\}\), namely the set of all stopping time after \(\theta\).

Let us recall the definition of the *essential supremum* of an arbitrary family of r.v. \((X_l, l \in \mathcal{L})\), as the unique r.v. \(X = \text{ess sup}_t X_t\) such that \(X \geq X_l\) a.s. for all \(l \in \mathcal{L}\) and \(X \leq Y\) for all r.v. \(Y \geq X_l\) a.s. \(\forall l \in \mathcal{L}\). So now we are ready to give the following proposition that characterizes the Snell envelope.

**Proposition 5.1 (Snell envelope).** Let the process \(U\) respecting the definitions and hypothesis given above. Then, there exists an \(\mathcal{F}\)-adapted \(\mathbb{R}\)-valued càdlàg process \(S := (S_t)_{0 \leq t \leq T}\) such that:

(i) \(S\) is the smallest super-martingale that dominates \(U\), that is if \((\overline{S}_t)_{0 \leq t \leq T}\) is another càdlàg supermartingale such that \(\overline{S}_t \geq U_t\) for all \(t \in [0, T]\) then \(\overline{S}_t \geq S_t\) for any \(t\).

(ii) For any \(\mathcal{F}\)-stopping time \(\theta\) we have that

\[
S_\theta = \text{ess sup}_{\tau \in \mathcal{T}_\theta} E[U_\tau | \mathcal{F}_\theta] \quad \text{and} \quad S_T = U_T.
\]

The process \(S\) is called the Snell envelope of \(U\).

**Remarks 5.1** 1) If \(U\) is a process with only positive jumps then \(S\) is a continuous process. To understand the continuity property of the Snell envelope \(S\) is sufficient to apply the *Doob-Meyer decomposition* to \(S = M - A\) for some \(\mathcal{F}\)-martingale \(M\) and \(\mathcal{F}\)-predictable non decreasing process \(A\). It is known (see Karatzas, Shreve (1998)) that if \(U\) has continuous paths, then \(A\) is continuous.

\(^2\) This is in fact a weaker condition of the typical stronger condition imposed on the payoff function, like linear or sub-linear growth.

\(^3\) For details, the main reference are Karatzas, Shreve (1998) and Peskir, Shirayev (2006).

\(^4\) In particular, we will see in the numerical part, that the Snell envelope is the probabilistic counterpart of the Bellmann recursive principle of optimality.
too. Therefore, given that the filtration \( (F_t) \) is Brownian, the martingale \( M \) and \( S \) must also be continuous. In general, if we want to preserve the continuity of the Snell envelope \( U \) can have only positive jump (the technical condition is that \( U \) be upper-semicontinuous from the left and uniformly integrable).

2) We note that for \( \theta = 0 \) the minimal optimal stopping time \( \tau^* \) for \( S_0 \) exists and is given by

\[
\tau^* = \inf\{t \geq 0 : S_t \leq U_t\}
\]

so that the Snell envelope is the smallest supermartingale

\[
S_0 = E[U_{\tau^*} | \mathcal{F}_\theta = 0] = \text{ess sup}_{\tau \geq 0} E[U_{\tau} | \mathcal{F}_0].
\]

3) We conclude with the following stability property: if \( (U^n)_{n \geq 0} \) and \( U \) are càdlàg and uniformly integrable and such that the sequence \( (U^n)_{n \geq 0} \) converges increasingly and pointwisely to \( U \) then \( (S^{U^n})_{n \geq 0} \) converges in the same way to \( S^U \); so \( S^{U^n} \) and \( S^U \) are the Snell envelopes of \( U_n \) and \( U \). Furthermore, if \( U \) belongs to \( \mathcal{K}^p \), then also \( S^U \) belongs to it.

### 5.3 Verification theorem, existence and uniqueness of the solution.

Now, recalled the properties of the Snell envelope, the objective is to show that the solution of the switching problem that is the determination of the optimal switching strategy coincides with the proof of the existence - via a verification theorem - of the vector process \((Y^z, Y^\zeta)\) solution of a system of Snell envelopes, where \(Y^z\) (resp. \(Y^\zeta\)) indicates the optimal expected costs when collateral is set to zero (resp. one).

Before enunciating the theorem, let us define, for two continuous \( \mathbb{F}_t \)-adapted and real valued processes \((\delta_t)_{0 \leq t \leq T}, (\delta'_t)_{0 \leq t \leq T}\) and a given switching time \(\tau\), the following relation

\[
D_\tau(\delta = \delta') := \inf\{s \geq \tau, \delta_s = \delta'_s\} \land T.
\]

**Theorem 5.2 (Verification theorem).** Suppose there exist two processes \(Y^z := (Y^z_t)_{0 \leq t \leq T}\) and \(Y^\zeta := (Y^\zeta_t)_{0 \leq t \leq T}\) belonging to \(\mathcal{K}^p\) such that for any \(t \in [0, T]\) and a generic switching time \(\tau := \tau_j\)

\[
Y^z_t = \text{ess sup}_{r \geq t} \mathbb{E}\left[ \int_t^T e^{-r(t-s)} F^z(s, Y_s) ds + (Y^\zeta_t - e^{-rT} c^z) \mathbb{1}_{\{\tau < T\}} | \mathcal{F}_t \right], \quad Y^\zeta_T = 0, \quad Y^z_T = 0.
\]

Then \(Y^z\) and \(Y^\zeta\) are unique. Therefore, the following is verified

\[(a) \quad Y^z_0 = V(u^*) = \sup_{u \in \mathbb{C}^{ad}} (-J(y, u)).\]

\[(b) \quad \text{Defining the stopping times sequence } (\tau_j)_{j \geq 1} \text{ as}
\]

\[
\tau_1 = D_0(Y^z = Y^\zeta - c^z) \quad \text{on } \{\tau_1 < T\} \quad \text{(and } \{z_0 = 1\});
\]

\[
\tau_2 = D_1(Y^\zeta = Y^z - c^\zeta) \quad \text{on } \{\tau_2 < T\} \quad \text{(and } \{z_0 = 1\});
\]

(5.4)

(5.5)
and in general for $j \geq 1$
\[
\begin{align*}
\tau_{2j+1} &= D_{\tau_{2j}}(Y^z = Y^\zeta - c^z) \quad \text{on } \{\tau_{2j+1} < T\} \quad (5.6) \\
\tau_{2j+2} &= D_{\tau_{2j+1}}(Y^\zeta = Y^z - c^\zeta) \quad \text{on } \{\tau_{2j+2} < T\}. \quad (5.7)
\end{align*}
\]

Then the switching control strategy $u^* := \{(\tau_j^*)_{j \geq 1}, (Z_j)_{j \geq 1}\}$ is optimal.

**Proof.** The proof follows the same lines of the one in the paper of Djehiche et al. (2008). Firstly, we note that the requirements on $Y^z$ and $Y^\zeta$ in $\mathcal{K}_p$ imply their continuity and, by the properties of the Snell envelope (see Remarks 1 and 3), the continuity of the process $(\int_0^t e^{-rt} F^z(s, y)ds + (Y^\zeta_t - e^{-rt} c^\zeta))_{t \leq T}$ that belongs also to $\mathcal{K}_p$. Therefore, the process $(Y^\zeta_t - e^{-rt} c^\zeta)_{t \leq T}$ is of course continuous on $[0, T)$ and has at most a positive jump in $T$, so (by remark 1) this confirms the continuity of $Y^z$. The same reasoning are true for $Y^\zeta$, so that both are well defined.

We start proving the statement (a) by noting that for any $t \in [0, T]$ is valid the following relation
\[
Y^z_t + \int_0^t e^{-rt} F^z(s, \mathcal{Y}_s)ds = \text{ess sup}_{r \geq 0} \mathbb{E}\left[\int_0^T e^{-rT} F^z(s, \mathcal{Y}_s)ds + (Y^\zeta_T - e^{-rT} c^\zeta) \mathbf{1}_{\{T < \tau\}} | \mathcal{F}_t\right].
\]

Now, taking $Y^\zeta_0$ that is a $\mathcal{F}_0$-measurable random variable and it is constant with $\mathbb{Q}$-a.s. which implies $Y^\zeta_0 = \mathbb{E}[Y^\zeta]$. Recalling from remark 2) and from the optimality of the first stopping time after $t = 0$, say $\tau_1$, we get
\[
Y^\zeta_0 = \mathbb{E}\left[\int_0^{\tau_1} e^{-r\tau_1} F^\zeta(s, \mathcal{Y}_s)ds + (Y^\zeta_{\tau_1} - e^{-r\tau_1} c^\zeta) \mathbf{1}_{\{\tau_1 < T\}} | \mathcal{F}_0\right]
\]

\[
= \mathbb{E}\left[\int_0^{\tau_1} e^{-r\tau_1} F^\zeta(s, \mathcal{Y}_s)ds + Y^\zeta_{\tau_1} \mathbf{1}_{\{\tau_1 < T\}} - e^{-r\tau_1} c^\zeta \mathbf{1}_{\{\tau_1 < T\}}\right].
\]

Now since $[Y^\zeta_{\tau_1} \mathbf{1}_{\{\tau_1 < T\}}]$ is $\mathcal{F}_{\tau_1}$-measurable, reapplying the Snell envelope and the definition of $Y^\zeta_{\tau_1}$, we get
\[
\mathbf{1}_{\{\tau_1 < T\}} Y^\zeta_{\tau_1} = \text{ess sup}_{r \in \tau_1} \mathbb{E}\left[\int_{\tau_1}^{\tau_2} e^{-(r\tau_2 - s)} F^\zeta(s, \mathcal{Y}_s)ds + (Y^z_{\tau_2} - e^{-r\tau_2} c^\zeta) \mathbf{1}_{\{\tau_2 < T\}} | \mathcal{F}_{\tau_1}\right]
\]

that substituted in $Y^z_0$ gives
\[
Y^z_0 = \mathbb{E}\left[\int_0^{\tau_1} e^{-r\tau_1} F^\zeta(s, \mathcal{Y}_s)ds + \int_{\tau_1}^{\tau_2} e^{-(r\tau_2 - s)} F^\zeta(s, \mathcal{Y}_s)ds + Y^\zeta_{\tau_2} \mathbf{1}_{\{\tau_2 < T\}} + e^{-r\tau_1} c^\zeta \mathbf{1}_{\{\tau_1 < T\}} - e^{-r\tau_2} c^\zeta \mathbf{1}_{\{\tau_2 < T\}}\right] ( \text{summing the integrals})
\]

\[
= \mathbb{E}\left[\int_0^{\tau_2} e^{-r\tau_2} \Psi(s, \mathcal{Y}_s; u_s)ds + Y^z_{\tau_2} \mathbf{1}_{\{\tau_2 < T\}} - e^{-r\tau_1} c^\zeta \mathbf{1}_{\{\tau_1 < T\}} - e^{-r\tau_2} c^\zeta \mathbf{1}_{\{\tau_2 < T\}}\right].
\]

Now, repeating the same recursive argument for $j$ increasing we get in general
\[
Y^z_0 = \mathbb{E}\left[\int_0^{\tau_2j} e^{-r\tau_2j} \Psi(s, \mathcal{Y}_s; u_s)ds + Y^z_{\tau_2j} \mathbf{1}_{\{\tau_2j < T\}} - \sum_{1 \leq k \leq j} (e^{-r\tau_{2k-1}} c^\zeta \mathbf{1}_{\{\tau_{2k-1} < T\}} + e^{-r\tau_{2k}} c^\zeta \mathbf{1}_{\{\tau_{2k} < T\}})\right].
\]
that is
\[ Y_0^\tau = V(u^*). \]

It’s left to show that \( Y_0^\tau \geq V(u) \) for any \( u \in C^{ad} \) that implies the proof of the statement (b). So, taking a finite strategy \( u = (\tau_j)_{j \geq 1}, Z_j \), starting from the optimal stopping \( \bar{\tau}_1 \) and following the same reasoning done above we have that
\[ Y_0^\tau \geq \mathbb{E}\left[ \int_0^{\bar{\tau}_1} e^{-r \tau} F^\tau(s, \mathcal{Y}_s) ds + (Y^\tau_{\bar{\tau}_1} - e^{-r \bar{\tau}_1} c^\tau) \mathbb{1}_{\{\bar{\tau}_1 < T\}} \right] \]
and following the same recursive procedure we get in the end
\[ Y_0^\tau \geq \mathbb{E}\left[ \int_0^{\bar{\tau}_j} e^{-r \tau} \Psi(s, \mathcal{Y}_s; u_s) ds + Y_{\bar{\tau}_j}^\tau \mathbb{1}_{\{\bar{\tau}_j < T\}} - \sum_{1 \leq k \leq j} (e^{-r \bar{\tau}_{jk-1}} c^\tau \mathbb{1}_{\{\bar{\tau}_{jk-1} < T\}} + e^{-r \bar{\tau}_{jk}} c^\tau \mathbb{1}_{\{\bar{\tau}_{jk} < T\}}) \right]. \]

By the finiteness of \( u \) we get that for \( j \to \infty \) the right hand side of the last equation converges to \( V(u) \). Therefore the following it’s true
\[ Y_0^\tau = V(u^*) \geq V(u) \]
which implies the optimality of the control strategy \( u^* \) and indeed the proof of (b).

The last thing left is to establish the uniqueness of the pair \((Y^\tau, \Psi^\tau)\) that satisfies the Snell envelope given in (a). The proof use the same recursive arguments for \( Y^\tau_j \) taking \( u \) over the set \( C^{ad} \) of admissible strategies and then assuming the existence of an other pair process \((\tilde{Y}^\tau, \tilde{Y}^\tau)\), but thanks to the uniqueness of the Snell envelope this pair must coincides with \((Y^\tau, \Psi^\tau)\). \diamond

The verification theorem just proved is a very important result but it assumes a priori the existence of the solution process \((Y^\tau, \Psi^\tau)\). The second important contribution of the work of Djehiche et al. (2008) is the proof of the existence and continuity of this pair process obtained through limit of sequences of the processes \((Y^{\tau,j})_{j \geq 0}\) and \((\Psi^{\tau,j})_{j \geq 1}\) defined for any \( t \in [0, T] \) as (recalling the Snell envelope)
\[ Y^{\tau,0}_t = \operatorname{ess} \sup_{\tau \geq t} \mathbb{E}[\int_t^\tau e^{-r \tau} F^\tau(s, \mathcal{Y}_s) ds | \mathcal{F}_t]. \]

for \( j = 0 \) (such that \( Y^{\tau,0}_t = 0 \) being defined for \( j \geq 1 \) and also \( c^\tau = 0 \) too), while for \( j \geq 1 \) we get
\[ Y^{\tau,j}_t = \operatorname{ess} \sup_{\tau \geq t} \mathbb{E}[\int_t^\tau e^{-r (\tau - s)} F^\tau(s, \mathcal{Y}_s) ds + (Y^{\tau,j-1}_t - e^{-r \tau} c^\tau) \mathbb{1}_{\{\tau < T\}} | \mathcal{F}_t]. \]
\[ Y^{\tau,j}_t = \operatorname{ess} \sup_{\tau \geq t} \mathbb{E}[\int_t^\tau e^{-r (\tau - s)} F^\tau(s, \mathcal{Y}_s) ds + (Y^{\tau,j-1}_t - e^{-r \tau} c^\tau) \mathbb{1}_{\{\tau < T\}} | \mathcal{F}_t]. \]

We now just state the properties of these sequences of processes that are necessary to show their existence and then that they satisfy the verification theorem 5.2.

\footnote{Note that the expression is not the same of the Def. 5) because here we have omitted the switching indicator \( Z_j \) because they can be considered inside the indicator of the switching costs for simplicity.}
The proof can be easily adapted from the already mentioned paper following the lines of the proof of the theorem 5.2 and using standard convergence results and properties of the Snell envelope.

**Proposition 5.3.** a) For each \( j \geq 1 \), the processes \( Y^{z,j} \) and \( Y^{\zeta,j} \) are continuous and belong to \( K^p \).

b) The sequences \( (Y^{z,j})_{j \geq 0} \) and \( (Y^{\zeta,j})_{j \geq 0} \) converge increasingly and point-wisely with \( \mathbb{Q} \)-a.s. for any \( t \in [0, T] \) and in \( \mathcal{M}^p \) to càdlàg processes \( \bar{Y}^z \) and \( \bar{Y}^\zeta \) (respectively). Moreover, these limit processes satisfy

1. \[ \mathbb{E}[\sup_{0 \leq t \leq T} |\bar{Y}^Z_t|^p] < \infty, \ Z = \{z, \zeta\}; \] (5.8)

2. for any \( t \in [0, T] \),

\[
\bar{Y}^z_t = \text{ess sup}_{\tau \geq t} \mathbb{E}[\int_t^\tau e^{-r(\tau-s)} F^z(s, Y_s)ds + (Y^{\zeta}_\tau - e^{-r\tau}c^z)1_{\{\tau < T\}}|\mathcal{F}_t] \\
\bar{Y}^\zeta_t = \text{ess sup}_{\tau \geq t} \mathbb{E}[\int_t^\tau e^{-r(\tau-s)} F^\zeta(s, Y_s)ds + (Y^z_\tau - e^{-r\tau}c^\zeta)1_{\{\tau < T\}}|\mathcal{F}_t]. \] (5.9)

In the end, proving that the limit processes \( (\bar{Y}^z_t, \bar{Y}^\zeta_t) \) are also continuous (through the *Dellacherie-Meyer decomposition* of the Snell envelope (see also remark 3) above) and that \( \mathbb{E}[\sup_{s \leq T} |Y^{Z,j}_s - Y^Z_s|^p] \to 0 \) for \( n \to \infty \), one gets the proof that they satisfies the verification theorem and, through this construction, that the solution exists\(^6\).

### 5.4 Solution connection with Reflected Backward SDEs and variational inequalities

The optimal solution \( (Y^z, Y^\zeta) \) of our switching control problem established via Snell envelope can be characterized - thanks to important results due to El Karoui et al. (1997) - also in terms of systems of reflected BSDEs that have deep connections with PDEs. In particular, it can be proved, under some specific conditions, that this vector process is also a viscosity solution of a system of PDE with interconnected obstacles. So, the objective of this section will be to show that \( Y^z = v^z(t, y) \) and \( Y^\zeta = v^\zeta(t, y) \) namely that the probabilistic solution for our problem is also a *viscosity solution* for a system of PDE with interconnected obstacles which is exactly the HJB system derived in section 4) for which we couldn’t say much about the existence of the assumed smooth solution. By doing this we can close the circle overcoming the technical difficulties of the analytical approach by employing other results from an other branch of mathematics, that is a quite usual task nowadays.

To show this important result we need to verify the technical conditions that allows to apply it to our switching control problem and to recall the main definitions for *viscosity solutions* and *reflected backward SDEs*. We start from the latter, recalling that the reflected backward SDEs (or

\(^6\) Further details on this part can be found in the already mentioned paper.
5. Snell envelope approach, RBSDE and connections with QVI

RBSDE) are just the probabilistic counterpart of the variational approach when the stochastic factor dynamic \((X_t)\) follows an Ito diffusion (that is markovian). By the choice done for the dynamics of our model (or also by property 2, section 4), this is surely true and RBSDE definition in this markovian case can be stated as follows.

**Definition 5.4 (RBSDE).** Suppose to have an Ito diffusion \(X_t\) with initial condition \(X_0 = x\) and that exists a vector process \((Y^x, N^x, A^x)\) adapted to \((F^X_t)\) and such that \(E[\sup_{0 \leq t \leq T} |Y^x_t|^2 + \int_0^T \|N^x_t\|^2 dt + |A_T|^2] < \infty\) (namely they belong to \(\mathcal{M}^2\)), with \(A\) continuous and increasing, a RBSDE is defined by the following system

\[
Y^x_t = \int_t^T F(s, X^x_s) ds - \int_t^T N_s dW_s + A_T - A_t, \quad (5.10)
\]
\[
Y^x_t \geq SW(t, Y^x_t) \quad (5.11)
\]
\[
0 = \int_0^T (Y^x_t - SW(t, Y^x_t)) dA_t, \quad A_0 = 0 \quad (5.12)
\]

As we know from the introduction, \(F(\cdot)\) is the generator function depending on the (markovian) stochastic vector dynamic \(X_t\), \(SW(\cdot)\) indicates in general the barrier that is, in our case, the switching boundary defined as \(SW(t, Y^x_t) = Y_t^{\xi, \zeta} - c^{\xi, \zeta}_t\) (also defined elsewhere as the "intervention operator"), while \(N\) is a conditional expectation/volatility process that helps \(Y_t\) to be \(\mathcal{F}_t\)-measurable and \(A\) a compensator process that increases only when \(Y\) hits the barrier. The condition (5.12) means that the action of the increasing process \(A\) is minimal in the sense that it is active only when the constraint is saturated, namely when \(Y^x_t = SW(t, Y^x_t)\).

We keep this notion of RBSDE in mind because it’s the bridge that we need to show the connection between the probabilistic solution derived via Snell envelope in the verification theorem and the viscosity solutions of the analytical approach.

For what concerns the assumptions that we need to check in order to proceed, they are the following.

**A.1** The dynamic of the model are described by markovian diffusions with associated infinitesimal generator \(\mathcal{A}\) whose drift and diffusion coefficient functions satisfy the standard conditions for the solution existence and uniqueness.

