Numerical PDE Methods for Pricing Bilateral XVA

by

Yuwei Chen

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Computer Science
University of Toronto

Abstract

Numerical PDE Methods for Pricing Bilateral XVA

Yuwei Chen
Doctor of Philosophy
Graduate Department of Computer Science
University of Toronto
2023

Adjusting derivative prices to take into account default risk has attracted the attention of several researchers and practitioners, especially after the 2007-2008 financial crisis. The thesis is a study, via numerical Partial Differential Equation (PDE) approaches, of the modeling and computation of valuation adjustment problems in financial derivative pricing if we consider the bilateral counterparty default intensities.

Under some assumptions, the valuation of financial derivatives, including a value adjustment to account for bilateral default risk (XVA), assuming constant default intensities gives rise to a nonlinear PDE [11]. We formulate a penalty iteration method to handle the nonlinearity in this PDE, study its convergence, and extend it to American derivatives, resulting in a double-penalty iteration. We also propose boundary conditions and their discretization for the XVA PDE problem in the cases of call, put options, and forward contracts. Numerical results demonstrate the effectiveness of our methods.

Then, we derive a novel PDE for derivative pricing including XVA, assuming that the default risk of one of the counterparties follows a Cox-Ingersoll-Ross (CIR) process, while the other party has constant default risk. The time-dependent PDE derived is of Black-Scholes type and involves two "space" variables, namely the asset price and the buyer default intensity, as well as a nonlinear source term. We formulate boundary conditions appropriate for the default intensity variable. The numerical solution of the PDE uses a penalty-like iteration for handling the nonlinearity. We also develop a novel asymptotic approximation formula for the adjusted price of derivatives, resulting in a very efficient, accurate, and convenient for practitioners formula. Numerical results indicate stable

second order convergence for the 2D PDE solution in terms of the discretization size and convergence of order at least 1.5 for the asymptotic approximation in terms of inverse of the mean-reversion rate.

We extend the PDE model we developed to price European derivatives including XVA, considering stochastic counterparty default intensity, to American derivatives. We also extend the asymptotic approximation to the American put option problem. A key step is to derive the asymptotic approximation to the free boundary. We present numerical experiments in order to study the accuracy and effectiveness of the 2D PDE and asymptotic approximations.

Acknowledgements

First and foremost, I have to express my most sincere gratitude to my supervisor, Professor Christina Christara, for her invaluable research guidance, inspiration, and feedback. This thesis could not have been accomplished without her insights and knowledge. Besides academic supervision, Christina has also provided me great encouragement and support during difficult times, especially during the pandemic period. She is not only a supervisor, but also a life mentor to me. No words can express my thanks to her over these past years.

I would also like to show my gratitude to my committee members, Professor Kenneth Jackson and Professor Wayne Enright, who have provided me insightful, invaluable comments and suggestions to my research ideas, published paper and this thesis. I would also like to thank University of A Coruña faculty member Professor Carlos Vázquez Cendón for agreeing to be external examiner, and University of Toronto faculty member Professor Kirill Serkh for joining my final oral examination.

I am also grateful to my friends and colleagues, in University of Toronto, for all the fun and friendship we have in the last years. I would especially like to thank Dr. Selwyn Yuen, for offering me the opportunity to engage in interesting projects at Canadian Pension Plan Investment Board (CPPIB).

Last but not least, I would like to thank my parents, Zhihui Chen and Liping Zhang, and my wife, Yongchu Liu for their unconditional love and supports.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Credit/Debt valuation adjustment

Counterparty credit risk [36, 28], also known as default risk, is the risk that, during the duration of an agreement, one party fails to make payments to another party. Derivatives also expose counterparties to credit risk. Since the 2007-2008 financial crisis, it has become a standard practice for derivatives' sellers to adjust the value of the derivatives transactions to reflect the possible losses from counterparty default. This adjustment is usually called credit valuation adjustment (CVA) [36, 28]. However, a more complete valuation adjustment includes debt valuation adjustment, funding valuation adjustment, etc.

Debt valuation adjustment (DVA) [36, 28] is an estimate of the costs to the counterparty of the seller's default risk. It can also be seen as the CVA from the point of view of the counterparty. Funding valuation adjustment (FVA) is an adjustment to the value of a derivative or a portfolio which is designed to make sure that a dealer recovers its average funding costs when he or she trades and hedges derivatives.

Although it is still controversial whether or not a seller should charge DVA or FVA to value derivatives [35, 12, 30], in this work, we leave the debates behind and mainly focus on the mathematical model of valuation adjustment and the numerical method to solve the arising partial differential equation (PDE). We now define the total valuation adjustment [28, 3] as

$$XVA = DVA - CVA + FVA.$$

If V is the unadjusted value of the derivative, e.g. calculated by Black-Scholes-Merton model, and $\hat{V}$ is the value after taking credit risk into account, then the valuation ad-

justment is

$$XVA = \hat{V} - V. \tag{1.1}$$

Note that, the CVA (or XVA) is normally calculated at a counterparty level and not the trade level. This means the CVA can be different to different counterparties even if they have the same derivatives.

The Basel III regulation framework was developed in response to the 2007-2008 financial crisis. The committee set rules for OTC market that bilateral trading needs to take default risk and funding costs into account. Once the rules were set, several methods and frameworks were developed for valuation of derivatives under counterparty risk. Piterbarg [42] derived valuation formulas for the price of derivatives, which incorporate funding cost and collateral agreements. The new model forces an adjustment to the discounting term of the Black-Scholes PDE. Burgard and Kjaer [10, 11, 13] generalized Piterbarg's model to include sellers and buyers' default risk. They applied the replication portfolio approach to derive the PDE representations of derivatives' values with bilateral risk and funding costs. Moreover, via the Feynman-Kac theorem, the solution to the resulting PDE can also be written in terms of expectation. Usually, simulation approaches such as Monte Carlo are a popular choice for CVA or XVA pricing. In [28, 29], authors showed detailed derivation of expectation formulae of CVA or bilateral CVA [1] (BCVA). The ease to implement makes expectation or simulation methods more widely used by practitioners, especially when the CVA or XVA involves more than 5 underlying assets in the corresponding portfolio. This approach suffers less from the curse of dimensionality. Brigo and his coauthors [7, 8, 9] also introduced a general valuation framework for calculating the bilateral CVA by an expectation representation of risk valuation including CVA, funding spread and collateralization. Simulation approaches are used to approximate the arbitrage-free valuation of bilateral counterparty risk under collateralization. In [7, 8, 9], the analysis focuses on credit default swaps (CDS) as underlying portfolios and claimed CVA can be viewed as an option, the so-called contingent credit default swap, on the clean value of the contract. Another popular approach to adjustment valuation is based on *backward stochastic differential equation* (BSDE) analysis. Crépey [17] introduced a BSDE approach to valuation and hedging of bilateral counterparty risk on OTC derivatives under funding constraints. Bichuch, Capponi and Strurm [6] also developed an arbitrage-free valuation framework with a BSDE for the price of European claim taking into account funding spread, the repo market, collateral servicing costs and counterparty credit risks. They showed that the XVA is unique in the absence of rate

[1] In some textbooks or literatures, people use BCVA to represent the bilateral valuation adjustments considering both CVA and DVA, i.e. BCVA = DVA - CVA.

asymmetries. A PDE representation of the BSDE was also shown and studied.

Regarding the PDE arising from the XVA problem, depending on how the derivative is valued in the case of either party of the contract defaults, a nonlinear term may be included [11]. Hence, appropriate numerical methods are needed to handle the nonlinear term, as well as the discretization of the PDE. Arregui, Salvador, and Vázquez consider the PDE models in [11] and price European options with XVA [3], as well as American options with XVA [4]. In both cases, they use the characteristics method for timestepping, a finite element method for the space discretization, and a fixed-point iteration scheme to handle the nonlinearity from the XVA. The experimental results indicate that their method is first order convergent. Arregui, Salvador, and Vázquez [5] also developed a Monte Carlo approach to American options pricing including XVA. They present the adaptation of Monte Carlo to numerically solve the nonlinear term in the XVA problem.

In all the PDE model works we mentioned above, the default intensities are assumed to be constant. In reality, default intensities exhibit stochasticity [18]. In addition, people also notice that, there is dependency between exposure (or underlying assets) and counterparty credit risk, which is usually called wrong/right way risk [34]. Therefore, a multi-stochastic factors' model becomes necessary to reflect these issues. In [19], Feng modeled CVA for European options under a Bates model with the stochastic intensity of the jump to default following a Cox-Ingersoll-Ross (CIR) process. A numerical PDE-based Monte Carlo framework was built, consisting of path simulation, independent exposure estimation and CVA computation. In [41, 2], the authors considered the stochastic short term credit default swap (CDS) spread and resulted in a multi-dimensional in space PDE. In [2], existence and uniqueness of the solution to the nonlinear PDE was proved as well. In this thesis, we derive the PDE assuming the default intensity of the counterparty follows a stochastic process, namely a CIR process. This also results in a multi-dimensional in space nonlinear PDE, which provides an effective and more direct framework to handle the correlation between underlying assets and default intensities. We also extend this novel PDE model to the American case.

1.2 The expectation representation

In this subsection, we briefly introduce the steps in Gregory's textbook [29] to derive the expectation representation for BCVA in the case of no wrong-way risk. No wrong-way risk means there is no dependency between exposure, default and maturity.

We assume the derivative value including CVA and DVA is denoted as $\hat{V}(t,T)$, with the maturity of derivative being T. In this subsection, we also assume the mark-to-

market value M to be the risk-free value $V(t,T)$. Then $V(s,T)$, $t < s \leq T$, denotes the future uncertain mark-to-market value accounting for the discounting effects. We refer to the seller as party B, to the counterparty as counterparty C or just party C, the default time of party B as τ_B, default time of party C as τ_C, and the first-to-default time as $\tau_f = \min(\tau_B, \tau_C)$. We also denote by R_B and R_C the recovery rates of two parties respectively. The positive and negative values of any security Z are denoted by $Z^+ = \max\{Z, 0\}$ and $Z^- = \min\{Z, 0\}$. Here are the four parts of payoffs of $\hat{V}$ we should consider:

- When **neither party B or C defaults before** T, the derivative value including CVA and DVA, $\hat{V}$, is equal to the risk-free value, V, the payoff can be written as

$$\mathbb{1}(\tau_f > T)V(t,T),$$

  where $\mathbb{1}(\tau_f > T)$ is an indicator function which returns 1 if first-to-default has not occurred before T, and zero otherwise.
- When **counterparty C defaults first and before** T, the party B will receive a recovery fraction R_C of the mark-to-market $V(\tau_C, T)$, if $V(\tau_C, T)$ is positive, and have to settle the amount $V(\tau_C, T)$, if negative. Hence, the default payoff becomes

$$\mathbb{1}(\tau_f \leq T)\mathbb{1}(\tau_f = \tau_C)(R_C V^+(\tau_f, T) + V^-(\tau_f, T)).$$

- When **counterparty B defaults first and before** T, the party B owe a recovery fraction R_B of the mark-to-market $V(\tau_B, T)$, if $V(\tau_B, T)$ is negative, and receive the mark-to-market value $V(\tau_B, T)$, if positive. Hence, the default payoff becomes

$$\mathbb{1}(\tau_f \leq T)\mathbb{1}(\tau_f = \tau_B)(V^+(\tau_f, T) + R_B V^-(\tau_f, T)).$$

- When **either party B or C defaults before** T, all the cashflow prior to first-to-default date τ_f are paid as

$$\mathbb{1}(\tau_f \leq T)V(t, \tau_f).$$

Combing all parts of payoff together, the derivative values including CVA and DVA can be written as

$$\begin{aligned}\hat{V}(t,T) = \mathbb{E}^{\mathbb{Q}}\Big[&\mathbb{1}(\tau_f > T)V(t,T) + \mathbb{1}(\tau_f \leq T)V(t,\tau_f) \\ &+\mathbb{1}(\tau_f \leq T)\mathbb{1}(\tau_f = \tau_C)(R_C V^+(\tau_f, T) + V^-(\tau_f, T)) \\ &+\mathbb{1}(\tau_f \leq T)\mathbb{1}(\tau_f = \tau_B)(V^+(\tau_f, T) + R_B V^-(\tau_f, T))\Big] \qquad (1.2)\end{aligned}$$

Since $V(\tau_f, T) = V^+(\tau_f, T) + V^-(\tau_f, T)$, we have

$$
\begin{aligned}
\hat{V}(t,T) &= \mathbb{E}^{\mathcal{Q}}\Big[\mathbb{1}(\tau_f > T)V(t,T) + \mathbb{1}(\tau_f \le T)V(t,\tau_f) \\
&\qquad +\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_C)(R_C V^+(\tau_f,T) + V(\tau_f,T) - V^+(\tau_f,T)) \\
&\qquad +\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_B)(V(\tau_f,T) - V^-(\tau_f,T) + R_B V^-(\tau_f,T))\Big] \\
&= \mathbb{E}^{\mathcal{Q}}\Big[\mathbb{1}(\tau_f > T)V(t,T) + \mathbb{1}(\tau_f \le T)V(t,\tau_f) \\
&\qquad +\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_C)V(\tau_f,T) \\
&\qquad +\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_B)V(\tau_f,T) \\
&\qquad +\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_C)(R_C V^+(\tau_f,T) - V^+(\tau_f,T)) \\
&\qquad +\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_B)(R_B V^-(\tau_f,T) - V^-(\tau_f,T))\Big] \\
&= \mathbb{E}^{\mathcal{Q}}\Big[\mathbb{1}(\tau_f > T)V(t,T) + \mathbb{1}(\tau_f \le T)V(t,\tau_f) \\
&\qquad +\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_C)V(\tau_f,T) \\
&\qquad +\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_B)V(\tau_f,T) \\
&\qquad -\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_C)(1 - R_C)V^+(\tau_f,T) \\
&\qquad -\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_B)(1 - R_B)V^-(\tau_f,T)\Big] \\
&= \mathbb{E}^{\mathcal{Q}}\Big[\mathbb{1}(\tau_f > T)V(t,T) + \mathbb{1}(\tau_f \le T)V(t,\tau_f) + \mathbb{1}(\tau_f \le T)V(\tau_f,T) \\
&\qquad -\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_C)(1 - R_C)V^+(\tau_f,T) \\
&\qquad -\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_B)(1 - R_B)V^-(\tau_f,T)\Big] \\
&= \mathbb{E}^{\mathcal{Q}}\Big[\mathbb{1}(\tau_f > T)V(t,T) + \mathbb{1}(\tau_f \le T)V(t,T) \\
&\qquad -\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_C)(1 - R_C)V^+(\tau_f,T) \\
&\qquad -\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_B)(1 - R_B)V^-(\tau_f,T)\Big] \\
&= \mathbb{E}^{\mathcal{Q}}\Big[V(t,T) - (1 - R_C)\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_C)V^+(\tau_f,T) \\
&\qquad -(1 - R_B)\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_B)V^-(\tau_f,T)\Big] \\
&= V(t,T) - \mathbb{E}^{\mathcal{Q}}\Big[(1 - R_C)\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_C)V^+(\tau_f,T) \\
&\qquad +(1 - R_B)\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_B)V^-(\tau_f,T).\Big] \qquad (1.3)
\end{aligned}
$$

Hence, we can identify the BCVA, CVA and DVA as

$$\begin{aligned}\mathbf{BCVA}(t,T) = -\mathbb{E}^{\mathcal{Q}}\Big[&(1-R_C)\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_C)V^+(\tau_f,T) \\ &+(1-R_B)\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_B)V^-(\tau_f,T)\Big],\end{aligned} \tag{1.4}$$

If we define the bilateral valuation adjustment to be $\mathbf{BCVA} = \mathbf{DVA} - \mathbf{CVA}$, then

$$\mathbf{CVA}(t,T) = (1-R_C)\mathbb{E}^{\mathcal{Q}}\Big[\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_C)V^+(\tau_f,T)\Big], \tag{1.5}$$

$$\mathbf{DVA}(t,T) = -(1-R_B)\mathbb{E}^{\mathcal{Q}}\Big[\mathbb{1}(\tau_f \le T)\mathbb{1}(\tau_f = \tau_B)V^-(\tau_f,T)\Big]. \tag{1.6}$$

Since the expectation is over all times before maturity, we can integrate over all possible default times, assuming no dependency between exposure and default events, and we obtain that

$$\mathbf{CVA}(t,T) = (1-R_C)\mathbb{E}^{\mathcal{Q}}\Big[\int_t^T B(t,u)V^+(u,T))d\mathbb{P}_{\mathbb{C}}(u)\Big] \tag{1.7}$$

$$\mathbf{DVA}(t,T) = -(1-R_B)\mathbb{E}^{\mathcal{Q}}\Big[\int_t^T B(t,u)V^-(u,T))d\mathbb{P}_{\mathbb{B}}(u)\Big] \tag{1.8}$$

where $B(t,u)$ is the risk-free discount factor, $\mathbb{P}_{\mathbb{C}}$ is the survival probability for counterparty C and, $\mathbb{P}_{\mathbb{C}}$ is the survival probability for party B.
Hence, if we further assume the discount factors and survival probabilities are deterministic or independent with exposures, we can have

$$\mathbf{BCVA}(t,T) = -(1-R_C)\int_t^T B(t,u)\mathbf{EE}(u,T)d\mathbb{P}_{\mathbb{C}}(u) \tag{1.9}$$

$$-(1-R_B)\int_t^T B(t,u)\mathbf{ENE}(u,T)d\mathbb{P}_{\mathbb{B}}(u) \tag{1.10}$$

where the expected positive exposure is $\mathbf{EE}(u,T) = \mathbb{E}^{\mathcal{Q}}[V^+(u,T)]$ and the expected negative exposure is $\mathbf{ENE}(u,T) = \mathbb{E}^{\mathcal{Q}}[V^-(u,T)]$.

1.3 The PDE representation

In this subsection, we briefly review some of the results of the paper by Burgard and Kjaer [11], which provides a PDE representation for the value of a financial derivative if the effects of bilateral default risk and the funding costs are considered. This extension of the Black-Scholes PDE is driven by using several hedging arguments. The strategy

is that, the seller could short the zero-coupon bond of the counterparty to hedge out counterparty credit risk and repurchase its own bonds to hedge out its own credit risk. This PDE model is fundamental for the thesis.

Following [11], we refer again to the seller as party B and to the counterparty as counterparty C or just party C. We use the following notations

- S: the price of underlying asset.
- K: the strike price in the case of option or forward contract.
- μ: the drift in S.
- σ: the volatility in S.
- γ: the dividend rate of S.
- q: the repo rate of underlying asset.
- M: the mark-to-market value (or close-out value) when one party defaults.
- r: the risk-free interest rate.
- P_B: the price of party B's bond.
- P_C: the price of counterparty C's bond.
- r_B: the yield on party B's bond.
- r_C : the yield on counterparty C's bond.
- λ_B: the default intensity of party B.
- λ_C: the default intensity of counterparty C.
- R_B: the recovery percentage on M when party B defaults.
- R_C: the recovery percentage on M when party C defaults.
- r_F: the party B's funding rate for borrowed cash, where $r_F = r$ if derivative can be used as collateral, and $r_F = r + (1 - R_B)\lambda_B$ if derivative cannot be used as collateral.
- $s_F = r_F - r$: the funding spread in the case where there is one issuer bond.

As in the Black-Scholes framework, a certain self-financing portfolio is considered such that it could hedge out all underlying risk factors in the model. Let's assume the portfolio Π consists of the following four traded assets:

- D units of cash deposit in risk-free bank.
- α units of risky, zero-recovery, zero-coupon bond of party B whose price is P_B
- β units of risky, zero-recovery, zero-coupon bond of counterparty C whose price is P_C
- δ units of the underlying asset whose price is S.

We assume the prices of these assets except the credit-risky derivative are modeled by

$$\frac{dD}{D} = rdt \tag{1.11}$$
$$\frac{dP_B}{P_B} = r_B dt - dJ_B \tag{1.12}$$
$$\frac{dP_C}{P_C} = r_C dt - dJ_C \tag{1.13}$$
$$\frac{dS}{S} = \mu dt + \sigma dW \tag{1.14}$$

where W is a Brownian motion, J_B and J_C are two independent Poisson processes that jump from 0 to 1 when the corresponding party defaults. This assumption of J_B and J_C implies that, considering only unilateral credit risk, we can hedge the stochasticity of one of J_B or J_C, using the respective bond P_B or P_C. It is also assumed that default of any party or counterparty will not affect the spot price S. Since P_B and P_C are defined as zero-recovery, zero-coupon bonds, the spreads $r_B - r$ and $r_C - r$ are the same as the respective default intensities λ_B and λ_C.

Now assume two parties B and C enter a derivative contract on the spot asset according to which C pays B the amount $H(S)$ at maturity time T. Thus $H(S) > 0$ means that the seller receives cash from the counterparty, while $H(S) < 0$ means the seller needs to pay cash to the counterparty. We denote by $\hat{V}(t, S, J_B, J_C)$ the value of the derivative at time t which also depends on the spot price S of underlying asset and default state of two parties. When any party B or C defaults, the derivative terminates immediately. Later,we use $\hat{V}(t, S)$ as a abbreviation for $\hat{V}(t, S, 0, 0)$, while $V(t, S)$ denotes the value of same derivatives without considering any credit risk.

Let $M(t,S)$ be the mark-to-market value (or close-out value) of the derivative when one party defaults. Generally, $M(t,S)$ will be close to $V(t,S)$. Therefore, the vast majority of research papers on valuation with counterparty risk use $M(t,S) = V(t,S)$. Another scenario, where $M(t,S) = \hat{V}(t,S,0,0)$, is worth considering as well. Let $R_B \in [0,1]$ and $R_C \in [0,1]$ be the recovery rates (of default) on the derivative positions of parties B and C, respectively.

Denote $M^+ = \max\{M,0\}$ and $M^- = \min\{M,0\}$. When the seller (or party B) defaults first, if the mark-to-market value is positive to party B then $M^-(t,S) = 0$, $\hat{V} = M^+ = M$ which means the counterparty needs to pay the full mark-to-market value to the seller. If mark-to-market value is negative, then $M^+(t,S) = 0$, $\hat{V} = R_B M^- = R_B M$ which means the recovery value of mark-to-market value should be paid to the counterparty. The default of counterparty C could be analyzed in the same manner. Therefore, we have following boundary conditions for J_B and J_C:

$$\hat{V}(t,S,1,0) = M^+(t,S) + R_B M^-(t,S) \quad \text{Party B defaults,} \tag{1.15}$$

$$\hat{V}(t,S,0,1) = R_C M^+(t,S) + M^-(t,S) \quad \text{Counterparty C defaults.} \tag{1.16}$$

The replicating portfolio $\Pi(t)$ should hedge out the value of the derivative contract to seller at any time t i.e. $\Pi(t) + \hat{V}(t) = 0$. Thus,

$$-\hat{V}(t) = \delta(t)S(t) + \alpha(t)P_B(t) + \beta(t)P_C(t) + D(t). \tag{1.17}$$

By the assumption of self-financing portfolio,

$$-d\hat{V}(t) = \delta(t)dS(t) + \alpha(t)dP_B(t) + \beta(t)dP_C(t) + d\bar{D}(t) \tag{1.18}$$

where the change in the cash account [2] is

$$d\bar{D}(t) = \delta(\gamma - q)Sdt + \{r(-\hat{V} - \alpha P_B)^+ + s_F(-\hat{V} - \alpha P_B)^-\}dt - r\beta P_C dt. \tag{1.19}$$

[2] More details about the mechanism of cash account can found in [11].

Therefore, (1.18) becomes

$$\begin{aligned}-d\hat{V}(t) =&\delta(\mu S dt + \sigma S dW) + \alpha P_B(r_B dt - dJ_B) + \beta P_C(r_C dt - dJ_C) \\ &+ \delta(\gamma - q)S dt + \{r(-\hat{V} - \alpha P_B) + s_F(-\hat{V} - \alpha P_B)^-\}dt - r\beta P_C dt \qquad (1.20)\\ =&\{\delta\mu S + \delta(\gamma - q)S + \alpha P_B(r_B - r) + \beta P_C(r_C - r) - r\hat{V} + s_F(-\hat{V} - \alpha P_B)^-\}dt \\ &+ (\delta\sigma S)dW + (-\alpha P_B)dJ_B + (-\beta P_C)dJ_C. \qquad (1.21)\end{aligned}$$

On the other hand, by Ito's lemma for jump diffusions, the derivative value $\hat{V}(t)$ moves by

$$d\hat{V} = (\frac{\partial \hat{V}}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + \mu S \frac{\partial \hat{V}}{\partial S})dt + \sigma S \frac{\partial \hat{V}}{\partial S} dW + \Delta\hat{V_B} dJ_B + \Delta\hat{V_C} dJ_C \qquad (1.22)$$

where

$$\Delta\hat{V_B} = \hat{V}(t, S, 1, 0) - \hat{V}(t, S, 0, 0) \qquad (1.23)$$

$$\Delta\hat{V_C} = \hat{V}(t, S, 0, 1) - \hat{V}(t, S, 0, 0) \qquad (1.24)$$

which can be computed from boundary conditions (1.15) and (1.16).

Substitute (1.22) into (1.21) and eliminate all the risk factors (or the uncertainty parts), by choosing δ, α and β as

$$\delta = -\frac{\partial \hat{V}}{\partial S} \qquad (1.25)$$

$$\alpha = \frac{\Delta\hat{V_B}}{P_B} \qquad (1.26)$$

$$\beta = \frac{\Delta\hat{V_C}}{P_C}. \qquad (1.27)$$

If we introduce the Black-Scholes differential operator $\mathcal{L}$ as

$$\mathcal{L}V \equiv \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \frac{\partial V}{\partial S}(q - \gamma)S - rV, \qquad (1.28)$$

then it follows that $\hat{V}$ is the solution of the PDE

$$\begin{cases} \frac{\partial \hat{V}}{\partial t} + \mathcal{L}\hat{V} = s_F(\hat{V} + \Delta\hat{V_B})^+ - \lambda_B\Delta\hat{V_B} - \lambda_C\Delta\hat{V_C} \\ \hat{V}(T, S) = H(S) \end{cases} \qquad (1.29)$$

where λ_B, λ_C are the default intensities of party and counterparty, respectively, which

are also the same as the respective differences between the bond yields and risk-free rate, that is, $\lambda_B = r_B - r$ and $\lambda_C = r_C - r$. If we substitute (1.23) and (1.24) and boundary conditions (1.15) and (1.16) into the above PDE, the final PDE is

$$\begin{cases} \frac{\partial \hat{V}}{\partial t} + \mathcal{L}\hat{V} = (\lambda_B + \lambda_C)\hat{V} + s_F M^+ - \lambda_B(R_B M^- + M^+) - \lambda_C(R_C M^+ + M^-) \\ \hat{V}(T,S) = H(S). \end{cases} \tag{1.30}$$

Mark-to-market value at default

As we discussed before, according to the ISDA [3] 2002 Master Agreement, in case either party defaults, the value of the derivative is determined by a mark-to-market value M of derivative. By convention, the positive values correspond to seller's assets and counterparty liabilities, while negative values correspond to seller's liabilities and counterparty assets. In PDE (1.30), we consider two cases:

- If $M = \hat{V}$, PDE (1.30) is written as a nonlinear PDE

$$\begin{cases} \frac{\partial \hat{V}}{\partial t} + \mathcal{L}\hat{V} = s_F \hat{V}^+ + \lambda_B(1 - R_B)\hat{V}^- + \lambda_C(1 - R_C)\hat{V}^+ \\ \hat{V}(T,S) = H(S) \end{cases} \tag{1.31}$$

- If $M = V$, PDE (1.30) is written as a linear PDE

$$\begin{cases} \frac{\partial \hat{V}}{\partial t} + \mathcal{L}\hat{V} - (\lambda_B + \lambda_C)\hat{V} = s_F V^+ - (\lambda_B R_B + \lambda_C)V^- - (\lambda_B + \lambda_C R_C)V^+ \\ \hat{V}(T,S) = H(S) \end{cases} \tag{1.32}$$

Note that, in (1.31) and (1.32), for several derivatives, V is considered known and can be computed by vanilla Black-Scholes formula directly.

PDEs (1.31) and (1.32) are used when we aim to compute the adjusted value $\hat{V} = V + U$, where V is risk-free derivative value and U is the value adjustment. If we aim to compute the total value adjustment U itself, then using $\hat{V} = V + U$, $\frac{\partial V}{\partial t} + \mathcal{L}V = 0$ and $V(T,S) = \hat{V}(T,S) = H(S)$, from (1.31) and (1.32), we get the PDEs, that U satisfies in each of two cases:

[3] International Swaps and Derivatives Association

- If $M = \hat{V}$, (1.31) is written for U as

$$\begin{cases} \frac{\partial U}{\partial t} + \mathcal{L}U = s_F(V+U)^+ + \lambda_B(1-R_B)(U+V)^- + \lambda_C(1-R_C)(U+V)^+ \\ U(T,S) = 0 \end{cases} \tag{1.33}$$

  In this case, if we apply the Feynman-Kac theorem, with the assumption of deterministic hazard and interest rates, the expectation representation of U can be written as the following nonlinear integral equation

$$\begin{aligned} U(t,S) = &-(1-R_B)\int_t^T \lambda_B(u)B_r(t,u)\mathbb{E}_t^{\mathcal{Q}}[(V(u,S(u))+U(u,S(u)))^-]du \\ &-(1-R_C)\int_t^T \lambda_C(u)B_r(t,u)\mathbb{E}_t^{\mathcal{Q}}[(V(u,S(u))+U(u,S(u)))^+]du \\ &-\int_t^T s_F(u)B_r(t,u)\mathbb{E}_t^{\mathcal{Q}}[(V(u,S(u))+U(u,S(u)))^+]du, \end{aligned} \tag{1.34}$$

  where $B_r(t,u)$ is the risk-free discount factor and S follows the standard geometric Brownian motion.