Check: these assumptions are clearly valid for our model for which the infinitesimal generators \(\mathcal{A}^\xi\) and \(\mathcal{A}^\zeta\) are defined in chapter three or as in 4.6 and 4.7.

**A.2** The switching cost functions \(l^i(t)\) have to be constant or deterministic.

Check: they cannot be state dependent and hence stochastic because it would prevent the continuity of the solution, that remains an open problem. In our case, the switching costs are \(l^i(t) = e^{-\gamma t} c^Z\) that are clearly deterministic.

**A.3** The running costs functions \(F^Z(\cdot)\) have to be continuous and of polynomial growth.

Check: this is true from the properties established above in proposition 4.3 and 4.4.
In addition, the following estimates - that agree with the polynomial growth condition - are assumed (as from Djehiche et al. (2008) paper):

\[
\mathbb{E}[\sup_{0 \leq s \leq T} |\mathcal{Y}_s|^\beta] \leq C(1 + |y|^\beta)
\]

for any \( \beta \geq 2 \) and

\[
\mathbb{E}[\sup_{0 \leq s \leq T} |\mathcal{Y}_s - \mathcal{Y}_s'|^2] \leq C(1 + |y|^2)(|y - y'|^2 + |t - t'|)
\]

for any \( t, t' \in [0, T] \) and \( y, y' \in \mathbb{R}^2 \).

With this set of assumptions, say \([A]\), in mind, we consider the following system of variational inequalities\(^7\) which is easy to see as just a reformulation (in general terms) of the HJB system of PDE of the fourth chapter:

\[
\begin{align*}
\min \{-\partial_t v^z(t, y) - \mathcal{A}^z v^z(t, y) + r v^z(t, y) - F^z(t, y), \quad v^z(t, y) - (v^z(t, y) - l^z(t))\} &= 0, \quad v^z(T, y) = 0 \\
\min \{-\partial_t v^\xi(t, y) - \mathcal{A}^\xi v^\xi(t, y) + r v^\xi(t, y) - F^\xi(t, y), \quad v^\xi(t, y) - (v^\xi(t, y) - l^\xi(t))\} &= 0, \quad v^\xi(T, y) = 0
\end{align*}
\]

Now, we need to recall the notion of viscosity solution for this system of PDEs.

**Definition 5.5 (Viscosity solution).** Let \((v^z, v^\xi)\) be a pair of continuous functions on \([0, T] \times \mathbb{R}^k\) with values in \(\mathbb{R}\) and such that \((v^z, v^\xi)(T, y) = 0\). The vector \((v^z, v^\xi)\) is defined

a) a viscosity supersolution of the above system of PDEs if for any \((t_0, x_0) \in [0, T] \times \mathbb{R}^k\) and any pair of smooth functions \((\phi^z, \phi^\xi) \in (C^{1,2}([0, T] \times \mathbb{R}^k))^2\) such that \((\phi^z, \phi^\xi)(t_0, y_0) = (v^z, v^\xi)(t_0, y_0)\) and \(((t_0, y_0)\) is a maximum point of \(\phi^z - v^z\) and \(\phi^\xi - v^\xi\), we have:

\[
\begin{align*}
\{ \min \{-\partial_t \phi^z(t_0, y_0) - \mathcal{A}^z \phi^z(t_0, y_0) + r \phi^z(t_0, y_0) - F^z(t_0, y_0), \quad v^z(t_0, y_0) - (v^z(t_0, y_0) - l^z(t_0))\} &\geq 0, \\
\min \{-\partial_t \phi^\xi(t_0, y_0) - \mathcal{A}^\xi \phi^\xi(t_0, y_0) + r \phi^\xi(t_0, y_0) - F^\xi(t_0, y_0), \quad v^\xi(t_0, y_0) - (v^\xi(t_0, y_0) - l^\xi(t_0))\} &\geq 0.
\end{align*}
\]

b) A viscosity subsolution of the above system of PDEs if for any \((t_0, x_0) \in [0, T] \times \mathbb{R}^k\) and any pair of smooth functions \((\phi^z, \phi^\xi) \in (C^{1,2}([0, T] \times \mathbb{R}^k))^2\) such that \((\phi^z, \phi^\xi)(t_0, y_0) = (v^z, v^\xi)(t_0, y_0)\) and \(((t_0, y_0)\) is a minimum point of \(\phi^z - v^z\) and \(\phi^\xi - v^\xi\), we have:

\[
\begin{align*}
\{ \min \{-\partial_t \phi^z(t_0, y_0) - \mathcal{A}^z \phi^z(t_0, y_0) + r \phi^z(t_0, y_0) - F^z(t_0, y_0), \quad v^z(t_0, y_0) - (v^z(t_0, y_0) - l^z(t_0))\} &\leq 0, \\
\min \{-\partial_t \phi^\xi(t_0, y_0) - \mathcal{A}^\xi \phi^\xi(t_0, y_0) + r \phi^\xi(t_0, y_0) - F^\xi(t_0, y_0), \quad v^\xi(t_0, y_0) - (v^\xi(t_0, y_0) - l^\xi(t_0))\} &\leq 0.
\end{align*}
\]

c) A viscosity solution if it is both a viscosity supersolution and subsolution.

\(^7\) Or, as already mentioned, system of PDE with interconnected obstacles.
Now, we have all the elements to "close the circle": before enunciating this central result that is based on two central theorems proved in the paper of El Karoui et al. (1997) on the connection between the solutions of RBSDE and PDEs, we describe the ideas and the logic steps underlying the proof that we do not give but details can be found in Djehiche et al. (2008) other than the already mentioned paper of El Karoui et al. (1997)

First thing, we need to recall the vector process \((Y_{z,j}^{s,y}, Y_{\zeta,j}^{s,y})_{0 \leq s \leq T}\) obtained via Snell envelope and that we proved to satisfy the verification theorem 5.2, the existence (and uniqueness) and its continuity (it’s in \(K^p\)). Taking the pair sequences \((Y_{z,j}^{s,y})_{s \leq T}\) and \((Y_{\zeta,j}^{s,y})_{s \leq T}\) defined recursively via Snell envelope, thanks to the results of El Karoui et al. paper (proposition 2.3), these processes are proved to be solutions of one obstacle RBSDE, that is (depending on the regimes \((z,\zeta)\))

\[
Y_{t}^{(z,\zeta)}j = \int_{s}^{T} F_{z,\zeta}^{(t,u,y)} du - \int_{s}^{T} N_{u}^{(z,\zeta)}j dW_u + A_{t}^{(z,\zeta)}j - A_{s}^{(z,\zeta)}j, \quad t \leq s \leq T \\
0 = \int_{t}^{T} (Y_{t}^{(z,\zeta)}j - (Y_{t}^{(\zeta,z)}j - l_{t}^{(\zeta)})dA_{t}^{(z,\zeta)}j; 
\]

In this way, it’s established the connection between Snell envelope and the solution of RBSDE. To prove the remaining analytical connection with the viscosity solution of PDEs with obstacles, the above vector process is associated to sequences of deterministic functions defined as follows \(Y_{z,j}^{s,y}(s,y)\) and \(Y_{\zeta,j}^{s,y}(s,y)\). Then, by induction and under the assumptions \([A]\) one is able to prove the existence and the convergence of these sequences to deterministic, continuous (in \((t,y)\)) and of polynomial growth functions \((v^z, v^\zeta) = (Y^z, Y^\zeta)\). In this way, the conditions to apply the theorem (8.5) of El Karoui’s work are satisfied and one can use it to conclude that \((v^z, v^\zeta)\) are the viscosity solutions of the system of PDE with interconnected obstacles defined above, namely the deep connection between the probabilistic and analytic solution that we wanted to highlight.

In the following, we conclude giving the statement of the fundamental result just described.

**Theorem 5.6 (Snell Envelope-RBSDE-QVI connection).** Under the assumptions \([A]\), there exist two deterministic functions \(v^z(t,y)\) and \(v^\zeta(t,y)\) defined on \([0, T] \times \mathbb{R}^k\) and real valued such that

a) \(v^z\) and \(v^\zeta\) are continuous in \((t,y)\), satisfy a polynomial growth condition and for each \(t \in [0, T]\) and \(s \in [t, T]\)

\[
Y_{s}^{z,y} = v^z(s,y) \quad (5.13) \\
Y_{s}^{\zeta,y} = v^\zeta(s,y) \quad (5.14)
\]

b) The vector pair \((v^z, v^\zeta)\) is a viscosity solution for the system of variational inequalities of definition 5.5.
5. **Snell envelope approach, RBSDE and connections with QVI**

5.5 *The optimal switching solution recast as an iterative optimal stopping procedure*

Before dealing with the implementation and the numerical solution of the problem we need to make a bridge between the theoretical results established in the last section in relation to the solution of our switching control problem and their practical application and implementation. As we already mentioned, given the characteristics of our problem that are dimensions greater than two \( d \geq 3 \), *non-linearity* and *path-dependency*, we are not going to solve numerically the system of *Hamilton-Jacobi-Bellman equations*. The idea is to build a Monte Carlo algorithm to determine the solution employing the Snell envelope tool and the related dynamic programming principle. In fact, already in the proof of the verification theorem 5.2 one can catch an idea for a suitable algorithm because of the recursive procedure over the stopping time set \( T \) used to derive the optimal control strategy and the solution process \((Y^z, Y^ζ)\). This recursive procedure and the related Snell envelope properties can be adopted to reformulate our optimal switching problem as a *multiple iterative optimal stopping*.

This approach has been largely applied in the literature to solve general *stochastic impulse control problem*. The main reference for the switching type problems is the already mentioned paper of Carmona and Ludkovski (2006). So we show here how to derive the iterative optimal stopping procedure on which will be based the implementation of the numerical part. Firstly, let us consider the special case in which the number of possible switches over \([0, T]\) is assumed to be fixed or user defined. By doing this, we define the set of admissible strategies/controls over \([t, T]\) \((t \in [0, T])\) with at most \( l \) switches as follows

\[
C_{ad}^l(t) := \{ \{T, Z\} \in C_{ad}(t) : \tau_j = T \text{ for } j \geq l + 1 \},
\]

so that one can get different optimization problem and value functions for the number of switches \( l \) varying. In particular, the optimization problems, for example, over \( C_{ad}^{l+1} \) and \( C_{ad}^l \) should be related to each other for any \( l \) (being \( C_{ad}^{l+1} \cup C_{ad}^l \) or contained). In fact, by the *Bellman principle of optimality* (that is the dynamic programming in discrete time) solving the problem with \( l + 1 \) switches is equivalent to find the optimal first switching time \( \tau \) that maximizes/minimizes the payoff until \( \tau \) plus the value function in \( \tau \) that solve optimally the problem with the remaining \( l \) switches. So these reasoning suggests a practical and simpler way to solve the optimal switching problem by reducing it to a recursive sequence of \( l \) optimal stopping problems that allows to keep trace of the optimal switching time too.

So, let us define the value function of our \( l \) switching control problem \( V^l(t, y, Z) \) \((Z = \{z, ζ\} \text{ and } y \text{ the realization of our markovian vector dynamic}^8)\), through the following recursive sequence of optimal stopping problem starting from \( l = 0, 1, 2, \ldots, \bar{l} \) and \( t \in [0, T] \) and going forward, that

---

8 We remind that the value function have to be intended always negative with a minus ahead, to give sense to the minimization problem.
5. **Snell envelope approach, RBSDE and connections with QVI**

is

\[
V^0(t, y, Z) := \mathbb{E} \left[ \int_t^T e^{-r(T-s)} F_Z(s, \mathcal{Y}_s) ds | \mathcal{Y}_t = y \right],
\]

\[
V^l(t, y, Z) := \sup_{\tau \in \mathcal{T}} \mathbb{E} \left[ \int_t^{\tau \wedge T} e^{-r(T-s)} F_Z(s, \mathcal{Y}_s) ds + SW^{l,Z}(\tau, \mathcal{Y}_\tau) | \mathcal{Y}_t = y \right]
\]

(5.15)

where the recursion is inside the switching/intervention operator already defined above for the RBSDE. The \( SW(.) \) operator takes the following expression

\[
SW^{l,Z=\zeta}(\tau, \mathcal{Y}_\tau) := \max_{Z \in \mathcal{Z}} \{ V^{l-1}(\tau, y, Z = \zeta) - c_{Z=\zeta} \}
\]

(5.16)

given that the maximum is trivial in our case with only two regimes. Now it is easier to see the sense and the utility of the iterative procedure in which at every calculation step one already know from the past step the value function \( V^{l-1}(.) \) with \( l - 1 \) switches and so every step is reduced to an optimal stopping problem easier to solve and implement. We will see this in the next section.

The well-posedness of this characterization and its equivalence to the value function with \( l \) switches allowed so defined

\[
\bar{V}^l(t, y, Z) := \esssup_{\tau \in \mathcal{T}} \mathbb{E} \left[ \int_t^\tau e^{-r(\tau-s)} F_Z(s, \mathcal{Y}_s) ds + SW^{l,Z}(\tau, \mathcal{Y}_\tau) | \mathcal{F}_t \right]
\]

can be easily shown recalling the Snell envelope property and the related recursion procedure to derive the solution (theorem 5.2). In fact, setting \( \Psi^Z_t = \int_t^0 e^{-rt} F_Z(s, \mathcal{Y}_s) ds \) and \( U^{l,Z}_t = \int_0^t e^{-rt} F_Z(s, \mathcal{Y}_s) ds + SW^{l,Z}(t, \mathcal{Y}_t) \), it is easy to see that \( U^{l,Z} \) satisfies all the measurability and regularity assumptions to apply the Snell envelope that would be given (keeping the same notation) by \( S^{l,Z}_t = \esssup_{\tau \in \mathcal{T}} \mathbb{E}[U^{l,Z}_\tau | \mathcal{F}_t] \). The Snell envelope compared with the value functions \( \bar{V}^l(t, y, Z) \) gives the claimed equivalence with the iterative procedure, that is

\[
\bar{V}^l(t, \mathcal{Y}_t, Z) = S^{l,Z}_t - \Psi^Z_t.
\]

Consequently, the optimal stopping time \( \tau^*_l \) corresponding to the Snell envelope defined by \( \bar{V}^l(t, y, Z) \) will be clearly defined as follows

\[
\tau^*_l = \inf \{ s \geq t : \bar{V}^l(s, \mathcal{Y}_s, Z) = SW^{l,Z}(s, \mathcal{Y}_s) \} \wedge T
\]

We conclude by stating that the existence of this optimal switching policy and the convergence of \( \bar{V}(.) \) for \( l \to \infty \) to the true value function \( V(.) \) derives from the results proved in section 5.1 through the Snell envelope. Further details on the proofs of the above results and statements can be found in Ludkovski (2005).
6. NUMERICAL SOLUTION APPROACH AND IMPLEMENTATION

Truth is much too complicated to allow anything but approximations.

J. von Neumann

6.1 Introduction

In this part of the work we build on the theoretical results and insights of the past two sections to define a suitable numerical procedure in order to solve our switching problem. As we already mentioned, the main characteristics of our valuation problem are the non-linearity and recursion of the value functions that depend (forward) - path by path - on the switching strategy, that is the agent control sequence \((\tau_j^\ast, (Z_j^\ast))_{j \geq 1}\). From a numerical point of view, the system of HJB equations that we showed (under certain conditions) to be non-linear second order (integral) PDE with obstacles, are hard to solve in dimension three and special finite differences for finite element schemes are needed.

Anyway, a more recent branch of algorithm has been developed for the numerical solution of backward stochastic differential equations which are, as already mentioned, the probabilistic counterpart for PDE problems with underlying markov diffusion process (like our ones). These algorithms are characterized by a two-stage procedure based on Monte Carlo simulation. The first consists of a time discretization of the BSDE\(^1\). The main difficulty here is that, on the one hand, the discretization quite naturally works backwards in time, because the terminal condition is given. On the other hand, the numerical solution should be adapted to the filtration (because the true solution is so). However, the information grows forwards in time. This problem can be solved by projecting the solution on the available information in each step while going backwards in time. However, ”projecting on the available information” means that in each time step a conditional (nested) expectation must be evaluated. As the conditional expectation cannot be calculated in closed form, in a second step one has to apply an approximation procedure for the conditional expectations and the most popular method in finance to make this is through the least square Monte Carlo method proposed by Longstaff, Schwartz (2001) \(^2\) This has been the starting point of a fruitful series of research papers intended to speed up the convergence and to reduce the discretization and projection errors of the method mainly due to the choice of the approximating

\(^1\) Bouchard and Touzi, Discrete-time approximation and Monte Carlo simulation of BSDE, (2002) is the main reference.

\(^2\) The idea was actually introduced in a paper of Carriere (1996).
basis function for the conditional expectation\(^3\).

For what concerns our switching control problem, we have already shown (in chapter five) the strict relation between reflected BSDE and the Snell envelope solution approach. So we have at disposal all the above mentioned numerical approach for the numerical solution of our problem. In particular, we follow the path indicated in section 5.5 in which we recast the problem as an iterative optimal stopping time whose main ingredients for the numerical solution are a Monte Carlo generator for the paths simulation, a backward induction procedure founded on the dynamic programming principle (or Bellman’s optimality in discrete time) and an approximation method (via least-square regression) of the nested conditional expectation at each step of the recursion. In the following, we start to highlight these and all the working hypothesis set for the numerical algorithm definition and its implementation (for a defaultable IRS) in order to derive numerically the solution namely the value function and the optimal switching strategy for our model.

6.2 Numerical solution: the basic underlying idea.

Let us start by describing the main issues related to the numerical solution of our switching control problem and the basic underlying idea of the chosen solution approach.

1) Firstly, reminding our problem formulation, the main issue that we need to tackle for its implementation is that the cost functions \( F^Z(s, \mathcal{Y}_s) \) are integral/summation of in general non linear and recursive function depending on future switching times \( \tau_j \) and on the stochastic factors \((X, \lambda)\). So, the problem is that evolving the stochastic factor dynamics is not enough to recover the paths of these cost functions, but a precalculation phase is needed to get the paths of the cost function building blocks which are mainly the contract \( NPV_t \) or equivalently the positive/negative exposures \( NPV_t^{\pm} \). Indeed, the basic idea is to simulate the dynamics in order to get a grid of simulated building blocks but to make this we remark that one need to calculate the value of the NPV and exposures at forward times not only at inception/valuation date, say \( t \). The solutions to this issue are possibly twofold:

\begin{itemize}
  \item[a)] to assume a specific and suitable dynamic for the \( NPV \) or the exposures in order to run forward in time its values; in this case, it would be as considering the price process of an exchange traded security for the \( NPV_t = NPV_t^{Exch} \), as a (partial) solution to the recursion problem that we highlighted in section 3.3;
  \item[b)] to simulate paths of the stochastic factor that drives the exposures and use the Longstaff-Schwartz (LS) method to determine the forward conditional expectation of these quantities assuming that they can be approximated through basis functions and projected
\end{itemize}

\(^3\) In particular, are worth of mention the paper of Pages et al. (2011) on the quantization approach, the work of Bouchard et al. (2004) on the Malliavin approach to Monte Carlo approximations of conditional expectations and other works based on non parametric regressions techniques are some few examples of a large body of research in this field.
6. Numerical solution approach and implementation

on the information available at forward time that is given from the simulated underlying paths. Is worth of mention that in general the LS method should be applied two times in our model: the first time, to estimate the (conditional) expected positive exposure $NPV^+$ and the second for the negative exposure $NPV^-$; but thanks to the "symmetry hypothesis" that we impose in the following section, it will be possible to make computations one time just for the NPV.

For the rest of the section we are intended to go for the second way.

2) Once we have decided in the first step how to get the values of the building blocks, what is left is to apply the optimal recursion procedure defined through Snell envelope and dynamic programming. Here, the problem is the same because we have the relevant paths to calculate the cost functions, but we need to determine the optimal switching strategy recursively via dynamic programming. In order to do this, we need a grid for the values of the quantities that enters our cost functions, but this means that the phase 1) described above characterized by the LS method, has to be run different times using a nested simulation approach in order to get this grid. As it is quite obvious, this implies a large computational costs that can be overcome mainly through parallel computing and programming methods (like CUDA programming and using GPCU hardware). Fortunately, a second strategy can help to reduce the computational costs. It needs just two Monte Carlo run thanks to an approximation and the orthogonality hypothesis between the driving stochastic factors $X \perp \lambda$: in the first run we simulate $X$ in order to get through LS method the expected NPV and exposures. From these, we determine the CVA/DVA and the collateral cost by running a Monte Carlo for the default intensities and times and for the funding costs. In fact, taking in count the recovery rate, multiplying these values for the ones obtained from the first run (with LS method) and summing up over time we can obtain the grid points of the CVA/DVA and collateral costs necessary for the implementation of the iterative stopping procedure in order to finally get the value function and the optimal control/switching strategy.

Remarks 6.2. Note that for funding we have mentioned a stochastic factor, but without loosing much generalization, we can also use a deterministic specification function to model the funding term (as it is in our functional). For what concerns the orthogonality condition, the procedure can be further generalized removing it and employing copula based simulation, but the computational costs returns to be high and cumbersome.

3) The last important issue is that of the choice of the relevant probability measure under which the computations should be carried on. Because of the fact that we always post the problem as a dynamic stochastic optimization problem of the agent we keep on holding separate it from the underlying pricing and hedging problem that we try to formally establish and takeover later. So, for these reasons, all the calculations are intended in risk management sense and indeed under the objective/historical measure.
6. Numerical solution approach and implementation

6.3 Numerical procedure definition and analysis

As mentioned in the last paragraph, we divide in two macro-phases the analysis of the numerical procedure for the solution of our switching control problem to which we dedicate the present section. This approach can be extended to cover more general payoff (typically, path dependent) and dynamics, but we choose to keep it tied to the model set in the previous sections (in which as payoff of the underlying contract we have made the choice of a simple interest rate swap signed between counterparty A and B).

1) Path precalculation phase and first "modified Longstaff-Schwartz" application.

a) Value function and system dynamic working choices.

Firstly, let us recall the main ingredients and assumptions of the problem we want to tackle numerically. In particular, we recall our value function recast as an iterative optimal stopping problem (as defined in section 5.3), that is

\[
\begin{align*}
V^0(t, y, Z) &:= \mathbb{E} \left[ \int_t^T e^{-r(T-s)} F^Z(s, Y_s) ds \mid Y_t = y \right], \\
V^j(t, y, Z) &:= \sup_{\tau \in T} \mathbb{E} \left[ \int_t^{T \wedge \tau} e^{-r(T-s)} F^Z(s, Y_s) ds + SW^j(\tau, Y_\tau) \mid Y_t = y \right]
\end{align*}
\]

where \( t \in [0, T] \), \( \tau := \tau^4_j \) and \( j = 0, 1, 2, \ldots, M \) represents a given number of switches allowed, \( Z = \{ z, \zeta \} \) the switching indicators, \( y \) indicates the realization of our markovian vector dynamic \( Y_t = (X_t, \lambda_t) \) set in section 3.1) intended to be defined under the real/objective measure \( Q \) while \( SW(.) \) is the switching/intervention operator which guide the recursion and takes the following expression

\[
SW^l,Z=z(\tau, Y_\tau) := \max_{Z \in Z} \{ V^{l-1}(\tau, y, Z = \zeta) - c^{Z=z}_\tau \} = \{ V^{l-1}(\tau, y, Z = \zeta) - c^{Z=z}_\tau \},
\]

given that the maximum is trivial in our case with only two regimes \( \{ z, \zeta \} \). For what concerns our modeling choice for the stochastic vector dynamic, because of we need numerous simulation of the exposures of the given contract (represented by an IRS) not only at the valuation date \( t = 0 \) but also at forward times and given that the exposures depend - for what concerns the interest rate modeling part - on the libor/forward rate evolution, a libor market model dynamic would be the natural modeling choice. But given the complexity of the problem and the high number of simulation and valuation date, using a LMM dynamic would be a cumbersome numerical task. So, without great loss of generality, the choice has been in favor of a more manageable short rate model. In chapter three we set a CIR for the stochastic dynamic of \( X \) factor, but to ensure a

\[\text{Here we make an abuse of notation, but is intended that we assume similar conditions on default times } \tau \text{ and switching times as in chapter five.}\]

\[\text{We remind that the value function have to be intended always negative with a minus ahead, to give sense to the minimization problem and that the allowed switching time are finite because of the presence of switching costs.}\]
grater variety to the simulated term structure we use for the computations a *shifted two factor gaussian* short rate model (say G2++). Indeed, the dynamics for \( X \) becomes

\[
X_t := r_t = y_t + z_t + \varphi_t, \quad r(0) = r_0
\]

where the processes \( \{ y_t : t \geq 0 \} \) and \( \{ z_t : t \geq 0 \} \) satisfy

\[
\begin{align*}
\mathrm{d}y_t &= -\mu y_t \mathrm{d}t + \sigma \mathrm{d}W^1_t, \quad y(0) = y_0 \\
\mathrm{d}z_t &= -\nu y_t \mathrm{d}t + \eta \mathrm{d}W^2_t, \quad z(0) = z_0 \\
\mathrm{d}(y, z)_t &= \mathrm{d}(W^1, W^2)_t = \rho_{y,z} \mathrm{d}t
\end{align*}
\]

where \((W^1, W^2)\) is a two-dimensional \( \mathcal{Q} \)-Brownian motion with instantaneous correlation \( \rho \in [-1, 1] \), \( \varphi_t \) is a deterministic function that ensures the fitting on the initial market term structure and \( \{ r_0, \mu, \nu, \sigma, \eta \} \) are positive constant parameters representing the drifts and instantaneous volatilities of \( r_t = X_t \). Because of diffusion parameters do not change under change of measure and we are not under the pricing measure we have chosen to calibrate it on the historical volatility of the reference rate underlying the contract, typically the *six months euribor*. The remaining parameters are instead determined by standard calibration on the cap-floor quotes.

For what concerns the model for the default intensity \( \lambda_t \), the choice has been in favor of a typical *Cox process* with stochastic intensity \( \lambda \), assumed \( \mathcal{F}_t \)-adapted, positive and right continuous \( (\mathcal{F}_t^\lambda = \sigma(\{\lambda_s : s \leq t\})) \) having dynamic of CIR type given below

\[
\begin{align*}
\mathrm{d}\lambda_t &= \kappa(\gamma - \lambda_t) \mathrm{d}t + \nu \sqrt{\lambda_t} \mathrm{d}W^\lambda_t \\
\lambda(0) &= \lambda_0
\end{align*}
\]

where \( W^\lambda_t \) is also a standard \( \mathcal{Q} \)-Brownian motion. We also recall that the *cumulated intensity* or *hazard process* is defined as \( \Lambda(T) = \int_0^T \lambda(t) \mathrm{d}t \) which is a random variable (absolutely) continuous and increasing. In this framework \( \Lambda(T) \) play a central role for jump time \( \tau \) simulation given that the first jump time of the process, transformed through its cumulated intensity, is an exponential random \( \xi \) variable independent of \( \mathcal{F}_t \) (being part of the defaultable filtration \( \mathcal{H}_t \)), that is

\[
\Lambda(\tau) = \xi \implies \text{(inverting)}
\]

\[
\tau := \Lambda^{-1}(\xi)
\]

The procedure to simulate the default times \( \tau \) is the standard one (see for example Brigo, Mercurio, 2006). The vector process parameters \( (\kappa, \gamma, \nu) \) of positive constants, can be in general calibrated to credit/bond market spreads or CDS spreads of specific names. In particular, we will distinguish

\footnote{The system of correlated SDE is implemented by simulating independently the two Brownian motions and correlating the dynamics of \((x_t, y_t)\) imposing

\[
\begin{align*}
\mathrm{d}W^1_t &= \mathrm{d}W^1_t \\
\mathrm{d}W^2_t &= \rho \mathrm{d}W^1_t + \sqrt{1-\rho^2} \mathrm{d}W^2_t
\end{align*}
\]}

6. **Numerical solution approach and implementation**
between two different cases using the market quotes of two different names (labeled as LOW and HIGH counterparty risk) for volatilities and drift of the default intensities in order to investigate the effects on the value function and the optimal switching strategy of the problem. In the following tables we summarize the SDEs parameters and those related to the perfect collateralization regime and funding, the borrowing $r_{borr}$, opportunity costs $r_{opp}$ and for the switching costs $c_Z$ that we need to set in computations. The parameter $\delta$ is set equal to zero and to meaningful value in order to check the impact on the solution and on the optimal switching strategy. Here follows a table of the variables and the related parameters of our model.

<table>
<thead>
<tr>
<th>$X_t$</th>
<th>$\lambda$</th>
<th>Other parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>$\kappa$</td>
<td>$r_{opp}$</td>
</tr>
<tr>
<td>$\nu$</td>
<td>$\gamma$</td>
<td>$r_{borr}$</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>$\nu$</td>
<td>$R_c$</td>
</tr>
<tr>
<td>$\eta$</td>
<td>$\gamma$</td>
<td>$c_z$</td>
</tr>
<tr>
<td>$\rho$</td>
<td>$\epsilon$</td>
<td>$c_{\zeta\alpha}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\delta$</td>
</tr>
</tbody>
</table>

b) Discretization and implementation procedure.

From the last paragraph, we can observe that although the assumed system dynamics is not complicated, is difficult to recover a closed form expression for the expected value of the cost functions involved given their recursive definition. In particular, the present step in which we discretize the domain and show the procedure adopted for the implementation is central to understand the algorithm used to search for the solution of our switching control problem. So, let us define for first the temporal discretization of the time domain of the problem that is take as given, depending on the underlying contract maturity $T$. Set in $t = 0$ the valuation time, on the time domain $[0, T]$ we define the equally spaced time discretization step $\Delta t = \frac{T}{N}$, where $N$ is the total number of steps, in order to get a good approximation for the (discretized) vector dynamic $(X, \lambda)$. In particular, given the number of variables involved, for a maturity $T \geq 5$ years is advisable to use at least a weekly step, while with $T < 5$ years we use a daily step. Formally let us define the time domain discretization grid as follows

$$DG_t^N := \left\{n \Delta t, n = 1, \ldots, N; \Delta t = \frac{T}{N}\right\}.$$ 

The other relevant variables that we need to consider in the discretization are the payment times set which is defined according to the underlying contract, the switching time set which can be taken different from the collateralization call dates set, and the default times set.

For what concerns the payment set, this in in general defined as a set of dates in which the contract fix a dividend/cashflow payment for the parties. Setting $N_{Pay} = T$ the last payment at maturity and each payment $p$ with equal tenor over $[0, T]$, we define formally

$$DG_{Pay}^N := \left\{\text{tenor} \ast p = 1, \ldots, N_{Pay}\right\},$$
which is clearly a set strictly contained in the latter, that is $DG_{Pay}^\Delta \subset DG_t^\Delta$.

As regards the collateralization dates set, this is generally a subspace of dates contained in the discretization grid set. They can be also taken as given because set within the CSA. In our model in which collateralization can be activated (and switched off) at any time over the contract horizon, this will coincide with the related switching time set, namely

$$DG_{Coll}^\Delta = SW_{\zeta}^\Delta := \left\{ \tau_j \in T, \zeta_j = 1 \right\}$$

that is also a subset of $DG_t^\Delta$. In order to simplify interpolation between the different time set but also to make finer the switching boundary, we adopt as working hypothesis that at every time step $t \in DG_t^\Delta$ the counterparty will take her switching decision so that we can state the following chain of relation

$$DG_{Coll}^\Delta \subseteq SW_{\zeta}^\Delta \subset SW^\Delta \subseteq DG_t^\Delta$$

where

$$SW_{\zeta}^\Delta := \left\{ \tau_j \in T, z_j \in Z, j = 1, \ldots, M \right\} = SW_{\tau}^\Delta + SW_{z}^\Delta$$

The last set to consider in the picture is that of the default time $\tau$. As we know, it affects only the "full CVA" cost function $F_z(y, \alpha)$ and in the literature the default times are usually simulated and bucketed on the payment time set. Anyway, in our model the counterparty can switch at any time between zero or full collateral so that the default intensities need to be simulated at the same time of $SW^\Delta$, that is of the time grid $DG_t^\Delta$. So given

$$DG^\lambda = \left\{ \tau \in [0, T] \right\}$$

the above set chain of relations becomes

$$DG_{Coll}^\Delta \subseteq SW_{\zeta}^\Delta \subset SW^\Delta \subseteq DG^\lambda \subseteq DG_t^\Delta. \quad (6.2)$$

Once we have defined the time discretization and the related random time modeling choice, we can pass to describe the whole procedure followed to calculated the paths for the forward (conditional expectation) exposures and NPV of the underlying claim.

b1) First step deals with the implementation of a discretization scheme for our stochastic vector $(X_t = r_t, \lambda_t)_{t \in [0, T]}$. We recommend to try both a (first order) Euler discretization and a Milestein (second order) scheme or other finer schemes to check the difference in the discretization error. In particular, to reduce the bias due to this type of error, one can try to simulate the exact transition density for the CIR process $\lambda_t$. Details on the algorithm used to implement it can be found in Glasserman (2007).

b2) As already highlighted, we start by simulating the short rate dynamics $r_t$ under the objective measure $Q$ (so that in the drift parameters is hidden also the market price of risk that we omit to explicit to ease notation). Once get the matrix of simulated paths at discrete times $DG_t^\Delta$, we would need to recover the expected exposures $NPV^+$ and $NPV^-$ at the possible
6. Numerical solution approach and implementation

future switching times in $SW^\Delta$ (and also default times which are a subset of it), that is formally

$$\mathbb{E}_t\left[(NPV_s)^\pm | \mathcal{F}_t\right], \ \forall s \in SW^\Delta$$

and $s > 0$ (where $s = t = 0$ is our valuation time).

This would impose to calculate backward over the paths the conditional expectation of the positive and negative exposures using the orthogonal projections over a set of approximating basis functions $\Phi_i \in L^2$, namely a least squares minimization as in Longstaff-Schwartz, but doing it two times with different basis function to regress on the exposures. Fortunately, thanks to the symmetry proposition that we state below at point b5) we can apply this recursive estimation procedure just one time for the $NPV$ reducing the computational costs.

In particular, we note that this conditional expectations that we need to calculate forward in time are the so called continuation values of the Longstaff-Schwartz algorithm, so that a modified and simpler procedure is needed.

b3) The procedure followed is quite standard. Because of it is relevant the libor rate $L_{t,t+\vartheta}$ for the calculation of the $NPV$ of our claim, from the simulated paths of $r_t$ of point a), we recover the discount factors $P(t,t+\vartheta)$ at the same date of the set $DG^\Delta_t$ using the available closed formula for the G2++ model\(^7\). After that, one can recover the forward libor rate for the tenor $\vartheta$ (defined by the contract dividend payments) by imposing

$$L(t,t+\vartheta) = \frac{1}{\vartheta} \left( \frac{P(t,t)}{P(t,t+\vartheta)} - 1 \right) \ \forall \ t \in DG^\Delta_t.$$

An alternative is to run the simulations paths for $r_t$ under the forward measure $\mathcal{Q}^T$ that allows to get the right quantities to use for the $NPV$ calculations.

b4) The subsequent step once all these paths have been simulated, is to initialize from the end/maturity $T$ the recursive procedure used to calculate the expected value of the $NPV$\(^8\) (default free) conditional on the information available at every time step of the recursion. Formally, we have to run the following program

$$\begin{cases}
N_{PV}(T,\omega) = (L(T,\omega) - k) = (X_T(\omega) - k), \ \forall \omega \in \mathbb{N}; \\
\mathbb{E}\left[N_{PV}(t_n)|\mathcal{F}_{t_{n-1}}\right] \cong \sum_{i=1}^{N_b} \beta_i \Phi_i(X_{t_{n-1}}), \ \forall t_n \in DG^\Delta_t 
\end{cases}$$

(and path $\omega$) where $\beta_i$ are $\mathbb{R}$-valued coefficients and $N_b$ represents the number of basis function $\Phi_i$, elements of the Hilbert space $L^2$, chosen to approximate the continuation value\(^9\) between two given discretization dates. This program takes to find the optimal parameters $\beta_i$ which minimizes for every step of the recursion the following least square optimization

$$\min_\beta \left\| Y(\omega_1) - \Phi(X(\omega_1))\beta_1, \ldots, Y(\omega_K) - \Phi(X(\omega_N))\beta_n \right\|^2$$

\(^7\) See Brigo, Mercurio for details about the formula.

\(^8\) Here we are making an abuse of notation with respect to the first chapter setting $NPV_t = S_t^{\text{rf}}$.

\(^9\) That is the expected value that derive from not switching and wait until the next switching time.
for $K \leq N$, where $Y(.)$ represents the continuation value of the NPV already calculated in the step before. Applying this procedure until time zero one gets the paths for the future expected values for the NPV of the contract.

We underline, in particular, the importance to keep trace of the set of payments dates in the recursion and the choice of suitable basis functions. Typical choices are the Laguerre, Legendre and Chebischev polynomials with order not greater than five/six to obtain good results. We also recommend the use of polynomials that recall the payoff of the underlying contract and the use of techniques like Tichonov regularization or singular value decomposition in order to deal with the problem of singularity of the inverse matrix that may arise in the regressions.

b5) The last step consists in calculating the simulated paths for the BCVA$_t$ process and the collateral and funding costs process which emerges when the counterparty switches to full collateralization. Here are central the hypothesis done above, when formulating the model, in relation to the orthogonality between the default intensities $\lambda_t$ and the (pre-default) price process of the claim which is governed by $X_t$ and in addition the symmetry hypothesis that helps to simplify calculations.

In order to proceed in this plan, we firstly run a Monte Carlo simulation for the default intensities and times as described in b1) and bucketing them all on $DG^\Delta_i$ set. Then, we multiply them path by path, thanks to orthogonality condition $X_t \perp \lambda_t \Rightarrow NPV_t \perp \lambda_t$, for the expected NPV paths calculated in b4). Now, recalling the expressions for the cost functions in the two switching regimes $\{z_j, \zeta_j\}$, having as payoff an in interest rate swap swap

$$ F_z(y, \alpha) := \left[ BCVA(t) - \delta \right]^2 = \left[ (CA(t) - DVA(t)) - \delta \right]^2 $$

$$ = \left[ \left( \int_{s=0}^{T\wedge r_j} (1 - R_c) \left[ \left( \sum_{s=1}^{T} B_u \xi_u(X_u - k) \right)^+ \right] \lambda_u ds + \right. \right. $$

$$ \left. \left. - \int_{s=0}^{T\wedge r_j} (1 - R_c) \left[ \left( \sum_{s=1}^{T} B_u \xi_u(X_u - k) \right)^- \right] \lambda_u ds \right) - \delta \right]^2 $$

$$ \left(6.3\right) $$

$$ F_{\zeta}(y', \alpha') := \left[ \left( \int_{s=0}^{T\wedge r} R(s)[NPV(u)] ds - NPV(s) \right) - \delta \right]^2 $$

$$ = \left[ \left( \int_{s=0}^{T\wedge r} R(s) \sum_{s=1}^{T} B_u \xi_u[X_u - k] ds - \sum_{s}^{T} B_s \xi_s[X_s - k] \right) - \delta \right]^2 $$

$$ \left(6.4\right) $$

which are proved to satisfy the properties stated in propositions 4.1 and 4.2, proceeding in the same backward manner and using the same symmetry proposition that we state below, we can show that for two consecutive time step $t_i, t_{i+1} \in DG^\Delta_i \cup DG^\lambda (i = n \leq N \in DG^\Delta)$ and with $s$ payments remaining until $T$, the following discretized relations can be implemented for $BCVA_t$

$$ BCVA^\Delta(t_i) = (1 - R_c) E \left[ NPV(t_{i+1}) | F_{t_i} \right] \Delta t \lambda(t_i, t_{i+1}) + BCVA(t_{i+1}) $$

$$ = (1 - R_c) \sum_s B_t \xi_s(L(t_i, t_{i+s}) - k) \big( t_{i+1} - t_i \big) \lambda(t_i, t_{i+1}) + BCVA(t_{i+1}) $$

$$ \left(6.5\right) $$
∀\(t_i\in DG_t^\Delta\) and paths \(\omega\), with \(\{t_i \leq \tau\}\) (namely means that the default time can be verified and bucketed at the valuation time \(t_i\) or not verified) and \(\{\tau_j \leq t_i < \tau_{j+1}\}\) and \(\{z_j = 1\}\). Of course, BCVA\((t_i+1)\) is already known from the precedent step of the recursion and the conditional expectation term is already known from the calculations done in step b4).

Similarly, the discretization for collateral and funding cost would be as follows (here no default intensities simulation are needed)

\[
\text{Coll}^c(t_i) = R_{\text{fund}}(t_i+1-t_i)\mathbb{E}\left[\text{NPV}(t_i+1)\big|\mathcal{F}_{t_i}\right] \Delta t + \text{NPV}(t_i+1)
\]

\[
= R_{\text{fund}}(t_i+1-t_i) \sum_s B_t \xi_t(L(t_i,t_{i+s})-k)(t_{i+1}-t_i) + \text{Coll}^c(t_i+1)
\]

∀\(t_i\in DG_t^\Delta\) and paths \(\omega\), with \(\{t_i \leq \tau\}\) and \(\{\tau_j \leq t_i < \tau_{j+1}\}\) and \(\{z_j = 0\}\). We conclude by stating here the already invoked symmetry proposition that has allowed to get the above discretized representations.

**Proposition 5.2.1 (Symmetry).** Assuming to use the same MC run for calculations, under the orthogonality hypothesis \(X \perp \lambda\) and the assumptions on the default intensities of counterparties such that \(\lambda^A_t = \lambda^B_t\) and \(R^A_c = R^B_c\), applying the Longstaff-Schwartz algorithm for the expected NPV values instead of its singular application to the positive and negative exposure is equivalent and enough to recover the BCVA paths\(^{10}\).

**Proof.** The proof relies on simple considerations on the structure of the simulated paths in the two cases and on BCVA definition. Given the hypothesis to use the same Monte Carlo run to estimate the expected exposure \(\text{NPV}^+\) and \(\text{NPV}^-\), by Definition 2.2.4 we know that these contribute to the \(CVA\) and \(DVA\) terms (respectively) whose sum (whatever is the sign that depends on the perspective of one of the counterparty) is, by definition 2.2.3 and 2.2.4 equal to BCVA. But it is easy to see that being \(\text{NPV} = \text{NPV}^+ + \text{NPV}^-\) using its paths instead of the split ones would give the same path by path contributions to BCVA so that the estimate would be equivalent given that, by assumptions \(\lambda^A_t = \lambda^B_t\) and \(R^A_c = R^B_c\), also the paths for the default intensities and the recovery rate are the same ⋄.