- If $M = V$, (1.32) is written for U as

$$\begin{cases} \frac{\partial U}{\partial t} + \mathcal{L}U - (\lambda_B+\lambda_C)U = s_F V^+ - \lambda_B(1-R_B)V^- - \lambda_C(1-R_C)V^+ \\ U(T,S) = 0 \end{cases} \tag{1.35}$$

  In this case, if we apply the Feynman-Kac theorem, with the assumption of deterministic hazard and interest rates, the expectation representation of U can be written as the following linear integral equation

$$\begin{aligned} U(t,S) = &-(1-R_B)\int_t^T \lambda_B(u)B_{r+\lambda_B+\lambda_C}(t,u)\mathbb{E}_t^{\mathcal{Q}}[V^-(u,S(u))]du \\ &-(1-R_C)\int_t^T \lambda_C(u)B_{r+\lambda_B+\lambda_C}(t,u)\mathbb{E}_t^{\mathcal{Q}}[V^+(u,S(u))]du \\ &-\int_t^T s_F(u)B_{r+\lambda_B+\lambda_C}(t,u)\mathbb{E}_t^{\mathcal{Q}}[V^+(u,S(u))]du, \end{aligned} \tag{1.36}$$

  where $B_{r+\lambda_B+\lambda_C}(t,u)$ is the discount factor, S follows the standard geometric Brownian motion.

In both cases, $S \in [0,+\infty)$ and $t \in [0,T]$.

Remark 1 *While we mentioned two cases for the mark-to-market value M at default, namely $M = \hat{V}$ and $M = V$, in this thesis, we consider the first case, primarily because this is more challenging to solve. In practice though, none of the two is considered the official resolution case. According to the ISDA agreement, a dealer poll is carried to decide the mark-to-market value M at default. Clearly, this value cannot be very different from V or $\hat{V}$. Therefore, given that valuation adjustments need to be computed ahead of the agreement, from the practical point of view, it is interesting to study both cases.*

1.4 Thesis outline

The thesis is organized as following. Chapter 2 is the study of the nonlinear PDE model for bilateral XVA pricing in European and American type derivatives when both default intensities are constant. Two different iteration methods for nonlinearity are discussed, with an emphasis on the penalty-like method, which is extended to the double-penalty iteration method. Chapter 3 develops the nonlinear PDE for bilateral XVA pricing, assuming the counterparty default intensity to be stochastic. Two computational approaches, finite difference and asymptotic approximation are studied. For the finite difference scheme, we use second-order finite difference on spatial dimensions, Crank-Nicolson-Rannacher smoothing scheme on time dimension, and penalty-like iterations on nonlinearity. The asymptotic approximation is a closed-form approximation, based on singular perturbation theory. Chapter 4 presents the extension of the nonlinear PDE for bilateral XVA pricing of Chapter 3 to American type derivatives, assuming the counterparty default intensity to be stochastic, as well as the extension of the two different computational approaches, finite difference and asymptotic approximation. Chapter 5 presents numerical experiments that demonstrate effectiveness of our solutions to corresponding problems developed in previous three chapters.

Chapter 2

Pricing bilateral XVA with constant default intensities

In this chapter, we present methods to solve the PDEs (1.31) or (1.33) from XVA problem, with focus on the treatment of the nonlinear term. In our work, we use second-order finite differences in space and Crank-Nicolson discretization in time, with Rannacher smoothing when needed. We impose boundary conditions that are appropriate for call and put options, as well as for the forward contract. The main contributions of this part of work are

- a fast iterative method for handling the nonlinearity in the XVA PDE problem,
- its convergence analysis, as well as
- its extension to American style derivatives' XVA pricing.

The chapter is organized as follows. In Section 2.1, we summarize the formulation of the (European option) XVA pricing problem as a PDE. In Section 2.2, we describe the numerical methods used for the discretization of the PDE and of the boundary conditions, and introduce two types of iteration methods for handling the nonlinearity. For the most efficient of these, we present and prove a theorem about its convergence. Section 2.3 includes the extension of the above methods to the American option XVA pricing problem. The numerical experiments to study the behavior of the proposed methods, for XVA pricing in European and American options are left for Chapter 5.

2.1 Formulation

In this section, we show the PDE representation for the value of a financial derivative, if the effects of bilateral default risk and the funding costs are considered [11]. The PDE

model is extended from Black-Scholes PDE, driven by using multiple hedging arguments, including underlying assets, the zero-coupon bonds of the two parties and risk-free zero-coupon bonds. Note that, in [11], the two parties are referred to as party B (seller, bank), and party C (counterparty, investor), and we adopt the same naming/notation.

As mentioned before, to the International Swaps and Derivatives Association (ISDA) 2002 Master Agreement, in case either party of a contract defaults, the value of the derivative is determined by a Mark-to-Market rule M, which is chosen to be either $\hat{V}$ or V. Here, we assume that $M = \hat{V}$ and consider the solution of (1.31) for the adjusted price $\hat{V}$ and (1.33) for the valuation adjustment U. Note also that riskless [1] option value V satisfies the Black-Scholes PDE

$$\begin{cases} \frac{\partial V}{\partial t} + \mathcal{L}V = 0, \\ V(T,S) = H(S). \end{cases} \tag{2.1}$$

Note that, in (1.33), V is considered known. It can be computed by the Black-Scholes formula, or numerically. Note also that $\hat{V}$ in (1.31) and U in (1.33) are unknown, and, since these PDEs involve unknowns in the max or min notation, they are nonlinear. It is also worth noting that, if $\lambda_B = \lambda_C = s_F = 0$, then PDE (1.31) simplifies to (2.1).

2.2 Numerical methods

In order to make PDE (1.33) convenient for the application of numerical PDE methods, we apply the variable transformation $\tau = T - t$ in the time dimension. Then (1.33) becomes the initial value problem

$$\begin{cases} \frac{\partial U}{\partial \tau} = \mathcal{L}U + f(U,V), \\ U(0,S) = 0 \end{cases} \tag{2.2}$$

where

$$f(U,V) \equiv -(\lambda_B(1-R_B)(U+V)^- + \lambda_C(1-R_C)(U+V)^+ + s_F(U+V)^+). \tag{2.3}$$

[1] "Riskless" derivative in this thesis means a financial derivative without considering counterparties' default risk.

2.2.1 Discretization

In this subsection, we present the discretization of (2.2). The semi-infinite space domain of spot price S is truncated into $[0, S_{max}]$, for sufficiently large S_{max}. In Section (5.1.1), there is a discussion about the effect of S_{max} on the accuracy of the numerical solution. Then, $[0, S_{max}]$ is divided into N sub-intervals, with the gridpoints positioned uniformly or non-uniformly. Standard second-order finite differences are used for the space discretization of (2.2).

We assume that the timestepping is handled by the ϑ-method, [2] which, for $\vartheta = \frac{1}{2}$ and $\vartheta = 1$ becomes the Crank-Nicolson (CN) and Backward Euler methods, respectively. Note that PDE (2.2) does not need Rannacher smoothing due to the smooth initial condition.

2.2.2 Boundary conditions

As we indicated before, we need to truncate the semi-infinite domain $[0, +\infty)$ into $[0, S_{max}]$. We also need to formulate boundary conditions when $S = 0$ and $S = S_{max}$. To obtain the left boundary condition, we substitute $S = 0$ into (2.2) and get

$$\frac{\partial U}{\partial \tau} = -rU + f(U, V) \tag{2.4}$$

which can be seen as a first-order ordinary differential equation (ODE), computing approximate values of $U(\tau, 0)$ used as Dirichlet conditions.

Regarding the far-side boundary condition, when S_{max} is large enough, the behavior of derivative price $\hat{V}$ after value adjustment is expected to be similar to the Black-Scholes derivative price V. A choice of far-side boundary condition often used in Black-Scholes PDE is the *linear boundary condition*

$$\lim_{S \to \infty} \frac{\partial^2 U}{\partial S^2} = 0. \tag{2.5}$$

If we apply the far-side boundary condition (2.5) directly to (2.2) as in [48], the PDE at $S = S_{max}$ becomes

$$\frac{\partial U}{\partial \tau} = (q - \gamma) S \frac{\partial U}{\partial S} - rU + f(U, V), \tag{2.6}$$

where the first derivative term can be discretized by first-order or second-order one-sided difference scheme. The implementation of the linear boundary condition by discretizing

[2] Note that the ϑ notation for the time-stepping method is different from the $\theta(t)$ notation in the CIR model in next chapters.

(2.6) applies to all derivatives considered in this thesis.

Alternatively, we can implement the linear boundary condition (2.5) by considering the PDE problem

$$\begin{cases} \frac{\partial \hat{V}}{\partial \tau} = \mathcal{L}\hat{V} - \lambda_B(1-R_B)\hat{V}^- - \lambda_C(1-R_C)\hat{V}^+ - s_F\hat{V}^+, \\ \hat{V}(0,S) = H(S). \end{cases} \tag{2.7}$$

arising from (1.31) with $\tau = T - t$. As explained in [48], we assume that, close to S_{max} and beyond, we have

$$\hat{V}(\tau, S) = \alpha(\tau)S + \beta(\tau). \tag{2.8}$$

Taking into account the asymptotic behavior of $\hat{V}$ as $S \to S_{max}$, allows us to know the sign of $\hat{V}(\tau, S_{max})$, which simplifies the right hand side $-\lambda_B(1-R_B)\hat{V}^- - \lambda_C(1-R_C)\hat{V}^+ - s_F\hat{V}^+$ to either the positive or the negative terms. Substituting $\hat{V}$ of (2.8) into the PDE (2.7) with the simplified right hand side, results in two ODEs for $\alpha(\tau)$ and $\beta(\tau)$, respectively, which can be solved to give rise to Dirichlet boundary conditions for $\hat{V}(\tau, S_{max})$, and thus for $U(\tau, S_{max})$. Clearly, the boundary conditions depend on the type of financial derivative studied. More specifically, for a European Call or Long Forward, we obtain

$$U(\tau, S_{max}) = (e^{(-\lambda_C(1-R_C)-s_F)\tau} - 1)(e^{(q-\gamma-r)\tau}S_{max} - e^{-r\tau}K), \tag{2.9}$$

while for a European Put and any financial derivative whose value decays to 0 as $S \to S_{max}$, we obtain

$$U(\tau, S_{max}) = 0. \tag{2.10}$$

2.2.3 Iteration methods for nonlinear PDE

In this subsection, we introduce two iteration methods, with emphasis on the second one, to handle the nonlinearity in PDE (2.2).

Let $\tau_j, j = 0, \ldots, N^t$, be the timesteps at which the solution is computed, with $\tau_0 = 0 < \tau_1 < \cdots < \tau_{N_t} = T$, and $\Delta\tau^j = \tau_j - \tau_{j-1}$ be the jth time stepsize. Let $u^j, j = 0, \ldots, N_t$, denote the computed solution vector arising from the approximate values of U at the spatial gridpoints at time τ_j, with u^0 being the initial condition vector. Since we use an iteration method to handle the nonlinearity, let $u^{j,k}, k = 0, \ldots, maxit$, denote the computed solution vector at iteration k of timestep j, with $maxit$ the maximum number of iterations allowed. Let $v^j, j = 0, \ldots, N_t$, be the vector of values of V at the spatial gridpoints, and at timestep j. Note that the values of V can be computed using the Black-Scholes formula directly. For generic vectors u and v arising from (approximate)

values of U and V, respectively, at the spatial gridpoints, let $f(u, v)$ denote the vector arising from evaluating f at the components of u and v. Let also A be the matrix arising from the space discretization of $\mathcal{L}U$, and $\mathbb{I}$ be the identity matrix of same order. For simplicity, we assume the spatial gridpoints remain the same at all timesteps.

The first iteration method is a fixed-point iteration method similar to the one in [3]. We remark that the fixed-point iteration method in [3] assumes that the timestepping is handled by the method of characteristics and the space discretization by a finite element method. In our thesis, we consider CN timestepping and centered differences in space. Thus, at each timestep we need to solve

$$(\mathbb{I}-\vartheta\Delta\tau^j A)u^j = (\mathbb{I}+(1-\vartheta)\Delta\tau^j A)u^{j-1}+\vartheta\Delta\tau^j f(u^j, v^j)+(1-\vartheta)\Delta\tau^j f(u^{j-1}, v^{j-1}). \quad (2.11)$$

Then the fixed-point iteration method at time τ^j can be described as Algorithm 1.

Algorithm 1 Fixed-point iteration for (2.2) at step j, with ϑ-timestepping

Require: Solve $(\mathbb{I}-\vartheta\Delta\tau^j A)u^j = g^j+\vartheta\Delta\tau^j f(u^j, v^j)$ where $g^j = (\mathbb{I}+(1-\vartheta)\Delta\tau^j A)u^{j-1}+(1-\vartheta)\Delta\tau^j f(u^{j-1}, v^{j-1})$

1: Initialize $u^{j,0} = u^{j-1}$
2: **for** $k = 1, \ldots, maxit$ **do**
3: Solve $(\mathbb{I} - \vartheta\Delta\tau^j A)u^{j,k} = g^j + \vartheta\Delta\tau^j f(u^{j,k-1}, v^j)$
4: **if** stopping criterion satisfied **then**
5: Break
6: **end if**
7: **end for**
8: Set $u^j = u^{j,k}$

The stopping criterion in Algorithm 1, motivated by [22], is

$$\max_i \frac{|u_i^{j,k} - u_i^{j,k-1}|}{\max(1, |u_i^{j,k}|)} \leq tol. \quad (2.12)$$

where *tol* is a user-chosen tolerance.

The above iteration method is considered to be *explicit*, in the sense that the matrix solved at each iteration is the same and only the right-hand side vector changes.

Regarding the convergence of the fixed-point iteration, we can write it as $u^{j,k} = \mathbf{G}(u^{j,k-1})$, where $\mathbf{G}(u^{j,k-1}) = (\mathbb{I} - \vartheta\Delta\tau^j A)^{-1}(g^j + \vartheta\Delta\tau^j f(u^{j,k-1}, v^j))$, and show that $\mathbf{G}$ is a contraction, for sufficiently small $\Delta\tau^j$. We do not go into details, as the focus of this chapter is not to study the convergence of the fixed-point iteration.

We next present another iteration method for handing the nonlinearity in (2.2). We refer to it as *discrete penalty-like iteration*, or, simply, *penalty iteration*, as it is motivated

by the same-name method in [22], designed to handle the nonlinear PDE arising from the linear complementarity problem (LCP) in American option pricing. For this reason, with generic vectors u and v, we define the diagonal penalty matrix $P = P(u, v)$ by

$$[P(u,v)]_{ii} \equiv \begin{cases} -\lambda_B(1-R_B) & \text{if } u_i + v_i < 0 \\ -\lambda_C(1-R_C) - s_F & \text{if } u_i + v_i \geq 0 \end{cases} \tag{2.13}$$

Note that, with the above matrix P, the vector arising from the discretized form of $f(U, V)$ can be written as

$$f(u,v) = P(u,v)(u+v) \tag{2.14}$$

where, since $P(u, v)$ depends on u, there is nonlinearity between $P(u, v)$ and u. Note also that, if $\lambda_B \geq 0$, $\lambda_C \geq 0$, and $s_F \geq 0$, we have $P_{i,j}(u, v) \leq 0$. With the help of the matrix P, we can also write (2.11) as

$$\begin{aligned} &(\mathbb{I} - \vartheta\Delta\tau^j(A + P(u^j, v^j)))u^j \\ &= (\mathbb{I} + (1-\vartheta)\Delta\tau^j A)u^{j-1} + \vartheta\Delta\tau^j P(u^j, v^j)v^j + (1-\vartheta)\Delta\tau^j f(u^{j-1}, v^{j-1}). \end{aligned} \tag{2.15}$$

We can view the penalty iteration method as generalized Newton's method applied to the nonsmooth nonlinear problem (2.11) with the nonlinear term given by (2.14), and the generalized Jacobian defined by

$$\frac{\partial[f(u,v)]_i}{\partial u_j} \equiv \begin{cases} [P(u,v)]_{ii} & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{2.16}$$

Let now $P^k = P(u^{j,k}, v^j)$, where for brevity, in the notation P^k, we omitted the superscript j. Using the penalty matrix P^k, the proposed discrete penalty iteration for (2.2) is described in Algorithm 2.
The stopping criterion in Algorithm 2, motivated again by [22] is

$$(P^k = P^{k-1}) \text{ or } (\max_i \frac{|u_i^{j,k} - u_i^{j,k-1}|}{\max(1, |u_i^{j,k}|)} \leq tol). \tag{2.17}$$

Note that the second iteration method is considered to be *implicit*, in the sense that the matrix solved at each iteration depends on the iteration index k. However, since the matrix solved at each iteration is adjusted by only a diagonal matrix, the sparsity structure of the matrix remains the same, which is also the same structure of the matrix

Algorithm 2 Discrete penalty-like iteration for (2.2) at step j, with ϑ-timestepping

Require: Solve $(\mathbb{I} - \vartheta\Delta\tau^j(A + P(u^j, v^j)))u^j = g^j + \vartheta\Delta\tau^j P(u^j, v^j)v^j$ where $g^j = (\mathbb{I} + (1-\vartheta)\Delta\tau^j A)u^{j-1} + (1-\vartheta)\Delta\tau^j f(u^{j-1}, v^{j-1})$

Initialize $u^{j,0} = u^{j-1}$ and $P^0 = P(u^{j,0}, v^j)$
for $k = 1, \ldots, maxit$ **do**
  Solve $(\mathbb{I} - \vartheta\Delta\tau^j(A + P^{k-1}))u^{j,k} = g^j + \vartheta\Delta\tau^j P^{k-1}v^j$
  Compute $P^k = P(u^{j,k}, v^j)$ by (2.13)
  if stopping criterion satisfied **then**
    Break
  end if
end for
Set $u^j = u^{j,k}$

corresponding to a linear PDE with same differential operator as in (2.2). Note also that, if $\lambda_B \geq 0$, $\lambda_C \geq 0$, the matrix P enhances the diagonal dominance of A.

It is also worth noticing that, in this penalty iteration, there is no large penalty parameter, as is in the case of American options [22]. However, we still use the term penalty iteration, because the method can be viewed as generalized Newton's method.

In order to study the convergence of the discrete penalty-like iteration, we make use of a monotonicity argument for the matrix A. In [16], sufficient conditions are derived under which the matrix arising from the discretization of $\mathcal{L}$ by finite differences and with Dirichlet boundary conditions on both ends is strictly diagonally dominant with negative diagonal entries and non-negative off-diagonal entries (Lemma 4.1 in [16]), therefore, monotone. To facilitate the monotonicity argument for the proof of Theorem 1, we consider that A is formed with Dirichlet conditions at both ends, such as those described in Section 2.2.2. We also assume that the conditions of Lemma 4.1 in [16] hold and, therefore, A is non-singular, monotone and an M-matrix. It is easy to see that, under the same conditions, $(\mathbb{I} - \vartheta\Delta\tau^j(A + P^{k-1}))$ is also non-singular, monotone and an M-matrix.

Theorem 1 *(Convergence of discrete penalty-like iteration). Under the assumption that the matrix of the linear system at each iteration in Algorithm 2, i.e.* $(\mathbb{I} - \vartheta\Delta\tau^j(A + P^{k-1}))$*, is monotone, then*

(i) The discrete penalty-like iteration converges to the unique solution.

(ii) The iteration converges monotonically.

(iii) The iteration has finite termination.

Proof. Recall that, during timestep j, the system to be solved at iteration k, for $k \geq 1$,

is

$$(\mathbb{I} - \vartheta\Delta\tau^j(A + P^{k-1}))u^{j,k} = (\mathbb{I} + (1-\vartheta)\Delta\tau^j A)u^{j-1} + \vartheta\Delta\tau^j P^{k-1}v^j + b^j. \tag{2.18}$$

where $b^j = (1-\vartheta)\Delta\tau^j f(u^{j-1}, v^{j-1})$, for simplicity. The proof follows the lines of [22], except that we need to distinguish cases for the values of $-\lambda_B(1-R_B)$ and $-\lambda_C(1-R_C) - s_F$.

First, let us prove the monotone property of the iteration.
Equation (2.18) can be rewritten as

$$(\mathbb{I}-\vartheta\Delta\tau^j(A+P^k))u^{j,k}+\vartheta\Delta\tau^j(P^k-P^{k-1})u^{j,k} = (\mathbb{I}+(1-\vartheta)\Delta\tau^j A)u^{j-1}+\vartheta\Delta\tau^j P^{k-1}v^j+b^j. \tag{2.19}$$

Furthermore, equation (2.18) can also be written for the version of iteration $k+1$ as

$$(\mathbb{I} - \vartheta\Delta\tau^j(A + P^k))u^{j,k+1} = (\mathbb{I} + (1-\vartheta)\Delta\tau^j A)u^{j-1} + \vartheta\Delta\tau^j P^k v^j + b^j. \tag{2.20}$$

Subtracting equation (2.19) from equation (2.20) gives

$$(\mathbb{I} - \vartheta\Delta\tau^j(A + P^k))(u^{j,k+1} - u^{j,k}) = \vartheta\Delta\tau^j(P^k - P^{k-1})(u^{j,k} + v^j) \tag{2.21}$$

We, firstly, work on the case that

$$-\lambda_C(1-R_C) - s_F \le -\lambda_B(1-R_B) \le 0.$$

Examining each component of the right hand side of equation (2.21), there are two possible subcases:

- If $(u^{j,k} + v^j)_i \ge 0$, then $P^k_{ii} = -\lambda_C(1-R_C) - s_F$.
  If $P^{k-1}_{ii} = -\lambda_C(1-R_C) - s_F$, then $(P^k_{ii} - P^{k-1}_{ii}) = 0$.
  If $P^{k-1}_{ii} = -\lambda_B(1-R_B)$, then $(P^k_{ii} - P^{k-1}_{ii}) \le 0$.
  Hence $\vartheta\Delta\tau^j[(P^k - P^{k-1})(u^{j,k} + v^j)]_i \le 0$.

- If $(u^{j,k} + v^j)_i < 0$, then $P^k_{ii} = -\lambda_B(1-R_B)$.
  If $P^{k-1}_{ii} = -\lambda_C(1-R_C) - s_F$, then $(P^k_{ii} - P^{k-1}_{ii}) \ge 0$.
  If $P^{k-1}_{ii} = -\lambda_B(1-R_B)$, then $(P^k_{ii} - P^{k-1}_{ii}) = 0$.
  Hence $\vartheta\Delta\tau^j[(P^k - P^{k-1})(u^{j,k} + v^j)]_i \le 0$.

Therefore, for $k \ge 1$, the vector $\vartheta\Delta\tau^j(P^k - P^{k-1})(u^{j,k} + v^j)$ is always non-positive. In equation (2.21), since $(\mathbb{I} - \vartheta\Delta\tau^j(A + P^k))$ is monotone, we have $u^{j,k+1} \le u^{j,k}$ for $k \ge 1$.

Thus, the iterates decrease monotonically.

The other case, i.e. $-\lambda_B(1-R_B) < -\lambda_C(1-R_C) - s_F \leq 0$, can be analyzed in a similar way. In this case, the iterates increase monotonically.

Second, we prove the iteration has finite termination to a solution of (2.15). We consider first the case $-\lambda_C(1-R_C) - s_F < -\lambda_B(1-R_B) \leq 0$. Let us define

$$S_1^k = \{i | P_{ii}^k = -\lambda_B(1-R_B)\} \tag{2.22}$$
$$S_2^k = \{i | P_{ii}^k = -\lambda_C(1-R_C) - s_F\}. \tag{2.23}$$

Therefore, $S_1^k \dot{\cup} S_2^k$ is the set of all the nodes in the discretization. Since in the case of $-\lambda_C(1-R_C) - s_F < -\lambda_B(1-R_B) \leq 0$, $u^{j,k}$ decreases monotonically, any node in S_1^k remains in S_1^l if $l \geq k \geq 1$. At the same time, if $S_1^k = S_1^{k-1}, S_2^k = S_2^{k-1}$, then $P^k = P^{k-1} = 0$, and the iteration terminates.

Therefore, during each iteration, at least one node will move into the set S_1^k and remain in S_1^{k+1} before the termination. Hence the iteration always terminates within at most $N+1$ iterations, where N is the size of the linear system, although, numerical experiments indicate that the iteration needs only one or two steps. Moreover, at termination, we have $(\mathbb{I} - \vartheta\Delta\tau^j(A+P^{k-1}))u^{j,k} = (\mathbb{I} + (1-\vartheta)\Delta\tau^j A)u^{j-1} + \vartheta\Delta\tau^j P^{k-1}v^j + b^j$, thus, $(\mathbb{I} - \vartheta\Delta\tau^j(A+P^k))u^{j,k} = (\mathbb{I}+(1-\vartheta)\Delta\tau^j A)u^{j-1} + \vartheta\Delta\tau^j P^k v^j + b^j$. Therefore, $u^{j,k}$ is a solution to (2.15). Additionally, the iterates are clearly bounded.

We now consider the case $-\lambda_C(1-R_C) - s_F = -\lambda_B(1-R_B) \leq 0$. In this case, it is easy to see that $P^k = P^{k-1}$. Thus, the iteration terminates in one step.

Finally, the case $-\lambda_B(1-R_B) < -\lambda_C(1-R_C) - s_F \leq 0$ can be handled similarly as the case $-\lambda_C(1-R_C) - s_F < -\lambda_B(1-R_B) \leq 0$.

Last, we demonstrate the uniqueness of the solution to the nonlinear problem (2.15).

Suppose there are two solutions u_1 and u_2 and $P_1 = P(u_1, v)$ and $P_2 = P(u_2, v)$ for any arbitrary v. Then, u_1 and u_2 must satisfy

$$(\mathbb{I} - \vartheta\Delta\tau^j(A+P_1))u_1 = (\mathbb{I} + (1-\vartheta)\Delta\tau^j A)u_{j-1} + \vartheta\Delta\tau^j P_1 v + b^j, \tag{2.24}$$
$$(\mathbb{I} - \vartheta\Delta\tau^j(A+P_2))u_2 = (\mathbb{I} + (1-\vartheta)\Delta\tau^j A)u_{j-1} + \vartheta\Delta\tau^j P_2 v + b^j. \tag{2.25}$$

Let us rewrite equation (2.24) as

$$(\mathbb{I} - \vartheta\Delta\tau^j(A+P_2))u_1 + \vartheta\Delta\tau^j[P_2 - P_1]u_1 = (\mathbb{I} + (1-\vartheta)\Delta\tau^j A)u^{j-1} + \vartheta\Delta\tau^j P_1 v + b^j. \tag{2.26}$$

Subtracting equation (2.25) from equation (2.26) gives

$$(\mathbb{I} - \vartheta\Delta\tau^j(A + P_2))(u_1 - u_2) = \vartheta\Delta\tau^j(P_1 - P_2)(u_1 + v). \tag{2.27}$$

By a similar argument we used in proving the monotone property of the iteration, we have that

$$\vartheta\Delta\tau^j(P_1 - P_2)(u_1 + v) \leq 0, \text{ if } -\lambda_B(1 - R_B) \geq -\lambda_C(1 - R_C) - s_F, \tag{2.28}$$
$$\vartheta\Delta\tau^j(P_1 - P_2)(u_1 + v) \geq 0, \text{ if } -\lambda_B(1 - R_B) < -\lambda_C(1 - R_C) - s_F. \tag{2.29}$$

Since the left hand side matrix of (2.27) $(\mathbb{I} - \vartheta\Delta\tau^j(A + P_2))$ is monotone, we have that

$$u_1 - u_2 \leq 0, \text{ if } -\lambda_B(1 - R_B) \geq -\lambda_C(1 - R_C) - s_F, \tag{2.30}$$
$$u_1 - u_2 \geq 0, \text{ if } -\lambda_B(1 - R_B) < -\lambda_C(1 - R_C) - s_F. \tag{2.31}$$

We can interchange the subscripts of u_1 and u_2, and obtain

$$u_1 - u_2 \geq 0, \text{ if } -\lambda_B(1 - R_B) \geq -\lambda_C(1 - R_C) - s_F, \tag{2.32}$$
$$u_1 - u_2 \leq 0, \text{ if } -\lambda_B(1 - R_B) < -\lambda_C(1 - R_C) - s_F. \tag{2.33}$$

Hence, in any case, $u_1 = u_2$, indicating that the iterations converge to a unique solution.

□

Remark 2 *The PDE problem (2.2), with the nonlinear term in (2.3), can be written as a control problem*

$$\frac{\partial U}{\partial \tau} = \mathcal{L}U + \max_{\xi_1 \in \{0,1\}} \min_{\xi_2 \in \{0,1\}} \left[-\xi_1 \lambda_B (1 - R_B)(U + V) - \xi_2 (\lambda_C (1 - R_C) + s_F)(U + V) \right], \tag{2.34}$$

where ξ_1 and ξ_2 are two controls corresponding to the cases that $U + V$ is negative or positive, respectively. Since the problem involves both max and min, it can be viewed as a Hamilton-Jacobi-Bellman-Isaacs (HJBI) problem. For such problems, in general, the Newton-type iteration scheme is not guaranteed to converge, unless the initial guess is close enough to the solution, or a more complex treatment is considered [21, 38]. However, in our case, the control problem (2.34) is such that it can be reduced to a single control with only max or min terms. More specifically, if $-\lambda_C(1-R_C) - s_F \leq -\lambda_B(1-R_B) \leq 0$,

problem (2.34) can be written as

$$\frac{\partial U}{\partial \tau} = \mathcal{L}U + \min_{\xi\in\{0,1\}} \left[-(1-\xi)\lambda_B(1-R_B)(U+V) - \xi(\lambda_C(1-R_C)+s_F)(U+V)\right] \quad (2.35)$$

while, if $-\lambda_B(1-R_B) < -\lambda_C(1-R_C) - s_F \leq 0$*, problem (2.34) can be written as*

$$\frac{\partial U}{\partial \tau} = \mathcal{L}U + \max_{\xi\in\{0,1\}} \left[-(1-\xi)\lambda_B(1-R_B)(U+V) - \xi(\lambda_C(1-R_C)+s_F)(U+V)\right]. \quad (2.36)$$

both of which can be viewed as Hamilton-Jacobi-Bellman (HJB) problems. For problems (2.35) and (2.36), the monotonicity argument can be used to show global convergence [22, 21, 38].

2.3 American derivatives

Just as the formulation of the American derivative pricing problem can be presented as a linear complementarity problem (LCP) [22, 47], the formulation of the XVA pricing problem in American derivatives can also be presented as an LCP [4].

We consider the case $M = \hat{V}$, and, as in [4], build the formulation of XVA pricing in American derivatives starting from Equation (1.33) and adding appropriate constraints that force the adjusted price $\hat{V}(t,S)$ to be above the payoff $V^*(t,S)$.

Then, the American derivative price $\hat{V}$, if taking risk into account, satisfies

$$\left\{ \begin{array}{r} \frac{\partial \hat{V}}{\partial t} + \mathcal{L}\hat{V} + f(\hat{V}) = 0 \\ \hat{V} - V^* \geq 0 \end{array} \right\} \vee \left\{ \begin{array}{r} \frac{\partial \hat{V}}{\partial t} + \mathcal{L}\hat{V} + f(\hat{V}) \geq 0 \\ \hat{V} - V^* = 0 \end{array} \right\} \quad (2.37)$$

where

$$f(\hat{V}) = -(\lambda_B(1-R_B)\hat{V}^- + \lambda_C(1-R_C)\hat{V}^+ + s_F\hat{V}^+), \quad (2.38)$$

$$V^*(S) = \hat{V}(0,S), \quad (2.39)$$

and the notation $\vee$ means "or". Typically at each time t, there is a particular asset price S^* which divides the asset price domain into two regions: one side that corresponds to the left condition of (2.37), where it is suggested to hold the option, and another side that corresponds to the right condition of (2.37), where it is suggested to early exercise the option. This particular value of S^* is unknown before the differential equation is solved, and is usually called *free boundary* of LCP.

Note that the value $V(t,S)$ of American derivatives without valuation adjustment satisfies the LCP [22]

$$\left\{ \begin{array}{c} \frac{\partial V}{\partial t} + \mathcal{L}V = 0 \\ V - V^* \geq 0 \end{array} \right\} \vee \left\{ \begin{array}{c} \frac{\partial V}{\partial t} + \mathcal{L}V \geq 0 \\ V - V^* = 0 \end{array} \right\} \tag{2.40}$$

Furthermore, note that equations (2.37) and (2.40) have the same terminal condition $\hat{V}(T,S) = V(T,S) = H(S)$ where $H(S)$ is payoff function at maturity.

The LCP problem (2.40) also has a free boundary. However it is not guaranteed that the two *free boundaries* of equations (2.37) and (2.40) are the same. Hence, it is impossible to find an individual PDE for the XVA value $U(t,S)$. In this section, we mainly focus on numerically solving (2.37).

2.3.1 Reformulation to penalty form

Forsyth and Vetzal [22] proposed the discrete penalty method to numerically solve the LCP arising from one-asset American option without XVA. They replace (2.40) by a nonlinear PDE via adding a large positive penalty term to the right hand side of the Black-Scholes equation.

In the American XVA problem, a similar penalty term can be added. If backward time $\tau = T - t$ is applied, the penalty form of (2.37) is written as

$$\begin{cases} \frac{\partial \hat{V}}{\partial \tau} = \mathcal{L}\hat{V} + f(\hat{V}) + p\max(V^* - \hat{V}, 0), \\ \hat{V}(0,S) = H(S), \end{cases} \tag{2.41}$$

where p is a large positive penalty factor. The penalty term forces the solution of (2.37) to approximately satisfy the obstacle condition $\hat{V} - V^* \geq 0$. We recall that $f(\hat{V})$ is given by (2.38).

As in Section 2.2.2, the boundary conditions for $S = 0$ are formed by solving

$$\frac{\partial \hat{V}}{\partial \tau} = -r\hat{V} + f(\hat{V}) + p\max(V^* - \hat{V}, 0), \tag{2.42}$$

while for $S = S_{max}$, the linear boundary condition

$$\lim_{S \to \infty} \frac{\partial^2 \hat{V}}{\partial S^2} = 0,$$

is implemented either by discretizing

$$\frac{\partial \hat{V}}{\partial \tau} = (q-\gamma)S\frac{\partial \hat{V}}{\partial S} - r\hat{V} + f(\hat{V}) + p\max(V^* - \hat{V}, 0), \tag{2.43}$$

or by the alternative that gives rise to Dirichlet conditions that depend on the type of financial derivative considered. For an American Call or Long Forward, we obtain

$$\hat{V}(\tau, S_{max}) = \max\{e^{(-\lambda_C(1-R_C)-s_F)\tau}(e^{(q-\gamma-r)\tau}S_{max} - e^{-r\tau}K), S_{max} - K\}, \tag{2.44}$$

while, for an American Put,

$$\hat{V}(\tau, S_{max}) = 0. \tag{2.45}$$

2.3.2 Double-penalty iteration method

Penalty term discretization

The timestepping method and the space discretization for $\mathcal{L}\hat{V}$ in (2.41) are similar to those for the European case. We now discuss the treatment of the penalty term $p\max(V^* - \hat{V}, 0)$ and the nonlinear term $f(\hat{V})$. Let $\hat{v}^j, j = 0, \ldots, N_t$, be the computed solution vector arising from the approximate values of $\hat{V}$ at the spatial gridpoints at time τ_j. Also let v^* be the vector of values of payoff V^* at the spatial gridpoints.

When computing the numerical solution $\hat{v}^j$ at step j, the penalty term $p\max(V^* - \hat{V}, 0)$ (arising from the American feature) is discretized as $P_A(\hat{v}^j)(v^* - \hat{v}^j)$, where $P_A(\hat{v}^j)$ is a diagonal matrix defined by

$$[P_A(\hat{v}^j)]_{ii} \equiv \begin{cases} p & \text{if } (\hat{v}^j)_i < (v^*)_i, \\ 0 & \text{otherwise.} \end{cases} \tag{2.46}$$

The nonlinear term $f(\hat{V})$ (arising from the XVA) is discretized as $P_X(\hat{v}^j)(\hat{v}^j)$, where $P_X(\hat{v}^j)$ is also a diagonal matrix defined by

$$[P_X(\hat{v}^j)]_{ii} \equiv \begin{cases} -\lambda_B(1-R_B) & \text{if } (\hat{v}^j)_i < 0, \\ -\lambda_C(1-R_C) - s_F & \text{if } (\hat{v}^j)_i \geq 0. \end{cases} \tag{2.47}$$

Therefore, to compute $\hat{v}^j$, given $\hat{v}^{j-1}$, the following system of algebraic equations needs

to be solved:

$$\begin{aligned}&[\mathbb{I}-\vartheta\Delta\tau^j(A+P_X(\hat{v}^j))]\hat{v}^j+P_A(\hat{v}^j)\hat{v}^j\\&=(\mathbb{I}+(1-\vartheta)\Delta\tau^jA)\hat{v}^{j-1}+(1-\vartheta)\Delta\tau^jP_X(\hat{v}^{j-1})\hat{v}^{j-1}+P_A(\hat{v}^j)v^*.\end{aligned}\tag{2.48}$$

Note that there are two sources of nonlinearity, namely $P_A(\hat{v}^j)$ and $P_X(\hat{v}^j)$, with respect to $\hat{v}^j$. Algorithm 2 and the penalty iteration method to value American option [22] are combined into one iteration method, which can be viewed as generalized Newton's method, to solve (2.48). In the construction of P_A, the large penalty factor p is chosen as

$$p=\frac{1}{tol}$$

where *tol* is the tolerance for the stopping criterion of the iteration method.

Double-penalty iteration

If we view the nonlinear term $f(\hat{V})$ or its discrete form $P_X(\hat{v}^j)(\hat{v}^j)$ as a second penalty term, this generalized Newton's iteration method to solve (2.48) at each timestep can be called *discrete double-penalty iteration* for American XVA pricing.

Algorithm 3 Discrete double-penalty iteration for (2.41) at step j, with ϑ-timestepping

Require: Solve $[\mathbb{I}-\vartheta\Delta\tau^j(A+P_X(\hat{v}^j))]\hat{v}^j+P_A(\hat{v}^j)\hat{v}^j=g^j+P_A(\hat{v}^j)v^*$
where $g^j=(\mathbb{I}+(1-\vartheta)\Delta\tau^jA)\hat{v}^{j-1}+(1-\vartheta)\Delta\tau^jP_X(\hat{v}^{j-1})\hat{v}^{j-1}$.

1: Initialize $\hat{v}^{j,0}=\hat{v}^{j-1}$, $P_A^0=P_A(\hat{v}^{j,0})$ and $P_X^0=P_X(\hat{v}^{j,0})$
2: **for** $k=1,\ldots,maxit$ **do**
3: Solve $[\mathbb{I}-\vartheta\Delta\tau^j(A+P_X^{k-1})+P_A^{k-1}]\hat{v}^{j,k}=g^j+P_A^{k-1}v^*$
4: Compute $P_A^k=P_A(\hat{v}^{j,k})$ by (2.46) and $P_X^k=P_X(\hat{v}^{j,k})$ by (2.47)
5: **if** stopping criterion satisfied **then**
6: Break
7: **end if**
8: **end for**
9: Set $\hat{v}^j=\hat{v}^{j,k}$

The stopping criterion in Algorithm 3, modified from (2.17), is

$$\big[(P_A^k=P_A^{k-1})\text{ and }(P_X^k=P_X^{k-1})\big]\text{ or }\big[\max_i\frac{|\hat{v}_i^{j,k}-\hat{v}_i^{j,k-1}|}{\max(1,|\hat{v}_i^{j,k}|)}\le tol\big].\tag{2.49}$$

Since the matrix solved at each iteration is adjusted by only two diagonal matrices, the sparsity structure of the matrix remains the same. Also since the penalty parameter p is positive, the matrix P_A enhances the diagonal dominance of $\mathbb{I}-\vartheta\Delta\tau^j(A+P_X^{k-1})$. It is

easy to see that the matrix $[\mathbb{I}-\vartheta\Delta\tau^j(A+P_X^{k-1})+P_A^{k-1}]$ is also non-singular, monotone and an M-matrix, under the conditions of Lemma 4.1 in [16], assuming Dirichlet conditions are used.

Remark 3 *The penalty form of the American XVA problem (2.41) can also be written in a control problem form as,*

$$\frac{\partial \hat{V}}{\partial \tau} = \mathcal{L}\hat{V} + \max_{\xi_1,\xi_3\in\{0,1\}} \min_{\xi_2\in\{0,1\}} \left[-\xi_1\lambda_B(1-R_B)\hat{V} - \xi_2(\lambda_C(1-R_C)+s_F)\hat{V} + \xi_3 p(V^*-\hat{V}) \right], \quad (2.50)$$

which involves three controls and both max and min.

When $-\lambda_C(1-R_C) - s_F \leq -\lambda_B(1-R_B) \leq 0$, *problem (2.50) cannot be reduced into a single max or min control problem, and the monotonicity argument fails in the convergence study. Note that* ξ_1 *and* ξ_2 *depend on the same value* $\hat{V}$, *but* ξ_3 *depends on another, namely* $(V^* - \hat{V})$, *arising from American constraints.*

When $-\lambda_B(1-R_B) < -\lambda_C(1-R_C) - s_F \leq 0$, *which is a less frequent and less interesting case, problem (2.50) will end up with a single max*

$$\frac{\partial \hat{V}}{\partial \tau} = \mathcal{L}\hat{V} + \max_{\xi_1,\xi_3\in\{0,1\}} \left[-(1-\xi_1)\lambda_B(1-R_B)\hat{V} - \xi_1(\lambda_C(1-R_C)+s_F)\hat{V} + \xi_3 p(V^*-\hat{V}) \right]. \quad (2.51)$$

To address the convergence problem in HJBI problem, the authors in [38], with appropriate assumptions and more complex arguments, show global convergence for a certain class of HJBI boundary value problems. In [21], other treatments, e.g. a relaxation scheme, or a piecewise constant policy timestepping, are proposed that guarantee global convergence. For our double-penalty method, we did not show global convergence, but, as demonstrated in the numerical experiments, the algorithm exhibits fast convergence. It is worth noting, that, in an initial value problem, a close enough to the solution initial guess is relatively easy to obtain for each timestep.

Chapter 3

Pricing bilateral XVA with stochastic default intensities

In this chapter, we investigate the formulation of pricing XVA when stochastic default intensities are modeled by a mean reversion stochastic process, and correlated with underlying risky assets, and develop computational methods for solving the resulting model. We assume one of the counterparties (usually the buyer) has stochastic default intensity, while another party (usually the seller) exhibits constant risk. The contributions of this part of thesis are:

- We derive a time-dependent nonlinear source term PDE in two space dimensions, namely the asset price S and the default intensity λ_C of the counterparty, and formulate boundary conditions for the problem. We emphasize that the boundary conditions, especially those for the λ_C variable, are critical for the success of the method. We discretize the PDE and boundary conditions, describe how to resolve the nonlinearity, and provide a numerical solution to the PDE problem.

- We develop an alternative solution technique to the two-dimensional (2D) PDE, namely an asymptotic approximation, assuming the mean reversion rate of the default intensity to the mean reversion level is large. The asymptotic approximation is based on the one-dimensional XVA PDE solution and its derivative, and is, therefore, very efficient and simple to implement. We analyze the accuracy of the asymptotic approximation as the mean reversion rate increases.

The outline of the chapter is as follows. In Sections 3.1 and 3.2, we derive and numerically solve the 2D PDE the adjusted derivative price $\hat{V}$ satisfies and formulate boundary conditions for the S and λ_C boundaries. Since the 2D PDE involves a nonlinear

source term, we also present the iteration method to handle the nonlinearity. In Section 3.3, we develop the asymptotic approximation to $\hat{V}$, first for the case that the correlation between S and λ_C is 0, then for the case of nonzero correlation. We also present an analysis of the accuracy of the asymptotic approximation. The numerical experiments to study the behavior of the computed 2D PDE and asymptotic solutions, in terms of various numerical and model parameters, and the comparison of the two solution techniques are left for Chapter 5. We note that many of the results of this chapter will also appear in [15].

3.1 Formulation with stochastic default intensities

We consider the pricing of a European-style derivative on a single risky asset, with price $S(t)$, with two counterparties B and C. The contingent claim value considering default risk can be replicated in a economy consisting of the following four traded assets: the underlying risky asset price $S(t)$, the risky bonds of two parties $P_B(t)$ and $P_C(t)$, and the risk-free bond P. The dynamics of these four assets are modeled as

$$dS(t) = \mu(t)S(t)dt + \sigma^S S(t)dW^S(t) \tag{3.1}$$

$$dP_B = P_B(r_B(t)dt - dJ_B) \tag{3.2}$$

$$dP_C = P_C(r_C(t)dt - dJ_C) \tag{3.3}$$

$$dP(t) = r(t)P(t)dt \tag{3.4}$$

where $W^S(t)$ is a Brownian motion, with μ and σ^S being drift and volatility of $S(t)$ respectively, J_B and J_C are two independent jump processes, that jump from 0 to 1 when default of B or C occurs, respectively, and $r(t)$, $r_B(t)$, $r_C(t)$ are the yields of bonds P, P_B, and P_C, respectively. Note that, generically, the bond yield rate Y of a party is related to its default intensity λ by $\lambda = \frac{Y-r}{1-R}$, where r is the risk-free rate and R is the recovery rate on the bond of the party if default happens. Since P_B and P_C are defined as zero-recovery, zero-coupon bonds, the spreads $r_B - r$ and $r_C - r$ are the same as the respective default intensities λ_B and λ_C. Therefore, in our case, $r_B = \lambda_B + r$ and $r_C = \lambda_C + r$.

3.1.1 Cox-Ingersoll-Ross(CIR) type risk model

We assume the default intensity, $\lambda_C(t)$, of counterparty C is stochastic, while self-party B has low and constant default intensity, i.e. $\lambda_B(t) = \lambda_B$. We assume $\lambda_C(t)$ follows a

CIR process, a type of mean reversion process. The hazard rate process $\lambda_C(t)$ can be formulated as

$$d\lambda_C(t) = \kappa(t)[\theta(t) - \lambda_C(t)]dt + \sigma^{\lambda_C}\sqrt{\lambda_C(t)}dW^{\lambda_C}(t) \tag{3.5}$$

where $\kappa(t)$ is the mean reversion rate, $\theta(t)$ is mean reversion level, σ^{λ_C} is the volatility of mean reversion process, and $W^{\lambda_C}(t)$ is a standard Brownian motion. We also assume $W^S(t)$ and $W^{\lambda_C}(t)$ are correlated with correlation ρ, which is in general nonzero, in order to reflect the dependency between exposure and default risk. Furthermore, assume the Feller condition $2\kappa\theta > (\sigma^{\lambda_C})^2$ is satisfied to ensure that $\lambda_C(t)$ is strictly positive.

In modelling default risk λ, the CIR process has been widely used, for example, in collateralized debt obligation (CDO) [18]. Suppose that each underlying obligor defaults at some expected arrival time. At each time t before default time τ, the default arrives at some "intensity" $\lambda(t)$ with probability $P_t(\tau < t + \Delta t) \approx \lambda(t)\Delta t$. A process $\lambda(t)$ is a stochastic default process, if for a stopping time (default time) τ, whenever $t < \tau$, the survival probability is

$$P_t(\tau > t + s) = \mathbb{E}_t[\exp(\int_t^{t+s} -\lambda(u)du)]$$

where $\mathbb{E}_t$ denotes conditional expectation given all information at time t, and s is length of the period over which survival is considered. Note that, the CIR process is also used for stochastic volatility [33] and stochastic correlation [40]. It also facilitates the derivation of asymptotic approximation method, see Section 3.3.

3.1.2 Formulation of PDE

To formulate the PDE for the XVA pricing problem considering stochastic default intensities, we use dynamic hedging techniques similar to [11]. However, in [11], constant default intensities on both parties are assumed. In this subsection, we show how to embed stochastic default intensity into the XVA PDE and result in a two-dimensional in space time-dependent PDE.

In stochastic default intensity XVA pricing problem, there are only three traded risky assets S, J_B and J_C to hedge out four random sources W^S, W^{λ_C}, J_B and J_C, since λ_C is not a traded risky asset. This is usually called incomplete market. We cannot build a perfectly replicating portfolio with only these three risky assets. One technique is to assume the existence of another benchmark option to complete the market. A similar assumption is used in the PDE derivation of the stochastic volatility option pricing model

[24] or the stochastic correlation option pricing model [40].

Suppose the derivative price $\hat{V}(t, S, \lambda_C, J_B, J_C)$ is totally hedged by a self-financing portfolio Π, such as $\hat{V} + \Pi = 0$ or $-\hat{V} = \Pi$. At time t, the portfolio Π consists of the following assets:

- $\gamma(t)$ [1] units of another option $\tilde{V}(t, S, \lambda_C, J_B, J_C)$ on the same underlying, and with the same maturity and payoff,
- $\delta(t)$ units of the underlying asset $S(t)$,
- $\alpha(t)$ units of bond $P_B(t)$,
- $\beta(t)$ units of bond $P_C(t)$,
- $D(t)$ units of cash deposit.

By the convention in [11], derivative value $\hat{V}$ is positive means that this is a positive asset to party B, while $\hat{V} < 0$ means this is a positive asset to party C. From the 2002 ISDA Master Agreement, the surviving party can receive the recovery potion of contract's mark-to-market value if this derivative contract is positive to this party, while the surviving party should pay full mark-to-market value to the defaulting party, if this derivative contract is negative to the surviving party. Therefore, the "boundary" conditions for $\hat{V}(t, S, \lambda_C, J_B, J_C)$ according to the default of parties B and C, respectively are given by

$$\hat{V}(t, S, \lambda_C, 1, 0) = M^+ + R_B M^-, \tag{3.6}$$

$$\hat{V}(t, S, \lambda_C, 0, 1) = R_C M^+ + M^-, \tag{3.7}$$

where R_B and R_C denote the recovery rates on the derivative's position of parties B and C, respectively, and M is the close-out mark-to-market value of the derivative. Recall that the positive and negative values of any asset Z are denoted as $Z^+ \equiv \max\{Z, 0\}$ and $Z^- \equiv \min\{Z, 0\}$. In this chapter, $\hat{V}(t, S, \lambda_C, 0, 0)$ is usually written as $\hat{V}(t, S, \lambda_C)$.

The value of Π is written as

$$-\hat{V} = \Pi = \delta(t)S(t) + \alpha(t)P_B + \beta(t)P_C + \gamma(t)\tilde{V} + D(t). \tag{3.8}$$

By the assumptions of self-financing, the infinitesimal change is

$$-d\hat{V} = d\Pi = \delta(t)dS(t) + \alpha(t)dP_B + \beta(t)dP_C + \gamma(t)d\tilde{V} + d\bar{D}(t) \tag{3.9}$$

[1] In Section 3.1, $\gamma(t)$ is not the same as the γ, the dividend rate of underlying asset, in the rest of the thesis.

where the change in the cash account [2] is

$$d\bar{D}(t) = \delta(t)(-r_R)S(t)dt$$
$$+\{r(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^+ + r_F(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^-\}dt - r\beta(t)P_C dt. \quad (3.10)$$

Then the first term [3] in (3.10) is corresponding to the cash change from underlying asset share position, combining dividend income and financing cost. The second term (in curly brackets) is corresponding to the cash change in the "funding" account. In this account, any surplus cash held by the seller after the own bonds and option $\tilde{V}$ have been purchased must earn risk-free rate r in order not to introduce any further credit risk. If no surplus cash, the seller needs to pay the rate r_F. The third term is corresponding to the cash changes due to the seller shorting the counterparty bond through a repurchase agreement, which incurs financing costs of rate r.

Applying Ito's lemma for jump diffusions to $\hat{V}$ and $\tilde{V}$ results in

$$d\hat{V} = \mathcal{M}\hat{V}dt + (\sigma^S)S\frac{\partial \hat{V}}{\partial S}dW^S + (\sigma^{\lambda_C})\sqrt{\lambda_C}\frac{\partial \hat{V}}{\partial \lambda_C}dW^{\lambda_C} + \Delta\hat{V}_B dJ_B + \Delta\hat{V}_C dJ_C, \quad (3.11)$$

$$d\tilde{V} = \mathcal{M}\tilde{V}dt + (\sigma^S)S\frac{\partial \tilde{V}}{\partial S}dW^S + (\sigma^{\lambda_C})\sqrt{\lambda_C}\frac{\partial \tilde{V}}{\partial \lambda_C}dW^{\lambda_C} + \Delta\tilde{V}_B dJ_B + \Delta\tilde{V}_C dJ_C, \quad (3.12)$$

where

$$\Delta\hat{V}_B = \hat{V}(t, S, \lambda_C, 1, 0) - \hat{V}(t, S, \lambda_C), \quad (3.13)$$
$$\Delta\tilde{V}_B = \tilde{V}(t, S, \lambda_C, 1, 0) - \tilde{V}(t, S, \lambda_C), \quad (3.14)$$
$$\Delta\hat{V}_C = \hat{V}(t, S, \lambda_C, 0, 1) - \hat{V}(t, S, \lambda_C), \quad (3.15)$$
$$\Delta\tilde{V}_C = \tilde{V}(t, S, \lambda_C, 0, 1) - \tilde{V}(t, S, \lambda_C), \quad (3.16)$$

and where the differential operator $\mathcal{M}$ is defined by

$$\mathcal{M}V = \frac{\partial V}{\partial t} + \frac{1}{2}(\sigma^S)^2 S^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2 V}{\partial \lambda_C^2}$$
$$+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2 V}{\partial S\partial\lambda_C} + \mu S\frac{\partial V}{\partial S} + \kappa(\theta - \lambda_C)\frac{\partial V}{\partial \lambda_C}. \quad (3.17)$$

[2] More details about the mechanism of cash account can found in [11]

[3] r_R is stock financing rate minus dividend income rate.

Combining equations (3.9)-(3.17), the following hedging equation is obtained,

$$
\begin{aligned}
&-\left\{\left[\frac{\partial \hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}+\mu S\frac{\partial\hat{V}}{\partial S}+\kappa(\theta-\lambda_C)\frac{\partial\hat{V}}{\partial\lambda_C}\right]dt\right.\\
&\left.+(\sigma^S)S\frac{\partial\hat{V}}{\partial S}dW^S+(\sigma^{\lambda_C})\sqrt{\lambda_C}\frac{\partial\hat{V}}{\partial\lambda_C}dW^{\lambda_C}+\Delta\hat{V}_BdJ_B+\Delta\hat{V}_CdJ_C\right\}\\
&=\delta(t)\left(\mu Sdt+\sigma^SSdW^S\right)+\alpha(t)P_B(r_Bdt-dJ_B)+\beta(t)P_C(r_Cdt-dJ_C)\\
&+\gamma(t)\left\{\left[\frac{\partial\tilde{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\tilde{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\tilde{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\tilde{V}}{\partial S\partial\lambda_C}+\mu S\frac{\partial\tilde{V}}{\partial S}+\kappa(\theta-\lambda_C)\frac{\partial\tilde{V}}{\partial\lambda_C}\right]dt\right.\\
&\left.+(\sigma^S)S\frac{\partial\tilde{V}}{\partial S}dW^S+(\sigma^{\lambda_C})\sqrt{\lambda_C}\frac{\partial\tilde{V}}{\partial\lambda_C}dW^{\lambda_C}+\Delta\tilde{V}_BdJ_B+\Delta\tilde{V}_CdJ_C\right\}\\
&-\delta(t)r_RS(t)dt+\{r(-\hat{V}-\alpha(t)P_B-\gamma(t)\tilde{V})^+ + r_F(-\hat{V}-\alpha(t)P_B-\gamma(t)\tilde{V})^-\}dt-r\beta P_Cdt. \qquad (3.18)
\end{aligned}
$$

In order to remove all the risk factors in (3.18), the following equations must be satisfied:

$$
\begin{aligned}
-(\sigma^{\lambda_C})\sqrt{\lambda_C}\frac{\partial\hat{V}}{\partial\lambda_C}dW^{\lambda_C}&=\gamma(t)(\sigma_C^{\lambda})\sqrt{\lambda_C}\frac{\partial\tilde{V}}{\partial\lambda_C}dW^{\lambda_C}, &(3.19)\\
-(\sigma^S)S\frac{\partial\hat{V}}{\partial S}dW^S&=\gamma(t)\left((\sigma^S)S\frac{\partial\tilde{V}}{\partial S}dW^S\right)+\delta(t)(\sigma^S)SdW^S, &(3.20)\\
-\Delta\hat{V}_BdJ_B&=-\alpha(t)P_BdJ_B+\gamma(t)\Delta\tilde{V}_BdJ_B &(3.21)\\
-\Delta\hat{V}_CdJ_C&=-\beta(t)P_CdJ_C+\gamma(t)\Delta\tilde{V}_CdJ_C. &(3.22)
\end{aligned}
$$

Hence, the portion of each asset in the portfolio is chosen as

$$
\begin{aligned}
\gamma(t)&=-\frac{\partial\hat{V}}{\partial\lambda_C}\Big/\frac{\partial\tilde{V}}{\partial\lambda_C}, &(3.23)\\
\delta(t)&=-\gamma(t)\frac{\partial\tilde{V}}{\partial S}-\frac{\partial\hat{V}}{\partial S}=(\frac{\partial\hat{V}}{\partial\lambda_C}\Big/\frac{\partial\tilde{V}}{\partial\lambda_C})\frac{\partial\tilde{V}}{\partial S}-\frac{\partial\hat{V}}{\partial S}, &(3.24)\\
\alpha(t)&=\frac{1}{P_B}\left[\Delta\hat{V}_B+\gamma(t)\Delta\tilde{V}_B\right]=\frac{1}{P_B}\left[\Delta\hat{V}_B-(\frac{\partial\hat{V}}{\partial\lambda_C}\Big/\frac{\partial\tilde{V}}{\partial\lambda_C})\Delta\tilde{V}_B\right], &(3.25)\\
\beta(t)&=\frac{1}{P_C}\left[\Delta\hat{V}_C+\gamma(t)\Delta\tilde{V}_C\right]=\frac{1}{P_C}\left[\Delta\hat{V}_C-(\frac{\partial\hat{V}}{\partial\lambda_C}\Big/\frac{\partial\tilde{V}}{\partial\lambda_C})\Delta\tilde{V}_C\right], &(3.26)
\end{aligned}
$$

where

$$\Delta\hat{V}_B = \hat{M}^+ + R_B\hat{M}^- - \hat{V}, \tag{3.27}$$
$$\Delta\hat{V}_C = R_C\hat{M}^+ + \hat{M}^- - \hat{V}, \tag{3.28}$$
$$\Delta\tilde{V}_B = \tilde{M}^+ + R_B\tilde{M}^- - \tilde{V}, \tag{3.29}$$
$$\Delta\tilde{V}_C = R_C\tilde{M}^+ + \tilde{M}^- - \tilde{V}, \tag{3.30}$$

$\hat{M}$ is close-out value for derivative $\hat{V}$ and $\tilde{M}$ is close-out value for derivative $\tilde{V}$.