2) Value function and optimal switching strategy calculation procedure.

Once we have obtained the simulated paths for the cost functions we are ready to apply the iterative procedure over the \(l\) switching times (assumed finite \(l \in \{1, \ldots, M = \bar{l}\}\) ) based on the dynamic programming principle in order to compute the (discretized) value function \(V^l(t,z,y)\) over the set \(DG_t^\Delta = SW^\Delta\). The program to implement is the following.

\(^{10}\) Actually also the factor \(R_{\text{fund}}\) is intended to be equal for both counterparties.
a) Let us consider for simplicity two calculation times $t_1$ and $t_2$, thanks to the stated optimality principle, the counterparty has to decide between immediate switch at $t_1$ to the other regime $z$ (or $\zeta$) versus no switching and wait until $t_2$. So, discretizing the recursive value function equation, the following program has to be run

\[
V^l(t_1, \mathcal{Y}_{t_1}, Z) = \max \left( \mathbb{E}\left[ \int_{t_1}^{t_2} F^Z(s, \mathcal{Y}_s) ds + V^l(t_2, \mathcal{Y}_{t_2}, Z)|\mathcal{F}_{t_1} \right], \mathbb{SW}^l Z(t_1, \mathcal{Y}_{t_1}) \right)
\]

\[
\simeq \max \left( F^Z(t_1, \mathcal{Y}_{t_1}) \Delta t + \mathbb{E}\left[ V^l(t_2, \mathcal{Y}_{t_2}, Z)|\mathcal{F}_{t_1} \right], \{V^{l-1}(t_1, \mathcal{Y}_{t_1}, Z = \zeta) - c^{Z=z}_{t_1}\} \right)
\]

formulation that becomes in our cost minimization problem as follows

\[
V^l(t_1, \mathcal{Y}_{t_1}, Z) \simeq \min \left( F^Z(t_1, \mathcal{Y}_{t_1}) \Delta t + \mathbb{E}\left[ V^l(t_2, \mathcal{Y}_{t_2}, Z)|\mathcal{F}_{t_1} \right], \{V^{l-1}(t_1, \mathcal{Y}_{t_1}, Z = \zeta) + c^{Z=z}_{t_1}\} \right)
\]

where here $BCVA^\Delta_{t_1} = \mathcal{Y}_{t_1,z}$ (and similarly $Coll^c_{t_1} = \mathcal{Y}_{t_1,\zeta}$), so that the last one can be made more explicit as follows

\[
\begin{cases}
\text{if in } t_2 \{Z = z\} \implies \quad V^l(t_1, \mathcal{Y}_{t_1}, Z) \simeq \min \left( (BCVA^\Delta_{t_1} - \delta)^2 \Delta t + \mathbb{E}\left[ V^l(t_2, \mathcal{Y}_{t_2}, z)|\mathcal{F}_{t_1} \right], \{V^{l-1}(t_1, \mathcal{Y}_{t_1}, \zeta) + c^z_{t_1}\} \right) \\
\text{if in } t_2 \{Z = \zeta\} \implies \quad V^l(t_1, \mathcal{Y}_{t_1}, Z) \simeq \min \left( (Coll^c_{t_1} - \delta)^2 \Delta t + \mathbb{E}\left[ V^l(t_2, \mathcal{Y}_{t_2}, \zeta)|\mathcal{F}_{t_1} \right], \{V^{l-1}(t_1, \mathcal{Y}_{t_1}, z) + c^\zeta_{t_1}\} \right)
\end{cases}
\]

(6.7) \, (6.8)

Everything in this iterative procedure is known from the calculations done in the latter phase except the conditional expectation of the cost function $\mathbb{E}[\cdot|\mathcal{F}_t]$ which represents the value of waiting - like the well known holding/continuation value - until the next switching time remaining in the same regime.

At this point, to calculate this conditional expectation we have chosen to reapply the Longstaff-Schwartz algorithm to the simulated paths $(BCVA^\Delta$ and $Coll^c)$ that enter our cost functions which we already know from the pre-calculation phase. Now, the procedure follows the same lines of the one described at point b4), what changes is the type of basis function -say $\Psi^l$ - used in the estimate of the expectation, given that the cost functions are of quadratic type, that is formally

\[
\begin{cases}
\mathbb{E}\left[ V^l(t_n, \mathcal{Y}_{t_n}, z)|\mathcal{F}_{t_{n-1}} \right] \simeq \sum_{i=1}^{N_h} \beta_i \Psi^{l,z}(\mathcal{Y}_{t_{n-1}}) , \forall t_n \in DG^\Delta_t \\
\mathbb{E}\left[ V^l(t_n, \mathcal{Y}_{t_n}, \zeta)|\mathcal{F}_{t_{n-1}} \right] \simeq \sum_{i=1}^{N_h} \beta_i \Psi^{l,\zeta}(\mathcal{Y}_{t_{n-1}}) , \forall t_n \in DG^\Delta_t
\end{cases}
\]
This is the most delicate part of the procedure in which a careful choice of the basis functions is needed to get meaningful results. Once this is done, the iterative procedure can be started firstly initializing the value function at the maturity setting \( V^l(T, Y_T, Z) = 0 \) for all \( l, z \) and then going backward computing iteratively the continuation values \( \mathbb{E}[V^l(t_n, Y_{t_n}, z)|F_{t_{n-1}}] \) along each paths. Then adding the term \( F^Z(t_n, Y_{t_n})\Delta t \) and taking the minimum as in equation 6.7 (or 6.8), averaging and discounting back to inception \( t = 0 \) one obtains the value function for the problem. By keeping trace of the paths in which is optimal to switch/not switch one can easily recover the optimal switching strategy \( \{\tau_j^*, Z_j^*\}_{j \geq 1} \).

Remarks 6.3. Is worth of mention that the above general solution representation set the number of switches allowed \( l \) as given/fixed. This complicate computations given that for every path and time step one has to compute forward for \( l = 0, \ldots, \bar{l} \) the values for \( V^{l-1}(.) \) that enters in the optimal backward recursion as the cost/reward that counterparty gets by switching to the other regime incorporating also the minor flexibility due to less switches allowed. Alternatively, one can set \( l = \infty \) that is the counterparty has the right to switch whenever she wants until the maturity of the underlying contract. This allows to flatten the recursion procedure and reduce sensibly the computational costs.

b) In addition or alternatively to the procedure shown above, one can keep trace of the smallest optimal switching time \( \tau^l(t_n, Y_{t_n}, Z) \) that is also computed in backward manner and represents for every path the smallest time in which is optimal to switch to an other regime. The formal (recursive) definition is the following:
\[
\tau^l(t_n, Y_{t_n}, Z) = \begin{cases} 
\tau^l(t_{n+1}, Y_{t_{n+1}}, Z), & \text{no switch} \\
n, & \text{switch}
\end{cases}
\]
and the value function will be given by the sum path by path of the future costs (discounted) until the optimal switching times, namely
\[
V^l(t_n\Delta t, Y_{t_n\Delta t}, Z) = \mathbb{E}\left[ \sum_{j=n}^{\tau^l} F^Z(j\Delta t, Y_{j\Delta t})\Delta t + SW^l, Z(\tau^l\Delta t, Y_{\tau^l\Delta t})|F_{t_n\Delta t}\right]. \tag{6.9}
\]
We conclude by noting that the last formulation for the value function allows faster computations because here the regressions are not stored, they are just used to update the switching times.

c) The last thing that we need to highlight is how to recover the optimal switching boundary for our problem. It represents the graph \( (t, Y_t) \) of the boundaries at which our optimal switching strategy changes regime (for each \( l, Z \)). Keeping track of the minimal switching times \( \tau^l \), one can reconstruct the switching boundary by drawing, at the end of the algorithm, the graph of \( \tau^l(0, Y_0, Z) \) against \( Y_t \). Formally, the set
\[
\{Y_{n\Delta t} : \forall n | \tau^l(0, Y_n, Z) = n\} \tag{6.10}
\]
is defined as the empirical region of switching from a given regime $Z$ at time $n\Delta t$. It is important to recover the switching boundaries especially in a risk management view, as one can construct and check the optimal switching policies by simulating paths of $Y$ and verifying the minor costs and benefits obtained by the implementation of the switching policy$^{12}$.

So, after the description of the algorithm that we have decided to adopt to solve our model we pass to show some of the results obtained from the numerical computations.

### 6.4 Numerical implementation and examples.

In this section we pass to show and analyze the main results of the implementation of the algorithm above defined for the solution of our stochastic switching control problem. From the theory of chapter 4 and 5, we are sure of the existence of a solution for our problem and also of the optimal switching strategy, although this strategy can also reveal to be of "banal type" (as stated in Definition 4.6).

As we have already highlighted in Proposition 4.5 and Proposition 4.8, the existence of "non banal switching solution" would depend on "how distant" are paths by paths over time the cost functions of our two regimes and, in particular, under the "symmetry hypothesis", the parameters and variables of relation 4.19 should be determinant.

More specifically, in addition to the switching costs whose level can determine the convenience to switch, we expect that, the volatility of the stochastic factors of our model, in particular the volatility and drift parameters of the default intensities would have a relevant impact on the switching strategy of the counterparty that is brought to switch more often.

In our model the funding costs are modeled as deterministic; of course in reality a counterparty which sees its default intensities increase see also the credit worthiness get lower with a likely increase also in the cost of funding. In this sense, one should need a stochastic model also for funding costs correlated with default intensities or eventually including also liquidity spreads issues. Here we work with a simpler deterministic hypothesis, which is enough to get some results, but with our numerical approach MC based it would not be difficult to generalize it adding an other stochastic factor (of course one should be prepared to deal with the increasing computational costs!); so we leave the investigations for further research.

In the following we focus the analysis on these variables and we report the results of the implementation of the algorithm defined in section 6.3 applied to a simple defaultable interest rate swap, EURIBOR6m vs fixed rate$^{13}$ traded at par and maturing in one year $T = 1\text{Y}$ from the inception $t$. The data used in computations are taken at the reference date of 2012/06/15 from the main international data provider (Bloomberg and Reuters).

---

$^{12}$ Also respect to the one that would have been implemented without the flexibility of the switching type mechanism

$^{13}$ For convenience the setting of the rate is in arrears.
Here we report the market yield curve build on monetary rates (EURIBOR RATES) before the year and on the swap rates after the year, and the ATM (flat) implied Cap volatilities.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Id</th>
<th>Spot Rates</th>
<th>Discount Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1m</td>
<td>EUR030</td>
<td>0.382%</td>
<td>0.9997</td>
</tr>
<tr>
<td>3m</td>
<td>EUR090</td>
<td>0.662%</td>
<td>0.9983</td>
</tr>
<tr>
<td>6m</td>
<td>EUR180</td>
<td>0.939%</td>
<td>0.9953</td>
</tr>
<tr>
<td>1y</td>
<td>EUR360</td>
<td>1.226%</td>
<td>0.9879</td>
</tr>
<tr>
<td>2y</td>
<td>EURS02</td>
<td>0.876%</td>
<td>0.9827</td>
</tr>
<tr>
<td>3y</td>
<td>EURS03</td>
<td>0.985%</td>
<td>0.9710</td>
</tr>
<tr>
<td>4y</td>
<td>EURS04</td>
<td>1.151%</td>
<td>0.9553</td>
</tr>
<tr>
<td>5y</td>
<td>EURS05</td>
<td>1.331%</td>
<td>0.9360</td>
</tr>
<tr>
<td>7y</td>
<td>EURS07</td>
<td>1.632%</td>
<td>0.8929</td>
</tr>
<tr>
<td>10y</td>
<td>EURS10</td>
<td>1.949%</td>
<td>0.8245</td>
</tr>
<tr>
<td>12y</td>
<td>EURS12</td>
<td>2.089%</td>
<td>0.7802</td>
</tr>
<tr>
<td>15y</td>
<td>EURS15</td>
<td>2.204%</td>
<td>0.7211</td>
</tr>
<tr>
<td>20y</td>
<td>EURS20</td>
<td>2.211%</td>
<td>0.6457</td>
</tr>
<tr>
<td>25y</td>
<td>EURS25</td>
<td>2.202%</td>
<td>0.5802</td>
</tr>
<tr>
<td>30y</td>
<td>EURS30</td>
<td>2.193%</td>
<td>0.5217</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Market implied vol</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Y</td>
<td>94.2%</td>
</tr>
<tr>
<td>1.5Y</td>
<td>93%</td>
</tr>
<tr>
<td>2Y</td>
<td>90.6%</td>
</tr>
<tr>
<td>3Y</td>
<td>67.7%</td>
</tr>
<tr>
<td>4Y</td>
<td>63.7%</td>
</tr>
<tr>
<td>5Y</td>
<td>61.2%</td>
</tr>
<tr>
<td>6Y</td>
<td>59.1%</td>
</tr>
<tr>
<td>7Y</td>
<td>57.5%</td>
</tr>
<tr>
<td>8Y</td>
<td>56.2%</td>
</tr>
<tr>
<td>9Y</td>
<td>54.8%</td>
</tr>
<tr>
<td>10Y</td>
<td>53.6%</td>
</tr>
<tr>
<td>12Y</td>
<td>51.6%</td>
</tr>
<tr>
<td>15Y</td>
<td>49.6%</td>
</tr>
<tr>
<td>20Y</td>
<td>48.8%</td>
</tr>
<tr>
<td>25Y</td>
<td>47.7%</td>
</tr>
<tr>
<td>30Y</td>
<td>46.4%</td>
</tr>
</tbody>
</table>

Being traded at par, the strike rate used is the one year par swap rate $S01 = 0.91\%$. In order to keep a good representation of the reality by the model and to deal also with computational costs,
Numerical solution approach and implementation

the discretization step used to simulate the vector dynamics and to update the switching strategy, has been chosen to be daily, that is $\Delta t = \frac{1}{N}$ where $N = 252$ (counting only the working days). The parameters of the $G2++$ process for $X_t$ calibrated on market implied (Black) volatilities\textsuperscript{14} are the following

<table>
<thead>
<tr>
<th>Calibrated parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>0.00013</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.06730</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.12924</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0.14014</td>
</tr>
<tr>
<td>$\rho$</td>
<td>$-0.99948$</td>
</tr>
</tbody>
</table>

while the daily historical volatility of the EURIBOR6m, calculated over the last year data sample, is $\sigma_{hist} = 0.12654$. Excluding the default intensities on which we have decided to focus the analysis of the value function behavior, in calculations we have set the remaining parameters as follows:

- the risk-free rate $r_{free}$ being near to zero, if one considers as reference the 1Y yield from German bond, is set simply as constant $r_{free} = 0$;

- the borrowing rate $r_{barr}$ should consider the particular balance-sheet and risk condition of the counterparty. For convenience we assume it as constant set to $r_{barr} = 0.01$;

- similar considerations are valid for the opportunity rate $r_{opp}$ which is set equal to $r_{opp} = 0.03$ by considering on average the rate of return of no risk free investments in the current stagnant market situation;

- the instantaneous switching costs $c_z$ and $c_\zeta$ are set as constants not greater than the 2% of the notional of the underlying contract. Anyway, we show their impact on the switching strategy by varying their values;

- the recovery rate parameter $R_c$ is assumed fixed and equal to 40% loss recovery if counterparty defaults;

- the parameter $\delta$ which enters in the cost functions is also set to zero\textsuperscript{15} $\delta = 0$ and then varied.

\textsuperscript{14}The calibration procedure is a standard minimization of sum of the square percentage difference between the model and the market price, namely:

$$\min_{\theta} \sum_{i=1}^{n_{\text{payments}}} \left\| \frac{P_{\text{mkt}}(t,i) - P_{\text{model}}(t,i,\theta)}{P_{\text{mkt}}(t,i)} \right\|^2 .$$

\textsuperscript{15}This means that counterparty would be adverse to every variation from zero of the running costs.
As already mentioned, for what concerns the default intensities parameters we have calibrated the model parameters \((\kappa, \gamma, \lambda_0)\) on the following CDS spreads quotes of two bank names (at the same date 15/6/2012)

<table>
<thead>
<tr>
<th>Bucket</th>
<th>Dexia</th>
<th>DB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Y</td>
<td>915.716</td>
<td>93.648</td>
</tr>
<tr>
<td>3Y</td>
<td>817.403</td>
<td>156.834</td>
</tr>
<tr>
<td>5Y</td>
<td>782.0214</td>
<td>196.917</td>
</tr>
<tr>
<td>7Y</td>
<td>825.052</td>
<td>206.326</td>
</tr>
<tr>
<td>10Y</td>
<td>796.494</td>
<td>214.172</td>
</tr>
</tbody>
</table>

given the intention to investigate the results in case of HIGH (Dexia Subordinated) and LOW (Deutsche Bank (Senior), DB) default risk level for the counterparty. The calibration results are reported in the following table with the volatility parameter \(\upsilon\) which is set taking the (daily) CDS historical volatility over the last two year.

<table>
<thead>
<tr>
<th>(\lambda_t) parametrization</th>
<th>level</th>
<th>LOW</th>
<th>HIGH</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\kappa)</td>
<td>1.03921</td>
<td>0.30821</td>
<td></td>
</tr>
<tr>
<td>(\gamma)</td>
<td>0.02120</td>
<td>0.11220</td>
<td></td>
</tr>
<tr>
<td>(\upsilon)</td>
<td>0.20122</td>
<td>0.44214</td>
<td></td>
</tr>
<tr>
<td>(\lambda_0)</td>
<td>0.04031</td>
<td>0.20316</td>
<td></td>
</tr>
</tbody>
</table>

**Results in HIGH CASE.** From the application of the algorithm described in the last section and with the parametrization set as follows:

- \(\lambda_t = \text{HIGH}\);
- \(c_z = 0\) and \(c_\zeta = 0\);
- \(\delta = 0\);
- \(Z = \zeta = 0\) in \(T\), at maturity;
- \(N_{\text{paths}} = 1000\) which indicates the number of simulated paths; these are enough to derive the optimal switching strategy and to understand the value function behavior. In fact the problem has been tackled from a risk management view so that from a pricing view a greater number of paths should be run (other than a change of measure under which make calculations).

Given the other parameters as defined above, we get the following main results.

In Fig 6.1 we have plotted the switching indicators \(Z = \{0, 1\}\) that are determined optimally by the application of the iterative switching algorithm (equation 6.7 and 6.8) over the time (discretized) domain \(t \in [0, 252]\). It shows almost four region in which the switching over the paths and time have a daily frequency but also a big mass given the high number of paths in which results to be optimal to switch and switch back often. This was expected given the higher volatility of the default intensities and the fact that the instantaneous switching costs \(c_Z\) are set to zero. This behavior can be seen also from Fig 6.2 in which we have plotted
6. Numerical solution approach and implementation

Fig. 6.1: Graph of the switching indicators $Z = \{0, 1\}$ over time $t \in [0, 252]$

Fig. 6.2: Graph of the minimum number of remaining switches for all the paths $N_{\text{paths}}$ over time $t \in [0, 252]$

the minimum number of remaining (optimal) switches registered over the paths and time (as time elapses). Being the total number of switches equal to the number of time steps (253), by time elapsing also the number of switches decreases and the plot clearly shows in correspondence of flat lines and (negative) jumps the passage from of a period of non activity in terms of switching to a period of activity.

We also report in Fig. 6.3 the graph of the optimal switching boundaries for our problem as defined in equation 6.9. The upper line shows the optimal values of the paths of $Y^z$ above which it becomes optimal to switch, while the lower line indicates the paths for $Y^x$ above
6. Numerical solution approach and implementation

which is not optimal to switch. So the area between the two boundaries can be defined as the "non action/switching area".

This boundaries are central in the risk management view, given that by confronting these values with that of simulated paths of the stochastic vector dynamic that guide a given underlying claim one can study and define an optimal strategy that maximizes the hedging portfolio wealth or minimizes the related costs.

In the following table we summarize the results of the value function (with initial condition $\zeta = 0$) for our problem, with a notional (of the underlying contract) $Notional = 1000$

<table>
<thead>
<tr>
<th>Results ($\lambda_t = HIGH, c_Z = 0$)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$V(t, u)^*$</td>
<td>0.0755</td>
</tr>
<tr>
<td>$V(t, u)^{Cva}$</td>
<td>0.2452</td>
</tr>
<tr>
<td>$V(t, u)^{Coll}$</td>
<td>0.1116</td>
</tr>
</tbody>
</table>

In the table we have also reported the value function values in the case of no switch allowed, namely $V(t, u)^{Cva}$ is referred to the case $z = 1$ (no collateralization) $\forall t \in [0, T]$ and $V(t, u)^{Coll}$ is referred to the complement $z = 0$ ($\zeta = 1$, full collateralization). As one could have expected, the CVA term $V(t, u)^{Cva}$ has the greatest value given the higher default intensities, while the value function $V(t, u)^*$ with the contingent switching collateralization allows to minimize the costs (as defined above) given the greater flexibility of the mechanism with consistent savings over the funding/collateralization and CVA costs, especially if one considers higher level of contract’s notional.