Therefore, (3.18) becomes

$$\begin{aligned}
&-\left[\frac{\partial\hat{V}}{\partial t} + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2} + \rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}\right]dt\\
&= \alpha(t)P_Br_Bdt + \beta(t)P_Cr_Cdt\\
&+\gamma(t)\left[\frac{\partial\tilde{V}}{\partial t} + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\tilde{V}}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\tilde{V}}{\partial\lambda_C^2} + \rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\tilde{V}}{\partial S\partial\lambda_C}\right]dt - \delta(t)r_RS(t)dt\\
&+\{r(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^+ + r_F(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^-\}dt - r\beta P_Cdt.
\end{aligned} \tag{3.31}$$

Rearranging some terms in (3.31) results in

$$\begin{aligned}
&\alpha(t)P_Br_Bdt + \beta(t)P_Cr_Cdt - r\beta P_Cdt\\
&\quad + \{r(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^+ + r_F(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^-\}dt\\
=&\alpha(t)P_Br_Bdt + \beta(t)P_Cr_Cdt - r\beta P_Cdt\\
&\quad + r(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})dt + s_F(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^-dt\\
=&\alpha(t)P_B(r_B - r)dt + \beta(t)P_C(r_C - r)dt\\
&\quad + r(-\hat{V} - \gamma(t)\tilde{V})dt + s_F(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^-dt.
\end{aligned} \tag{3.32}$$

Taking into account (3.32), equation (3.31) becomes

$$\begin{aligned}
&-\left[\frac{\partial\hat{V}}{\partial t} + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2} + \rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}\right]dt\\
=&\gamma(t)\left[\frac{\partial\tilde{V}}{\partial t} + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\tilde{V}}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\tilde{V}}{\partial\lambda_C^2} + \rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\tilde{V}}{\partial S\partial\lambda_C}\right]dt\\
&- \delta(t)r_RSdt + \alpha(t)P_B(r_B - r)dt + \beta(t)P_C(r_C - r)dt\\
&+ r(-\hat{V} - \gamma(t)\tilde{V})dt + s_F(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^-dt.
\end{aligned} \tag{3.33}$$

If we substitute all the portions into (3.33), we obtain

$$
\begin{aligned}
&-\left[\frac{\partial \hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}-r\hat{V}\right]dt\\
=&\gamma(t)\left[\frac{\partial \tilde{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\tilde{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\tilde{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\tilde{V}}{\partial S\partial\lambda_C}-r\tilde{V}\right]dt\\
&+[\gamma(t)\frac{\partial\tilde{V}}{\partial S}+\frac{\partial\hat{V}}{\partial S}]r_RSdt+(\Delta\hat{V}_B+\gamma(t)\Delta\tilde{V}_B)(r_B-r)dt+(\Delta\hat{V}_C+\gamma(t)\Delta\tilde{V}_C)(r_C-r)dt\\
&+s_F(-\hat{V}-(\Delta\hat{V}_B+\gamma(t)\Delta\tilde{V}_B)-\gamma(t)\tilde{V})^-dt. \qquad (3.34)
\end{aligned}
$$

Then, if we rearrange (3.34) and switch the sign on both hand sides, we have

$$
\begin{aligned}
&\left[\frac{\partial \hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}-r\hat{V}+r_RS\frac{\partial\hat{V}}{\partial S}\right.\\
&\left.+\lambda_B\Delta\hat{V}_B+\lambda_C\Delta\hat{V}_C\right]\\
=&(-\gamma(t))\left[\frac{\partial \tilde{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\tilde{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\tilde{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\tilde{V}}{\partial S\partial\lambda_C}-r\tilde{V}+r_RS\frac{\partial\tilde{V}}{\partial S}\right.\\
&\left.+\lambda_B\Delta\tilde{V}_B+\lambda_C\Delta\hat{V}_C\right]+s_F((\Delta\hat{V}_B+\tilde{V})+\gamma(t)(\Delta\tilde{V}_B+\tilde{V}))^+. \qquad (3.35)
\end{aligned}
$$

If we substitute (3.27)-(3.30) into (3.35), we obtain

$$
\begin{aligned}
&\left[\frac{\partial \hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}-r\hat{V}+r_RS\frac{\partial\hat{V}}{\partial S}\right.\\
&\left.-\lambda_B(\hat{M}^+ + R_B\hat{M}^- - \hat{V})-\lambda_C(R_C\hat{M}^+ + \hat{M}^- - \hat{V})\right]\\
=&(-\gamma(t))\left[\frac{\partial \tilde{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\tilde{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\tilde{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma_C^{\lambda}S\sqrt{\lambda_C}\frac{\partial^2\tilde{V}}{\partial S\partial\lambda_C}-r\tilde{V}+r_RS\frac{\partial\tilde{V}}{\partial S}\right.\\
&\left.-\lambda_B(\tilde{M}^+ + R_B\tilde{M}^- - \tilde{V})-\lambda_C(R_C\tilde{M}^+ + \tilde{M}^- - \tilde{V})\right]\\
&+s_F((\hat{M}^+ + R_B\hat{M}^-)+\gamma(t)(\tilde{M}^+ + R_B\tilde{M}^-))^+. \qquad (3.36)
\end{aligned}
$$

If assume the mark-to-market values of the two options move in the same way, i.e. they are positive and negative at the same time, then we have that $((\hat{M}^+ + R_B\hat{M}^-)+\gamma(t)(\tilde{M}^+ +$

$R_B\tilde{M}^-))^+ = \hat{M}^+ + \gamma(t)\tilde{M}^+$, and

$$
\begin{aligned}
&(1\Big/\frac{\partial\hat{V}}{\partial\lambda_C})\Big[\frac{\partial\hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}-r\hat{V}+r_RS\frac{\partial\hat{V}}{\partial S}\\
&+\lambda_B(\hat{M}^+ + R_B\hat{M}^- - \hat{V})+\lambda_C(R_C\hat{M}^+ + \hat{M}^- - \hat{V})-s_F\hat{M}^+\Big]\\
=&(1\Big/\frac{\partial\tilde{V}}{\partial\lambda_C})\Big[\frac{\partial\tilde{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\tilde{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\tilde{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\tilde{V}}{\partial S\partial\lambda_C}-r\tilde{V}+r_RS\frac{\partial\tilde{V}}{\partial S}\\
&+\lambda_B(\tilde{M}^+ + R_B\tilde{M}^- - \tilde{V})+\lambda_C(R_C\tilde{M}^+ + \tilde{M}^- - \tilde{V})-s_F\tilde{M}^+\Big] \qquad (3.37)
\end{aligned}
$$

Equation (3.37) holds for any derivatives $\hat{V}$ and $\tilde{V}$. Assuming that both sides of (3.37) are equal to the market price of stochastic default intensity, which is usually set as the drift term of intensity dynamics, $-\kappa(t)[\theta(t)-\lambda_C(t)]$, then the PDE becomes

$$
\begin{aligned}
&\frac{\partial\hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}\\
&\qquad +r_RS\frac{\partial\hat{V}}{\partial S}+\kappa[\theta-\lambda_C]\frac{\partial\hat{V}}{\partial\lambda_C}-(r+\lambda_B+\lambda_C)\hat{V}\\
&\qquad =-\lambda_B(\hat{M}^+ + R_B\hat{M}^-)-\lambda_C(R_C\hat{M}^+ + \hat{M}^-)+s_F\hat{M}^+\\
&\qquad =(s_F-\lambda_B-R_C\lambda_C)\hat{M}^+ + (-R_B\lambda_B-\lambda_C)\hat{M}^-. \qquad (3.38)
\end{aligned}
$$

Assuming the mark-to-market value is equal to the risky price, i.e $\hat{M}=\hat{V}$, the derivative price considering bilateral risk satisfies the nonlinear PDE

$$
\begin{aligned}
&\frac{\partial\hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}\\
&\qquad +r_RS\frac{\partial\hat{V}}{\partial S}+\kappa[\theta-\lambda_C]\frac{\partial\hat{V}}{\partial\lambda_C}-r\hat{V}\\
&\qquad =(s_F+(1-R_C)\lambda_C)\hat{V}^+ + (1-R_B)\lambda_B\hat{V}^-. \qquad (3.39)
\end{aligned}
$$

If backward time $\tau = T-t$ is applied, (3.39) becomes

$$
\frac{\partial\hat{V}}{\partial\tau} = \mathcal{L}\hat{V} + f(\lambda_C,\hat{V}), \qquad (3.40)
$$

where the differential operator $\mathcal{L}$ and the nonlinear term $f(\cdot)$ are defined by

$$\begin{aligned}\mathcal{L}\hat{V} \equiv & \frac{1}{2}(\sigma^S)^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_C})^2 \lambda_C \frac{\partial^2 \hat{V}}{\partial \lambda_C^2} + \rho \sigma^S \sigma^{\lambda_C} S \sqrt{\lambda_C} \frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C} \\ & + r_R S \frac{\partial \hat{V}}{\partial S} + \kappa[\theta - \lambda_C] \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V}, \end{aligned} \tag{3.41}$$

$$f(\lambda_C, \hat{V}) \equiv -(s_F + (1 - R_C)\lambda_C)\hat{V}^+ - (1 - R_B)\lambda_B \hat{V}^-. \tag{3.42}$$

The initial (or terminal) condition to (3.40) is the payoff function $H(S)$ of the derivative. The same initial (or terminal) condition holds for the original Black-Scholes equation for the price of the derivative without considering credit risk. It is worth noting that, if $\lambda_B = \lambda_C = s_F = 0$, then PDE (3.40) does not simplify to (2.1) due to the mean-reversion process where the mean θ cannot be zero.

3.1.3 PDE with constant default intensity

For later reference and for comparison purposes, we present the PDE satisfied by the adjusted for default risk derivative price $\hat{V}^c$ when the default intensity λ_C is a given constant:

$$\frac{\partial \hat{V}^c}{\partial \tau} = \frac{1}{2}(\sigma^S)^2 S^2 \frac{\partial^2 \hat{V}^c}{\partial S^2} + r_R S \frac{\partial \hat{V}^c}{\partial S} - r\hat{V}^c - (s_F + (1 - R_C)\lambda_C)(\hat{V}^c)^+ - (1 - R_B)\lambda_B(\hat{V}^c)^-. \tag{3.43}$$

PDE (3.43) is derived in [11] and numerically solved in Chapter 2 of this thesis. Note that PDE (3.43) is the same is the PDE (1.31), taking into account that the transformation $\tau = T - t$. However, we use the notation $\hat{V}^c$ in order to distinguish the constant default intensity derivative price from the stochastic default intensity derivative price.

3.2 Numerical methods

PDE (3.40) for the price $\hat{V}$ of the derivative considering stochastic default intensity for party C is defined in the domain

$$(\tau, S, \lambda_C) \in (0, T] \times [0, \infty) \times [0, \infty),$$

which is unbounded in the two spatial variables.

While the implementation of finite differences is straightforward for (3.40), special care is still needed to introduce appropriate boundary conditions and to deal with the

nonlinearity in the source term.

3.2.1 Discretization

In this subsection, we present the discretization of (3.40). The semi-infinite space domain of spot price S is truncated into $[0, S^{max}]$, for sufficiently large S^{max}, while the semi-infinite space domain of party C spot default intensity λ_C is truncated into $[0, \lambda_C^{max}]$, for sufficiently large λ_C^{max}. In Section (5.2.1), there is a discussion about the effect of S_{max} and λ_C^{max} on the accuracy of the numerical solution. Then, $[0, S^{max}]$ is divided into N subintervals, with the gridpoints $S_0 = 0 < S_1 < \ldots < S_N = S^{max}$ positioned uniformly or nonuniformly, while $[0, \lambda_C^{max}]$ is divided into M subintervals, with the gridpoints $(\lambda_C)_0 = 0 < (\lambda_C)_1 < \ldots < (\lambda_C)_M = \lambda_C^{max}$ positioned uniformly or nonuniformly. The details of the nonuniform positioning of the S- or the λ_C-gridpoints and possible advantages thereof are discussed in Chapter 5. Standard second-order centered finite differences are used for the space discretization of (3.40) except at the boundary points. The details of the boundary conditions and their discretization, as well as the handling of the nonlinear term are discussed in the following two subsections.

For the time-stepping, we employ the ϑ-method [4], which, for $\vartheta = \frac{1}{2}$ and $\vartheta = 1$ becomes the Crank-Nicolson (CN) and Backward Euler (BE) methods, respectively. We also use Rannacher smoothing, which consists of first applying few BE timesteps, then applying CN timestepping. Let $\tau_j, j = 0, \ldots, N_t$, be the timesteps at which the solution is computed, with $\tau_0 = 0 < \tau_1 < \cdots < \tau_{N_t} = T$, and let $\Delta\tau^j = \tau_j - \tau_{j-1}$ be the jth time stepsize. If uniform timesteps are used, then $\Delta\tau = T/N_t$, and Rannacher smoothing first applies four BE timesteps with stepsize $\Delta\tau/2$, then switches to CN with stepsize $\Delta\tau$ for the remaining timesteps, resulting in a total of $N_t + 2$ timesteps.

3.2.2 Boundary conditions

We consider the bounded spatial domain $[0, S^{max}] \times [0, \lambda_C^{max}]$, where S^{max} and λ_C^{max} are sufficiently large, and setup boundary conditions as follows:

- On the $S = 0$ boundary, i.e. on $\{(S, \lambda_C) \in \{S = 0\} \times [0, \lambda_C^{max}]\}$, substitute $S = 0$ into (3.40). This results in a one-dimensional time-dependent PDE,

$$\frac{\partial \hat{V}}{\partial \tau} = \frac{1}{2}(\sigma^{\lambda_C})^2 \lambda_C \frac{\partial^2 \hat{V}}{\partial \lambda_C^2} + \kappa[\theta - \lambda_C]\frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{3.44}$$

[4] Note that the ϑ notation for the time-stepping method is different from the $\theta(t)$ notation in the CIR model.

PDE (3.44) is numerically solved and its computed solution used as Dirichlet boundary condition for solving (3.40). Standard centered differences are used for the discretization of (3.44). However, in order to numerically solve (3.44), appropriate boundary conditions are needed. On the corner point $(0, 0)$, we substitute $\lambda_C = 0$ into (3.44), and get

$$\frac{\partial \hat{V}}{\partial \tau} = \kappa\theta \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{3.45}$$

On the corner point $(0, \lambda_C^{max})$, we substitute $\frac{\partial^2 \hat{V}}{\partial \lambda_C^2} = 0$ into (3.44), and get

$$\frac{\partial \hat{V}}{\partial \tau} = \kappa[\theta - \lambda_C] \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{3.46}$$

Relations (3.45) and (3.46) provide the near and far field boundary conditions, respectively, for (3.44). The first derivative terms of (3.45) and (3.46) are discretized by one-sided finite differences (forward or backward, respectively).

PDE (3.44) together with boundary conditions (3.45) and (3.46) provide approximations to $\hat{V}(\tau, 0, \lambda_C)$ to be used as Dirichlet boundary condition for solving (3.40).

- On the $S = S^{max}$ boundary, i.e. on $\{(S, \lambda_C) \in \{S = S^{max}\} \times [0, \lambda_C^{max}]\}$, we impose the linear boundary condition

$$\frac{\partial^2 \hat{V}}{\partial S^2} = 0. \tag{3.47}$$

Instead of discretizing this condition directly, we substitute $\frac{\partial^2 \hat{V}}{\partial S^2} = 0$ into the PDE (3.40), and get

$$\frac{\partial \hat{V}}{\partial \tau} = \frac{1}{2}(\sigma^{\lambda_C})^2 \lambda_C \frac{\partial^2 \hat{V}}{\partial \lambda_C^2} + \rho\sigma^S \sigma^{\lambda_C} S\sqrt{\lambda_C} \frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C} + r_R S \frac{\partial \hat{V}}{\partial S} + \kappa[\theta - \lambda_C] \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{3.48}$$

Relation (3.48) forms the boundary condition at $S = S^{max}$. The first derivative term $\frac{\partial \hat{V}}{\partial S}$ in (3.48) is discretized by backward differences. For all non-boundary λ_C-points, $\frac{\partial^2 \hat{V}}{\partial \lambda_C^2}$ and $\frac{\partial \hat{V}}{\partial \lambda_C}$ are discretized by standard centered differences, and $\frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C}$ by the Cartesian product of backward differences in S and centered differences in λ_C. On the corner point $(S^{max}, 0)$, relation (3.48) becomes

$$\frac{\partial \hat{V}}{\partial \tau} = r_R S \frac{\partial \hat{V}}{\partial S} + \kappa\theta \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}), \tag{3.49}$$

with $\frac{\partial \hat{V}}{\partial S}$ discretized by backward and $\frac{\partial \hat{V}}{\partial \lambda_C}$ by forward differences.

On the corner point $(S^{max}, \lambda_C^{max})$, besides condition (3.47), we also impose the linear boundary condition on λ_C

$$\frac{\partial^2 \hat{V}}{\partial \lambda_C^2} = 0. \tag{3.50}$$

With (3.47) and (3.50), PDE (3.40) becomes

$$\frac{\partial \hat{V}}{\partial \tau} = \rho \sigma^S \sigma^{\lambda_C} S \sqrt{\lambda_C} \frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C} + r_R S \frac{\partial \hat{V}}{\partial S} + \kappa[\theta - \lambda_C] \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{3.51}$$

For the discretization of $\frac{\partial \hat{V}}{\partial S}$ and $\frac{\partial \hat{V}}{\partial \lambda_C}$ in (3.51), one-sided differences are used, while for $\frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C}$ a Cartesian product of one-sided differences in S and λ_C.

- On the $\lambda_C = 0$ boundary, i.e. $\{(S, \lambda_C) \in (0, S^{max}) \times \{\lambda_C = 0\}\}$, if the Feller condition $2\kappa\theta > (\sigma^{\lambda_C})^2$ is satisfied, this is outflow boundary by Fichera theory, as in the case of mean reversion stochastic volatility or correlation. While the boundary condition is not necessary from the mathematical point of view, we impose an equation at the boundary in order to obtain a square (uniquely solvable) linear system.

  More specifically, we substitute $\lambda_C = 0$ into (3.40), and get

$$\frac{\partial \hat{V}}{\partial \tau} = \frac{1}{2}(\sigma^S)^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + r_R S \frac{\partial \hat{V}}{\partial S} + \kappa\theta \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{3.52}$$

  This PDE is actually the same as the original PDE (3.40) on the $\lambda_C = 0$ boundary. Thus, we do not introduce a new condition, but just apply the PDE itself. In (3.52), the derivatives $\frac{\partial^2 \hat{V}}{\partial S^2}$ and $\frac{\partial \hat{V}}{\partial S}$ are discretized by standard centered differences, while $\frac{\partial \hat{V}}{\partial \lambda_C}$ is discretized by one-sided finite differences.
- On the $\lambda_C = \lambda_C^{max}$ boundary, i.e. $\{(S, \lambda_C) \in (0, S^{max}) \times \{\lambda_C = \lambda_C^{max}\}\}$, we impose the condition $\frac{\partial^2 \hat{V}}{\partial \lambda_C^2} = 0$, substitute this into (3.40), and get

$$\frac{\partial \hat{V}}{\partial \tau} = \frac{1}{2}(\sigma^S)^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + \rho \sigma^S \sigma^{\lambda_C} S \sqrt{\lambda_C} \frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C} + r_R S \frac{\partial \hat{V}}{\partial S} + \kappa[\theta - \lambda_C] \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{3.53}$$

  In (3.53), the derivatives $\frac{\partial^2 \hat{V}}{\partial S^2}$ and $\frac{\partial \hat{V}}{\partial S}$ are discretized by standard centered differences, while $\frac{\partial \hat{V}}{\partial \lambda_C}$ is discretized by one-sided finite differences, and $\frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C}$ by the Cartesian product of centered differences in S and one-sided differences in λ_C. This boundary condition is inspired partly by the fact that, in the constant default in-

tensity calls or puts cases, λ_C acts as a discounting rate (exponential decay rate) – see formula (5.1) in Chapter 5; and also by the fact that numerical experiments for the constant default intensity forward case have shown that λ_C acts similarly, in that, as it increases, the price flattens (at some possibly positive or negative value). Therefore, we expect that homogeneous Neumann conditions would be appropriate for the far side λ_C boundary. Such conditions have been used by many in the literature [50, 37, 49, 31], when considering mean reversion stochastic asset volatility dynamics. Condition (3.53) is more general, in that it covers homogeneous Neumann as well as strictly linear boundary conditions.

It is worth mentioning that the authors of [49], when considering the mean reversion stochastic volatility problem, include an elaborate discussion on how to obtain equations on points where the stochastic volatility tends to 0 or to infinity, for the European and the American cases. The conditions obtained are either homogeneous Dirichlet or Neumann (which linear boundary conditions cover) or non-homogeneous Dirichlet (which are usually problem/product dependent). Taking into account that the Dirichlet conditions are computed at $S = 0$, in advance of the main simulation of (3.40), the total number of unknowns in each timestep of the main simulation is $(M + 1)N$. We number the spatial gridpoints first bottom-up, then left-to-right. Thus, the index i that runs over all spatial gridpoints, is related to the indices i_1 and i_2 that run over all S- and λ_C-gridpoints, respectively, by $i = (i_1 - 1)(M + 1) + i_2$, $i_1 = 1, \ldots, N$, $i_2 = 0, \ldots, M$. With this numbering the arising matrix is block-tridiagonal with tridiagonal blocks.

3.2.3 Iteration methods for nonlinear PDE

In this subsection, we present an iterative method for handling the nonlinearity in (3.40). We refer to it as *discrete penalty-like iteration*, or, simply, *penalty iteration*, motivated by the similarly named method in [22] designed to resolve the nonlinear PDE arising from the linear complementarity problem (LCP) in American option pricing. The first introduction of this method was in [14], for the one-dimensional XVA problem. The method is also presented in Chapter 2. This chapter extends the ideas in Chapter 2 to the case of the multi-dimensional PDE arising from the XVA problem with stochastic default intensity λ_C.

Let $\hat{v}^j, j = 0, \ldots, N_t$, denote the vector of approximate solution values of $\hat{V}$ at the two-dimensional spatial gridpoints at time τ_j, while $\hat{v}^0$ is the initial condition vector. Since we use an iteration method to handle the nonlinearity, let $\hat{v}^{j,k}, k = 0, \ldots, maxit$, denote the computed solution vector at iteration k of timestep j, with $maxit$ the maximum

number of iterations allowed per timestep. Let $f(\hat{v})$ denote the vector arising from evaluating f at the components of $\hat{v}$. This means that $(f(\hat{v}))_i = f((\lambda_C)_{i_2}, \hat{v}_i)$, where $i_2 = i - \lfloor \frac{i}{M+1} \rfloor (M+1)$. Let also A be the matrix arising from the space discretization of $\mathcal{L}\hat{V}$, and $\mathbb{I}$ be the identity matrix of compatible order. For simplicity, we assume the spatial gridpoints remain the same at all timesteps.

For some generic vector v, we define the diagonal penalty matrix $P = P(v)$ by

$$[P(v)]_{i,i} \equiv \begin{cases} -\lambda_B(1 - R_B) & \text{if } v_i < 0 \\ -(\lambda_C)_{i_2}(1 - R_C) - s_F & \text{if } v_i \geq 0. \end{cases} \tag{3.54}$$

Thus, in contrast with the method in Chapter 2, if $v_i \geq 0$, the entries of the penalty matrix are variable. The vector arising from the discretized form of $f(\lambda_C, \hat{V})$ is written as

$$f(\hat{v}) = P(\hat{v})\hat{v}, \tag{3.55}$$

and note that there is nonlinearity between $P(\hat{v})$ and $\hat{v}$.

With the help of the matrix P, the linear system that needs to be solved in each timestep is

$$[\mathbb{I} - \vartheta\Delta\tau^j(A + P(\hat{v}^j))]\hat{v}^j = (\mathbb{I} + (1-\vartheta)\Delta\tau^j A)\hat{v}^{j-1} + (1-\vartheta)\Delta\tau^j P(\hat{v}^{j-1})\hat{v}^{j-1}. \tag{3.56}$$

The proposed discrete penalty-like iteration for (3.40) is described in Algorithm 4.

Algorithm 4 Discrete penalty iteration for (3.40) at step j, with ϑ-timestepping

Require: Solve $[\mathbb{I} - \vartheta\Delta\tau^j(A + P(\hat{v}^j))]\hat{v}^j = g^j$
where $g^j = (\mathbb{I} + (1-\vartheta)\Delta\tau^j A)\hat{v}^{j-1} + (1-\vartheta)\Delta\tau^j P(\hat{v}^{j-1})\hat{v}^{j-1}$.

1: Initialize $\hat{v}^{j,0} = \hat{v}^{j-1}$, and $P^{j,0} = P(\hat{v}^{j,0})$
2: **for** $k = 1, \ldots, maxit$ **do**
3: Solve $[\mathbb{I} - \vartheta\Delta\tau^j(A + P^{j,k-1})]\hat{v}^{j,k} = g^j$
4: Compute $P^{j,k} = P(\hat{v}^{j,k})$ by (3.54)
5: **if** stopping criterion satisfied **then**
6: Break
7: **end if**
8: **end for**
9: Set $\hat{v}^j = \hat{v}^{j,k}$

The stopping criterion in Algorithm 4 is

$$(P^{j,k} = P^{j,k-1}) \text{ or } (\max_i \frac{|\hat{v}_i^{j,k} - \hat{v}_i^{j,k-1}|}{\max(1, |\hat{v}_i^{j,k}|)} \leq tol). \tag{3.57}$$

Note that the matrix solved at each iteration may change, but the change does not affect the sparsity pattern or the computational complexity of the solution. Also note that, if $\lambda_B \geq 0$, $(\lambda_C)_i \geq 0$, and $s_F \geq 0$, we have $P_{i,i}(v) \leq 0$, which enhances the diagonal dominance of A.

For the nonlinearity in the one-dimensional PDE (3.44), a similar penalty iteration is applied, except that, since the problem is along the λ_C dimension only, the diagonal matrix P is defined by

$$[P(v)]_{i_2,i_2} \equiv \begin{cases} -\lambda_B(1 - R_B) & \text{if } v_{i_2} < 0 \\ -(\lambda_C)_{i_2}(1 - R_C) - s_F & \text{if } v_{i_2} \geq 0, \end{cases} \tag{3.58}$$

where v is now a vector of approximate values of $\hat{V}$ at the gridpoints $(0, (\lambda_C)_{i_2})$, $i_2 = 0, \ldots, M$.

3.3 Asymptotic solution

Solving the time-dependent two-dimensional PDE derived in Section 3.1 involves a heavy computational cost. An asymptotic approximation formulae allows us to obtain reasonably accurate results in a more efficient way, namely by a closed-form formula based on one-dimensional PDE results and some additional terms. This asymptotic approximation, built on singular perturbation theory, has been used in the literature, especially when some stochastic variables evolve according to some mean-reverting stochastic processes, [26, 44, 40, 23]. In this section, we derive such an asymptotic approximation for the price of derivatives considering mean-reverting stochastic default intensity.

For convenience, we repeat the stochastic differential equation that the CIR process for λ_C follows:

$$d\lambda_C(t) = \kappa[\theta - \lambda_C(t)]dt + \sigma^{\lambda_C}\sqrt{\lambda_C(t)}dW^{\lambda_C}(t). \tag{3.59}$$

Although one cannot derive a closed-form solution for (3.59), the conditional distribution

is a non-central chi-square distribution with

$$\mathbb{E}[\lambda_C(t))] = e^{-\kappa t}\lambda_C(0) + \theta(1 - e^{-\kappa t}), \tag{3.60}$$

$$\text{Var}[\lambda_C(t)] = \frac{(\sigma^{\lambda_C})^2}{\kappa}\lambda_C(0)(e^{-\kappa t} - e^{-2\kappa t}) + \frac{\theta(\sigma^{\lambda_C})^2}{2\kappa}(1 - e^{-\kappa t})^2. \tag{3.61}$$

As t goes to infinity, we have that the long-run mean level and variance are θ and $\frac{\theta(\sigma^{\lambda_C})^2}{2\kappa}$, respectively.

The invariant distribution λ of a CIR process can be shown to follow a Gamma distribution, with shape parameter $\alpha = \frac{2\kappa\theta}{(\sigma^{\lambda_C})^2}$ and scale parameter $\beta = \frac{(\sigma^{\lambda_C})^2}{2\kappa}$. While not shown here, the invariant distribution is given by [44]

$$\Phi(\lambda) = \frac{e^{-\lambda/\beta}\lambda^{\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}}. \tag{3.62}$$

Following [26, 40], we assume $\kappa = 1/\epsilon$, where $\epsilon > 0$ is small. We also keep the variance ν^2 of the λ_C process invariant distribution constant, thus we scale σ^{λ_C} as $\sigma^{\lambda_C} = \frac{\nu}{\sqrt{\epsilon}}$.

The partial differential equation (3.40), which $\hat{V}(\tau, S, \lambda_C)$ satisfies becomes

$$(\frac{1}{\epsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 + \mathcal{L}_2)\hat{V} = 0, \tag{3.63}$$

where

$$\mathcal{L}_0 \equiv \frac{1}{2}\nu^2\lambda_C\frac{\partial^2}{\partial\lambda_C^2} + (\theta - \lambda_C)\frac{\partial}{\partial\lambda_C}, \tag{3.64}$$

$$\mathcal{L}_1 \equiv \rho\sigma^S\nu S\sqrt{\lambda_C}\frac{\partial^2}{\partial S\partial\lambda_C}, \tag{3.65}$$

$$\mathcal{L}_2 \equiv (-\frac{\partial}{\partial\tau}) + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2}{\partial S^2} + r_R S\frac{\partial}{\partial S} - r\mathcal{I} + f(\lambda_C, \hat{V}), \tag{3.66}$$

with $\mathcal{I}$ being the identity operator.

Remark 4 *For a given problem, κ is given and, therefore, ϵ is calculated as $\frac{1}{\kappa}$. Note that the Feller condition imposes a lower bound for κ, which is translated to an upper bound of ϵ. This means that κ cannot be very small, and ϵ cannot be very large. However, in practical situations, ϵ need not be very small either. For example, as we will be seeing in the numerical experiments, typical values of κ could include $\kappa = 1$, therefore, $\epsilon = 1$. In practice, κ could be calibrated from the bond price of the counterparty. In this study, we assume that $\epsilon < 1$, so that the powers of ϵ converge to zero.*

3.3.1 Asymptotic approximation for zero correlation

When $\rho = 0$, the partial differential equation (3.63), which $\hat{V}(\tau, S, \lambda_C)$ satisfies, becomes

$$(\frac{1}{\epsilon}\mathcal{L}_0 + \mathcal{L}_2)\hat{V} = 0. \tag{3.67}$$

Let $\hat{V}^\epsilon$ be an expansion of $\hat{V}$ in terms of powers of ϵ:

$$\hat{V}^\epsilon = \hat{V}_0 + \epsilon\hat{V}_1 + \epsilon^2\hat{V}_2 + \dots \tag{3.68}$$

We will approximate $\hat{V}$ by the first two terms i.e. by $\hat{V}^{\epsilon,1} = \hat{V}_0 + \epsilon\hat{V}_1$. For $\hat{V}_0$, we impose the same initial condition as for the solution of (3.40), that is, $\hat{V}_0(0, S, \lambda_C) = H(S)$. Let $\langle\cdot\rangle$ denote expectation with respect to the invariant distribution of λ_C.
Substituting (3.68) into (3.67) and equating the lower order terms to zero, we have:

$$\mathcal{O}(\frac{1}{\epsilon}) : \mathcal{L}_0\hat{V}_0 = 0 \tag{3.69}$$

$$\mathcal{O}(1) : \mathcal{L}_0\hat{V}_1 + \mathcal{L}_2\hat{V}_0 = 0 \tag{3.70}$$

$$\mathcal{O}(\epsilon) : \mathcal{L}_0\hat{V}_2 + \mathcal{L}_2\hat{V}_1 = 0. \tag{3.71}$$

Equation (3.69) implies that $\hat{V}_0$ is independent of λ_C, i.e. $\hat{V}_0 = \hat{V}_0(\tau, S)$. Equation (3.69) is a Poisson equation with respect to the operator $\mathcal{L}_0$ in the variable λ_C, which implies the centering condition

$$\langle\mathcal{L}_2\hat{V}_0\rangle = 0. \tag{3.72}$$

As $\hat{V}_0$ is independent of λ_C, the centering condition implies $\langle\mathcal{L}_2\rangle\hat{V}_0 = 0$, where $\langle\mathcal{L}_2\rangle$ is the operator with default intensity being the long-run mean level θ of λ_C under its invariant distribution expectation. Therefore, $\hat{V}_0$ is the solution of the Black-Scholes equation (3.43) taking default risk into account, with constant default intensity $\lambda_C = \theta$, and with terminal condition $\hat{V}_0(0, S) = H(S)$, as derived in [11]. To compute $\hat{V}_0$, for certain derivatives (e.g. European Call and Put), analytic formulae are available, while, for other contingent claims, an one-dimensional parabolic PDE needs to be numerically solved, see, for example, the case of Long Forward contract in Section 5.1.1.

Now let's try to find $\hat{V}_1$. Combining (3.70) and the fact that $\langle\mathcal{L}_2\hat{V}_0\rangle = \langle\mathcal{L}_2\rangle\hat{V}_0 = 0$, we have

$$\mathcal{L}_0\hat{V}_1 = -\mathcal{L}_2\hat{V}_0 = -(\mathcal{L}_2\hat{V}_0 - \langle\mathcal{L}_2\rangle\hat{V}_0) = -(\mathcal{L}_2 - \langle\mathcal{L}_2\rangle)\hat{V}_0, \tag{3.73}$$

where

$$(\mathcal{L}_2 - \langle\mathcal{L}_2\rangle)\hat{V}_0 = f(\lambda_C, \hat{V}_0) - f(\theta, \hat{V}_0) = (1 - R_C)(\theta - \lambda_C)\hat{V}_0^+. \tag{3.74}$$

Suppose the function $\phi(\lambda_C)$ is the solution to $\mathcal{L}_0\phi = (\theta - \lambda_C)$. It is easy to verify that $\phi(\lambda_C) = (\lambda_C - \theta) + \tilde{C}(\tau, S)$, where $\tilde{C}$ independent of λ_C. Hence, we can write $\hat{V}_1$ as

$$\hat{V}_1 = -(1 - R_C)[(\lambda_C - \theta) + \tilde{C}(\tau, S)]\hat{V}_0^+. \tag{3.75}$$

Without loss of generality of $\tilde{C}(\tau, S)$, we can also rewrite $\hat{V}_1$ as

$$\hat{V}_1 = -(1 - R_C)(\lambda_C - \theta)\hat{V}_0^+ + C(\tau, S) = (1 - R_C)(\theta - \lambda_C)\hat{V}_0^+ + C(\tau, S), \tag{3.76}$$

where $C(\tau, S) = -(1-R_C)\tilde{C}(\tau, S)\hat{V}_0^+$. To find the form of $C(\tau, S)$, we derive an equation that $C(\tau, S)$ satisfies and obtain a solution for it. The Poisson equation (3.71) implies the centering condition

$$\langle\mathcal{L}_2\hat{V}_1\rangle = 0. \tag{3.77}$$

Now consider

$$\begin{aligned}
&\langle\mathcal{L}_2\hat{V}_1\rangle = \langle\mathcal{L}_2(-(1 - R_C)(\lambda_C - \theta)\hat{V}_0^+ + C(\tau, S))\rangle \\
\implies\ &\langle\mathcal{L}_2 C(\tau, S)\rangle = \langle\mathcal{L}_2(1 - R_C)(\lambda_C - \theta)\hat{V}_0^+\rangle \\
\implies\ &\langle\mathcal{L}_2\rangle C(\tau, S) = \langle(\mathcal{L}_2 - \langle\mathcal{L}_2\rangle)(1 - R_C)(\lambda_C - \theta)\hat{V}_0^+\rangle + \langle\mathcal{L}_2\rangle(1 - R_C)(\lambda_C - \theta)\hat{V}_0^+.
\end{aligned}$$

Note that $\langle\mathcal{L}_2\rangle\hat{V}_0^+ = 0$, thus

$$\begin{aligned}
\langle\mathcal{L}_2\rangle C(\tau, S) &= \langle(\mathcal{L}_2 - \langle\mathcal{L}_2\rangle)(1 - R_C)(\lambda_C - \theta)\hat{V}_0^+\rangle \\
&= \langle(1 - R_C)^2(\lambda_C - \theta)(\theta - \lambda_C)\hat{V}_0^+\rangle \\
&= \langle(\lambda_C - \theta)(\theta - \lambda_C)\rangle(1 - R_C)^2\hat{V}_0^+ \\
&= -(1 - R_C)^2\frac{\theta\nu^2}{2}\hat{V}_0^+.
\end{aligned} \tag{3.78}$$

Note that $C(\tau, S) = \tau(1 - R_C)^2\frac{\theta\nu^2}{2}\hat{V}_0^+$ is a solution to (3.78). Thus, (3.76) gives

$$\hat{V}_1 = (1 - R_C)(\theta - \lambda_C)\hat{V}_0^+ + \tau(1 - R_C)^2\frac{\theta\nu^2}{2}\hat{V}_0^+. \tag{3.79}$$

Therefore, we obtain the approximation

$$\hat{V}^{\epsilon,1} = \hat{V}_0 + \epsilon(1-R_C)(\theta-\lambda_C)\hat{V}_0^+ + \epsilon\tau(1-R_C)^2\frac{\theta\nu^2}{2}\hat{V}_0^+, \tag{3.80}$$

where $\hat{V}_0$ is computed as explained above. Note that this approximation is only for the case of zero correlation.

3.3.2 Asymptotic approximation for general correlation

Recall that the partial differential equation which $\hat{V}(\tau, S, \lambda_C)$ satisfies is (3.63). Let $\hat{V}^\epsilon$ be a power series expansion of $\hat{V}$ in $\sqrt{\epsilon}$:

$$\hat{V}^\epsilon = \hat{V}_0 + \sqrt{\epsilon}\hat{V}_{1/2}{}^5 + \epsilon\hat{V}_1 + \epsilon\sqrt{\epsilon}\hat{V}_{3/2} + \dots \tag{3.81}$$

We approximate $\hat{V}$ by the first three terms i.e. by $\hat{V}^{\epsilon,1} = \hat{V}_0 + \sqrt{\epsilon}\hat{V}_{1/2} + \epsilon\hat{V}_1$. For $\hat{V}_0$, we impose the same initial condition as for the case of zero correlation and as for (3.40), that is, $\hat{V}_0(0, S, \lambda_C) = H(S)$.
Substituting (3.81) into (3.63) and equating the lower order terms to zero, we have:

$$\mathcal{O}(\frac{1}{\epsilon}) : \mathcal{L}_0\hat{V}_0 = 0 \tag{3.82}$$

$$\mathcal{O}(\frac{1}{\sqrt{\epsilon}}) : \mathcal{L}_0\hat{V}_{1/2} + \mathcal{L}_1\hat{V}_0 = 0 \tag{3.83}$$

$$\mathcal{O}(1) : \mathcal{L}_0\hat{V}_1 + \mathcal{L}_1\hat{V}_{1/2} + \mathcal{L}_2\hat{V}_0 = 0 \tag{3.84}$$

$$\mathcal{O}(\sqrt{\epsilon}) : \mathcal{L}_0\hat{V}_{3/2} + \mathcal{L}_1\hat{V}_1 + \mathcal{L}_2\hat{V}_{1/2} = 0. \tag{3.85}$$

Equation (3.82) implies that $\hat{V}_0$ is independent of λ_C, i.e. $\hat{V}_0 = \hat{V}_0(\tau, S)$. Hence, $\mathcal{L}_1\hat{V}_0 = 0$. In (3.83), this results in $\mathcal{L}_0\hat{V}_{1/2} = 0$, which implies $\hat{V}_{1/2}$ is independent of λ_C as well, i.e. $\hat{V}_{1/2} = \hat{V}_{1/2}(\tau, S)$. Coming to the $\mathcal{O}(1)$ term, Equation (3.84), given $\mathcal{L}_1\hat{V}_{1/2} = 0$, reduces to $\mathcal{L}_0\hat{V}_1 + \mathcal{L}_2\hat{V}_0 = 0$. This is a Poisson equation with respect to the operator $\mathcal{L}_0$ in the variable λ_C, which implies the centering condition

$$\langle\mathcal{L}_2\hat{V}_0\rangle = 0. \tag{3.86}$$

Similarly as in the case of zero correlation, as $\hat{V}_0$ is independent of λ_C, the centering

[5]In the notations $\hat{V}_{1/2}$ and $\hat{V}_{3/2}$, the subscripts are consistent with the powers of the associated ϵ coefficients. In this way, $\hat{V}_1$ of the general correlation case, is derived to be the same as that for the zero correlation case.

condition (3.86) becomes $\langle\mathcal{L}_2\rangle\hat{V}_0 = 0$. Therefore, $\hat{V}_0$ is the solution to the Black-Scholes equation (3.43) computed as explained in the zero correlation case.

Now let's try to find an expression for $\hat{V}_1$, then for $\hat{V}_{1/2}$. As mentioned, the $\mathcal{O}(1)$ term (3.84) can be reduced to $\mathcal{L}_0\hat{V}_1 + \mathcal{L}_2\hat{V}_0 = 0$, which is exactly same as Equation (3.70). Hence, the formula for $\hat{V}_1$ is given in the previous subsection, and is

$$\hat{V}_1 = (1-R_C)(\theta-\lambda_C)\hat{V}_0^+ + \tau(1-R_C)^2\frac{\theta\nu^2}{2}\hat{V}_0^+. \tag{3.87}$$

For equation (3.85), again, the solvability of this Poisson equation requires

$$\langle\mathcal{L}_1\hat{V}_1 + \mathcal{L}_2\hat{V}_{1/2}\rangle = 0, \tag{3.88}$$

which gives

$$\begin{aligned}\langle\mathcal{L}_2\rangle\hat{V}_{1/2} &= -\langle\mathcal{L}_1\hat{V}_1\rangle \\ &= -\langle\rho\sigma^S\nu S\sqrt{\lambda_C}\frac{\partial^2}{\partial S\partial\lambda_C}[(1-R_C)(\theta-\lambda_C)\hat{V}_0^+ + \tau(1-R_C)^2\frac{\theta\nu^2}{2}\hat{V}_0^+]\rangle \\ &= \langle\rho\sigma^S\nu S\sqrt{\lambda_C}(1-R_C)\frac{\partial\hat{V}_0^+}{\partial S}\rangle = \rho\sigma^S\nu S(1-R_C)\langle\sqrt{\lambda_C}\rangle\frac{\partial\hat{V}_0^+}{\partial S}.\end{aligned} \tag{3.89}$$

Because $\langle\mathcal{L}_2\rangle\hat{V}_0^+ = 0$, we can verify that the solution to (3.89) is

$$\hat{V}_{1/2}(\tau,S) = -\tau\rho\sigma^S\nu S(1-R_C)\langle\sqrt{\lambda_C}\rangle\frac{\partial\hat{V}_0^+}{\partial S}. \tag{3.90}$$

Therefore, the approximation $\hat{V}^{\epsilon,1}$ to $\hat{V}$ is obtained as

$$\hat{V}^{\epsilon,1} = \hat{V}_0 - \sqrt{\epsilon}\tau\rho\sigma^S\nu S(1-R_C)\langle\sqrt{\lambda_C}\rangle\frac{\partial\hat{V}_0^+}{\partial S} + \epsilon(1-R_C)(\theta-\lambda_C)\hat{V}_0^+ + \epsilon\tau(1-R_C)^2\frac{\theta\nu^2}{2}\hat{V}_0^+. \tag{3.91}$$

More details

- Function $\hat{V}_0^+$ might be a nonsmooth function, whose derivative with respect to S on few points may be undefined. From a financial interpretation, we define

$$\frac{\partial\hat{V}_0^+}{\partial S} = \begin{cases}\frac{\partial\hat{V}_0}{\partial S} & \hat{V}_0 > 0 \\ 0 & \hat{V}_0 \le 0\end{cases}. \tag{3.92}$$

- In approximation (3.91), by the definition of $\langle\cdot\rangle$, the quantity $\langle\sqrt{\lambda_C}\rangle$ can be com-

puted by

$$\langle\sqrt{\lambda_C}\rangle := \int_0^\infty \sqrt{\lambda_C}\Phi(\lambda_C)d\lambda_C, \tag{3.93}$$

where $\Phi(\lambda_C)$ is the probability density function of stationary distribution given in (3.62). Numerical quadrature, e.g. MATLAB's `integral`, can be used if an analytical formula cannot be easily obtained.

3.3.3 Accuracy analysis of asymptotic approximation

In this subsection, we analyze the order of convergence of the approximation $\hat{V}^{\epsilon,1}$ to $\hat{V}$ in terms of ϵ, under the assumption that there exists a positive lower bound δ_C to variable $(s_F + (1 - R_C)\lambda_C)$. We consider the case of general correlation. The case of zero correlation is briefly considered at the end. In order to analyze the quality of the approximation $\hat{V}^{\epsilon,1}$ of (3.91) to $\hat{V}$, we begin by recalling (3.63), and defining

$$\mathcal{L}^\epsilon\hat{V}^\epsilon \equiv (\frac{1}{\epsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 + \mathcal{L}_2)\hat{V}^\epsilon = 0, \tag{3.94}$$

$$\mathcal{E} \equiv \hat{V}^\epsilon - \hat{V}_0 - \sqrt{\epsilon}\hat{V}_{1/2} - \epsilon\hat{V}_1 - \epsilon\sqrt{\epsilon}\hat{V}_{3/2}, \tag{3.95}$$

where we also recall that $\hat{V}^\epsilon = \hat{V}_0 + \sqrt{\epsilon}\hat{V}_{1/2} + \epsilon\hat{V}_1 + \epsilon\sqrt{\epsilon}\hat{V}_{3/2} + \epsilon^2\hat{V}_2 + \ldots$. The initial condition for $\mathcal{E}$ is

$$\mathcal{E}(0, S, \lambda_C) = -\epsilon(1 - R_C)(\theta - \lambda_C)\hat{V}_0^+(0, S) - \epsilon\sqrt{\epsilon}\hat{V}_{3/2}(0, S, \lambda_C), \tag{3.96}$$

since, at $\tau = 0$, $\hat{V}^\epsilon = \hat{V}_0$, which is also equal to the payoff function.

In addition, we have

$$\begin{aligned}
\mathcal{L}^\epsilon\mathcal{E} =&(\frac{1}{\epsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 + \mathcal{L}_2)(\hat{V}^\epsilon - \hat{V}_0 - \sqrt{\epsilon}\hat{V}_{1/2} - \epsilon\hat{V}_1 - \epsilon\sqrt{\epsilon}\hat{V}_{3/2}) \\
=& -\frac{1}{\epsilon}(\mathcal{L}_0\hat{V}_0) - \frac{1}{\sqrt{\epsilon}}(\mathcal{L}_0\hat{V}_{1/2} + \mathcal{L}_1\hat{V}_0) - (\mathcal{L}_0\hat{V}_1 + \mathcal{L}_1\hat{V}_{1/2} + \mathcal{L}_2\hat{V}_0) \\
&- \sqrt{\epsilon}(\mathcal{L}_0\hat{V}_{3/2} + \mathcal{L}_1\hat{V}_1 + \mathcal{L}_2\hat{V}_{1/2}) - \epsilon(\mathcal{L}_1\hat{V}_{3/2} + \sqrt{\epsilon}\mathcal{L}_2\hat{V}_{3/2}) \\
=& -\epsilon(\mathcal{L}_1\hat{V}_{3/2} + \sqrt{\epsilon}\mathcal{L}_2\hat{V}_{3/2}),
\end{aligned} \tag{3.97}$$

taking also into account (3.82)-(3.85). The Feynman-Kac probabilistic representation

formula for the solution of (3.97) is given as

$$\begin{aligned}\mathcal{E}(\tau,S,\lambda_C) =& -\mathbb{E}^Q\left[e^{-r\tau}(\epsilon(1-R_C)(\theta-\lambda_C)\hat{V}_0^+ + \epsilon^{3/2}\hat{V}_{3/2} \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right]\\ &+\mathbb{E}^Q\left[\int_0^\tau e^{-r(\tau-s)}(\epsilon\mathcal{L}_1\hat{V}_{3/2}+\epsilon^{3/2}\mathcal{L}_2\hat{V}_{3/2})ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right]\\ &+\mathbb{E}^Q\left[\int_0^\tau e^{-r(\tau-s)}f(\lambda_C,\mathcal{E})ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right]. \end{aligned} \tag{3.98}$$

By (3.85), we have

$$\begin{aligned}\mathcal{L}_0\hat{V}_{3/2} &= -\mathcal{L}_1\hat{V}_1-\mathcal{L}_2\hat{V}_{1/2}\\ &= \rho\sigma^S\nu S(1-R_C)(\sqrt{\lambda_C}-\langle\sqrt{\lambda_C}\rangle)\frac{\partial\hat{V}_0^+}{\partial S}+f(\theta,\hat{V}_{1/2})-f(\lambda_C,\hat{V}_{1/2}).\end{aligned} \tag{3.99}$$

Hence, following the arguments in [40], $|\hat{V}_{3/2}|\le C_1|\frac{\partial\hat{V}_0^+}{\partial S}|$, where C_1 is a positive constant. Then we have

$$\left|\mathbb{E}^Q\left[\epsilon^{3/2}\hat{V}_{3/2} \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right]\right|\le C_2\epsilon^{3/2}, \tag{3.100}$$

$$\left|\mathbb{E}^Q\left[\int_0^\tau e^{-r(\tau-s)}\epsilon^{3/2}\mathcal{L}_2\hat{V}_{3/2}ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right]\right|\le C_3\epsilon^{3/2}, \tag{3.101}$$

where C_2 and C_3 are positive constants. We also notice that

$$\begin{aligned}&\left|\mathbb{E}^Q\left[\int_0^\tau e^{-r(\tau-s)}\epsilon\mathcal{L}_1\hat{V}_{3/2}ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right]\right|\\ =&\left|\mathbb{E}^Q\left[\int_0^\tau e^{-r(\tau-s)}\epsilon\rho\sigma^S\nu S\sqrt{\lambda_C}\frac{\partial^2\hat{V}_{3/2}}{\partial S\partial\lambda_C}ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right]\right|\\ =&\left|\mathbb{E}^Q\left[\int_0^\tau e^{-r(\tau-s)}\epsilon^{3/2}\rho\sigma^S S\sigma^{\lambda_C}\sqrt{\lambda_C}\frac{\partial^2\hat{V}_{3/2}}{\partial S\partial\lambda_C}ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right]\right|\\ \le& C_4\epsilon^{3/2},\end{aligned} \tag{3.102}$$

where C_4 is a positive constant.

With similar arguments as in [23, 40, 44], and taking into account that $\langle\theta-\lambda_C\rangle=0$, we have, with C_5 and C_6 positive constants,

$$\begin{aligned}&\left|\mathbb{E}^Q\left[e^{-r\tau}((1-R_C)(\theta-\lambda_C)\hat{V}_0^+ \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right]\right|\\ \le& C_5\left|\mathbb{E}^Q\left[(\theta-\lambda_C) \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right]\right| \le C_5e^{-C_6\frac{1}{\epsilon}},\end{aligned} \tag{3.103}$$

which converges exponentially fast as $\epsilon \to 0$.

Now let's consider the last term of the right-hand-side of (3.98):

$$\begin{aligned}
&\mathbb{E}^Q\left[\int_0^\tau e^{-r(\tau-s)} f(\lambda_C, \mathcal{E}) ds \mid S^\epsilon_{T-\tau} = S, (\lambda_C)^\epsilon_{T-\tau} = \lambda_C\right] \\
=&\mathbb{E}^Q\left[\int_0^\tau e^{-r(\tau-s)}[-(s_F + (1-R_C)\lambda_C)\mathcal{E}^+ - (1-R_B)\lambda_B\mathcal{E}^-] ds \mid S^\epsilon_{T-\tau} = S, (\lambda_C)^\epsilon_{T-\tau} = \lambda_C\right] \\
=&-\mathbb{E}^Q\left[\int_0^\tau e^{-r(\tau-s)}[(s_F + (1-R_C)\lambda_C)\mathcal{E}^+ + (1-R_B)\lambda_B\mathcal{E}^-] ds \mid S^\epsilon_{T-\tau} = S, (\lambda_C)^\epsilon_{T-\tau} = \lambda_C\right].
\end{aligned} \tag{3.104}$$

Taking into account (3.100)-(3.104), equation (3.98) results in

$$\begin{aligned}
&\left|\mathcal{E}(\tau, S, \lambda_C) + \mathbb{E}^Q\Big[\int_0^\tau e^{-r(\tau-s)}[(s_F + (1-R_C)\lambda_C)\mathcal{E}(\tau, S, \lambda_C)^+ \right. \\
&\left. +(1-R_B)\lambda_B\mathcal{E}(\tau, S, \lambda_C)^-] ds \mid S^\epsilon_{T-\tau} = S, (\lambda_C)^\epsilon_{T-\tau} = \lambda_C\Big]\right| \le C\epsilon^{3/2},
\end{aligned} \tag{3.105}$$

where C is a positive constant.

We consider two cases for the sign of $\mathcal{E}(\tau, S, \lambda_C)$. If $\mathcal{E}(\tau, S, \lambda_C) \le 0$, equation (3.105) results in

$$\left|\mathcal{E}(\tau, S, \lambda_C) + \mathbb{E}^Q\Big[\int_0^\tau e^{-r(\tau-s)}(1-R_B)\lambda_B\mathcal{E}(\tau, S, \lambda_C) ds \mid S^\epsilon_{T-\tau} = S, (\lambda_C)^\epsilon_{T-\tau} = \lambda_C\Big]\right| \le C\epsilon^{3/2}, \tag{3.106}$$

where both terms inside the absolute value are negative or zero. We also have

$$\begin{aligned}
&\left|\mathbb{E}^Q\Big[\int_0^\tau e^{-r(\tau-s)}(1-R_B)\lambda_B\mathcal{E}(\tau, S, \lambda_C) ds \mid S^\epsilon_{T-\tau} = S, (\lambda_C)^\epsilon_{T-\tau} = \lambda_C\Big]\right| \\
\ge&\left|(1-R_B)\lambda_B e^{-r\tau}\mathbb{E}^Q\Big[\int_0^\tau \mathcal{E}(\tau, S, \lambda_C) ds \mid S^\epsilon_{T-\tau} = S, (\lambda_C)^\epsilon_{T-\tau} = \lambda_C\Big]\right| \\
=&\left|(1-R_B)\lambda_B e^{-r\tau}\mathbb{E}^Q\Big[\mathcal{E}(\tau, S, \lambda_C) \mid S^\epsilon_{T-\tau} = S, (\lambda_C)^\epsilon_{T-\tau} = \lambda_C\Big]\right| \\
=&\left|(1-R_B)\lambda_B e^{-r\tau}\tau\mathcal{E}(\tau, S, \lambda_C)\right|.
\end{aligned} \tag{3.107}$$

With (3.107), relation (3.106) leads to

$$\left|\mathcal{E}(\tau, S, \lambda_C) + (1-R_B)\lambda_B e^{-r\tau}\tau\mathcal{E}(\tau, S, \lambda_C)\right| \le C\epsilon^{3/2}, \tag{3.108}$$

from which we get

$$|1+(1-R_B)\lambda_B e^{-r\tau}\tau||\mathcal{E}(\tau,S,\lambda_C)| \leq C\epsilon^{3/2} \implies$$
$$|\mathcal{E}(\tau,S,\lambda_C)| \leq \frac{C}{|1+(1-R_B)\lambda_B e^{-r\tau}\tau|}\epsilon^{3/2}. \tag{3.109}$$

If $\mathcal{E}(\tau,S,\lambda_C)>0$, equation (3.105) results in

$$\Big|\mathcal{E}(\tau,S,\lambda_C)+\mathbb{E}^Q\Big[\int_0^\tau e^{-r(\tau-s)}[(s_F+(1-R_C)\lambda_C)\mathcal{E}(\tau,S,\lambda_C)^+]ds \\ \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\Big| \leq C\epsilon^{3/2}, \tag{3.110}$$

where both terms inside the absolute value are positive. We also have

$$\begin{aligned}
&\Big|\mathbb{E}^Q\Big[\int_0^\tau e^{-r(\tau-s)}[(s_F+(1-R_C)\lambda_C)\mathcal{E}(\tau,S,\lambda_C)^+]ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\Big| \\
\geq&\Big|\mathbb{E}^Q\Big[\int_0^\tau e^{-r(\tau)}\delta_C\mathcal{E}(\tau,S,\lambda_C)ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\Big| \\
=&\Big|e^{-r\tau}\delta_C\mathbb{E}^Q\Big[\int_0^\tau \mathcal{E}(\tau,S,\lambda_C)ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\Big| \\
=&\Big|e^{-r\tau}\delta_C\tau\mathcal{E}(\tau,S,\lambda_C)\Big|,
\end{aligned} \tag{3.111}$$

where $\delta_C>0$ is a constant lower bound of $(s_F+(1-R_C)\lambda_C)$. With (3.111), relation (3.110) leads to

$$\Big|\mathcal{E}(\tau,S,\lambda_C)+e^{-r\tau}\delta_C\tau\mathcal{E}(\tau,S,\lambda_C)\Big| \leq C\epsilon^{3/2}, \tag{3.112}$$

from which we get

$$|1+e^{-r\tau}\delta_C\tau||\mathcal{E}(\tau,S,\lambda_C)| \leq C\epsilon^{3/2} \implies$$
$$|\mathcal{E}(\tau,S,\lambda_C)| \leq \frac{C}{|1+e^{-r\tau}\delta_C\tau|}\epsilon^{3/2}. \tag{3.113}$$

Combining (3.109) and (3.113), we have that the accuracy of the asymptotic solution $\hat{V}^{\epsilon,1}=\hat{V}_0+\sqrt{\epsilon}\hat{V}_{1/2}+\epsilon\hat{V}_1$ of (3.91) is at least of order $\mathcal{O}(\epsilon^{3/2})$.

For the case of zero correlation, since the analysis technique is similar, and since the general correlation case is more interesting, we do not show the details, but present the final result. We can show that the approximation $\hat{V}^{\epsilon,1}$ of (3.80) is at least of order $\mathcal{O}(\epsilon^2)$.