In the following we report similar results as shown above highlighting the variation of the instantaneous switching costs $c_Z$ and the effects on value function.
6. Numerical solution approach and implementation

Fig. 6.4: Graph of the switching indicators $Z = \{0, 1\}$ over time $t \in [0, 252]$, with $c_z = c_\zeta = 0.01$

<table>
<thead>
<tr>
<th>Results ($\lambda_t = \text{HIGH}, c_z = c_\zeta = 0.01$)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$V(t, u)^*$</td>
<td>0.1014</td>
</tr>
<tr>
<td>$V(t, u)_{Cva}$</td>
<td>0.2452</td>
</tr>
<tr>
<td>$V(t, u)_{Coll}$</td>
<td>0.1116</td>
</tr>
</tbody>
</table>

Fig. 6.5: Graph of the switching indicators $Z = \{0, 1\}$ over time $t \in [0, 252]$, with $c_z = c_\zeta = 0.05$
6. Numerical solution approach and implementation

<table>
<thead>
<tr>
<th>Results ( $\lambda_t = \text{HIGH}, c_z = c_\zeta = 0.05$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V(t, u)^*$</td>
</tr>
<tr>
<td>$V(t, u)^{Cva}$</td>
</tr>
<tr>
<td>$V(t, u)^{Coll}$</td>
</tr>
</tbody>
</table>

Fig. 6.6: Graph of the switching indicators $Z = \{0, 1\}$ over time $t \in [0, 252]$, with $c_z = 0.002$ $c_\zeta = 0.001$

From these graphs and results, we can see that by increasing the instantaneous costs of switching reduces the optimality of switching from a regime to the other as shown in Fig. 6.4 and 6.5 in which the vertical lines that indicates the outcome of switching indicators over paths and time are very rare or an horizontal line which indicates zero switches; in particular the tables shows how the value functions tends to increase and converge to the case that we have defined in chapter four as banal switching strategy, that corresponds also to the value function of the case of not switching $V(t, u)^{Coll} = V(t, u)^*$. In Fig. 6.6 we have reported a special case in which $c_z > c_\zeta$ that lower more the value function of our problem, even respect to the precedent case with $c_z = c_\zeta = 0$; in fact one would expect a value function greater than the case with zero switching costs, but due to its nonlinearity we have verified cases like that.

We made also some trial by increasing the value of $\delta$; in general this increases the cost functions and reduce the convenience to switch but with different setting of the switching costs the effect on the value function is not obvious. So, it is important to study each case carefully. From a risk management point of view, it is relevant to know or to have an estimate of the instantaneous switching costs in order to determine the optimal savings through the flexibility of switching mechanism.
6. Numerical solution approach and implementation

<table>
<thead>
<tr>
<th>Results ( $\lambda_t = \text{HIGH}, c_z = 0.002, c_\zeta = 0.001$)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$V(t, u)^*$</td>
<td>0.0628</td>
</tr>
<tr>
<td>$V(t, u)^{Cva}$</td>
<td>0.2452</td>
</tr>
<tr>
<td>$V(t, u)^{Coll}$</td>
<td>0.1116</td>
</tr>
</tbody>
</table>

**Results in LOW CASE.** In this case the parametrization set is the following:

- $\lambda_t = \text{LOW}$;
- $c_z = 0$ and $c_\zeta = 0$;
- $\delta = 0$;
- $Z = \zeta = 0$ in $T$, at maturity;
- $N_{\text{paths}} = 1000$.

Keeping the other parameters as set above, we get the following main results.

In this case, although the switching costs $c_Z$ are null, we can see from Fig. 6.7 that the lower default intensities $\lambda_t$ determine the optimality for counterparty to switch only in the last part of the time domain where the expected cost functions of the two regimes turn to be really near path by path. This can be also easily checked by the plot of the minimum number of remaining switches through time (registered over all the paths) (Fig. 6.8).

If one tries to set significantly positive instantaneous switching costs $C_Z > 0$, no switches take place anymore getting similar results as shown in Fig. 6.5.

This behavior can be also verified from the results of the value function in this case.
6. **Numerical solution approach and implementation**

![Graph of the minimum number of remaining switches for all the paths $N_{\text{paths}}$ over time $t \in [0, 252]$](image)

**Fig. 6.8:** Graph of the minimum number of remaining switches for all the paths $N_{\text{paths}}$ over time $t \in [0, 252]$

<table>
<thead>
<tr>
<th>Results ($\lambda_t = \text{LOW}$, $c_Z = c_\zeta = 0$)</th>
<th>0.0117</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V(t, u)^*$</td>
<td></td>
</tr>
<tr>
<td>$V(t, u)^{\text{Cva}}$</td>
<td>0.0153</td>
</tr>
<tr>
<td>$V(t, u)^{\text{Coll}}$</td>
<td>0.1116</td>
</tr>
</tbody>
</table>

From the table, we note that the model value function $V(t, u)^*$ and that without possibility to switch from the CVA regime $V(t, u)^{\text{Cva}}$ are almost equal, while the running costs of the collateralized regime are higher and equal to the case $c_Z = 0$ of the HIGH case results. Here in fact what changes are just the default intensities which have much lower drift and volatility than the latter case. Similar reasoning and careful check has to be done by changing the other parameters of the model.

From the results just showed, we can say that:

- the value function results show a clear dependence on counterparty default intensities which impact on CVA cost function, given the other parameters; of course varying also the collateral/funding costs we would get different values of $V(t, u)^*$, as we expected from chapter four (Proposition 4.8);

- we have also shown that the number of switching times depends on how far are over the paths the cost functions and the related conditional expectation from continuation. If these terms are near we verify period of switching activity especially if the instantaneous switching costs are null; instead by increasing these costs (both) we get that the value function tends towards a banal solution (as from definition 4.6);

- we also mention that the higher number of switching increase in general the value of the contingent collateralization of switching type as one would expect in the sense that a similar
type of optionality should be valued more as long as the switches are expected to be -
convenient over time. This also implies that this type of contingency should be valued more
and can be an optimal way to manage counterparty risk in cases in which counterparty
default rate is high or in non standard market conditions.

6.5 Numerical issues and further developments.

We conclude this chapter with some words on the results and on the solution approach proposed.
The algorithm proposed is based on some simplifying assumptions like that of symmetry hypothesis
that allowed us to speed up computations. But thanks to the Monte Carlo engine, one can
generalize it with ease, using more general processes/dynamics, introduce other stochastic factors,
like for the funding issue, modeling differently the default intensities allowing them to happen at
non standard times and also changing the switching time set, allowing the calculation for the value
functions for a fixed number of switching.

All these possible developments and improvements of the model have a common delicate issue
which is the estimate of the forward conditional expectations over future times in which the choice
of the basis functions is not unique and has to be done with great care. One should also investigate
alternative technique and methods to pursue this task as the ones described in the introduction:
quantization method, Malliavin calculus technique or non parametric regressions.

Therefore one should also try to solve the model using (if available) a different approach like the
analytical pde methods (finite difference, finite elements) in order to get a term to compare the
results obtained.

In conclusion, we underline the importance of a study of the stability/convergence and of the error
analysis of the MC based algorithm proposed. We have tried to tackle the problem but given the
high nonlinearity and recursion of terms involved it remains an open problem that we leave for
future research.
GENERALIZATION TO STOCHASTIC GAMES AND THE PRICE-HEDGE PROBLEM.

The enchanting charms of this sublime science reveal themselves in all their beauty only to those who have the courage to go deeply into it.

C.F. Gauss

7.1 Introduction.

In this chapter we build on the analysis and results of our model and on the main assumptions imposed throughout the chapters in order to define the main lines of possible generalizations. To begin, we can say that the perimeter of the application of our model is quite wide given the objective, namely the valuation and risk management of general defaultable OTC contracts in presence of CVA, collateral and funding issue, that is claims that pay dividend flows over time and for which a CSA between the parties has been signed. For what concerns the model framework of reduced form type, it has been thought for bilateral contracts (typically swap contracts) but it is enough general and can be easily extended to include other agents like the funder or a third reference entity like in CDS contracts.

As regards the modeling choices, the main lines of development are the ones that we get by solving the problem eliminating the main assumptions that we have imposed, which we recall here:

**Hp 1)** The analysis is done under the "symmetry hypothesis": the parties of the deal has symmetric objectives and the same default intensities $\lambda_A = \lambda_B$.

**Hp 2)** Both the parties are counterparty risk averse but we assume no strategic interaction for the moment.

**Hp 3)** The analysis is focussed on the full/perfect collateralization case;

**Hp 4)** The CVA cost process is not funded, $C^{\text{fund}} = 0$;

**Hp 5)** All the processes considered are intended to be pre-default value processes (as from lemma 2.2.1.)

Mainly, one can generalize the model by assuming a more general dynamic or adding stochastic factors for example to the funding costs (which are deterministic) or setting a different and more complicated objective functional - for example by removing the orthogonality hypothesis $X \perp \lambda$ -
or also modeling differently the CSA cash-flows including the partial collateralization case, which implies a modification of the control set definition for the problem. Anyway, given our main interest in the optimal contract design and risk management of this type of contingent mechanism, we focus on the following two main issues:

a) on the problem of removing the first symmetry hypothesis Hp 1) that, as already mentioned, will take us to reformulate the problem as a special stochastic differential game;

b) on the problem of formulating the pricing and hedge problem for a general defaultable contract with switching mechanism and the related solution.

As regards the first issue, the body of literature related to stochastic differential games is very vast given that the roots of the theory are in the pioneering works of von Neumann\(^1\) - for (mainly cooperative) zero sum game - and Nash\(^2\) - for (non cooperative) non-zero sum game - and the work of Isaacs\(^3\) who studied for first the differential games in a deterministic setting. In the stochastic framework it is worth of mention the seminal work of Eugene Dynkin who firstly analyzed stochastic differential games where the agent control set is given by stopping times, so that these "games on stopping" are known as Dynkin games in his honor.

Let us briefly remind that stochastic differential games are a family of dynamic, continuous time versions of differential games (as defined by Isaacs) incorporating randomness in both the states and the rewards. The states are random, described typically by an adapted diffusion process whose dynamics are known. To play a game, a player receives a running reward cumulated at some rate till the end of the game and a terminal reward granted at the end of the game. The rewards are related to both the state process and the controls at the choice of the players, as deterministic or random functions or functionals of them. A control represents a player action in attempt to influence his rewards. Assuming his rationality, a player should certainly act in the most profitable way to his knowledge. Since the rewards can be random, they are usually measured in the expectation, or some other more sophisticated criteria, for example in the variance as a measure of risk.

Certainly, it is impossible to mention here all the numerous important contribution in this field of research and to give a systematic account of the theory and of the literature. For a complete treatment of the different type of games we refer to the book of Isaacs on differential games. So we give in the following just a simplified classification restricting ourselves to the literature more related to our stochastic switching control problem.

This classification is based on the following main categories:

a) game and equilibrium type: it includes zero and non-zero-sum games whose solution can be searched mainly in terms of cooperative or non-cooperative (Nash) equilibrium. This depends also on the characteristics of the game which are mainly the system dynamic - that can be markovian/non markovian - and controls which can be state controls, stopping controls (as

---

\(^1\) "Theory of games and economic behavior" (1944), Princeton Press

\(^2\) "Equilibrium Points in N-person Games" (1950), Proceedings of the National Academy of Sciences

\(^3\) "Differential games" (1999), Dover.
in Dynkin games) or both which are called mixed control and stopping. Also the number of player is relevant, here we focus on the case \( p = 2 \).

b) solution approaches: these are mainly the analytical approach that allows - typically under a markovian framework - to formulate the stochastic differential game (SDG) as a system of (second order) Hamilton-Jacobi-Bellman equations or variational inequalities to solve, proving existence and (possibly) uniqueness of the solution, namely of the equilibrium of the game. The main solution techniques are the ones related to PDE theory, the dynamic programming principle and viscosity solution.

Worth of mention, from the analytical point of view, are the important works of Bensoussan and Friedman (1977) that for first showed the existence of a Nash equilibrium for a non-zero-sum SDG with stopping \( \{\tau_1, \tau_2\} \) as controls, formulating the problem as a system of quasi-variational inequalities (solved through fixed point methods), assuming continuous and bounded running rewards and terminal rewards; instead Fleming and Souganidis (1989) for first showed the existence and uniqueness of the solution/equilibrium for zero-sum SDG through dynamic programming and viscosity solution approach. These techniques have become very popular and used in the recent literature given the deep connection with probabilistic tools, as we have already mentioned in chapter five.

In fact, the probabilistic approach is the other more general one\(^4\) that makes use of the martingale (also via duality methods) and Snell envelope theory and, in addition, of the deep results of the forward-backward SDE theory in order to derive existence and uniqueness of the optimal control/stopping strategy for the game.

The main works worth of mention - other that the already mentioned work of Cvitanic and Karatzas (1996) that for first highlights the connection between the solution of zero-sum Dynkin game and that of doubly reflected BSDE (other than its analytical solution) - are that of: Hamadene (1998) that shows how the solution of a SDG is related to that of a backward-forward SDE; Hamadene and Lepeltier (2000) that extend the analysis through reflected BSDE to ”mixed game” problems; El Karoui and Hamadene (2003) that generalize the existence and uniqueness results for zero and non-zero-sum game with ”risk sensitive” controls; Hamadene and Zhang (2008) that use Snell envelope technique to show the existence of a Nash equilibrium for non-zero-sum Dynkin game in a non-markovian framework and Hamadene and Zhang (2010) that tackle the solution of a general switching control problem via systems of interconnected (nonlinear) RBSDE (with oblique reflection).

Much of these literature and works have been inspired also by valuation problems in the financial industry. We refer mainly to the american game option problem as defined in Kifer (2000) (also known as israeli option). This has given impulse to the literature related mainly to convertible (and switchable) bond valuation whose solution can be related to that of a zero-sum Dynkin game.

\(^4\) In fact, it allows to deal also with general non markovian dynamics for the system state variables.
7. Generalization to stochastic games and the price-hedge problem.

In the following section we concentrate on the analysis of the issues related to the formulation and the possible solution approach for our problem that we will model as a generalized Dynkin game of switching type. In order to do this, let us recall (in a markovian framework) the formal definition for both a non-zero-sum and zero-sum Dynkin game (with a game of switching type). In order to do this, let us recall (in a markovian framework) the formal definition for both a non-zero-sum and zero-sum Dynkin game (with a game of switching type). In order to do this, let us recall (in a markovian framework) the formal definition for both a non-zero-sum and zero-sum Dynkin game (with a game of switching type).

- Non-zero-sum Dynkin game: Given a standard probability space represented by the triple \((\Omega, \mathcal{F}, \mathbb{P})\) where we define \(W = (W_t)_{0 \leq t \leq T}\) be a standard \(d\)-dimensional Brownian motion adapted to the space filtration. Assuming as true the usual conditions on the drift function \(\mu(.)\) and volatility function \(\sigma(.)\), such that the following SDE admits a unique solution

\[
\begin{align*}
    dy(t) &= \mu(y(t), t)dt + \sigma(y(t), t)dW(t), \quad t \in [0, T] \\
    y(0) &= y_0.
\end{align*}
\]

Let \((p = 1, 2)\) \(f_p(y, t)\) the running reward function and \(\phi_p(y, t), \psi_p(y, t)\) the reward function obtained by the players upon stopping the game be continuous and bounded in \(\mathbb{R}^d \times [0, T]\), with \(f_p \in \mathbb{L}^2\) square integrable and \(\psi_p \leq \phi_p\) for all \((y, t) \in \mathbb{R}^d \times [0, T]\). Then let \(g_p(y(T))\) the terminal reward function also continuous and bounded. In a game of this kind, the two players have to decide optimally when to stop the game finding the optimal control given by the stopping times \((\tau_1, \tau_2)\) that give the maximum expected reward. So let us set the payoff functional for the two players of this Dynkin game as follows

\[
J^p(y, \tau_1, \tau_2) = \mathbb{E} \left[ \int_{\tau_1 \wedge \tau_2 \wedge T} f_p(y(s), s)ds + \mathbb{1}_{\{\tau_1 < \tau_2\}} \phi_p(y(\tau_1), \tau_1) + \mathbb{1}_{\{\tau_1 \geq \tau_2, T > \tau_1\}} \psi_p(y(\tau_2), \tau_2) + \mathbb{1}_{\{\tau_1 = \tau_2 = T\}} g_p(y(T)) \right] \quad \text{for} \quad j \neq i(\in \{1, 2\}).
\]

and for \((t \leq \tau_1 \leq T)\). Being in a non-zero-sum game with the players that aims to maximize their payoff \(J^p(.)\) without cooperation, the problem here is to find a Nash equilibrium point (NEP) for the game, that is to determine the couple of optimal stopping times \((\tau^*_1, \tau^*_2)\) such that

\[
\begin{align*}
    J^1(y, \tau^*_1, \tau^*_2) &\geq J^2(y, \tau_1, \tau_2), \quad \forall \tau_1, \tau_2 \in [t, T] \\
    J^2(y, \tau^*_1, \tau^*_2) &\geq J^1(y, \tau_1, \tau_2), \quad \forall \tau_1, \tau_2 \in [t, T]
\end{align*}
\]

namely the supremum of the payoff functional over the stopping time set. In other words, the NEP implies that every player has no incentive to change his strategy given that the other one has already defined optimally his strategy.

This type of game, as shown in Bensoussan and Friedman (1977), has an analytical representation given by a system of variational inequalities but it admits also a stochastic counterpart through system of BSDE with reflecting barrier. We return to its formal definition in the next section in relation to our problem.\(^5\). The Nash equilibrium defined above can be fairly

\(^5\) One can also refer to chapter five and definition 5.4.
7. Generalization to stochastic games and the price-hedge problem. We show this below in relation to zero-sum games.

- **Zero-sum mixed game**: A zero-sum game is characterized by the antagonistic interaction of the players that in this case has the same payoff functional but their objective are different because for one player the payoff is a reward (let’s think typically at the buyer of a convertible bond) that he wants to maximize, while for the other one is a cost that he intends to minimize.

In the generalized case of mixed games of control and stopping, the set of control will be enriched by the $F_t$-progressively measurable processes $(\alpha_t)_{t \leq T}$ and $(\beta_t)_{t \leq T}$ that are the intervention function namely the state controls respectively for the player $p_1$ and $p_2$. In addition, the players have to decide optimally when to stop the game setting the stopping times $\tau$ (for $p_1$) and $\sigma$ (for $p_2$). Indeed, the system dynamic being controlled by the agents can be expressed as the following controlled diffusion (remaining in a markovian framework):

$$dy(t)^{\alpha,\beta} = \mu(t, y_t^{\alpha,\beta}, \alpha_t, \beta_t)dt + \eta(t, y_t^{\alpha,\beta}, \alpha_t, \beta_t)dW(t), \quad t \in [0, T]$$

$$y(0)^{\alpha,\beta} = y_0.$$

The zero-sum game payoff being the same for both the players will be

$$\Gamma(\alpha, \tau; \beta, \sigma) := E \left[ \int_{\tau \wedge \sigma}^T f(s, y_s^{\alpha,\beta}, \alpha_s, \beta_s)ds + 1_{\{\tau \leq \sigma < T\}}\phi(\tau, y^{\alpha,\beta}_\tau) + 1_{\{\sigma < \tau\}}\psi(\sigma, y^{\alpha,\beta}_\sigma) + 1_{\{\tau = \sigma = T\}}g(y^{\alpha,\beta}(T)) \right] \quad (t \leq \tau \leq \sigma < T)$$

where the running and reward functions are intended to be the same as in the non-zero-sum case but clearly now they are the same for both players.

The solution of this SDG is typically tackled by studying the upper and lower value function of the players, which are

$$U(t, y) := \sup_{\alpha} \inf_{\beta} \sup_{\tau} \inf_{\sigma} \Gamma(\alpha, \tau; \beta, \sigma) \quad (upper \ value \ p1)$$

$$L(t, y) := \inf_{\beta} \sup_{\alpha} \inf_{\sigma} \Gamma(\alpha, \tau; \beta, \sigma) \quad (lower \ value \ p2)$$

Under some standard condition on the reward function and on controls, the problem has been tackled analytically representing the lower and upper value of the game as a system of nonlinear PDE with two obstacles/barriers, defined as follows

$$\begin{align*}
\min \left\{ u(t, y) - \phi(t, y), \max \left\{ \frac{\partial u}{\partial t}(t, y) - H^- (t, y, u, Du, D^2 u), u(t, y) - \psi(t, y) \right\} \right\} &= 0 \\
u(T, y) &= g(y),
\end{align*}$$

$$\begin{align*}
\min \left\{ v(t, y) - \phi(t, y), \max \left\{ \frac{\partial v}{\partial t}(t, y) - H^+ (t, y, v, Dv, D^2 v), v(t, y) - \psi(t, y) \right\} \right\} &= 0 \\
v(T, y) &= g(y),
\end{align*}$$

generalized in the case of *mixed game of control and stopping*. We show this below in relation to zero-sum games.
where $H^+(\cdot)$ and $H^-(\cdot)$ are the Hamiltonian operators (as defined in chapter four) associated to the upper and lower value function of the SDG. To solve the system, the unknown solution function $u$ and $v$ can be shown (under some technical assumptions) to be viscosity solutions of the above two PDE with obstacles and to coincide with the value functions $\mathcal{U}(t,y)$ and $\mathcal{L}(t,y)$ of the game.