3.4 Consideration for bilateral stochastic default intensities

In this section, we give some brief remarks how the 2D PDE derived with stochastic counterparty default intensity can be extended to a 3D PDE with stochastic default intensities for both parties. We could reasonably assume that each party follows a CIR stochastic process:

$$d\lambda_B(t) = \kappa_B(t)[\theta_B(t) - \lambda_B(t)]dt + \sigma^{\lambda_B}\sqrt{\lambda_B(t)}dW^{\lambda_B}(t), \tag{3.114}$$

$$d\lambda_C(t) = \kappa_C(t)[\theta_C(t) - \lambda_C(t)]dt + \sigma^{\lambda_C}\sqrt{\lambda_C(t)}dW^{\lambda_C}(t), \tag{3.115}$$

where κ_B and κ_C are the rates of mean-reversion of parties B and C respectively, θ_B and θ_C are the levels of mean-reversion of parties B and C respectively, σ^{λ_B} and σ^{λ_C} are the volatilities of mean-reversion of parties B and C respectively, and $W^{\lambda_B}(t)$ and $W^{\lambda_C}(t)$ are independent of each other, assuming independence between the two parties. However, we also assume that $W^{\lambda_B}(t)$ and $W^{\lambda_C}(t)$ are correlated W^S with correlation ρ_B and ρ_C respectively. We expect the following PDE can be derived if backward time $\tau = T - t$ is applied:

$$\begin{aligned}\frac{\partial \hat{V}}{\partial \tau} =& \frac{1}{2}(\sigma^S)^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_B})^2 \lambda_B \frac{\partial^2 \hat{V}}{\partial \lambda_B^2} + \frac{1}{2}(\sigma^{\lambda_C})^2 \lambda_C \frac{\partial^2 \hat{V}}{\partial \lambda_C^2} \\ &+ \rho_B \sigma^S \sigma^{\lambda_B} S\sqrt{\lambda_B}\frac{\partial^2 \hat{V}}{\partial S \partial \lambda_B} + \rho_C \sigma^S \sigma^{\lambda_C} S\sqrt{\lambda_C}\frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C} \\ &+ r_R S\frac{\partial \hat{V}}{\partial S} + \kappa_B[\theta_B - \lambda_B]\frac{\partial \hat{V}}{\partial \lambda_B} + \kappa_C[\theta_C - \lambda_C]\frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} \\ &- (s_F + (1 - R_C)\lambda_C)\hat{V}^+ - (1 - R_B)\lambda_B \hat{V}^-. \end{aligned} \tag{3.116}$$

This 3D time-dependent PDE is much heavier to numerically solve than the 2D time-dependent PDE (3.40).

The asymptotic approach can also be extended to bilateral stochastic default intensities. For this purpose, we need to assume that, $\kappa_B = \frac{1}{\epsilon_B}$ and $\kappa_C = \frac{1}{\epsilon_C}$, where $\epsilon_B, \epsilon_C > 0$ are small. We also keep the variance ν_B and ν_C of λ_B and λ_C processes invariant distribution constant, hence we apply the scalings $\sigma^{\lambda_B} = \frac{\nu_B}{\sqrt{\epsilon_B}}$ and $\sigma^{\lambda_C} = \frac{\nu_C}{\sqrt{\epsilon_C}}$ respectively. Then, the 3D time-dependent PDE above, which $\hat{V}(\tau, S, \lambda_B, \lambda_C)$ satisfies, becomes

$$(\frac{1}{\epsilon_B}\mathcal{L}_{0,B} + \frac{1}{\epsilon_C}\mathcal{L}_{0,C} + \frac{1}{\sqrt{\epsilon_B}}\mathcal{L}_{1,B} + \frac{1}{\sqrt{\epsilon_C}}\mathcal{L}_{1,C} + \mathcal{L}_2)\hat{V} = 0, \tag{3.117}$$

where

$$\mathcal{L}_{0,B} \equiv \frac{1}{2}\nu_B^2\lambda_B\frac{\partial^2}{\partial\lambda_B^2} + (\theta_B - \lambda_B)\frac{\partial}{\partial\lambda_B}, \tag{3.118}$$

$$\mathcal{L}_{0,C} \equiv \frac{1}{2}\nu_C^2\lambda_C\frac{\partial^2}{\partial\lambda_C^2} + (\theta_C - \lambda_C)\frac{\partial}{\partial\lambda_C}, \tag{3.119}$$

$$\mathcal{L}_{1,B} \equiv \rho_B\sigma^S\nu_B S\sqrt{\lambda_B}\frac{\partial^2}{\partial S\partial\lambda_B}, \tag{3.120}$$

$$\mathcal{L}_{1,C} \equiv \rho_C\sigma^S\nu_C S\sqrt{\lambda_C}\frac{\partial^2}{\partial S\partial\lambda_C}, \tag{3.121}$$

$$\mathcal{L}_2 \equiv (-\frac{\partial}{\partial\tau}) + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2}{\partial S^2} + r_R S\frac{\partial}{\partial S} - r\mathcal{I} + f(\hat{V},\lambda_B,\lambda_C), \tag{3.122}$$

$$f(\hat{V},\lambda_B,\lambda_C) = -(s_F + (1-R_C)\lambda_C)\hat{V}^+ - (1-R_B)\lambda_B\hat{V}^- \tag{3.123}$$

with $\mathcal{I}$ being the identity operator.

The double asymptotic expansion in terms of ϵ_B, and ϵ_C would be

$$\hat{V}^{\epsilon_B,\epsilon_C} = \sum_{j\geq 0}\sum_{i\geq 0}(\epsilon_B)^{i/2}(\epsilon_C)^{j/2}\hat{V}_{(i/2,j/2)}. \tag{3.124}$$

This is a formal series expansion, for which we need to find $\hat{V}_{(i/2,j/2)}$, for $i+j \leq 2$ explicitly, to have the double asymptotic approximation. The details are certainly more involved, but we could reasonably expect the approximation, because of the independence between the two parties, would be

$$\begin{aligned}
\hat{V}^{\epsilon_B,\epsilon_C} \approx & \hat{V}^{(\epsilon_B,\epsilon_C),(1,1)} = \hat{V}_0 + \sqrt{\epsilon_B}\hat{V}_{(1/2,0)} + \sqrt{\epsilon_C}\hat{V}_{(0,1/2)} + \epsilon_B\hat{V}_{(1,0)} + \epsilon_C\hat{V}_{(0,1)} \\
= & \hat{V}_0 - \sqrt{\epsilon_B}\tau\rho_B\sigma^S\nu_B S(1-R_B)\langle\sqrt{\lambda_B}\rangle\frac{\partial\hat{V}_0^+}{\partial S} \\
& - \sqrt{\epsilon_C}\tau\rho_C\sigma^S\nu_C S(1-R_C)\langle\sqrt{\lambda_C}\rangle\frac{\partial\hat{V}_0^+}{\partial S} \\
& + \epsilon_B(1-R_B)(\theta_B-\lambda_B)\hat{V}_0^+ + \epsilon_B\tau(1-R_B)^2\frac{\theta_B\nu_B^2}{2}\hat{V}_0^+ \\
& + \epsilon_C(1-R_C)(\theta_C-\lambda_C)\hat{V}_0^+ + \epsilon_C\tau(1-R_C)^2\frac{\theta_C\nu_C^2}{2}\hat{V}_0^+
\end{aligned} \tag{3.125}$$

where $\hat{V}_0$ is the solution of the Black-Scholes equation (3.43) taking default risk into account, with constant default intensities $\lambda_B = \theta_B$ and $\lambda_C = \theta_C$. It is useful to mention the asymptotic approach requires to numerically solve a 1D time-dependent PDE and apply an asymptotic closed-form approximation, either in the case of unilateral stochastic default intensity (2D time-dependent PDE) or in the case of bilateral stochastic default

intensities (3D time-dependent PDE). Therefore, the cost of the asymptotic approach for either in the case of unilateral stochastic default intensity or in the case of bilateral stochastic default intensities, is about the same.

Chapter 4

American XVA with stochastic default intensities

In this chapter, we investigate the formulation of pricing XVA in American derivatives if one counterparty exhibits stochastic default intensity. We also develop two different computation methods for accurately and efficiently solving the model. We only use American put option as an example, but the ideas and techniques can be extended similarly to other financial derivatives. The first computational approach is the numerical PDE approach. We extend the double-penalty iteration method in Chapter 2 and the discretization scheme in Chapter 3 to numerically approximate the numerical solution of the two-dimensional PDE. The second approach is the asymptotic approach similar to the one presented in Chapter 3, designed to work for American derivatives. The asymptotic approximation uses the solution to the 1D XVA pricing PDE and some correction terms we develop from the 2D XVA PDE without solving it. Such approximation technique is usually challenging in linear complementarity problems, because the location of the boundary to the PDE region is unknown. In [46], the authors applied singular perturbation theory and gave expansions of the price in terms of volatility, which is considered a small parameter. They use this technique to price vanilla European, American and Barrier options. In their American options asymptotic approximations, they also expand the value of the free boundary in terms of volatility, and come up with multiple correction terms added to the strike price. A similar treatment is also used in another asymptotic expansion in [25], which establishes a robust correction to the Black-Scholes American derivatives prices with stochastic volatility. The authors of [25] assume fast mean-reversion volatility and apply asymptotic expansion in terms of the mean-reversion speed. As a result, the two-dimensional free-boundary problem is asymptotically approximated by a one-dimensional American Black-Scholes problem plus the solution to

a fixed-boundary-value problem. Thus, the solution to the two-dimensional American problem is approximated by the solution of one-dimensional problem plus a correction term. At the same time, the free boundary of the two-dimensional American problem is approximated by the free-boundary of the one-dimensional American Black-Scholes problem. The authors claim that the asymptotic accuracy of the free boundary approximation is only order of half. In later work [1], they introduce correction terms and slightly improve the approximation of free boundary. In [1], the authors consider the stochastic volatility is driven by a fast mean-reversion or by other slow fluctuation factors. The idea of dealing with the free boundary in this chapter is similar to the one in [25, 1].

The contributions this part of thesis are:

- We formulate a time-dependent PDE in two spatial dimensions, with multiple nonlinear source terms. One dimension is for the underlying asset and another one is for the counterparty default intensity. One nonlinear term models the American constraint, and the other nonlinear terms model the XVA adjustments. We also develop the appropriate boundary conditions for the problem. We present a finite difference scheme for the discretization. We extend the double-penalty method in Chapter 2 to handle the different nonlinear source terms.
- We extend the asymptotic approximation we developed for European derivatives in Chapter 3 to American derivatives in XVA pricing problem. The asymptotic approximation also includes an asymptotic expansion to the free boundary. We formulate an accurate approach to approximate all terms of the free boundary expansion. We study how this approximate free boundary affects our asymptotic approximated prices.

The outline of this chapter is as follows. In Section 4.1, we present the formulation of the American type XVA pricing problem as a linear complementarity PDE. We also present the penalty form to the corresponding problem. In Section 4.2, we describe the numerical methods used for discretizations, the appropriate boundary conditions, and the double-penalty method to deal with the multiple nonlinearity. In Section 4.3, we present the asymptotic approximations to the free boundary and to the solution of the problem. The numerical experiments to compare the numerical PDE and asymptotic solutions for American put options, and study their accuracy and effectiveness are left for Chapter 5.

4.1 Formulation for American underlying asset

In this section, we show the PDE formulation of XVA pricing problem for American derivatives, with the stochastic counterparty default risk considered. The PDE model can be developed from dynamic hedging strategies, by using multiple hedging arguments, including underlying assets, the zero-coupon bonds of the two parties, risk-free zero-coupon bonds and another similar financial derivative [24]; see also Section 3.3. Similarly to the original American option PDE, the XVA pricing for American derivatives is also a linear complementarity problem (LCP) - see Section 2.3 - even when the stochastic counterparty default intensity is considered.

In this section, we will use the same notations as those in Chapter 2 and Chapter 3, we just repeat them: Let S be the underlying stock value, t the forward time, T the expiry time of the contract, σ the volatility in S, r the risk-free interest rate, γ the dividend yield of S, q the stock repo rate, r_B the seller's bond yield, r_C the counterparty's bond yield, λ_B (bank hazard rate, instantaneous probability of bank default), λ_C (counterparty hazard rate), R_B the recovery percentage on $M(=\hat{V})$ if seller defaults, R_C the recovery percentage on $M(=\hat{V})$ if counterparty defaults, r_F the seller's funding rate for borrowed cash, where $r_F = r+(1-R_B)\lambda_B$ if derivative cannot be used as collateral, and $s_F = r_F - r$. Let also $r_R = q - \gamma$.

If the backward time $\tau = T - t$ is applied, the American derivative price $\hat{V}(\tau, S, \lambda_C)$, if taking stochastic risk into account, satisfies

$$\left\{ \begin{array}{r} -\frac{\partial \hat{V}}{\partial \tau} + \mathcal{L}\hat{V} + f(\hat{V}) > 0 \\ \hat{V} - V^* = 0 \end{array} \right\} \vee \left\{ \begin{array}{r} -\frac{\partial \hat{V}}{\partial \tau} + \mathcal{L}\hat{V} + f(\hat{V}) = 0 \\ \hat{V} - V^* \geq 0 \end{array} \right\} \tag{4.1}$$

where

$$\begin{aligned} \mathcal{L}\hat{V} \equiv & \frac{1}{2}(\sigma^S)^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_C})^2 \lambda_C \frac{\partial^2 \hat{V}}{\partial \lambda_C^2} + \rho \sigma^S \sigma^{\lambda_C} S \sqrt{\lambda_C} \frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C} \\ & + r_R S \frac{\partial \hat{V}}{\partial S} + \kappa[\theta - \lambda_C] \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V}, \end{aligned} \tag{4.2}$$

$$f(\lambda_C, \hat{V}) \equiv -(s_F + (1 - R_C)\lambda_C)\hat{V}^+ - (1 - R_B)\lambda_B \hat{V}^-, \tag{4.3}$$

$$V^* \equiv \hat{V}(0, S, \lambda_C), \tag{4.4}$$

and the notation $\vee$ means "or". Problem (4.1) is a free boundary problem, with the free boundary being a surface which can be written as $S_{fb}(\tau, \lambda_C)$. The free boundary separates the domain into two regions. We refer to the region where the left part of (4.1) holds as exercise region and the region where the right part of (4.1) holds as hold region.

Without loss of generality, we consider the American put case. In this case, in the hold region in (4.1), we have $S \geq S_{fb}$, while in the exercise region, we have $S < S_{fb}$. We also have the initial condition

$$\hat{V}(0, S, \lambda_C) = (K - S)^+, \tag{4.5}$$

and the free boundary conditions

$$\hat{V}(\tau, S_{fb}, \lambda_C) = K - S_{fb}, \tag{4.6}$$

$$\frac{\partial \hat{V}}{\partial S}(\tau, S_{fb}, \lambda_C) = -1, \tag{4.7}$$

$$\frac{\partial \hat{V}}{\partial \lambda_C}(\tau, S_{fb}, \lambda_C) = 0. \tag{4.8}$$

4.1.1 Reformulation to penalty form

For one-asset American option pricing without considering default risk, Forsyth and Vetzal [22] proposed the discrete penalty method to numerically solve the LCP. A large positive penalty term is added to the Black-Scholes equation, resulting in a nonlinear PDE. In [22], a rigorous study of equivalence between the LCP and the penalized nonlinear PDE is presented. In [14] (see also Section 2.3), the double-penalty method is firstly introduced to solve the LCP arising from American type XVA with constant bilateral default risk, which has multiple nonlinear terms.

In the current multi-dimensional American XVA problem, with stochastic counterparty default risk, a similar penalty term can be added. The penalty form of (4.1) is written as

$$\begin{cases} \frac{\partial \hat{V}}{\partial \tau} = \mathcal{L}\hat{V} + f(\lambda_C, \hat{V}) + p \max(V^* - \hat{V}, 0), \\ \hat{V}(0, S, \lambda_C) = (K - S)^+, \end{cases} \tag{4.9}$$

where p is a large positive penalty factor. The penalty term forces the solution of (4.9) to approximately satisfy the obstacle condition $\hat{V} - V^* \geq 0$.

Remark 5 *In Chapter 2, we formulated the XVA valuation of American derivative, when the default intensity λ_C is given by a constant, in penalty form as*

$$\begin{aligned} \frac{\partial \hat{V}^c}{\partial \tau} &= \frac{1}{2}(\sigma^S)^2 S^2 \frac{\partial^2 \hat{V}^c}{\partial S^2} + r_R S \frac{\partial \hat{V}^c}{\partial S} - r\hat{V}^c \\ -(s_F + (1 - R_C)\lambda_C)(\hat{V}^c)^+ - (1 - R_B)\lambda_B(\hat{V}^c)^- &+ p \max\{V^* - (\hat{V}^c)^-, 0\}, \end{aligned} \tag{4.10}$$

where $\hat{V}^c$ is the adjusted American derivative price when taking constant default intensity λ_C into account. The notation here $\hat{V}^c$ is used in order to distinguish the constant default intensity American derivative price from the stochastic default intensity American derivative price. The numerical method to approximate the value of PDE (4.10) is introduced in Chapter 2.

4.2 Numerical methods

4.2.1 Discretization

The domain of the PDE (4.1), which prices the American type derivative considering stochastic default intensity is semi-infinite in the two spatial variables:

$$(\tau, S, \lambda_C) \in (0, T] \times [0, \infty) \times [0, \infty).$$

For computational purposes, the semi-infinite spatial domain of spot price S is truncated into $[0, S^{max}]$, for sufficiently large S^{max}, while the semi-infinite spatial domain of party C spot default intensity λ_C is truncated into $[0, \lambda_C^{max}]$, for sufficiently large λ_C^{max}. Then, $[0, S^{max}]$ is divided into N subintervals, with the gridpoints $S_0 = 0 < S_1 < \ldots < S_N = S^{max}$ positioned uniformly or nonuniformly, while $[0, \lambda_C^{max}]$ is divided into M subintervals, with the gridpoints $(\lambda_C)_0 = 0 < (\lambda_C)_1 < \ldots < (\lambda_C)_M = \lambda_C^{max}$ positioned uniformly or nonuniformly. Standard second-order centered finite differences are used for the space discretization of (4.1) except at the boundary points. The details of the boundary conditions and their discretization, as well as the handling of the nonlinear term and penalty terms are discussed in the following two subsections.

For the time-stepping, we employ the ϑ-method [1], which, for $\vartheta = \frac{1}{2}$ and $\vartheta = 1$ becomes the Crank-Nicolson (CN) and Backward Euler (BE) methods, respectively. We also use Rannacher smoothing, which consists of first applying few BE timesteps, then applying CN timestepping. Let $\tau_j, j = 0, \ldots, N_t$, be the timesteps at which the solution is computed, with $\tau_0 = 0 < \tau_1 < \cdots < \tau_{N_t} = T$, and let $\Delta\tau^j = \tau_j - \tau_{j-1}$ be the jth time stepsize. If uniform timesteps are used, then $\Delta\tau = T/N_t$, and Rannacher smoothing first applies four BE timesteps with stepsize $\Delta\tau/2$, then switches to CN with stepsize $\Delta\tau$ for the remaining timesteps, resulting in a total of $N_t + 2$ timesteps.

[1] Note that the ϑ notation for the time-stepping method is different from the $\theta(t)$ notation in the CIR model

4.2.2 Boundary conditions

The grid-based numerical methods need boundary conditions on the truncated spatial domain $[0, S^{max}] \times [0, \lambda_C^{max}]$ to approximate the solution to PDE (4.9). We follow the ideas of boundary conditions in Section 3.2.2, and set up boundary conditions as follows:

- On the $S = 0$ boundary, i.e. on $\{(S, \lambda_C) \in \{S = 0\} \times [0, \lambda_C^{max}]\}$, substitute $S = 0$ into (2.41). This results in a one-dimensional time-dependent PDE,

$$\frac{\partial \hat{V}}{\partial \tau} = \frac{1}{2}(\sigma^{\lambda_C})^2 \lambda_C \frac{\partial^2 \hat{V}}{\partial \lambda_C^2} + \kappa[\theta - \lambda_C]\frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}) + p\max(V^* - \hat{V}, 0), \tag{4.11}$$

  which is numerically solved and its computed solution used as Dirichlet boundary condition for (4.9).

- On the $S = S^{max}$ boundary, i.e. on $\{(S, \lambda_C) \in \{S = S^{max}\} \times [0, \lambda_C^{max}]\}$, we impose the linear boundary condition $\frac{\partial^2 \hat{V}}{\partial S^2} = 0$, substitute this into the PDE (2.41), and get

$$\begin{aligned} \frac{\partial \hat{V}}{\partial \tau} &= \frac{1}{2}(\sigma^{\lambda_C})^2 \lambda_C \frac{\partial^2 \hat{V}}{\partial \lambda_C^2} + \rho\sigma^S \sigma^{\lambda_C} S\sqrt{\lambda_C}\frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C} \\ &+ r_R S\frac{\partial \hat{V}}{\partial S} + \kappa[\theta - \lambda_C]\frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}) + p\max(V^* - \hat{V}, 0). \end{aligned} \tag{4.12}$$

- On the $\lambda_C = 0$ boundary, i.e. $\{(S, \lambda_C) \in (0, S^{max}) \times \{\lambda_C = 0\}\}$, substitute $\lambda_C = 0$ into (2.41), and get

$$\frac{\partial \hat{V}}{\partial \tau} = \frac{1}{2}(\sigma^S)^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + r_R S\frac{\partial \hat{V}}{\partial S} + \kappa\theta\frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}) + p\max(V^* - \hat{V}, 0). \tag{4.13}$$

- On the $\lambda_C = \lambda_C^{max}$ boundary, i.e. $\{(S, \lambda_C) \in (0, S^{max}) \times \{\lambda_C = \lambda_C^{max}\}\}$, we impose the linear boundary condition $\frac{\partial^2 \hat{V}}{\partial \lambda_C^2} = 0$, substitute this into PDE (2.41), and get

$$\begin{aligned} \frac{\partial \hat{V}}{\partial \tau} &= \frac{1}{2}(\sigma^S)^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + \rho\sigma^S \sigma^{\lambda_C} S\sqrt{\lambda_C}\frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C} \\ &+ r_R S\frac{\partial \hat{V}}{\partial S} + \kappa[\theta - \lambda_C]\frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}) + p\max(V^* - \hat{V}, 0). \end{aligned} \tag{4.14}$$

On the corners, we choose compatible boundary conditions, the details of which are omitted for brevity. Note that, these conditions differ from the respective European

boundary conditions only by the penalty term $p\max(V^* - \hat{V}, 0)$. In Section 3.2.2, there is more detailed discussion on the European boundary conditions including corners as well as the discretization of the conditions.

4.2.3 Double-penalty iteration for multi-dimensional problem

We now discuss the treatment of the penalty term $p\max(V^* - \hat{V}, 0)$ and the nonlinear term $f(\hat{V})$. Let $\hat{v}^j, j = 0, \ldots, N_t$, be the computed solution vector arising from the approximate values of $\hat{V}$ at the spatial gridpoints at time τ_j, and $\hat{v}^0$ be the initial condition vector. Also let v^* be the vector of values of pay-off V^* at the spatial gridpoints. Since we use an iteration method to handle the nonlinearity, let $\hat{v}^{j,k}, k = 0, \ldots, maxit$, denote the computed solution vector at iteration k of timestep j, with $maxit$ the maximum number of iterations allowed per timestep. Let $f(\hat{v})$ denote the vector arising from evaluating f at the components of $\hat{v}$. This means that $(f(\hat{v}))_i = f((\lambda_C)_{i_2}, \hat{v}_i)$, where $i_2 = i - \lfloor \frac{i}{M+1} \rfloor (M+1)$. Let also A be the matrix arising from the space discretization of $\mathcal{L}\hat{V}$, and $\mathbb{I}$ be the identity matrix of compatible order. For simplicity, we assume the spatial gridpoints remain the same at all timesteps.

When computing the numerical solution $\hat{v}^j$ at step j, the penalty term $p\max(V^* - \hat{V}, 0)$ (arising from the American feature) is discretized as $P_A(\hat{v}^j)(v^* - \hat{v}^j)$, where $P_A(\hat{v}^j)$ is a diagonal matrix defined by

$$[P_A(\hat{v}^j)]_{i,i} \equiv \begin{cases} p & \text{if } (\hat{v}^j)_i < (v^*)_i, \\ 0 & \text{otherwise.} \end{cases} \tag{4.15}$$

The nonlinear term $f(\hat{V})$ (arising from the XVA) is discretized as $P_X(\hat{v}^j)(\hat{v}^j)$, where $P_X(\hat{v}^j)$ is also a diagonal matrix defined by

$$[P_X(\hat{v}^j)]_{i,i} \equiv \begin{cases} -\lambda_B(1 - R_B) & \text{if } (\hat{v}^j)_i < 0, \\ -(\lambda_C)_{i_2}(1 - R_C) - s_F & \text{if } (\hat{v}^j)_i \geq 0. \end{cases} \tag{4.16}$$

Therefore, to compute $\hat{v}^j$, given $\hat{v}^{j-1}$, the following system of algebraic equations needs to be solved:

$$\begin{aligned} &[\mathbb{I} - \theta\Delta\tau^j(A + P_X(\hat{v}^j))]\hat{v}^j + P_A(\hat{v}^j)\hat{v}^j \\ &= (\mathbb{I} + (1-\theta)\Delta\tau^j A)\hat{v}^{j-1} + (1-\theta)\Delta\tau^j P_X(\hat{v}^{j-1})\hat{v}^{j-1} + P_A(\hat{v}^j)v^*. \end{aligned} \tag{4.17}$$

Note that there are two sources of nonlinearity, namely $P_A(\hat{v}^j)$ and $P_X(\hat{v}^j)$, with respect

to $\hat{v}^j$. The iteration method to value XVA and the penalty iteration method to value American option [22] are combined into one iteration method, which can be viewed as generalized Newton's method, to solve (4.17). In the construction of P_A, the large penalty factor p is chosen as

$$p = \frac{1}{tol}$$

where *tol* is the tolerance for the stopping criterion of the iteration method.

If we view the nonlinear term $f(\hat{V})$ or its discrete form $P_X(\hat{v}^j)(\hat{v}^j)$ as a second penalty term, this generalized Newton's iteration method to solve (4.17) at each timestep can be called *discrete double-penalty iteration* for American XVA pricing.

Algorithm 5 Discrete double-penalty iteration for (4.9) at step j, with θ-timestepping

Require: Solve $[\mathbb{I} - \theta\Delta\tau^j(A + P_X(\hat{v}^j))]\hat{v}^j + P_A(\hat{v}^j)\hat{v}^j = g^j + P_A(\hat{v}^j)v^*$
where $g^j = (\mathbb{I} + (1-\theta)\Delta\tau^j A)\hat{v}^{j-1} + (1-\theta)\Delta\tau^j P_X(\hat{v}^{j-1})\hat{v}^{j-1}$.
Initialize $\hat{v}^{j,0} = \hat{v}^{j-1}$, $P_A^0 = P_A(\hat{v}^{j,0})$ and $P_X^0 = P_X(\hat{v}^{j,0})$
for $k = 1, \ldots, maxit$ **do**
  Solve $[\mathbb{I} - \theta\Delta\tau^j(A + P_X^{k-1}) + P_A^{k-1}]\hat{v}^{j,k} = g^j + P_A^{k-1}v^*$
  Compute $P_A^k = P_A(\hat{v}^{j,k})$ by (4.15) and $P_X^k = P_X(\hat{v}^{j,k})$ by (4.16)
  if stopping criterion satisfied **then**
    Break
  end if
end for
Set $\hat{v}^j = \hat{v}^{j,k}$

The stopping criterion in Algorithm 5, is

$$\big[(P_A^k = P_A^{k-1}) \text{ and } (P_X^k = P_X^{k-1})\big] \text{ or } \big[\max_i \frac{|\hat{v}_i^{j,k} - \hat{v}_i^{j,k-1}|}{\max(1, |\hat{v}_i^{j,k}|)} \leq tol\big]. \tag{4.18}$$

Since the matrix solved at each iteration is adjusted by only two diagonal matrices, the sparsity structure of the matrix remains the same. Since $[P_X]_{i,i}$ is negative, the diagonal of A is enhanced. Also since the penalty parameter p is positive, the matrix P_A enhances the diagonal of $\mathbb{I} - \theta\Delta\tau^j(A + P_X^{k-1})$.

4.2.4 Numerical approximation of free boundary of multi-dimensional American XVA problem

In this subsection, we show how to numerically approximate the free boundary of the multi-dimensional American XVA problem on a specific desired λ_C point, assuming we already have numerical solutions on the grid points, $(S_i, (\lambda_C)_j), i = 1, \ldots, N, j = 0, \ldots, M$.

The naive way is that, for a fixed λ_C, we choose the free boundary point as the point S_k to the left of the first grid point S_{k+1}, where the numerical solution is $2tol = \frac{2}{p}$ or more above the payoff function. Following this naive way, S_k maybe different for each λ_C, but it is always chosen among the grid points S_i, $i = 1, \ldots, N$.