In particular, when the Isaacs condition holds, namely for $(t,y,u,q,X) \in [0,T] \times \mathbb{R}^d \times \mathbb{R} \times (\mathcal{S}^d)$

$$H^-(t,y,u,q,X) = H^+(t,y,u,q,X)$$

then the two solution coincide and the SDG has a value namely

$$V := \sup_\alpha \inf_\beta \sup_\tau \inf_\sigma \Gamma(\alpha,\tau;\beta,\sigma) = \inf_\beta \sup_\alpha \inf_\sigma \sup_\tau \Gamma(\alpha,\tau;\beta,\sigma)$$

which is called the saddle point equilibrium of the mixed zero-sum game.

We mention also that in this case, as in the non-zero-sum case, the SDG has a stochastic representation which is expressed in terms of a doubly reflected BSDE (2RBSDE)$^6$. Recalling the definition of (single) reflected BSDE in chapter five, and setting - for a matter of notation - the early exercise rewards $\phi_t = U_t$ and $\psi_t = L_t$, the terminal reward $\xi$, $f$ the generator function and $Y$ the value process of the following 2RSBDE

$$\begin{cases}
    Y_t = \xi + \int_t^T f(s,Y_s,Z_s) + (K^+_T - K^+_s) - (K^-_T - K^-_s) - \int_t^T Z_sdW_s & \forall t \leq T \\
    L_t \leq Y_t \leq U_t, \forall t \leq T, \int_0^T (Y_s - L_s)dK^+_s = \int_0^T (U_s - Y_s)dK^-_s = 0
\end{cases}$$

that provides the value process $V$ of the associated zero sum (Dynkin) game. We underline that if we restrict to the case of data not depending on controls $\alpha$ and $\beta$ and consider the players criterion as the expectation of their payoff under the risk neutral measure, the game become a zero-sum Dynkin game whose solution is involved when dealing with the pricing of American game options.

For what concerns the issue related to point b) of the price/hedge issue of a contract of our kind, this is a recent topic in the literature and we will see that it also has deep connections with BSDE theory and the general theory of semi-martingale. For the analysis of this issue we refer directly to section 7.3.

7.2 On defaultable stochastic game option of switching type.

The main objective of this section is to generalize our model by removing the hypothesis $\lambda_A = \lambda_B$ and the symmetry between the counterparties that have signed the contract with contingent

---

$^6$ Its connection with the analytical representation and viscosity solution of PDE with obstacles, as already mentioned, has been established in the work of Cvitanic and Karatzas (1996).
collateralization of switching type. By doing this, we have been naturally lead to think at the problem as a stochastic/Dynkin (continuous) game, which is a stochastic differential game with stopping times as controls.

So, the objective here is to try to formulate and analyze the generalization of our model by introducing the strategic interaction between the counterparties and the main issues related to the conditions under which the problem admits a solution. In particular, in the search for an equilibrium for this special type of game, we will focus on the following essential points:

a) the rules of the game, of the moves alternation and of the switching strategies;

b) the definition of the switching control set and regimes for the problem;

c) the objective functional formulation for both the parties;

Firstly, let us keep as working framework the same defined in the second chapter which has been thought to work with the bilateral possibility to default for both the counterparties, but - to simplify and reduce problem dimensions - we assumed they had the same default intensities restricting the analysis on just one party switching strategy, as the contingent CSA was set unilaterally. So, we proceed to recall the main definitions and assumptions eventually modified in order to formulate the generalization.

a) Framework and assumptions: As already mentioned, the framework remains the same with the counterparties \{A, B\} both defaultable with now different default intensities \( \lambda_A \neq \lambda_B \). Both are assumed to behave rationally and have the objective to minimize the overall costs related to counterparty default - quantified through the BCVA - and those related to collateral and funding. The information flow is assumed symmetric. To ease the generalized formulation and analysis we keep also the hypothesis Hp 3), Hp 4) and Hp 5).

b) BCVA, CSA and contingent CSA: the definitions and propositions 2.2.3, 2.3.2 e 2.3.3 related to BCVA, CSA and contingent CSA processes are still valid given that we have already stated them for the bilateral case and for their symmetrical nature, which is expressed by the following relation

\[ BCVA_t^A = -BCVA_t^B \quad \forall \, t \in [0, T]. \]

This is valid also for the other processes, which means that only the sign changes (it depends on the counterparty point of view).

c) Funding: for funding, we generalize the setting in order to include the case of different funding rate for counterparties. In particular, we assume the existence of the following funding asset

\[ \text{\textsuperscript{7}Actually, the framework should be generalized in order to consider also the simultaneous default event of both the counterparties A and B, but we do not consider here this technical issue being not much relevant in the game analysis.} \]

7. Generalization to stochastic games and the price-hedge problem.

$B_{it}^{opp}, B_{it}^{borr}, B_{it}^{rem}$ and $\bar{B}_{it}^{rem}$ as defined in equations 2.28 – 2.31 for $A$ and $B$ such that

\[
\begin{align*}
\ dB_{it}^{opp,A} & \gg dB_{it}^{opp,B} \\
\ dB_{it}^{borr,A} & \gg dB_{it}^{borr,B} \\
\ dB_{it}^{rem,A} & \gg dB_{it}^{rem,B} \\
\ dB_{it}^{rem,A} & \gg dB_{it}^{rem,B}
\end{align*}
\]

**d) System dynamic:** for what concerns the stochastic vector dynamic that guides the system we keep the same setting of equations 3.1-3.3, in addition it will get the term for the default intensity dynamic of counterparty $B$ $d\lambda_t^B$ which is now possibly correlated with $\lambda_t^A$, so that the generalized system dynamic becomes (in vectorial form)

\[
d Y_{it}^{C_{ad}}(t) := 
\begin{bmatrix}
\ dt \\
\ dX_t \\
\ d\lambda_t^A \\
\ d\lambda_t^B \\
\ dZ
\end{bmatrix}, \quad Y_{C_{ad}}(0) = 
\begin{bmatrix}
\ t = 0 \\
\ x_0 \\
\ \lambda_0^A \\
\ \lambda_0^B \\
\ Z_0 = 1
\end{bmatrix}
\]

**e) Control set:** the generalized control set for our problem will be made up of all the finite sequences of indicators and switching times optimally defined by the counterparties, that is

\[
C^A = \{T^A, Z^A\} = \{\tau_j^A, z_j^A\}_{j=1}^M, \ \forall \ tau_j^A \in [0, T], \ z_j^A \in \{0, 1\} \quad (7.1)
\]

\[
C^B = \{T^B, Z^B\} = \{\tau_j^B, z_j^B\}_{j=1}^M, \ \forall \ tau_j^B \in [0, T], \ z_j^B \in \{0, 1\} \quad (7.2)
\]

**f) Payoff functionals and game formulation:** In order to formulate the game, firstly, we recall that in our model for the contingent CSA of switching type we have assumed no fixed times or other rules for switching, that is the counterparty can switch optimally every time until contract maturity $T$ in order to minimize its objective functional. But the functionals now are assumed different between the parties (no symmetry hypothesis): in particular, both players can be assumed to remain risk averse to the variance of CVA and collateral and funding costs, but depending on the different parametrization of the functionals (that we show below) and instantaneous switching costs other than the difference in default intensities, the problem can be naturally represented by a ”special” non-zero-sum stochastic differential game. This is special in the sense that now the controls are not just simple stopping times but sequences of random times that define the optimal times to switch from a regime to an other one (together with switching indicators). Now, given that the ”right to switch" is bilateral and no other rules are set, what the other party will do becomes relevant to define

\[\text{Typically, it is used a CIR or CIR++ process to model the } \lambda_t^B \text{ dynamic}\]
the own optimal switching strategy.
Let us be more formal and recalling the main notations of chapter three, we define our model’s SDG as a generalized Dynkin game of switching type as follows.

Definition 7.2.1 (Dynkin game of switching type definition). Let us consider two players/counterparts \( \{A, B\} \) that have signed a general contract with a contingent CSA of switching type. Given the respective payoff functionals \( F^i_Z(\cdot) \) (or running reward) where \( i \in \{A, B\} \) and \( Z \in \{z, \zeta\} \) (being different also by switching regimes), terminal reward \( G^i(\cdot) \) and instantaneous switching costs \( c^i_{Z_j}(\cdot) \), counterparties (namely the players) are rational and interact strategically in a non cooperative way and they aim to minimize the following objective functional

\[
J^i(y, u^i) = \inf_{u^i \in \mathcal{C}^i_{ad}} \mathbb{E} \left[ \sum_{j} \int_{t}^{\tau^j \wedge \tau^i - \wedge T} B_s \left[ F^j_Z(y_s, u^i, b^{-i}(u^i)) \right] ds \right. \\
+ \left. \sum_{j \geq 1} c^j_{Z_j}(t) \mathbb{1}_{\{\tau^j \wedge \tau^i - < T\}} + G^i(y(T)) \mathcal{F}_t \right] \text{ for } i \in \{A, B\}
\]

where the system dynamic and controls are defined at point d) and e) and we have set for notational convenience \( u^i := \{T^i, Z^i\} \text{ for } i \in \{A, B\} \).

We underline in definition 7.2.1 that the cost functions \( F^i_Z(\cdot) \) are here generalized in order to take into account the strategic interaction with the other player represented formally by a response function \( b^{-i}(u^i) \) to specify. In particular, the payoffs are intended to be defined as in section 3.3, but they can also differ here (between \( A \) and \( B \)) for the following terms - other than \( \lambda_t^A \neq \lambda_t^B -
\]

\[
\delta^A \lesssim \delta^B \\
R^A_t \lesssim R^B_t \\
c^A_{t, z} \lesssim c^B_{t, z}
\]

where \( R^i(\cdot) \) is defined in 3.7, \( \delta^i \) is a parameter of the objective functional and \( c^i_{t, z} \) are the instantaneous costs from switching.

Remarks 7.2.1. The game as formulated above in 7.3 (from the point of view of one of the player) is fairly general; in addition, one could also introduce the possibility for the players to stop the game adding a stopping time (and the related reward/cost function) to the set of controls made up of switching times and indicators. From the financial point of view, this can be justified by a early termination clause set in the contingent CSA defined by the parties. Anyway, given the problem recursion, this would add greater complications that we leave for further research.
Actually the game 7.3 is already complicated by the fact that, differently from the (non-zero-sum) Dynkin game as formulated in the introduction, in our generalized game on switching, the players
control strategies affect also each other payoffs. In fact, given that in our general formulation the players can switch optimally whenever over the life of the underlying contract, it is clear that - without setting any other rules for the game - the decision of one player to switch to a certain regime impose a different cost function $F_Z(.)^i$ also for the other player. So if A switches but for B the decision is not optimal, he is able to immediately switch back, taking in account the instantaneous switching costs\footnote{Note that from 7.3 the indicator $\mathbb{1}_{\{\tau_j^A \wedge \tau_j^B < T\}}$ the instantaneous switching costs enter in the functional whoever of the players decides to switch}. In this sense, the relative difference between players’ payoff (in the different regimes) and the strategic interaction between them over time become central in order to understand and analyze the problem solution/equilibrium.

We return on this points later, here is important to mention that in order to highlight this strategic dependence in the game - that is assumed to be played by rational and non-cooperative players - we have enriched the running cost function $F_Z(.)^i$ by a response function $b^{-i}(u^i)$, which can be intended mainly in two way:

a) as the "classical" best response function to the other player strategy, which implies the complete information assumption in the game, that is the players have the same information set about the system dynamic and they know each other payoff;

b) if the game information is not complete and there is a degree of uncertainty over the players payoff and their switching strategy, the function $b^{-i}(u^i)$ can be intended in generalized terms as a probability distribution assigned by a player to the optimal response of the other one.

We discuss further on the game information flows below. Now, the main issue to tackle is to understand the condition under which this generalized game (7.3) have sense and it will be played, which means that it will be signed by counterparties. This takes to the problem definition of an equilibrium for this game and to the condition under which its existence and uniqueness are ensured.

Before giving the formal definition of the game equilibrium, let us highlight the game pure strategies at a given time $\{\tau_{j-1}^i < t \leq \tau_j^i\}$ (under the assumption of simultaneous move for the players).

**Definition 7.2.2 (Pure strategies of the game of switching type).** For any given initial condition $z_0^A, z_0^B$ and $\forall z_j^A \in u^A$ and $z_j^B \in u^B$ and $\{\tau_{j-1}^i < t \leq \tau_j^i\}$, the pure strategies of our Dynkin game of switching type are defined as follows

---

Note that from 7.3 the indicator $\mathbb{1}_{\{\tau_j^A \wedge \tau_j^B < T\}}$ the instantaneous switching costs enter in the functional whoever of the players decides to switch.
7. Generalization to stochastic games and the price-hedge problem.

\[ \text{if} \{z_{j-1} = 0\} \Rightarrow \]
\[ \{z_j^A = 0, z_j^B = 0\} \Rightarrow \text{"no switch"} \]
\[ \{z_j^A = 0, z_j^B = 1\} \Rightarrow \text{"switch to 1"} \]
\[ \{z_j^A = 1, z_j^B = 0\} \Rightarrow \text{"switch to 1"} \]
\[ \{z_j^A = 1, z_j^B = 1\} \Rightarrow \text{"switch to 1"} \].

\[ \text{while if} \{z_{j-1} = 1\} \Rightarrow \]
\[ \{z_j^A = 0, z_j^B = 0\} \Rightarrow \text{"switch to 0"} \]
\[ \{z_j^A = 0, z_j^B = 1\} \Rightarrow \text{"switch to 0"} \]
\[ \{z_j^A = 1, z_j^B = 0\} \Rightarrow \text{"switch to 0"} \]
\[ \{z_j^A = 1, z_j^B = 1\} \Rightarrow \text{"no switch"} \].

In the table below we represent the standard game form at a given decision time with the possible (pure) strategies (namely the switching indicators) and the related random payoff between parenthesis.

<table>
<thead>
<tr>
<th>( A, B )</th>
<th>Switch</th>
<th>No Switch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switch</td>
<td>( 1, 1 ) ((J^A, J^B))</td>
<td>( 1, 0 ) ((J^A, J^B))</td>
</tr>
<tr>
<td>No Switch</td>
<td>( 0, 1 ) ((J^A, J^B))</td>
<td>( 0, 0 ) ((J^A, J^B))</td>
</tr>
</tbody>
</table>

Remarks 7.2.2. Clearly, from a static point of view the NEP for the game just shown can be easily found once the payoff \( J^i \) are known. But the problem is that game configurations like these has to be played over time taking in account as key factors:

a) the payoff value that derives from switching at a given time;

b) the expected value from waiting until the next switching time;

c) the optimal responses, namely the other party optimal switching strategy which implies also to know the same points a) and b) for the other party.

The resulting equilibrium is an optimal sequence of switching over time for both players that needs a (backward) dynamically recursive valuation. On an heuristic base, we expect that if the relative difference (over time) in players payoff functionals - mainly due to different function parametrization, default intensities \( \lambda^i \) or switching costs \( c^i_Z \) - remains low, it is more likely that the switching strategies on the main diagonal of the game \{1,1;0,0\} will be played (given that both players would have also similar best responses). This should ease the search for the equilibrium of the game but this makes more likely to incur in banal solutions. Otherwise, one should observe a more complicated strategic behavior that needs a careful study depending also on the type of
Generalization to stochastic games and the price-hedge problem.

Indeed, given the characteristics of our game namely a non-zero-sum game in which the agents are assumed rational and act in a non-cooperative way in order to minimize their objective function knowing that also the other part do the same, the equilibrium/solution of this type of games is the celebrated Nash equilibrium point (NEP). Actually, as shown in our case the equilibrium is characterized by a sequence of optimal switching through time and being game 7.3 a generalization of a Dynkin game, by similarity, we can state the following definition of a Nash equilibrium point for a Dynkin game of switching type.

Definition 7.2.3 (NEP for Dynkin game of switching type). Let us define the switching control sets for the player \( \{A, B\} \) of the generalized Dynkin game 7.3 as follows

\[
u_A := \{(\tau_j^A, z_j^A)_{j=1}^M, \forall \tau_j^A \in [0, T], z_j^A \in \{0, 1\}\);
\]

\[
u_B := \{(\tau_j^B, z_j^B)_{j=1}^M, \forall \tau_j^B \in [0, T], z_j^B \in \{0, 1\}\}.
\]

A Nash equilibrium point for game 7.3 is given by the pair of sequences of switching times and indicators \( \{u^*_A, u^*_B\} \) such that for any control sequences \( \{u_A, u_B\} \) the following conditions are satisfied

\[
J_A(y; u^*_A, u^*_B) \leq J_A(y; u_A, u_B) \tag{7.4}
\]

and

\[
J_B(y; u^*_A, u^*_B) \leq J_B(y; u_A, u_B) \tag{7.5}
\]

(the signs will be reversed in the maximization case).

A formal and rigorous proof of the existence (and uniqueness) of a NEP for the game 7.3 such that it is non trivial or banal in the sense of definition 4.6, is the big issue here. In order to approach the solution of the game, as we know from the introduction and past chapters, one has in general two way: analytic or probabilistic which are deeply interconnected (working in a markovian framework).

Referring to chapter five for the notations and definitions (5.4 in particular), from the theory of BSDE with reflection one can formulate and associate the solution of our game 7.3 to that of a system of coupled (non-linear) reflected BSDE with the following representation.

Definition 7.2.4 (RBSDE representation for game of switching type). Let us define the vector triple \((Y^i, Z^i, N^i, K^i, Z^i)\) for \(i \in \{A, B\}\) and \(Z \in \{z, \zeta\}\), with the same technical condition of definition 5.4. Then, given the standard Brownian motion vector \(W^y_t\), the terminal reward \(\xi^i\), the obstacles \(c^i_t\) and the generator functions \(F^i_Z(.; Y^{-i})\) both interconnected between the players,
we can represent the Dynkin game of switching type formulated in 7.3 through the following system of interconnected non-linear reflected BSDE

\[ \begin{align*}
    Y_t^{A,z} &\in K^2; N_t^{A,z} \in M^2; K_t^{A,z} \in K^2, K \text{ non decreasing and } K_0 = 0, \\
    Y_t^{A,z} &= \xi^A + \int_s^T F^A_Z(y_s, u^A, N^A_s, Y^B_s) ds - \int_s^T N^{A,z}_s dW_s + K^{A,z}_T - K^{A,z}_s, \quad t \leq s \leq T, \quad Z \in \{z, \zeta\} \\
    Y_t^{A,z} &\geq (Y_t^{A,z} - c_t^{A,z}); \int_0^T [Y_t^{A,z} - (Y_t^{A,z} - c_t^{A,z})] dK_t^{A,z} = 0; \\
    Y_t^{A,\zeta} &\geq (Y_t^{A,\zeta} - c_t^{A,\zeta}); \int_0^T [Y_t^{A,\zeta} - (Y_t^{A,\zeta} - c_t^{A,\zeta})] dK_t^{A,\zeta} = 0; \\
    Y_t^{B,z} &\in K^2; N_t^{B,z} \in M^2; K_t^{B,z} \in K^2, K \text{ non decreasing and } K_0 = 0, \\
    Y_t^{B,z} &= \xi^B + \int_s^T F^B_Z(y_s, u^B, N^B_s, Y^A_s) ds - \int_s^T N^{B,z}_s dW_s + K^{B,z}_T - K^{B,z}_s, \quad t \leq s \leq T, \quad Z \in \{z, \zeta\} \\
    Y_t^{B,z} &\geq (Y_t^{B,z} - c_t^{B,z}); \int_0^T [Y_t^{B,z} - (Y_t^{B,z} - c_t^{B,z})] dK_t^{B,z} = 0; \\
    Y_t^{B,\zeta} &\geq (Y_t^{B,\zeta} - c_t^{B,\zeta}); \int_0^T [Y_t^{B,\zeta} - (Y_t^{B,\zeta} - c_t^{B,\zeta})] dK_t^{B,\zeta} = 0
\end{align*} \]

From definition 7.2.4, it is evident that the system of RBSDE is a non-standard one given the characteristics of the generator functions - which are the cost function in our game - that are inter-dependent and this is highlighted by the presence of the other player value process inside \( F^i_i(\cdot) \). This makes hard to show the existence and uniqueness of the solution of the system that should be given by a two-dimensional vector made up by the triple \( (Y^*, Z, N^*, K^*) \) where its dimension is given by the two switching regimes and the optimal switching sequences defined by the crossing of the barriers from the value processes (as set in the third and sixth line of the RBSDE). Therefore, one should also verify that the solution value processes \( Y^*, Z \) coincide with the value functions of the two players of the non-zero-sum game of switching type (7.3).

As far as we know, these issues have been tackled - in relation to switching problems - in the already mentioned work of Hamadene and Zhang (2010). They study general system of m-dimensional BSDE called with "oblique reflection", which are RBSDE with both generator and barrier interconnected as in our case, showing existence and uniqueness of the solution while the optimal strategy in general does not exist but an "approximating optimal strategy" is constructed (through some technical estimates).

In particular, the main technical assumptions\(^\text{10}\) that are imposed in order to derive these results are:

a) **square integrability for both the generator function** \( F^i(\cdot) \) **and terminal reward** \( \xi \) **while the obstacles function are continuous and bounded**;

b) **Lipschitz continuity of the generator function respect to its terms**;

\(^{10}\)Here we just give a sketch of the main assumptions.
7. Generalization to stochastic games and the price-hedge problem.

c) both the generator and the obstacles are assumed to be increasing function of the other players utility/value process.

As also mentioned in the paper, the condition c) implies from a game point of view that the players are "partners", namely the interaction and the impact of the other players value processes has a unique positive sign. This is not the case of our non-zero-sum game in which the interaction allowed between the two players is antagonistic and more complicate. Therefore also the Lipschitz condition b) is hard to verify for our non-linear and recursive cost functions.

Indeed, we can say that - as far as we know - the existence and uniqueness solution of the game formulated in definition 7.2.4 is an open problem whose solution is out of the scope of the present work.

Even though one could assume to simplify the problem in order to work under the same assumptions a)- c) that would ensure the existence and uniqueness of the solution, it would remain to verify that the solution of the system of RBSDE is the *Nash equilibrium point* for the game, which is complicated by the fact that the optimal control strategy may not exist. Anyway, we end this section formalizing this issue in the following theorem whose proof may be an other open problem.

**Theorem 7.2.5 (NEP and RBSDE system solution).** Let us assume the existence and uniqueness of the solution for the system of definition 7.2.4, under the assumption a)- c). Then the system RBSDE value processes $Y^*_A$, $Y^*_B$ coincide with the player value functions of the game of switching type, that is

$$Y^*_A = J^A(y, u^*_A, u^*_B)$$
$$Y^*_B = J^B(y, u^*_B, u^*_A)$$

and are such that condition 7.4 and 7.5 are satisfied, which implies the existence and uniqueness of a Nash equilibrium point for the game (7.3).

**Remark 7.2.3.** As we already know, in the markovian framework - thanks to El-Karoui et al. results - the solution of the system of RBSDE is connected with the viscosity solution of a generalized system of non-linear PDE like the one defined defined in (4.8 - 4.9) with generator and obstacles different and interconnected between the two players, which is harder to study analytically. The main alternative is to try to approach numerically the problem, searching for the conditions under which one can find the equilibrium. A possibility is to apply Snell envelope and iterative optimal stopping techniques that we adopted in chapter six opportunely adapted to study the stochastic game 7.3 solution. In this case, the algorithm and implementation are more complex than the unilateral case and a careful study has to be undertook, so we leave this issue for future research.

7.3 Game solution in a "special case" and further analysis.

Before passing to deal with the price-hedge problem for contract with contingent CSA, let us make some further reasoning on the game characteristics in order to possibly simplify or modify
7. Generalization to stochastic games and the price-hedge problem.

the general formulation 7.3.

In particular, we focus the analysis on the following three main points - already underlined above - that have impact on the equilibrium characterization and existence:

- information set between the players/counterparties;
- rules of the game;
- differences in the objective functionals of the players/counterparties;

a) Firstly, a careful analysis of the game information set is fundamental to characterize and understand the game itself and its equilibrium. In our game formulation 7.3 we have assumed symmetry in the information available for the players which helps to simplify the analysis, but in general one needs to specify what is the information available to them at all the stage of the game. Given that we have been working under the market filtration \((\mathcal{F}_t)_{t \geq 0}\), under symmetry we get that both player knows \(\forall t \in [0,T]\) the values of the market variables and processes that enter the valuation problem, namely

\[
\mathcal{F}^A_t = \mathcal{F}^B_t = \mathcal{F}_t \quad \forall t \in [0,T]
\]

So, both players are able to calculate the outcomes/payoff of the game through time. This implies that the players know each other cost functions so that the game is said information complete\(^{11}\) and it is easier to solve for a NEP knowing the best response function. It is important to underline that the game is played simultaneously at the decision times but it is dynamic and recursive because the optimal strategy played today will depend not only on the initial condition (that is usually common knowledge) but on future decisions taken by both the players. So, these complicate characteristics of the game imposing a backward induction procedure to search for an equilibrium point.

Of course, the assumption to know the counterparty cost function is quite strong for our problem in which the parties of the underlying contract can operate in completely different markets or industries, but it can be not uncommon to verify a cooperative behavior between them. In particular, in cooperative games the players aim to maximize or minimize the sum of the values of their payoff over times, namely

\[
J^{coop}(y, u^*) := \left[ J^A(y, u^A) + J^B(y, u^B) \right].
\]

This type of equilibrium, depending on the type of game considered, is much more difficult to study in the stochastic framework, given the necessity to find the condition under which players cooperate over time (and paths) and have no incentives "to cheat" playing a different (non cooperative) strategy. This can be an interesting further generalization for our game model that would be important to examine in major depth.

\(^{11}\) They know strategies and payoff at every stage of the game.
b) Also the rules of the game are important in order to simplify the search for the equilibrium. In our model both the counterparties are able to switch optimally every time over the contract life. Discretizing the time domain, we have been lead to think at the game as played simultaneously through time over the switching time set that can be predefined in the contract or model specific. In terms of game theory, this means that a given decisional node of the game the players make their optimal choice based on the information available (which is common knowledge) at that node and at the subsequent node they observe the outcome of the last interaction and update their strategy.

An other possibility that may help to simplify things, is to assume that - by contract specifications - the counterparties can switch only at predefined times and that the two sets have null intersection, namely

\[ \{ \tau^A_j \cap \tau^B_j \}_{j=1}^M = \emptyset. \]

This happen for example if the right to switch is set as sequential. So, also in terms of game theory, the strategic interaction and the game become sequential: under the assumption of incomplete information, this type of games are generally solved via backward induction procedure - as it is necessary in our case - and one can search for a weaker type of Nash equilibrium. Clearly, we remind that this type of equilibrium should be studied under a stochastic framework which is analytically more difficult and also numerically it can be a cumbersome task.\(^{12}\)

A strategic sequential interaction like that, can be also obtained by contract setting in the CSA some time rules like the so called "grace periods" namely a delta time \(\Delta t\) that has to pass after a switching time before the other party can make its optimal switching decision.

c) The last point really relevant in our game analysis concerns the relative differences between counterparties objective functionals. As already mentioned above in relation to our model, the variables/parameters that have impact in this sense are:

- differences in the default intensities processes \(\lambda^A_t, \lambda^B_t\);
- differences in the cost function thresholds \(\delta^A, \delta^B\);
- differences in funding/opportunity costs \(R^A(t), R^B(t)\);
- differences in the (instantaneous) switching costs \(c^{z,A}_t, c^{z,B}_t\).

In particular, let us simplify things by considering the special case of our game 7.3 assuming symmetry between the party of the contract, and then showing the impact of just the difference in the cost functions of the threshold \(\delta\). Under symmetry it is easy to show that the game solution coincide with that of our control problem on which we have focussed the analysis in past chapters. In fact, considering the problem from just one side is equivalent to a game played by symmetric players with objective functional with the same parameters.

The reason for these modeling choices in chapter three was to reduce the problem dimensions and to simplify the analysis considering the problem from just one party point of view.

\(^{12}\) As it is already in the unilateral case.
In economic terms, considering a game problem between two "symmetric" players can be justified if one thinks to two institutions with similar business characteristics other than risk worthiness, that operate in the same country/region/market with the objective to optimally manage the counterparty risk and the collateral and funding costs by signing a contingent CSA in which are defined all the relevant parameters necessary to know each other objective functional.

So, let us consider a game played under these "special" conditions, it is not difficult to see that the game payoffs will be the same for both the player: in fact, being $\delta^A = \delta^B = 0$, the square of BCVA and collateral costs functions is the same and also the instantaneous costs are assumed equal. This implies that also the best response functions will be equal for both the player, which means they play the same switching strategy, however the game is played simultaneously or sequentially. So, on the basis of this chain of thoughts we can state the following result.

**Proposition 7.2.3 (NEP Existence and uniqueness in "symmetry" case.)** Assume as true the conditions and hypothesis imposed for the formulation of the symmetry (unilateral) case for a Dynkin game of switching type. In addition, assume the same technical condition imposed in section 5.2 to solve our control problem. Then exists and is unique a Nash equilibrium for this game and it coincides with the value function (5.1-5.2) of the stochastic control problem derived in the unilateral case, that is

$$J^A(y, u^*_A; u^*_B) = J^B(y, u^*_A; u^*_B) = V^*(y, u^*).$$  \hspace{1cm} (7.6)

**Proof.** The proof is easy by the above reasonings. In fact, under the "symmetry" conditions and recalling the notation from the general game formulation 7.3 we have that the following relations hold:

$$F^A_Z(y_s, u_A, b^A(u^A)) = F^B_Z(y_s, u_B, b^A(u^B))$$
$$c^A_z(t)1_{\{\tau^A_j \wedge \tau^B_j < T\}} = c^B_z(t)1_{\{\tau^B_j \wedge \tau^A_j < T\}}$$
$$G^A(y_T) = G^B(y_T)$$

and setting $\tau^*_j := \tau^A_j = \tau^B_j$ which implies $1_{\{\tau^A_j \wedge \tau^B_j < T\}} = 1_{\{\tau_j < T\}}$, being the optimal control sequence optimal for both players, namely $u^*_A = u^*_B$, which implies also the equality of the best response functions $b^A(u_B) = b^B(u_A)$. This means also that the strategic interaction becomes irrelevant and the game solution can be reduced to that of an optimal switching control problem equivalent for both the players, so that game 7.3 is reduced to the following problem

$$J(y, u) := J^A(y, u^A) = J^A(y, u^B)$$
$$J(y, u) = \inf_{u \in \mathcal{U}_{ad}} \mathbb{E}\left[ \sum_j \int_{\tau^*_j}^{\tau^*_j \wedge T} B_s[F_Z(y_s, u)] ds + \sum_{j \geq 1} c_z(t)1_{\{\tau_j < T\}} + G(y_T) \bigg| \mathcal{F}_t \right]$$
7. Generalization to stochastic games and the price-hedge problem.

which is equivalent to the control problem formulated in chapter three. Its solution is given by the value function $V^*(y,u^*)$ (known from chapter five) to which is associated the optimal sequence of switching times and indicators $u^* = \{T^*, Z^*\}$. By the optimality of this sequence, conditions 7.3 and 7.4 of NEP definition 7.2.1 are satisfied, being the optimal strategy for both players (by the symmetry hypothesis). So being the two problems representation actually the same, the NEP exists and is unique (from the existence and uniqueness of $V^*(y,u^*)$) and equation 7.6 is true, as we wanted to show $\diamondsuit$.

In general with different function parametrization between players and incomplete or asymmetric information, the equilibrium is much harder to find and different strategies has to be checked. To give an idea of this, let us consider just a slight modification of the symmetric case conditions, setting for example the cost functions threshold $\delta^A = \delta^B > 0$, and keeping the information incomplete and the game play simultaneous. By the symmetry of BCVA and (running) collateral/funding costs, we know that a positive value of one term for $A$, is negative for $B$ and viceversa. So introducing the threshold create different payoffs for the player, as we can easily see below

$$(BCVA^A - \delta)^2 \geq (BCVA^B - \delta)^2$$

given that if $BCVA^A_t > 0$ then $BCVA^B_t < 0$ and viceversa$^{13}$. So, even though they knew each other objective functional, there would be some paths and periods in which the strategic behavior of the players is in contrast, say of war type and others of peace type, making the analysis more complicate.

From the results of past chapters, is worth of mention the possibility that the game is not played, namely it reveals to be never optimal to switch for both the players. This singular game solution corresponds in the unilateral/symmetric case to the banal switching strategies stated in Definition 4.6. A possible example of this singular solution can be given thinking to our simultaneous game as a zero sum game. This can happen by considering linear objective functionals with threshold $\delta \approx 0$, in fact - by symmetry of the BCVA and of funding cost function - a positive outcome for one player is negative for the counterpart$^{14}$. Assuming instantaneous switching costs $c^z > 0$ for both players and - to simplify - that both know each other cost functions, we get that this game will never be played. The reason is that by the game zero sum structure, the optimal strategy for one is not optimal for the other, so every switch will be followed by the opposite switch at every switching time, like the sequence

$$\{z_1 = 1, z_2 = 0, z_3 = 1, \ldots, z_M = 1\}$$

But by rationality and taking in account the positive cost of switching, one can conclude that the game will never be played by a rational agent.

$^{13}$ This is true also for the other switching costs related to collateral and funding. Of course one should consider also the weight of the expected payoff value from keeping the strategy for an other period.

$^{14}$ Obviously, when the value of the cost functions compensate each other, no switch take place
So, let us summarize this last logic chain of thoughts in the following proposition.

**Proposition 7.2.4. (Game banal solution in the zero-sum case).** Let us assume that game 7.3 be a zero-sum game with linear functionals set for both the counterparties. Assuming in addition the same funding costs for both players, \( \delta \approx 0 \) and positive switching costs \( c^2 > 0 \) (for both), then the optimal strategy is to never play this game (that is a banal solution of the game).

**Remarks 7.3.1** Let us recall that in the special case of zero-sum games the game equilibrium, is expressed as follows

\[
J^{A=B}(y, u_A^*, u_B) \leq J^*(y, u_A^*, u_B^*) \leq J^{A=B}(y, u_A^*, u_B^*)
\]  

(7.7)

for all control sequences of \((u_A, u_B)\). In particular if

\[
\inf_{u_A} \sup_{u_B} J^{A=B}(y, u_A, u_B) = \sup_{u_B} \inf_{u_A} J^{A=B}(y, u_A, u_B).
\]  

(7.8)

then the zero-sum game equilibrium - called a saddle point of the game - is said to have a value.

As we know this depends on the verification of the Isaacs condition namely the equality between the lower and upper value functions for the players of the game. This type of equilibrium is central from the arbitrage free pricing point of view. In fact, the solution of the zero-sum game under the martingale pricing measure, that is

\[
J^*(y, u_A^*, u_B^*) := E^Q_*[\Pi(t, y)],
\]

implies also the existence and uniqueness of the arbitrage free price of the underlying claim which admits this stochastic zero-sum game representation.

We end the section by remarking the importance of the points highlighted in the construction of some kind of equilibrium for a Dinkyn game of switching type as our one. Although mainly theoretical, the existence of the equilibrium and the definition of the conditions under which a non banal solution is available, are relevant economically and in the contract design phase.

Hence, the main task to pursue in future research are a rigorous proof of the existence and of the equilibrium for this type of game, and the definition of an efficient algorithm to check the model solutions in the bilateral (namely, stochastic game) setting.
7. Generalization to stochastic games and the price-hedge problem.

7.4 Price and hedging issue for a general contract with contingent collateralization.

The left important issue on which we need to focus is the pricing and hedging problem for general OTC contracts in which a contingent CSA of switching type, like the one analyzed, has been set between the parties. The main idea of the model was to study the effects of a similar mitigation mechanism in a risk management view assuming a specific form (of variance type) for the cost function of the agents, that are counterparty risk adverse.

But what if one wants to know the price and the relative hedging strategy for a similar contract with CVA (bilateral), contingent collateral and funding issue? This is an hard problem mainly from a theoretical point of view. From the numerical one the issues involved are more or less the same we have already underlined in chapter six. From the theoretical one, the main reasons that make valuation problematic and complex are the following.

a) Recursive valuation problem: this is the main technical issue to tackle in order to get the price-hedge of the contract. The presence of a CSA in the contract and clearly of CVA whose value depends on future contract exposures impose to solve the recursion problem through dynamic programming and Snell envelope which are connected, as we already know, with theory of backward stochastic differential equations. In our case, the presence of the contingent CSA makes the value of the contract, and in particular of CVA and the collateral/funding term, dependent from the counterparty optimal switching strategy.

b) Market incompleteness and pricing measure choice: in pricing models with counterparty risk one typically works under the martingale pricing measure $Q^*$, but from market incompleteness - due mainly to the unheadgeable default risk - we know that this probability measure is not unique. This obviously implies that no perfect hedge can be build and alternative approaches like mainly the mean-variance hedging or utility/preference based has to be used.

So, from point b) we underline that solving our model under the pricing measure would give one price for the switching contingent CSA under quadratic type of preferences, as we have defined our cost functions $F_Z(\cdot)$. Indeed, the price obtained under this structure of preferences and assuming a priori the existence of the relative pricing measure does not allow to get an effectively arbitrage-free price if the processes and the measure under which are performed calculations is not the martingale one. Therefore, it remains open the problem of how to price the whole contract (including the contingent CSA).

As far as we know, the work of Crepey (2011) has been one of the first rigorous approach to price/hedge general defaultable contract with CSA and funding.

The main results of this important work are the derivation, in a reduced form setting under pre default valued ($\mathcal{F}$-adapted or predictable) processes and markovian dynamics for the driving stochastic factors, of a markovian BSDE as the main tool to price and hedge the general CVA.

---

15 In section 3.4 we discussed about this issue, the double order of recursion.
16 From the well known second fundamental theorem of arbitrage pricing.
17 Actually they derive the BSDE also for the whole contract price, but once one knows the CVA term can derive the other and viceversa.
process of the contract (including CSA and funding).
In addition, although the existence of the solution of this (non standard) BSDE is assumed (not proved), thanks to well-known results of El Karoui et al. (1997) on BSDE representation the author is able to derive the corresponding PIDE (because of the presence also of the jump process $N_t$) representation and the explicit form of the hedge process, which is optimal in min-variance sense, namely the hedge minimizes the market risk variance due to the hedging error/costs related to CVA.
A similar approach should be used for our pricing and hedging problem, which is actually complicated by the fact that the CSA is contingent, namely its value depends on the optimal switching strategy of the counterparty through time. So, let us show the main points of the price and hedge issue in our problem.
Firstly, we can keep the same reduced form framework of chapter two in which we assume the existence of martingale pricing measure $Q^*$ not necessarily unique. The underlying claim is defaultable and pays a dividend flow through time. We keep the perspective of one of the counterparties of the contract, given that we are tackling a pricing problem in the arbitrage-free context. The same definitions and propositions of chapter two related to:

1. clean price process and dividend process $S_t^f, D_t^f$;
2. bilateral risky dividend and price process, $S_t, D_t$;
3. bilateral CVA, $BCVA_t$, and the related contingent case $BCVA_t^C$;
4. collateralization $Coll_t$ and funding costs $C_t^{Fund}$,

still state.
In particular, we recall from definitions 2.5.1 and 2.5.2 (equations 2.26-2.28) on our contingent CSA that, depending on switching regime $z_j$, we have or zero collateral and full bilateral CVA or perfect collateralization (and a null CVA term), namely

$$BCVA_t^C + Coll_t^C := BCVA_t 1\{z_j = 1\} 1\{\tau_j \leq t < \tau_{j+1}\} + S_t^f 1\{z_j = 0\} 1\{\tau_j \leq t < \tau_{j+1}\}$$

for all $\tau_j, z_j \in u$ and $\tau_j \in [0, T \land \tau]$. This implies, from the contingent price process $S_t^C$ and $BCVA_t^C$ relations that

$$S_t^C = S_t = \begin{cases} S_t^f + BCVA_t & \text{if } z_j = 1 \text{ and } \{\tau_j \leq t < \tau_{j+1}\} \\ S_t^f & \text{if } z_j = 0 \text{ and } \{\tau_j \leq t < \tau_{j+1}\} \end{cases}$$

reminding that when collateralization is active also the funding issue $C_t^{Fund}$ enters the picture.
This suggests us that the pricing and hedging problem needs an opportune decomposition in order to find the price-hedge related to the switching regime defined optimally by the counterparty whose objective becomes now - in this pricing context - to minimize the hedging error/cost (or its variance) from the value of the whole contract.
Remarks 7.4.1. The main theoretical issue here is the existence of a martingale pricing measure $Q^*$ such that, for the assumed decomposition of the price process depending on the possible switching regimes, it ensures that the price of the contract remain arbitrage-free. As already mentioned, this measure is not unique given the impossibility to build a perfect replicating strategy, but a minimization of the hedging error or its variance is the strategy typically followed, with the pricing measure existence (not unique) that is assumed as given. We also make the same assumption here and the basic solution idea is to solve the pricing problem recursively in order to define the optimal switching strategy that now has to minimize the mean/variance error/costs of the hedging strategy of the whole contract with contingent CSA.

More specifically, to determine the hedge process/strategy we need also to assume - as in Crepey - the existence of a family of primary market assets $P^\text{asset}$ to be used as hedging assets\footnote{In this family is central the assumption on the presence of CDS (freshly emitted or par CDS) necessary to hedge the CVA term and the counterparty default risk through a rolling strategy.} in addition to the funding asset as defined in section 2.6. These assets are then used to build the self-financing hedge portfolio, say $W$ in which enter all the returns, the costs, gain/loss related to the hedging assets which are used to hedge (minimizing) all the risks of our whole contract:

- market risk, say $R^\text{mkt}$;
- counterparty/default risk, say $R^\text{def}$;
- funding risk, say $R^\text{Fund}$.

Without specifying formally (in relation to a given claim) all the hedging assets and the relative dynamics that enter the portfolio wealth process $W$, we simplify its definition by considering the whole risks defined above $R^\text{mkt}$, $R^\text{Fund}$ and $R^\text{def}$ and assuming that - depending on the active switching regime - a different hedging strategy will be set up or set off in order to minimize the risk that will be so decomposed in some way between the two switching regimes. To be more formal, we define as follows the decomposition of the wealth process related to the hedge set up by the counterparty of our defaultable contract with contingent switching CSA.