The above approximation is improved in accuracy following [45]. Along each λ_C line, for which we want to approximate the free boundary, we construct an interpolant $V_\partial(S)$ of the first (with respect to S) derivative values and find for which S it satisfies $V_\partial(S) = -1$. The interpolant is constructed using three points $\{(S_i, \frac{\partial \hat{V}}{\partial S}(S_i, \lambda_C)), i = k+2, k+3, k+4\}$ away from the naive free boundary S_k and in the PDE region (hold region), as the error on S_k is contaminated with errors from the early exercise region that are not smooth, and the finite difference approximations involving S_k are inaccurate. See Algorithm 6 for details. To solve the nonlinear equation $V_\partial(S) + 1 = 0$ we can use the standard quadratic root formula or Newton's method with initial guess S_{k+1}. An alternative would be to construct a quadratic interpolant $V_I(S)$ of $\hat{V}(S)$ using $\{(S_i, \hat{V}(S_i, \lambda_C), i = k+2, k+3, k+4\}$, and to solve $V_I(S) - V^*(S) = 0$ for S. This nonlinear equation has a double root, so Newton's method is expected to be slow. But the quadratic formula could still work well. In our code, we used the MATLAB function `fsolve` on $V_\partial(S) + 1 = 0$ with initial guess S_{k+1}, and tolerance 10^{-9}.

In the following algorithm, we assume we have already computed numerical approximations to $\hat{V}(S_i, (\lambda_C)_j)$, $i = 1, 2, \ldots, N, j = 0, 1, \ldots, M$. For simplicity, we denote the approximations by $\hat{V}(S_i, (\lambda_C)_j)$ as well.

Algorithm 6 Approximation of the free boundary for a given λ_C.

1: If λ_C is not a grid point, compute approximations $\hat{\hat{V}}(S_i, \lambda_C)$ to $\hat{V}(S_i, \lambda_C), i = 1, 2, \ldots, N$, by cubic spline interpolation on the values $\hat{V}(S_i, (\lambda_C)_j)$, $i = 1, 2, \ldots, N, j = 0, 1, \ldots, M$. Let $\hat{V}(S_i) = \hat{\hat{V}}(S_i, \lambda_C)$, $i = 1, 2, \ldots, N$.
2: If λ_C is grid point $(\lambda_C)_j$, let $\hat{V}(S_i) = \hat{V}(S_i, (\lambda_C)_j)$, $i = 1, 2, \ldots, N$.
3: Find the leftmost point S_{k+1}, such that $\hat{V}(S_{k+1}) - V^*(S_{k+1}) \geq 2tol$.
4: Compute finite difference approximations $\frac{\partial \hat{V}}{\partial S}(S_i)$, to $\frac{\partial \hat{V}}{\partial S}(S_i, \lambda_C), i = k+2, k+3, k+4$.
5: Construct the quadratic interpolant $V_\partial(S)$ of $\frac{\partial \hat{V}(S)}{\partial S}$ using $\{(S_i, \frac{\partial \hat{V}}{\partial S}(S_i)), i = k+2, k+3, k+4\}$.
6: Solve $V_\partial(S) + 1 = 0$ for S with initial guess S_{k+1}, to get the free boundary $S_{fb}^{\lambda_C}$ for the given λ_C.

4.3 Asymptotic solution and asymptotic approximation for the free boundary

Numerically solving the time-dependent multi-dimensional PDE, especially including multiple nonlinear source terms, is computationally expensive. An asymptotic approximation can ease this problem [24, 40] (see also Section 3.3). An asymptotic approximation requires the solution of a lower-dimensional PDE and applies some correction terms to it by a closed-form formula. However, asymptotic approximation for path-dependent derivatives is more complicated due to the lack of explicit formula for the corrections terms. For example, for American put, this problem is more involved due to the singularities at the free boundaries, because, usually at free boundary points, the solutions are only $\mathbb{C}^1$. In [25, 1], the authors applied asymptotic expansion to both the value function and the free boundary and find an asymptotic correction to the free boundary, which can transform the LCP into a fixed point problem.

For convenience, we repeat the stochastic differential equation that the CIR process for λ_C follows:

$$d\lambda_C(t) = \kappa[\theta - \lambda_C(t)]dt + \sigma^{\lambda_C}\sqrt{\lambda_C(t)}dW^{\lambda_C}(t). \tag{4.19}$$

Following [26, 40] (see also Section 3.3), we assume $\kappa = 1/\epsilon$, where $\epsilon > 0$ is small. We also keep the variance ν^2 of the λ_C process invariant distribution constant, thus we scale σ^{λ_C} as $\sigma^{\lambda_C} = \frac{\nu}{\sqrt{\epsilon}}$. Hence we can rewrite the equation in the hold region of (2.37) as

$$(\frac{1}{\epsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 + \mathcal{L}_2)\hat{V} = 0, \tag{4.20}$$

where

$$\mathcal{L}_0 \equiv \frac{1}{2}\nu^2\lambda_C\frac{\partial^2}{\partial\lambda_C^2} + (\theta - \lambda_C)\frac{\partial}{\partial\lambda_C}, \tag{4.21}$$

$$\mathcal{L}_1 \equiv \rho\sigma^S\nu S\sqrt{\lambda_C}\frac{\partial^2}{\partial S\partial\lambda_C}, \tag{4.22}$$

$$\mathcal{L}_2 \equiv (-\frac{\partial}{\partial\tau}) + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2}{\partial S^2} + r_RS\frac{\partial}{\partial S} - r\mathcal{I} + f(\lambda_C, \hat{V}), \tag{4.23}$$

with $\mathcal{I}$ being the identity operator.

We apply asymptotic expansion to both $\hat{V}$ and S_{fb} with respect to ϵ,

$$\hat{V} \equiv \hat{V}^{\epsilon} = \hat{V}_0 + \sqrt{\epsilon}\hat{V}_{1/2}{}^2 + \epsilon\hat{V}_1 + \epsilon\sqrt{\epsilon}\hat{V}_{3/2} + \dots, \tag{4.24}$$

$$S_{fb} \equiv S_{fb}^{\epsilon} = S_{f0} + \sqrt{\epsilon}S_{f1/2} + \dots, \tag{4.25}$$

which converge to the respective exact solutions if $\epsilon \to 0$. In this work, we will use the first three terms of (4.24) to estimate $\hat{V}$ such that $\hat{V} \approx \hat{V}^{\epsilon,1} = \hat{V}_0 + \sqrt{\epsilon}\hat{V}_{1/2} + \epsilon\hat{V}_1$ and the first two terms of (4.25) to estimate the free boundary S_{fb} such that $S_{fb} \approx S_{fb}^{\epsilon,1/2} = S_{f0} + \sqrt{\epsilon}S_{f1/2}$. Let $\langle\cdot\rangle$ denote expectation with respect to the invariant distribution of λ_C.

The expansion of the partial differential equation (4.20) (hold region) gives

$$\mathcal{O}(\frac{1}{\epsilon}) : \mathcal{L}_0\hat{V}_0 = 0 \tag{4.26}$$

$$\mathcal{O}(\frac{1}{\sqrt{\epsilon}}) : \mathcal{L}_0\hat{V}_{1/2} + \mathcal{L}_1\hat{V}_0 = 0 \tag{4.27}$$

$$\mathcal{O}(1) : \mathcal{L}_0\hat{V}_1 + \mathcal{L}_1\hat{V}_{1/2} + \mathcal{L}_2\hat{V}_0 = 0 \tag{4.28}$$

$$\mathcal{O}(\sqrt{\epsilon}) : \mathcal{L}_0\hat{V}_{3/2} + \mathcal{L}_1\hat{V}_1 + \mathcal{L}_2\hat{V}_{1/2} = 0. \tag{4.29}$$

We can also expand the free boundary conditions (4.6) (4.7) and (4.8), only keeping terms up to ϵ, as

$$\begin{aligned}\hat{V}_0(\tau, S_{f0}, \lambda_C) + \sqrt{\epsilon}(S_{f1/2}\frac{\partial\hat{V}_0}{\partial S}(\tau, S_{f0}, \lambda_C) + \hat{V}_{1/2}(\tau, S_{f0}, \lambda_C))\\ +\epsilon(S_{f1}\frac{\partial\hat{V}_0}{\partial S}(\tau, S_{f0}, \lambda_C) + S_{f1/2}\frac{\partial\hat{V}_{1/2}}{\partial S}(\tau, S_{f0}, \lambda_C) + \hat{V}_1(\tau, S_{f0}, \lambda_C))\\ = K - S_{f0} - \sqrt{\epsilon}S_{f1/2} - \epsilon S_{f1},\end{aligned} \tag{4.30}$$

$$\begin{aligned}\frac{\partial\hat{V}_0}{\partial S}(\tau, S_{f0}, \lambda_C) + \sqrt{\epsilon}(S_{f1/2}\frac{\partial^2\hat{V}_0}{\partial S^2}(\tau, S_{f0}, \lambda_C) + \frac{\partial\hat{V}_{1/2}}{\partial S}(\tau, S_{f0}, \lambda_C))\\ +\epsilon(S_{f1}\frac{\partial^2\hat{V}_0}{\partial S^2}(\tau, S_{f0}, \lambda_C) + S_{f1/2}\frac{\partial^2\hat{V}_{1/2}}{\partial S^2}(\tau, S_{f0}, \lambda_C) + \frac{\partial\hat{V}_1}{\partial S}(\tau, S_{f0}, \lambda_C)) = -1,\end{aligned} \tag{4.31}$$

$$\begin{aligned}\frac{\partial\hat{V}_0}{\partial\lambda_C}(\tau, S_{f0}, \lambda_C) + \sqrt{\epsilon}(S_{f1/2}\frac{\partial^2\hat{V}_0}{\partial S\partial\lambda_C}(\tau, S_{f0}, \lambda_C) + \frac{\partial\hat{V}_{1/2}}{\partial\lambda_C}(\tau, S_{f0}, \lambda_C))\\ +\epsilon(S_{f1}\frac{\partial^2\hat{V}_0}{\partial S\partial\lambda_C}(\tau, S_{f0}, \lambda_C) + S_{f1/2}\frac{\partial^2\hat{V}_{1/2}}{\partial S\partial\lambda_C}(\tau, S_{f0}, \lambda_C) + \frac{\partial\hat{V}_1}{\partial\lambda_C}(\tau, S_{f0}, \lambda_C)) = 0.\end{aligned} \tag{4.32}$$

[2]In the notations $\hat{V}_{1/2}$, $S_{1/2}$ and $\hat{V}_{3/2}$, the subscripts are consistent with the powers of the associated ϵ coefficients.

From Equations (4.30) - (4.32), we have

$$\hat{V}_0(\tau, S_{f0}, \lambda_C) = K - S_{f0}, \tag{4.33}$$

$$\frac{\partial \hat{V}_0}{\partial S}(\tau, S_{f0}, \lambda_C) = -1, \tag{4.34}$$

$$\frac{\partial \hat{V}_0}{\partial \lambda_C}(\tau, S_{f0}, \lambda_C) = 0, \tag{4.35}$$

$$S_{f1/2}\frac{\partial^2 \hat{V}_0}{\partial S^2}(\tau, S_{f0}, \lambda_C) + \frac{\partial \hat{V}_{1/2}}{\partial S}(\tau, S_{f0}, \lambda_C) = 0, \tag{4.36}$$

$$S_{f1}\frac{\partial^2 \hat{V}_0}{\partial S^2}(\tau, S_{f0}, \lambda_C) + S_{f1/2}\frac{\partial^2 \hat{V}_{1/2}}{\partial S^2}(\tau, S_{f0}, \lambda_C) + \frac{\partial \hat{V}_1}{\partial S}(\tau, S_{f0}, \lambda_C) = 0. \tag{4.37}$$

In the exercise region, the condition $\hat{V} = (K-S)^+$ suggests that $\hat{V}_0 = (K-S)^+$, $\hat{V}_{1/2} = 0$ and $\hat{V}_1 = 0$.

Equation (4.26) implies that $\hat{V}_0$ is independent of λ_C, i.e. $\hat{V}_0 = \hat{V}_0(\tau, S)$, in the hold region. In the exercise region, it's obvious that $\hat{V}_0$ is also independent of λ_C. Hence, $\mathcal{L}_1\hat{V}_0 = 0$ in either region.

Equation (4.27) results in $\mathcal{L}_0\hat{V}_{1/2} = 0$, which implies $\hat{V}_{1/2}$ is independent of λ_C as well, i.e. $\hat{V}_{1/2} = \hat{V}_{1/2}(\tau, S)$, in the hold region. In the exercise region, it's also obvious that $\hat{V}_{1/2}$ is independent of λ_C.

In the hold region, coming to the $\mathcal{O}(1)$ term, Equation (4.28), given $\mathcal{L}_1\hat{V}_{1/2} = 0$, reduces to $\mathcal{L}_0\hat{V}_1 + \mathcal{L}_2\hat{V}_0 = 0$. This is a Poisson equation with respect to the operator $\mathcal{L}_0$ in the variable λ_C, which implies the centering condition

$$\langle \mathcal{L}_2\hat{V}_0 \rangle = 0. \tag{4.38}$$

The centering condition (4.38) becomes PDE $\langle \mathcal{L}_2 \rangle \hat{V}_0 = 0$ in the hold region. In addition to this PDE, Equations (4.33), (4.34) and (4.35) show us that S_{f0} is the free boundary of this PDE. In the exercise region, we have $\hat{V}_0 = (K-S)^+$. Therefore, $\hat{V}_0$ is the solution to the one-dimensional American Black-Scholes XVA equation taking default risk into account, with constant default intensity being the mean of the CIR model, and with the terminal condition $\hat{V}_0(0, S) = (K-S)^+$. To compute the solution to this problem, no analytical formula is available. An advanced numerical method, such a penalty method, can be used to approximate this one-dimensional parabolic linear-complementarity PDE.

Combining the fact that $\hat{V}_{1/2}$ is independent of λ_C with the $\mathcal{O}(1)$ term in (4.28), we

have

$$\mathcal{L}_0\hat{V}_1 + \mathcal{L}_2\hat{V}_0 = 0 \Longrightarrow \mathcal{L}_0\hat{V}_1 = -\mathcal{L}_2\hat{V}_0. \tag{4.39}$$

For equation (4.29), the solvability of this Poisson equation requires the condition that

$$\langle\mathcal{L}_1\hat{V}_1 + \mathcal{L}_2\hat{V}_{1/2}\rangle = 0 \Longrightarrow \langle\mathcal{L}_2\rangle\hat{V}_{1/2} = -\langle\mathcal{L}_1\hat{V}_1\rangle. \tag{4.40}$$

In Chapter 3.3, solutions to (4.39) and (4.40) are offered in terms of $\hat{V}_0$ as

$$\hat{V}_{1/2}(\tau, S) = -\tau\rho\sigma^S\nu S(1-R_C)\langle\sqrt{\lambda_C}\rangle\frac{\partial\hat{V}_0^+}{\partial S}, \tag{4.41}$$

$$\hat{V}_1(\tau, S, \lambda_C) = (1-R_C)(\theta-\lambda_C)\hat{V}_0^+ + \tau(1-R_C)^2\frac{\theta\nu^2}{2}\hat{V}_0^+, \tag{4.42}$$

where $\hat{V}_0$ is solution to American put option including XVA in the hold region PDE. Hence, in the hold region, the asymptotic approximation to $\hat{V}$ is

$$\hat{V}^{\epsilon,1} = \hat{V}_0 - \sqrt{\epsilon}\tau\rho\sigma^S\nu S(1-R_C)\langle\sqrt{\lambda_C}\rangle\frac{\partial\hat{V}_0^+}{\partial S} + \epsilon(1-R_C)(\theta-\lambda_C)\hat{V}_0^+ + \epsilon\tau(1-R_C)^2\frac{\theta\nu^2}{2}\hat{V}_0^+. \tag{4.43}$$

From (4.36) and (4.37), we also have correction to the free boundary

$$S_{f1/2} = \left.\frac{-\frac{\partial\hat{V}_{1/2}}{\partial S}}{\frac{\partial^2\hat{V}_0}{\partial S^2}}\right|_{S=S_{f0}} = \left.\frac{\tau\rho\sigma^S\nu(1-R_C)\langle\sqrt{\lambda_C}\rangle(\frac{\partial\hat{V}_0^+}{\partial S} + S\frac{\partial^2\hat{V}_0^+}{\partial S^2})}{\frac{\partial^2\hat{V}_0}{\partial S^2}}\right|_{S=S_{f0}} \tag{4.44}$$

where S_{f0} is the free boundary of the LCP, representing the one-dimensional American Black-Scholes XVA equation taking default risk into account, with constant default intensity being the mean of the CIR model, and with the terminal condition $\hat{V}_0(0, S) = (K-S)^+$. From the view of computation, given $\hat{V}_0$, S_{f0} is computed by an algorithm similar to Algorithm 6, but adjusted to work on the one-dimensional problem; see Algorithm 7.

Hence the asymptotic approximation to free boundary S_{fb} is

$$S_{fb}^{\epsilon,1/2} = S_{f0} + \sqrt{\epsilon}S_{f1/2} \tag{4.45}$$

where $S_{f1/2}$ is defined in (4.44). Note that, for Equations (4.41), (4.43) and (4.44), we

define

$$\frac{\partial \hat{V}_0^+}{\partial S} \equiv \begin{cases} \frac{\partial \hat{V}_0}{\partial S} & \hat{V}_0 > 0 \\ 0 & \hat{V}_0 \leq 0 \end{cases} \quad \text{and} \quad \frac{\partial^2 \hat{V}_0}{\partial S^2} \equiv \begin{cases} \frac{\partial^2 \hat{V}_0}{\partial S^2} & \hat{V}_0 > 0 \\ 0 & \hat{V}_0 \leq 0 \end{cases}, \tag{4.46}$$

and, in Equations (4.44) and (4.43), the partial derivatives at $S = S_{f0}$ are taken to be one-sided derivatives into the hold region; see Algorithm 7.

In the following, we describe the algorithm we apply for the approximation of S_{f0} and of the derivative values $\frac{\partial \hat{V}_0^+}{\partial S}(S_{f0})$ and $\frac{\partial^2 \hat{V}_0^+}{\partial S^2}(S_{f0})$, that are needed in (4.44) and (4.45).

In Algorithm 7, we assume we have already computed numerical approximations to $\hat{V}_0(S_i)$, $i = 1, 2, \ldots, N$. For simplicity, we denote the approximations by $\hat{V}_0(S_i)$ as well.

Algorithm 7 Approximation of S_{f0}, $\frac{\partial^2 \hat{V}_0^+}{\partial S^2}(S_{f0})$ and $\frac{\partial^2 \hat{V}_0}{\partial S^2}(S_{f0})$.

1: Find the leftmost point S_{k+1}, such that $\hat{V}_0(S_{k+1}) - V^*(S_{k+1}) \geq 2tol$.
2: Compute finite difference approximations to $\frac{\partial \hat{V}_0}{\partial S}(S_i), i = k+2, k+3, k+4$.
3: Construct the quadratic interpolant $V_{\partial,0}(S)$ of $\frac{\partial \hat{V}_0(S)}{\partial S}$ using $\{(S_i, \frac{\partial \hat{V}_0}{\partial S}(S_i)), i = k+2, k+3, k+4\}$.
4: Solve $V_{\partial,0}(S) + 1 = 0$ for S with initial guess S_{k+1}, to get the free boundary S_{f0}.
5: Evaluate $V_{\partial,0}(S_{f0})$ and set $\frac{\partial \hat{V}_0^+}{\partial S}(S_{f0}) = V_{\partial,0}(S_{f0})$, or, equivalently, set $\frac{\partial \hat{V}_0^+}{\partial S}(S_{f0}) = -1$.
6: Compute finite difference approximations to $\frac{\partial^2 \hat{V}_0}{\partial S^2}(S_i), i = k+2, k+3, k+4$.
7: Construct the quadratic interpolant $V_{\partial\partial,0}(S)$ of $\frac{\partial^2 \hat{V}_0(S)}{\partial S^2}$ using $\{(S_i, \frac{\partial^2 \hat{V}_0}{\partial S^2}(S_i)), i = k+2, k+3, k+4\}$.
8: Evaluate $V_{\partial\partial,0}(S_{f0})$ and set $\frac{\partial^2 \hat{V}_0^+}{\partial S^2}(S_{f0}) = V_{\partial\partial,0}(S_{f0})$.

It is important to note that the interpolants $V_{\partial,0}(S)$ and $V_{\partial\partial,0}(S)$ are constructed using values into the hold region and away from the penalty region, and that the approximations to $\frac{\partial \hat{V}_0^+}{\partial S}(S_{f0})$ and $\frac{\partial^2 \hat{V}_0^+}{\partial S^2}(S_{f0})$ are constructed using extrapolation of the interpolants.

To conclude, our asymptotic approximation to American put option pricing including XVA, with stochastic default intensity, is summarized in Algorithm 8.

Algorithm 8 Asymptotic approximation to American put option pricing including XVA

1: Compute $\hat{V}_0$, by numerically solving the 1D American put option XVA PDE (4.10) with $\lambda_C = \theta$, using the double-penalty method in Chapter 2.
2: Compute the free boundary S_{f0} of 1D American put option XVA PDE, using $\hat{V}_0$.
3: Compute correction terms $\hat{V}_{1/2}$ and $\hat{V}_1$ by Equations (4.41) and (4.42).
4: Compute correction term, $S_{f1/2}$, to free boundary, by Equation (4.44).
5: Obtain the free boundary approximation $S_{fb}^{\epsilon,1/2} = S_{f0} + \sqrt{\epsilon} S_{f1/2}$ as Equation (4.45).
6: **if** Spot price $S \geq S_{fb}^{\epsilon,1/2}$ **then**
7: $\quad \hat{V} \approx \hat{V}^{\epsilon,1} = \hat{V}_0 + \sqrt{\epsilon}\hat{V}_{1/2} + \epsilon \hat{V}_1$, as Equation (4.43).
8: **else if** Spot price $S < S_{fb}^{\epsilon,1/2}$ **then**
9: $\quad \hat{V} \approx V^*$, which is pay-off function.
10: **end if**

Remark 6 *Similar as [25], the basic strategy for constructing an asymptotic approximation is to asymptotically expand both the value $\hat{V}$ and the free boundary S_{fb} in terms of ϵ, and obtain an approximated free boundary. Then, using the approximated free boundary, we divide the spot price region into the hold and exercise region, and apply a different formula for the price at each region. However, the approximated free boundary is $O(\epsilon)$ from the true free boundary S_{fb}. Hence, when the spot price S is close to the exercise boundary, the contract might move to the exercise region, in which case, the derivatives do not exist long enough for the mean-reverting effects of fast mean-reverting stochastic counterparty default intensity. In this case, the asymptotic approximation is not expected to be accurate. This technique is more effective when it is applied to approximate the value away from the approximated free boundary $S_{fb}^{\epsilon,1/2}$.*

Chapter 5

Numerical experiments

5.1 Numerical results of XVA with constant default intensities

We present numerical results from applying Algorithms 1 and 2 for the pricing of XVA in European Put, Call and Long Forward, as well as Algorithm 3 for the pricing of American Put, Call and Long Forward including XVA.

5.1.1 Examples of XVA in European derivatives

Table 5.1 presents the values of parameters used in the experiments in this section. These parameters have the same values as those used in the examples in [3, 11], except the S_{max} value, which, in [3, 11], is $4K$. A discussion on localization is found later in the section.

We discretize the spatial domain into N subintervals, so that the (nonuniform) gridpoints are concentrated around the strike K. We discuss later how we generate nonuniform gridpoints and what the effect of nonuniform grids is compared to uniform ones. The spatial derivatives are discretized by standard second-order centered differences, except the first derivative in the far-side condition (2.6), which is discretized by backward differences. It is worth mentioning that we experimented with the alternative implementation of the far-side condition, which gives rise to Dirichlet conditions (2.9) or (2.10), and the numerical results obtained were identical across all points (within about 7 digits of precision) with those obtained by discretizing (2.6). The number of timesteps is denoted by N_t, and $\Delta\tau = T/N_t$. Note that Rannacher smoothing is not needed (and therefore not applied) for the numerical solution of (1.33), since the initial condition is smooth.

Parameter	Value
Domain of S	$[0, 12K]$
Strike Price, K	15
Time to maturity, T	5
Volatility, σ	0.25
Repo rate minus dividend, $q - \gamma$	0.015
Interest rate, r	0.03
Default intensity of party B, λ_B	0.02
Default intensity of party C, λ_C	0.05
Recovery rate of party B, R_B	0.4
Recovery rate of party C, R_C	0.4
Funding spread, s_F	$(1 - R_B)\lambda_B$

Table 5.1: Model parameters for bilateral XVA in European derivatives with constant default intensities.

In all tables below, "iter tot 1" and "iter avg 1" are total and average (per timestep) number of iterations for Algorithm 1, and "iter tot 2" and "iter avg 2" are total and average (per timestep) number of iterations for Algorithm 2. The tolerance for both iteration schemes was set to 10^{-7}.

Call and put options

We present results from pricing the XVA of European options with the parameter settings in Table 5.1. For European call and put options, we have an exact solution formula as given in [11]

$$U(\tau, S) = -(1 - \exp(-((1 - R_B)\lambda_B + (1 - R_C)\lambda_C)\tau))V(\tau, S), \tag{5.1}$$

where the $V(\tau, S)$ can be calculated in closed-form. Thus we can calculate exact errors and the associated convergence orders.

N	N_t	error in U	order	iter tot 1	iter avg 1	iter tot 2	iter avg 2
50	100	1.41e-03	–	300	3.00	119	1.19
100	200	3.54e-04	2.00	600	3.00	242	1.21
200	400	8.86e-05	2.00	1200	3.00	415	1.04
400	800	2.21e-05	2.00	2400	3.00	814	1.02
800	1600	5.54e-06	2.00	3934	2.46	1614	1.01

Table 5.2: Results from solving (2.2) for a European Put using Algorithms 1 and 2 with the parameters in Table 5.1. Nonuniform grids are used.

N	N_t	error in U	order	iter tot 1	iter avg 1	iter tot 2	iter avg 2
50	100	1.41e-03	–	400	4.00	105	1.05
100	200	3.54e-04	2.00	600	3.00	214	1.07
200	400	8.86e-05	2.00	1200	3.00	430	1.08
400	800	2.22e-05	2.00	2400	3.00	842	1.05
800	1600	5.54e-06	2.00	4800	3.00	1635	1.02

Table 5.3: Results from solving (2.2) for a European Call using Algorithms 1 and 2 with the parameters in Table 5.1. Nonuniform grids are used.

In Tables 5.2 and 5.3, we report, for several discretization sizes, the discrete L^∞-norm error over all gridpoints at time $\tau = T$ $(t = 0)$, the associated convergence orders, and the total and average per timestep number of iterations. We notice that the average number of penalty-like iterations is very close to 1, irrespectively of the grid size, while the fixed-point iteration algorithm needs 2 to 3 average number of iterations to reach the same tolerance, where, again, the number of iteration seems independent of grid size. Therefore, the penalty iteration method is about 2 to 3 times more efficient than the fixed-point iteration method. Note that the number of penalty iterations is close to optimal. Essentially, in each of the first few timesteps two or three iterations are needed, and in each of the remaining timesteps just one iteration. We also mention that the two iteration schemes produced exactly the same errors for each of the discretization sizes, therefore, we only report one column of errors. The order of convergence is stably 2.

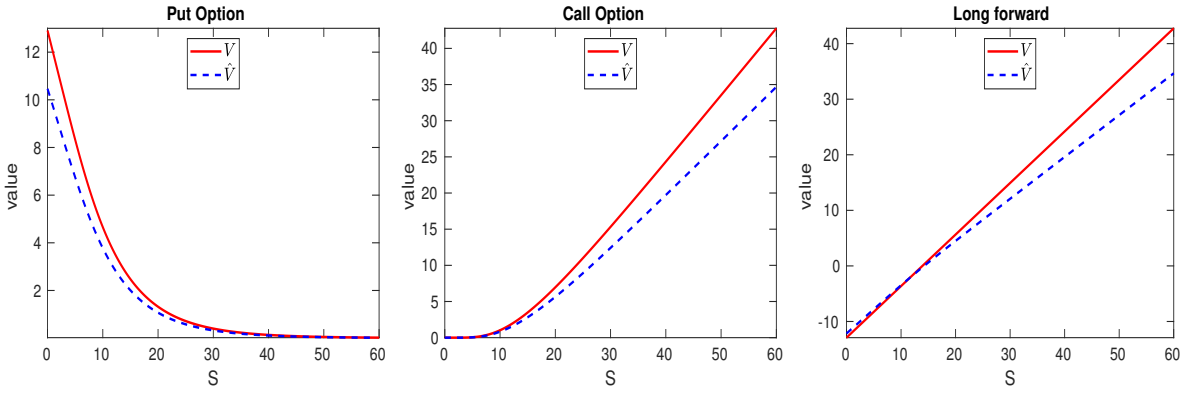

Figure 5.1: A visualization of various financial derivative values with $(\hat{V})$ or without (V) XVA with the parameters in Table 5.1. The difference between these two curves is XVA.

In Figure 5.1, left and centre plots, we visualize V and $\hat{V} = V + U$ for European call and put options. For these options, the XVA is negative and V is above $\hat{V}$.

Discussion on localization

The choice of S_{max} is crucial for an accurate solution. In [39], for European vanilla option, the far-side boundary is chosen so that

$$S_{max} > K \exp(\sqrt{2\sigma^2 T |\ln(tol)|}), \tag{5.2}$$

where *tol* is the computing tolerance. An alternative far-side boundary is suggested in [48]

$$S_{max} > K \exp((r - q - \frac{\sigma^2}{2})T + \sigma\nu\sqrt{T}), \tag{5.3}$$

where $\nu = 3$ is the usual choice for a European option to have small error around $S = K$. If smaller error along the whole price domain is required, ν can be set as 5. In our experiments, we have used $S_{max} = 12K$ for the 5-yr long European options, which seems to agree more with (5.3), with $\nu = 4.6$. Our experiments indicate that this choice of S_{max} balances the domain truncation and discretization errors for the finest grid resolution.