Definition 7.4.1 (Hedging strategy decomposition). Let us define as $(\phi^z, \phi^\zeta)$ the $\mathcal{F}_t$-predictable hedge vector process and with $(\epsilon^z, \epsilon^\zeta)$ the hedge error/cost process. Given the contingent characteristic of the CSA with two switching regimes, the wealth produced by the set up of the hedge admits the following decomposition

$$dW_t = \begin{cases} \phi^z_t (dR^\text{mkt} + dR^\text{def}) & \text{if } z_j = 1 \text{ and } \{\tau_j \leq t < \tau_{j+1}\} \\ \phi^\zeta_t (dR^\text{mkt} + dR^\text{Fund}) & \text{if } z_j = 0 \text{ and } \{\tau_j \leq t < \tau_{j+1}\}. \end{cases}$$

$\forall \tau_j, z_j \in \mathbb{T}, \mathbb{Z}, (j = 1, \ldots, M)$. This becomes, taking in account the hedge error/cost process $(\epsilon^z, \epsilon^\zeta)$,
7. Generalization to stochastic games and the price-hedge problem.

\[ dS^*_t = \begin{cases} 
\phi^*_z(dR^{mkt} + dR^{def}) + dc^*_t & \text{if } z_j = 1 \text{ and } \{ \tau_j \leq t < \tau_{j+1} \} \\
\phi^*_z(dR^{mkt} + dR^{fund}) + dc^*_t & \text{if } z_j = 0 \text{ and } \{ \tau_j \leq t < \tau_{j+1} \}.
\]

with initial condition \( W_0 = S^*_0 \) (the contract fair value) and \( \epsilon^*_0 = \epsilon^*_0 = 0 \).

So, from this representation we have a portfolio hedge which is decomposed in two part differentiated for the asset (and the relative process) used to hedge the different risks contingent on the activation of the collateral. We underline that risks related to funding can be also present in CVA hedging (namely when \( z_j = 1 \)) but in section 2.6 we have simplified things (Hp 4), CVA is not funded) taking them separate and we keep this assumption also here.

The \( dc_t \) term represents the error process deriving from the hedging strategy which is not perfect due mainly to default risk, funding risk and switching costs, all depending recursively. In fact one has to take in account also the instantaneous costs due to switching \( c^*_t \), which can represent the cost to set off and unwind an hedging strategy and pass to the other one related to the other switching regime. Although their presence, we assume that the hedge portfolio remains self-financing and that they contribute to the hedging error term (which has to be minimized).

The problem of finding the price-hedge, represented by the triple \((S^*, \tilde{\phi}, \tilde{\epsilon})\) in presence of contingent CSA of switching type is a more delicate task but the tool needed is the same that we have used to find the solution of the related control problem, that is Snell envelope and backward SDE (which are connected from the well known results of El Karoui et al. (1997)). In particular, given the presence of the already mentioned instantaneous switching costs due to hedging, these represent the barriers that determine the convenience to switch to the other regime, so actually one needs to solve a system of two (as state regimes) reflected BSDE, that can be expressed in general terms under the same condition on the processes of definition 5.4, as follows

\[
\begin{cases}
Y^z_t = \int_t^T \Psi^z(s, W_s)ds - \int_t^T N^z_sdW_s + A^z_T - A^z_t, \\
Y^z_t \geq (Y^z_T - c^z_T), \ \forall \ t \in [0, T] \\
\int_0^T (Y^z_t - (Y^z_T - c^z_T))dA^z_t = 0, \ A^z_0 = 0, \ \forall \ t \in [0, T];
\end{cases}
\]

\[
\begin{cases}
Y^\zeta_t = \int_t^T \Psi^\zeta(s, W_s)ds - \int_t^T N^\zeta_sdW_s + A^\zeta_T - A^\zeta_t, \\
Y^\zeta_t \geq (Y^\zeta_T - c^\zeta_T), \ \forall \ t \in [0, T] \\
\int_0^T (Y^\zeta_t - (Y^\zeta_T - c^\zeta_T))dA^\zeta_t = 0, \ A^\zeta_0 = 0, \ \forall \ t \in [0, T].
\end{cases}
\]

where now the generator functions \( \Psi(.^z), \Psi(.)^\zeta \) are given by the criterion chosen to minimize the hedging errors and are functions of the wealth dynamics \( W \) and \( N \) that represent the hedge vector.

Let us underline that the value and the hedge processes \((Y, N)\) depending on wealth are actually hidden inside \( W \).

The RBSDE existence and uniqueness of the solution \((Y^z_t, Y^\zeta_t)\) can be showed under the same
assumptions of the already mentioned paper whose other important result is the characterization of the solution through Snell envelope as follows

\[ Y^z_t = \text{ess sup}_{\tau \geq t} \mathbb{E}^*[\int_t^\tau \Psi^z(s, W_s)ds + (Y^z_{\tau} - c^z_\tau)1_{\{\tau < T\}}|\mathcal{F}_t], \quad Y^z_T = 0, \]

\[ Y^\zeta_t = \text{ess sup}_{\tau \geq t} \mathbb{E}^*[\int_t^\tau \Psi^\zeta(s, W_s)ds + (Y^\zeta_{\tau} - c^\zeta_\tau)1_{\{\tau < T\}}|\mathcal{F}_t], \quad Y^\zeta_T = 0. \]

where now the expectation is taken under the martingale pricing measure \( Q^* \). So, the two dimensional vector solution \((\bar{Y}^z, \bar{N}^z, \bar{A}^z)\) of the RBSDE system represents the price and hedge with minimized error for our problem, namely \((\bar{S}^z, \bar{\phi}^z, \bar{e}^z)\) for the two switching regimes.

Setting apart the problem of the error minimization and the related criterion which involves technical aspects of martingale and duality theory and a deeper analysis of the error process \( \epsilon_t \), we underline that also the assumed decomposition of the hedge portfolio based on the switching regime, which allows to get the price of the contract as the decomposition of the two solutions, is not in general unique and self-financing. Clearly, especially the self-financing condition is central to get an arbitrage free price for the claim and for the whole construction based on RBSDE.

So, one should impose in order to meet these conditions that:

**HP 1)** the portfolio hedging strategy set up at every switch remains self-financing;

**HP 2)** the funding is separate and does not contribute to the risk of other terms.

The first assumption can be interpreted as if one assumes that over time the hedging costs/revenues and the switching costs \( c^z_\tau \) compensate each other and no other inflow of money are posted in the strategy. The second hypothesis can be easily met in our twofold regime, but with different hypothesis on funding or with a different switching set it would not be valid anymore.

We are now in the conditions to state the following general theorem for pricing and hedging of a general contract with switching type CSA. Its rigorous proof relies mainly on the theory of BSDE and martingales.

**Theorem 7.4.1 (Price-hedge for switching type CSA defaultable contract).**

Let us assume a defaultable contract with switching type CSA and a stochastic framework as defined in chapter two. In addition, assume that:

1. exists (but is not unique) a martingale measure \( Q^* \) equivalent to the real/objective one \( Q \);

2. exists a primary market rich enough to set up/off over time a self-financing strategy which admits, under HP 1 and HP 2, a decomposition like the one proposed in definition 7.4.1 for \( dW \) and \( dS \);

3. are valid the conditions for the existence of the solution and its Snell envelope characterization, of a system of reflected BSDE, as set in definition 5.4 and El Karoui et al. (1997).
Generalization to stochastic games and the price-hedge problem.

Under this conditions the solution of the RBSDE system defined above exists and it gives us the price and hedge for our problem, that is

\[
\begin{align*}
\tilde{N}^* &= \tilde{\phi}^* \\
\tilde{A}^* &= \tilde{\epsilon}^* \\
\tilde{Y}^* &= \tilde{S}^*
\end{align*}
\]

where the error \( \tilde{\epsilon}^* \) minimizes a certain criterion, typically the variance, \( \tilde{\phi}^* \) is the optimal hedging strategy and \( \tilde{S}^* \) is the price of the contract which admit, by decomposition (definition 7.4.1), the following representation

\[
S_t^* = S_t^{*z} + S_t^{*\zeta}
\]

Remarks 7.4.2. Let us recall that under the assumption of markovian dynamics for the system, the problem admits also in this case a PDE representation in form of variational inequalities whose solution in viscosity sense coincide with the probabilistic one, as stated in theorem 5.6.

The price-hedge solution proposed in this section is clearly funded on strong assumptions, but if they had been proved to be necessary and/or sufficient for the theorem 7.4.1 (which is actually a conjecture), it would represent a relevant result in pricing claims with switching type contingencies (for which as far as I know, no similar results are known). Of course, also from the practical point of view, it would be important to implement the approach and apply it to different specific contingent claim in order to test numerically the solution.
8. APPLICATIONS OF SWITCHING CONTROL MODEL IN FINANCE.

To doubt everything or to believe everything are two equally convenient solutions; both dispense with the necessity of reflection.

J.H. Poincaré

8.1 Some interesting cases.

Switching type mechanisms like the one we have analyzed can find different applications in the wild finance world. The basic underlying idea is to ensure flexibility to agents investments decisions over time which is a usual objective in real option theory. Our problem has been thought mainly in a risk management view but with the development of new techniques and algorithms, also the related pricing problem will be tackled efficiently and more financial contract would find useful and convenient - in a optimal risk management view - this type of contingent mechanisms.

As regards just a possible further application in risk management, it would be important to deepen the analysis of a switching type collateralization from a portfolio perspective (taking for example also the view of a central clearing). In particular, it would be relevant to show possibly analytically but mainly with numerical examples, the greater convenience of the switching/contingent solution respect to a non contingent/standard collateral agreement like the partial or full one (including clauses like netting and others).

In a pricing view, we remind the example - of the fixed income market - some particular bonds called flippable or switchable, that are characterized by options to switch the coupon from fix to floating rate. Clearly, in this case the valuation is easier given that this securities have a market and are not traded OTC so one does not need to include in the picture counterparty risk, funding and CSA cashflows. Anyway, it would be interesting to delve into the valuation of an OTC contract in witch also the dividend flows can be subject to contingent switching mechanism. As regards similar case, a problem that can be very interesting and difficult to tackle is the valuation of a flexi swap in presence of a contingent collateralization like our one.

The main characteristics of a flexi-swap are:

a) the notional of the flexible swap at period $n$ must lie (inclusively) between predefined bounds $L_n$ and $U_n$;
b) the notional of the flexible swap at period $n$ must be less than or equal to the notional at the previous period $n - 1$;

c) The party paying fixed has the option at the start of each period $n$ to choose the notional, subject to the two conditions above.

In other words, we deal with a swap with multiple embedded option that allows one party to change the notional under certain constraints defined in the contract. This kind of interest rate swaps are usually used as hedging instruments of other swaps having notional linked to loans, especially mortgages. The underlying idea is that the fixed-rate payer (the option holder) will amortize as much as allowed if interest rates are very low, and will amortize as little as allowed if interest rates are very high.

Given a payment term structure $\{T_n\}_{n=0}^N$ and a set of coupons $X_n$ (with unit notional) fixing in $T_n$ e paying in $T_{n+1}$ ($n = 0, 1, \ldots, N - 1$), the *flexi swap* is a fixed vs floating swap where the fixed payer has to pay a net coupon $X_n R_n$ in $T_{n+1}$, the notional $R_0$ is fixed upfront at inception and for every $T_n$, $R_n$ can be amortized if it respects some given constraints defined as follows:

1. deterministic constraints : $R_n \in [g_{n_{\text{low}}}, g_{n_{\text{high}}}]$;

2. local constraints function of the current notional: $R_n \in [l_{n_{\text{low}}}(R_{n-1})]^{n_{\text{high}}}(R_{n-1})]$;

3. market constraints (libor, swap denoted with $X_n$): $R_n \in [m_{n_{\text{low}}}(X_n), m_{n_{\text{high}}}(X_n)]$.

The valuation procedure of this type of swap involves a backward recursion keeping track of the notional in every payment date. But introducing also the switching collateralization, the valuation becomes an intricate puzzle given the recursive relation between the optimal switching strategy, the price process of the claim which in addition depends on the optimal notional choice over time. Surely, one needs simplifying assumptions to break the curse of recursion.

Different other applications are possible to more complex and exotic products, but we focus the attention to the interesting problem of finding of close/exit from a long term swap in presence of a consultant in order to take in count the information asymmetries.

### 8.2 Optimal swap closure/stopping time.

The switching control problem can be a fruitful approach to tackle an other application related to the optimal stopping/closing of a financial contract like an interest rate swap or a more complex derivative claim. In this specific case the switching is set to be unique, so the problem is reduced to an optimal stopping time with single regime that is the contract closure/extinction.

More specifically, given some conditions defined within a contract between the counterparties, it

---

1. In this sense is like there were a third reference represented by the pool of loans.
2. Here note that the problem become relevant when the other party that play the is a third party like a financial consultant.
could be important and financially relevant to find the optimal stopping/closing time (if it exists) of the claim that maximizes the revenues of the operation or minimizes the costs. This implies the optimality of the debt management strategy via interest rate swap, whatever the swap was of hedging or speculative type.

One of the most interesting case that we have in mind is that of a typical interest rate swap written with the help of a consultant, say $C$, by a public sector agent $A$ with a bank as counterparty $B$. We, therefore, assume information asymmetries between $A$ and $B$, $A$ and $C$ but not between $B$ and $C$ so that we analyze the problem from $A$ and $C$ perspective because we assume that $B$ has not the interest to close the contract. This is what happens in reality, because the bank has a portfolio of derivatives that manages with general portfolio strategies, so for her is not relevant the single position. For what concerns $C$, we will consider her perspective too because we will highlight the big conflict between the $C$ objective which is exactly the opposite to that of $A$.

**Problem formulation.**

Assume a markovian setting (without counterparty risk) and a continuous time model characterized by a probability space described by the triple $(\Omega, \mathcal{F}_t, \mathbb{Q})$ in which the market filtration is generated by a Brownian motion $W$ which characterize the diffusive part of the stochastic factor $X$ that we set to model the term structure and the price process (NPV) of a given interest rate swap (a typical short rate markovian model). For convenience, the counterparty risk is assumed negligible, so that $\lambda_t = 0$.

We assume as given the condition of the swap defined between the parties. So from the contract we know the payment dates set, say $D = \{t_1, \ldots, t_n \leq T \text{ with } D \subseteq [0,T]\}$ that is the whole duration of the contract. Now, because at every payment date we have a net coupon to pay or receive, we define the *cumulated net coupons* $\text{CNC}_t$ as the stochastic process adapted to the market filtration which can be seen (for a given realization) as a (typically square integrable and continuous) function $\phi(t, X, \alpha) : [D \times \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}]$ of time (for $t \in D$) of the stochastic factor $X$ and of a set of given parameters $\alpha$ (so that it’s a vector of constants).

In the simpler case of fixed for floating IRS with unitary notional and $K$ as the swap rate of the fixed leg, we have for $t = 0$

$$\text{CNC}_t = \sum_{s=t_1}^{T} (X_s - K), \quad \text{CNC}_{t=0} = 0 \quad (8.1)$$

By the other side, the counterparty $A$ has the faculty to choice the optimal time to exit because with the help of $C$ she can monitor the "mark to market" that is the price process $NPV_t$. This is a typical càdlàg process $\mathcal{F}_t$-adapted, described by the function $\psi(t, X, \alpha) : [0, T] \times \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$. As in the latter case, given the discount factors $B_t = \exp(-r t)$ (that are determined by simulating $X$), the price process would be defined by (in $t = 0$)

$$NPV(t) = \sum_{t=0}^{T} B_s[X_s - K], \quad NPV(0) = 0 \quad (8.2)$$
Given the above definitions, the objective of the agent A would be - for a given valuation time \( t \in [0,T] \) in which we assume to know the initial condition that are the values of the net cumulated coupon \( CNC_t \) payed until \( t \) and the mark to market \( NPV_t \) of the swap - to find the optimal stopping time \( \tau \in T \) (where \( T \subseteq [0,T] \) is the set of the stopping times \( \tau \)) to exit from the contract that maximizes the expected incomes given by the cumulated expected coupons from \( t \) to \( \tau \) and the \( NPV(\tau) \). Symmetrically, if the initial condition are negative, the problem can be set in terms of costs minimization, searching for the infimum over the stopping time set \( T \) that minimizes the expected costs for the agent.

Formally, to complete the formulation of the problem, we need to consider even the costs that emerge from stopping, namely the switching costs that derive from closing the activity. These are typically the value of the mark to market in \( \tau \) that can be positive or negative, plus the hedging costs, say \( c_{\text{hedge}}^\tau \) and possible ”hidden commission”, say \( c_{\text{hidden}}^\tau \) that the bank charges, but because of \( C \) these can be made explicit and known to \( A \).

So, setting \( NPV(\tau) \) within the \textit{instantaneous switching} term and defining the other switching costs as

\[
C(\tau)^{\text{Switch}} = c_{\text{hedge}}^\tau + c_{\text{hidden}}^\tau
\]  

we are ready to give the complete formulation of this optimal switching/stopping control model, defining the value function as follows

\[
J^T(y) = \sup_{\tau \in T, T} \mathbb{E}^y \left[ \int_t^\tau e^{-\beta(s-t)} F(Y_s)ds + e^{-\beta \tau} G(Y_\tau) \right], \quad y = (t,x)
\]  

that becomes, substituting the functions defined above,

\[
J^T(y) = \sup_{\tau \in T, T} \mathbb{E}^y \left[ \int_t^\tau e^{-r(s-t)} (CNC(s))ds + e^{-rT}[NPV(\tau) + C^{\text{Switch}}(\tau)] \right], \quad y = (t,x)
\]  

such that

\[
\begin{align*}
\text{d}X_t &= b(t,X_t)dt + \sigma(t,X_t)dW^x_t \quad \text{for} \ t \in [0,T] \\
X(0) &= x_0 \\
\mathcal{C}_{ad} &= \{T_t \}, \text{ finite} \\
J^T(y) &= : NPV(T) = CNC(T) \Rightarrow "\text{Terminal condition}"
\end{align*}
\]  

where \( \mathcal{C}_{ad} \) indicates the admissible control se (assumed finite) while the dynamic for \( X \) is a general \textit{Ito diffusion} (that admits infinitesimal generator \( A(x) \)) and can be further specified, typically through a CIR/CIR++ process or even a two factor short rate process (like G2++).

\textbf{Remarks 8.1.} We remark here that we are actually considering no costs deriving from the monitoring activity of the consultant \( C \) that would complicate the objective functional and the model. Therefore, for what concerns the role of counterparty \( C \), if we assume that \( C \) is just interested in making profit from monitoring the \( A \) swap position, we can assume the same objective for \( C \) as \( A \). But if \( C \) is interested too in its consultancy service to \( A \) to help her in the recovery
of the losses incurred during the life of the contract (gaining, for example, a percentage of sums recovered) is clear that the objective would be the opposite of $A$, namely wait for negative NPV level in order to get a great loss on which works to recover and get a greater gain.

Remarks 8.2. The solution of this kind of problem can be approached using the same dynamic programming principle and the Snell Envelope technique and the numerical algorithm used for our model (defining the extinction value and the continuation value going backward). The dependence on the state variable of the switching cost pose the problem of the continuity of the solution and its uniqueness.

From the numerical point of view, different approach can be also employed other than least squares Monte Carlo (Longstaff-Schwartz approach), mainly PDE discretization methods, markov chain approximation, or also lattices and trees methods.
9. CONCLUSIONS.

The present research work has been focussed on the study and the analysis of general *defaultable OTC contract* in presence of the so called *contingent CSA*, that is a counterparty risk mitigation mechanism of switching type. The underlying economic idea is to show that collateralization is useful to reduce the default losses but it imposes to take in account the cost of collateral and funding, in contrast, by the other side, to the (bilateral) CVA costs. In this sense we have been lead to analyze this problem from the *risk management and optimal design* point of view, taking the perspective of just one party of the contract. By tackling this task, we have formulated a *stochastic switching control model* in which the agent/counterparty can optimally decide to switch from zero to perfect collateralization over the life of the underlying contract. We have analyzed the main possible approaches - analytical/PDE and probabilistic/BSDE - in order to solve this problem, finding in the probabilistic one the most adapt to tackle our highly non linear and recursive model.

We have also studied the numerical approaches for the solution coming to the definition of an algorithm based on an *iterative optimal stopping* approach combined with the *Longstaff-Schwartz* simulation method. The results and tests in the specific case of a defaultable interest rate swap, show clearly the relevance of the switching mechanism in reducing the overall costs respect to the perfect and zero collateral, that is in absence of contingent CSA. Of course further tests and alternative numerical solutions are needed and we have left them for future research.

Therefore, we have generalized also the model to the bilateral case allowing the strategic interaction between the agents. The problem in this case become a stochastic differential game (a generalized Dynkin game) of switching type, for which - as far as we know - no solution for a *Nash equilibrium point* is known. We have proved its existence under strong condition (in the so called *symmetric case*) but obviously further research in the field of stochastic games is needed to delve and attack the problem.

Then, we have passed to analyze the problem of pricing and hedging a general contract with contingent CSA, proposing a solution of the recursive problem through a decomposition of the hedging strategy on the base of the switching times and indicator and recurring to the tool of the reflected backward SDE. In particular, theorem 7.4.1 if proved to be true, it would represent a relevant result in pricing claims with this type of switching contingencies.

In the last chapter we have highlighted some of the possible generalizations and variants of the model proposed that can be interesting to analyze focussing in the end on the valuation of the optimal time to unwind a swap in presence of asymmetries between three agents.
REFERENCES


10. References


10. References