To show the sensitivity of the accuracy of the solution to the choice of S_{max}, in Table 5.4, we present results from the XVA valuation of a European call option with the parameters of Table 5.1, except that we set $S_{max} = 10K$ in one experiment and $S_{max} = 8K$ in another. We only report results from Algorithm 2, as the focus of this discussion is on the error of the computed approximations. These results should be compared to those in Table 5.3, which has $S_{max} = 12K$.

N	N_t	error in U	order	iter tot 2	iter avg 2
$S_{max} = 10K$					
50	100	1.19e-03	–	106	1.06
100	200	2.98e-04	2.00	217	1.09
200	400	7.45e-05	2.00	431	1.08
400	800	1.86e-05	2.00	842	1.05
800	1600	9.84e-06	0.92	1631	1.02
$S_{max} = 8K$					
50	100	9.30e-04	–	106	1.06
100	200	2.34e-04	1.99	219	1.10
200	400	5.85e-05	2.00	442	1.11
400	800	5.24e-05	0.16	846	1.06
800	1600	5.26e-05	-0.00	1635	1.02

Table 5.4: Results from solving (2.2) for a European Call using Algorithm 2 with the parameters in Table 5.1, except that S_{max} is as indicated. Nonuniform grids are used.

In Table 5.4, we notice that, although the penalty iteration with different choices of

S_{max} converges as fast as with $S_{max} = 12K$, the error with $S_{max} = 10K$ and $N = 800$ does not exhibit second order convergence. This phenomenon is more obvious for $S_{max} = 8K$ and $N \geq 400$. It is worth noting, that the errors for $S_{max} = 10K$ are smaller than the respective errors for $S_{max} = 12K$. This is because the localization error for $S_{max} = 10K$ is smaller than the discretization one, and the latter one is smaller than the discretization error for $S_{max} = 12K$.

Discussion on nonuniform space discretizations

In the derivative pricing problem, nonuniform discretization, with denser points around the strike, is often preferred, as the area around the strike exhibits more nonlinearity, and is more "interesting" to practitioners.

Smooth mappings of equally spaced gridpoints are often used to generate appropriate nonuniform discretizations. Suppose the domain is $[0, S_{max}]$ and it is equally divided into N subintervals, with x_i being the gridpoints, i.e. $x_i = ih, h = \frac{S_{max}}{N}$, where $i = 0, 1, \ldots, N$. The smooth mapping $w(x)$ generates the nonuniform grid as

$$S_i \equiv w(x_i) = (1 + \frac{\sinh(\beta - (x_i/x_N - \alpha))}{\sinh(\beta\alpha)})K \tag{5.4}$$

where K is the strike. This mapping produces denser gridpoints around K. Larger parameter α increases the density of the points. The purpose of parameter β is to ensure the last gridpoint is S_{max}. In practice, α is set around 0.4, so that the strike is aligned with a gridpoint.

In Tables 5.5 and 5.6, we report, for several discretization sizes, the discrete L^∞-norm error over all gridpoints at time $\tau = T$ $(t = 0)$, the associated convergence orders, and the total and average per timestep number of iterations, if uniform discretization is used. Compared to the results with nonuniform grid as shown in Tables 5.2 and 5.3, it is obvious that, if the same number of grid points and time steps are used (same computation cost), finer gridpoints around the strike result in smaller maximum error. It is worth noting that all these methods have second order convergence, as expected.

Forward contract

In this example, we present results from pricing the XVA of a 5-year Long Forward contract with the parameters in Table 5.1. For this problem, we don't have an exact solution to calculate exact errors and corresponding convergence orders. Thus the errors at one resolution are estimated from the difference to the previous (coarser) resolution.

N	N_t	error in U	order	iter tot 2	iter avg 2
50	100	2.64e-03	–	135	1.35
100	200	6.33e-04	2.06	250	1.25
200	400	1.57e-04	2.01	427	1.08
400	800	3.95e-06	1.99	822	1.03
800	1600	9.87e-06	2.00	1614	1.01

Table 5.5: Results from solving (2.2) for a European Put using Algorithm 2 with the parameters in Table 5.1. Uniform discretization is used.

N	N_t	error in U	order	iter tot 2	iter avg 2
50	100	2.64e-03	–	104	1.04
100	200	6.33e-04	2.06	208	1.04
200	400	1.57e-04	2.01	420	1.05
400	800	3.95e-05	1.99	846	1.06
800	1600	9.88e-06	2.00	1665	1.04

Table 5.6: Results from solving (2.2) for a European Call using Algorithm 2 with the parameters in Table 5.1. Uniform discretization is used.

In Table 5.7, we report, for several discretization sizes, the differences on the coarser gridpoints at time $\tau = T$ ($t = 0$), the associated convergence orders, and the total and average per timestep number of iterations. We observe similar performance in convergence and number of iterations as in the call and put option examples. The average number of iterations with Algorithm 2 is still very close to 1, independently of the grid size. The average number of iterations with Algorithm 1 is about 2 to 3, independently of the grid size as well. The order of convergence is stably 2.

N	N_t	diff in U	order	iter tot 1	iter avg 1	iter tot 2	iter avg 2
50	100	–	–	400	4.00	102	1.02
100	200	7.58e-04	–	600	3.00	206	1.03
200	400	1.90e-04	2.00	1200	3.00	412	1.03
400	800	4.76e-05	2.00	2400	3.00	821	1.03
800	1600	1.19e-05	2.00	4800	3.00	1644	1.03

Table 5.7: Results from solving (2.2) for a Long Forward contract using Algorithms 1 and 2 with the parameters in Table 5.1. Nonuniform grids are used.

The right part of Figure 5.1 shows the Long Forward contract value with and without XVA. The difference between the two curves is XVA. In this figure, the two curves cross each other. To the left of the crossing point, the value with XVA is higher than the value without XVA, which means that, when the contract is made, party B needs to charge the XVA to party C. On the contrary, to the right of the crossing point, where the value

with XVA is less than the value without XVA, the party B needs to pay the XVA to party C.

5.1.2 Examples of XVA in American derivatives

We consider the pricing of American Put, Call and Long Forward including XVA, with the parameter settings in Table 5.8, and the Mark-to-Market value $M = \hat{V}$. The parameter settings are the same as in [5] (except that the approach in [5] is a Monte Carlo one and there is no preset S_{max}). Tables 5.9 and 5.10 show the results from solving (2.41), using centered differences in space, Crank-Nicolson with Rannacher smoothing in time, and Algorithm 3 for the nonlinearity. The tolerance for Algorithm 3 was set to 10^{-7}, while the penalty parameter p was set to 10^7. We experimented with the far-side condition (2.43), discretized by backward differences, and with its alternative implementation which gives rise to Dirichlet conditions (2.44) or (2.45), and the numerical results obtained were identical (for the points shown and within the digits of precision displayed).

Parameter	Value
Domain of S	$[0, 10K]$
Strike Price, K	15
Time to maturity, T	0.5
Volatility, σ	0.25
Repo rate minus dividend, $r_R = q - \gamma$	0.06
Interest rate, r	0.04
Default intensity of party B, λ_B	0.04
Default intensity of party C, λ_C	0.04
Recovery rate of party B, R_B	0.3
Recovery rate of party C, R_C	0.3
Funding spread, s_F	$(1 - R_B)\lambda_B$

Table 5.8: Model parameters for American derivatives' pricing including bilateral XVA with constant default intensities.

For the American derivatives valuation including XVA, the exact solution is not available. To calculate approximate errors, for the American Put case, we pick three points around or on the strike $S = K = 15$, one at the money, one in the money, and one out of the money. For the American Call and Long Forward cases, we only show results at the money. At each point, the error at one resolution is estimated from the difference to the previous (coarser) resolution. The numerically observed orders of convergences are around two.

Looking at the number of iterations, we easily see that the average number of penalty

iterations in each timestep is around 1.2, which is slightly larger than in the European case. We expect an increase in the number of iterations, since (2.41) has more nonlinear terms than (2.2). It is also important to mention that the convergence of penalty iteration for American derivatives including XVA is still independent of the discretization size.

N	N_t	iter tot	iter avg	$\hat{V}$ value for $S = 15$		
				value	diff in $\hat{V}$	order
50	42	50	1.19	0.86422480	–	–
100	82	98	1.20	0.86679812	2.57e-03	–
200	162	196	1.21	0.86751795	7.20e-04	1.84
400	322	396	1.23	0.86771336	1.95e-04	1.88
800	642	801	1.25	0.86776884	5.55e-05	1.82

N	N_t	$\hat{V}$ value for $S = 14$			$\hat{V}$ value for $S = 16$		
		value	diff in $\hat{V}$	order	value	diff in $\hat{V}$	order
50	42	1.37720716	–	–	0.51568505	–	–
100	82	1.37913317	1.93e-03	–	0.51844504	2.76e-03	–
200	162	1.37960142	4.68e-04	2.04	0.51909720	6.52e-04	2.08
400	322	1.37973315	1.32e-04	1.83	0.51928378	1.87e-04	1.81
800	642	1.37976510	3.19e-05	2.04	0.51933352	4.97e-05	1.91

Table 5.9: Results on three points from solving (2.41) for American Put option valuation including bilateral XVA using Algorithm 3 with the parameters in Table 5.8. Nonuniform grids are used.

In order to compare the American and European option prices with and without XVA, we consider again the parameter settings in Table 5.8, and solve the respective problems with a Put initial condition. Figure 5.2 plots the prices. The centre plot is a zooming of the large box in left plot, while the right plot is a zooming of the small box in left plot. We observe that both V and $\hat{V}$ of American type options are above the pay-off and above V and $\hat{V}$ of European type. European V and $\hat{V}$ cross the pay-off. In both American and European type options, the values of $\hat{V}$ are lower than the values of V, as expected.

N	N_t	iter tot	iter avg	$\hat{V}$ value for Call Option		
				value	diff in $\hat{V}$	order
50	42	43	1.02	1.25145331	–	–
100	82	84	1.02	1.25384939	2.40e-03	–
200	162	165	1.02	1.25445008	6.01e-04	2.00
400	322	330	1.02	1.25460036	1.50e-04	2.00
800	642	654	1.02	1.25463794	3.76e-05	2.00
N	N_t	iter tot	iter avg	$\hat{V}$ value for Long Forward		
				value	diff in $\hat{V}$	order
50	42	47	1.12	0.42849973	–	–
100	82	92	1.12	0.42848609	1.36e-05	–
200	162	183	1.13	0.42848264	3.46e-06	1.98
400	322	366	1.14	0.42848177	8.64e-07	2.00
800	642	726	1.13	0.42848156	2.16e-07	2.00

Table 5.10: Results from solving (2.41) for other American derivatives including bilateral XVA using Algorithm 3 with the parameters in Table 5.8 when S is at the money ($S = K = 15$). Nonuniform grids are used.

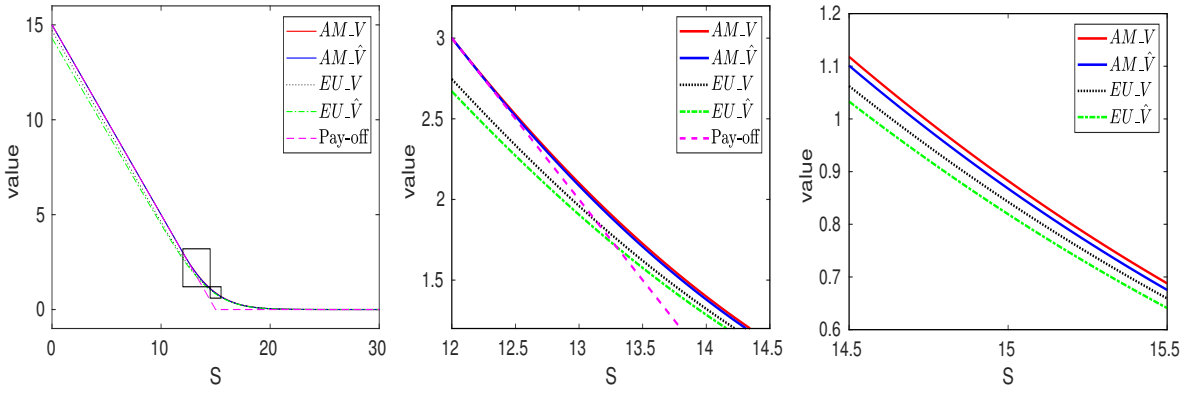

Figure 5.2: A visualization of European and American Put V and $\hat{V}$. The center plot is a zooming of the large box in the left plot, while the right plot is a zooming of the small box in left plot. AM_V and $AM\_\hat{V}$ (EU_V and $EU\_\hat{V}$) are abbreviations for American (European) derivative prices not including and including XVA, respectively.

5.2 Numerical results of XVA with stochastic default intensity in European derivatives

We present results of numerical experiments of European XVA with stochastic default intensities from applying the two proposed methods (numerical 2D PDE and asymptotic) on several financial derivatives. The market and numerical parameters used in the

experiments in this section are listed in Table 5.11. Whenever different parameters are used for the purpose of testing, we state their values explicitly. Note that, the choice of mean reversion rate κ requires a lower bound because of the Feller condition, $\kappa > \frac{(\sigma^{\lambda_C})^2}{2\theta}$. With the parameters in Table 5.11, the lower boundary is $\kappa > 0.4$, which $\kappa = 1$ does not violate.

Parameter	Value
Domain of S	$[0, 8K]$
Domain of λ_C	$[0, 6.05]$
Strike Price, K	15
Time to maturity, T	5
Volatility of asset, σ^S	0.4
Volatility of intensity of party C, σ^{λ_C}	0.2
Correlation between S and λ_C, ρ	0.3
Mean reversion level of intensity of party C, θ	0.05
Mean reversion rate of intensity of party C, κ	1
Repo rate minus dividend, $r_R = q - \gamma$	0.015
Interest rate, r	0.03
Default intensity of party B, λ_B	0.02
Recovery rate of party B, R_B	0.4
Recovery rate of party C, R_C	0.3
Funding spread, s_F	$(1 - R_B)\lambda_B$

Table 5.11: Model parameters for bilateral XVA with stochastic default intensity in European derivatives.

5.2.1 Numerical 2D PDE with penalty iterations

The spatial domain of S and λ_C are discretized into N and M subintervals respectively, and the (nonuniform) gridpoints on S are concentrated around the strike K, while the (nonuniform) gridpoints on λ_C are concentrated towards 0. We could let the λ_C points concentrate around the long-run mean θ, but since θ is very close to 0, we chose to concentrate around 0. The same technique is used in [37] for option pricing in the Heston model with correlation.

The nonuniform gridpoints on S are generated in the following way. Let $x_i = ih, i = 0, 1, \ldots, N$, with $h = S^{max}/N$, be uniform points in $[0, S^{max}]$. The smooth mapping $w(x)$

generates the S nonuniform grid as

$$S_i \equiv w(x_i) = (1 + \frac{\sinh(\beta - (x_i/x_N - \alpha))}{\sinh(\beta\alpha)})K \tag{5.5}$$

where K is the strike. This mapping produces denser grid points around K. Larger parameter α increases the density of the points. The purpose of parameter β is to ensure the last grid point is S^{max}. In practice, α is set to 0.39.

The nonuniform gridpoints on λ_C are generated in following way. Let $y_i = ih, i = 0, 1, \ldots, M$, with $h = \lambda_C^{max}/M$, be uniform points in $[0, \lambda_C^{max}]$. The smooth mapping $u(y)$ generates the λ_C nonuniform grid as

$$(\lambda_C)_i \equiv u(y_i) = \frac{\sinh(y_i)}{\sinh(y_M)} y_M. \tag{5.6}$$

The spatial derivatives are discretized by standard second-order centered differences, except the first derivatives in the boundary conditions, which is discretized by first-order forward or backward differences. The discretization of boundary conditions is discussed in details in Chapter 4. The number of timesteps is denoted by N_t, and $\Delta\tau = T/N_t$. In all cases, the timestepping is Crank-Nicolson-Rannacher, as explained in Chapter 4.

Algorithm 4 is used at each timestep. In all tables below, "iter tot" and "iter avg" are total and average (per timestep) number of iterations. The tolerance *tol* is set to 10^{-7}.

Call and put options

We present results from pricing the XVA with stochastic counterparty default intensity of European options with parameter settings in Table 5.11. For European call and put options, the XVA with stochastic counterparty default intensity does not have analytical solution. The error at one resolution is estimated by the difference from the previous (coarser) resolution. In Table 5.12, we show the results at-the-money with different resolutions. In Tables 5.13 and 5.14, we also list numerical results for several spot prices and different levels of default risk. We notice that the average number of penalty-like iterations is just a bit more than 1, irrespectively of the grid size, which is very close to optimal. The numerical results do not show any problem in terms of stability, and a second order convergence is achieved. From Tables 5.13 and 5.14, we also notice that, for both call and put options, large current counterparty default intensities result in valuation reduction.

N	M	N_t	iter tot	iter avg	$\hat{V}$ value for call option		
					value	diff in $\hat{V}$	order
16	8	10	11	1.10	3.8936802	–	–
32	16	18	19	1.06	3.9462626	5.26e-02	–
64	32	34	36	1.03	3.9585267	1.23e-02	2.10
128	64	66	67	1.02	3.9616227	3.10e-03	1.99
256	128	130	131	1.01	3.9623942	7.72e-04	2.00
512	256	258	259	1.00	3.9625865	1.92e-04	2.00
Richardson extrapolated value: 3.9626505							
N	M	N_t	iter tot	iter avg	$\hat{V}$ value for put option		
					value	diff in $\hat{V}$	order
16	8	10	14	1.40	3.2636520	–	–
32	16	18	22	1.22	3.3121738	4.85e-02	–
64	32	34	42	1.24	3.3235623	1.14e-02	2.09
128	64	66	81	1.23	3.3264560	2.89e-03	1.98
256	128	130	169	1.30	3.3271792	7.23e-04	2.00
512	256	258	331	1.18	3.3273597	1.81e-04	2.00
Richardson extrapolated value: 3.3274199							

Table 5.12: Results from solving (3.40) for European derivatives including bilateral XVA with stochastic default intensities on counterparty using Algorithm 4 with the parameters in Table 5.11 when S is at-the-money ($S = K = 15$) and $\lambda_C = \theta$. Nonuniform grids are used.

N, M, N_t	(7.5, 0.025)	(7.5, 0.05)	(7.5, 0.1)	(30, 0.01)	(30, 0.05)	(30, 0.01)
128, 66, 68	0.9039126	0.8863963	0.8530686	13.1325300	12.8901198	12.4246852
256,128,130	0.9044305	0.8869035	0.8535563	13.1337767	12.8913283	12.4258311
512,256,258	0.9045596	0.8870299	0.8536779	13.1340870	12.8916295	12.4261169
order	2.01	2.00	2.00	2.01	2.00	2.00

Table 5.13: Results from solving (3.40) for European call option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 4 with the parameters in Table 5.11 at various points. Nonuniform grids are used.

N, M, N_t	(7.5, 0.025)	(7.5, 0.05)	(7.5, 0.1)	(30, 0.01)	(30, 0.05)	(30, 0.01)
128, 66, 68	5.7760725	5.6809067	5.4942784	1.4177604	1.3956006	1.3517181
256,128,130	5.7764957	5.6813247	5.4946841	1.4190217	1.3968406	1.3529159
512,256,258	5.7766015	5.6814292	5.4947855	1.4193367	1.3971503	1.3532150
order	2.00	2.00	2.00	2.00	2.00	2.00

Table 5.14: Results from solving (3.40) for European put option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 4 with the parameters in Table 5.11 at various points. Nonuniform grids are used.

Remark 7 *While we present results for call and put options only, the PDE model as well as the numerical methods are directly applicable to other financial derivatives, for example, the forward contract, in which case the price can become negative, and both parties need to worry about defaults. Some results on bilateral XVA pricing of forwards with constant default intensity are found in Section 5.1.1.*

Effect of truncated boundaries

To show the sensitivity of the accuracy of the solution to the choice of S^{max} and λ_C^{max}, we present results from XVA valuation considering stochastic counterparty default risk with the parameters of 5.11, except that in Table 5.15, we set $S_{max} = 4K, 8K, 6K$ and $10K$, and in Table 5.16, we set $\lambda_C^{max} = 4.05, 6.05$ and 8.05. In Table 5.15, note that, for the same values of N, M, N_t, smaller S_{max} results in smaller difference, since the spatial stepsize is smaller with smaller S_{max}. However, $S_{max} = 8K$ makes the convergence order even smoother, and equal to 2. This allows us to claim that the extrapolated value obtained by $S_{max} = 8K$ is more accurate. It is then clear that $S_{max} = 8K$ gives the most accurate and reliable values. In Table 5.9, we see that $\lambda_C^{max} = 6.05$ and $\lambda_C^{max} = 8.05$ give the same accuracy results (including extrapolated values), and smoother order of convergence than $\lambda_C^{max} = 4.05$. Thus, the choices $S_{max} = 8K$ and $\lambda_C^{max} = 6.05$ chosen in Table 5.11 are appropriate and close to optimal for the quality of the numerical PDE approximation of the particular problem.

N	M	N_t	iter tot	iter avg	$\hat{V}$ value when $S^{max} = 4K$		
					value	diff in $\hat{V}$	order
16	8	10	11	1.10	3.9293825	–	–
32	16	18	19	1.06	3.9531642	2.38e-02	–
64	32	34	35	1.03	3.9583044	5.14e-03	2.21
128	64	66	67	1.02	3.9598170	1.51e-03	1.76
256	128	130	131	1.01	3.9601880	3.71e-04	2.03
512	256	258	259	1.00	3.9602776	8.97e-05	2.05
Richardson extrapolated value: 3.9603075							

N	M	N_t	iter tot	iter avg	$\hat{V}$ value when $S^{max} = 6K$		
					value	diff in $\hat{V}$	order
16	8	10	11	1.10	3.9122552	–	–
32	16	18	19	1.06	3.9504179	3.82e-02	–
64	32	35	35	1.03	3.9594910	9.07e-03	2.07
128	64	66	67	1.02	3.9618227	2.33e-03	1.96
256	128	130	131	1.01	3.9624022	5.79e-04	2.01
512	256	258	259	1.00	3.9625458	1.44e-04	2.01
Richardson extrapolated value: 3.9625937							

N	M	N_t	iter tot	iter avg	$\hat{V}$ value when $S^{max} = 8K$		
					value	diff in $\hat{V}$	order
16	8	10	11	1.10	3.8936802	–	–
32	16	18	19	1.06	3.9462626	5.26e-02	–
64	32	34	35	1.03	3.9585267	1.23e-02	2.10
128	64	66	67	1.02	3.9616227	3.10e-03	1.99
256	128	130	131	1.01	3.9623942	7.72e-04	2.00
512	256	258	259	1.00	3.9625865	1.92e-04	2.00
Richardson extrapolated value: 3.9626505							

N	M	N_t	iter tot	iter avg	$\hat{V}$ value when $S^{max} = 10K$		
					value	diff in $\hat{V}$	order
16	8	10	11	1.10	3.8763460	–	–
32	16	18	19	1.06	3.9426609	6.63e-02	–
64	32	34	35	1.03	3.9576239	1.50e-02	2.15
128	64	66	67	1.02	3.9613985	3.77e-03	1.99
256	128	130	131	1.01	3.9623399	9.41e-04	2.00
512	256	258	259	1.00	3.9625747	2.35e-04	2.00
Richardson extrapolated value: 3.9626530							

Table 5.15: Results from solving (3.40) for European call option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 4 with the parameters in Table 5.11, except S^{max} varying as indicated, when S is at-the-money ($S = K = 15$) and $\lambda_C = \theta$. Nonuniform grids are used.

N	M	N_t	iter tot	iter avg	$\hat{V}$ value when $\lambda_C^{max} = 4.05$		
					value	diff in $\hat{V}$	order
16	8	10	11	1.10	3.9177337	–	–
32	16	18	19	1.06	3.9496221	3.19e-02	–
64	32	34	35	1.03	3.9585993	8.98e-03	1.83
128	64	66	67	1.02	3.9615316	2.93e-03	1.61
256	128	130	131	1.01	3.9623705	8.39e-04	1.81
512	256	258	259	1.00	3.9625815	2.11e-04	1.99
Richardson extrapolated value: 3.9626518							
N	M	N_t	iter tot	iter avg	$\hat{V}$ value when $\lambda_C^{max} = 6.05$		
					value	diff in $\hat{V}$	order
16	8	10	11	1.10	3.8936802	–	–
32	16	18	19	1.06	3.9462626	5.26e-02	–
64	32	34	35	1.03	3.9585267	1.23e-02	2.10
128	64	66	67	1.02	3.9616227	3.10e-03	1.99
256	128	130	131	1.01	3.9623942	7.72e-04	2.00
512	256	258	259	1.00	3.9625865	1.92e-04	2.00
Richardson extrapolated value: 3.9626505							
N	M	N_t	iter tot	iter avg	$\hat{V}$ value when $\lambda_C^{max} = 8.05$		
					value	diff in $\hat{V}$	order
16	8	10	11	1.10	3.8964305	–	–
32	16	18	19	1.06	3.9463962	5.00e-02	–
64	32	34	35	1.03	3.9585381	1.21e-02	2.04
128	64	66	67	1.02	3.9616237	3.09e-03	1.98
256	128	130	131	1.01	3.9623942	7.70e-04	2.00
512	256	258	259	1.00	3.9625864	1.92e-04	2.00
Richardson extrapolated value: 3.9626505							

Table 5.16: Results from solving (3.40) for European call option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 4 with the parameters in Table 5.11, except λ_C^{max} varying as indicated, when S is at-the-money ($S = K = 15$) and $\lambda_C = \theta$. Nonuniform grids are used.

5.2.2 Asymptotic solution and effect of rate of mean reversion

In order to investigate how the speed of mean reversion κ affects the accuracy of the asymptotic approximation, we show the derivative values considering counterparty de-

fault risk with the parameters of Table 5.11, except that κ varies from 1 to 10. We present two Tables 5.17 and 5.18 (κ varying from 1 to 3), one with correlation ρ being zero and one with correlation given in Table 5.11. When κ varies, in order to keep ν constant, we vary σ^{λ_C} as $\frac{\nu}{\sqrt{\epsilon}}$. Since the exact solutions to the problems considered are unknown, we compute highly accurate approximations by the numerical 2D PDE approach, and use those to compare with the ones obtained by the asymptotic approach. To obtain highly accurate 2D PDE prices, the 2D PDE approximations from the two finest grids are extrapolated by one-level Richardson extrapolation, with second order convergence. We do not use more levels of Richardson extrapolation as it is known that higher levels of Richardson extrapolation are numerically unstable. The numerical results show good agreement among solutions under different approaches.

In Tables 5.17, 5.18, and Figure 5.3 (κ varying from 1 to 10), we can see that the differences between the PDE extrapolated solution and the asymptotic solutions are decreasing with increasing κ, the order of convergence with respect to κ^{-1} being approximately 1.5 for the non-zero correlation case, and more than 2 for the zero correlation one.

	(7.5, 0.05)	(7.5, 0.1)	(15, 0.05)	(15, 0.1)	(30, 0.05)	(30, 0.1)
const. intens.	5.6250695	5.6250695	3.2759704	3.2759704	1.3662239	1.3662239
$\kappa = 1, \sigma^{\lambda_C} = 0.2, \nu = 0.2$						
PDE FDM	5.6345450	5.4444596	3.2814455	3.1707436	1.3682827	1.3221227
PDE extrap	5.6345792	5.4444926	3.2815060	3.1708021	1.3683868	1.3222232
asymptotic	5.6388509	5.4419735	3.2839966	3.1693376	1.3695712	1.3217533
$\kappa = 2, \sigma^{\lambda_C} = 0.2828, \nu = 0.2$						
PDE FDM	5.6308259	5.5338160	3.2792796	3.2227830	1.3673795	1.3438218
PDE extrap	5.6308601	5.5338496	3.2793401	3.2228424	1.3674836	1.3439241
asymptotic	5.6319602	5.5335215	3.2799835	3.2226540	1.3678976	1.3439886
$\kappa = 3, \sigma^{\lambda_C} = 0.3464, \nu = 0.2$						
PDE FDM	5.6291365	5.5641450	3.2782957	3.2404460	1.3669693	1.3511868
PDE extrap	5.6291707	5.5641788	3.2783562	3.2405057	1.3670733	1.3512897
asymptotic	5.6296633	5.5640375	3.2786458	3.2404262	1.3673397	1.3514004

Table 5.17: Values by different approaches for European put option including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 5.11, except that κ and σ^{λ_C} vary as indicated, and $\rho = 0$, at several points. The grid size for the PDE solution is $N = 512, M = 256$, and extrapolation takes place between $N = 256, M = 128$ and $N = 512, M = 256$.

Comparing the differences in Tables 5.17 and 5.18, and looking at Figure 5.3, the zero correlation cases of the asymptotic solutions result in smaller differences from the extrapolated PDE values, than the nonzero correlation cases, while the 2D PDE solutions give approximately the same differences for $\rho = 0$ and $\rho \neq 0$.

	(7.5, 0.05)	(7.5, 0.1)	(15, 0.05)	(15, 0.1)	(30, 0.05)	(30, 0.1)
const. intens.	5.6250695	5.6250695	3.2759704	3.2759704	1.3662239	1.3662239
$\kappa = 1, \sigma^{\lambda_C} = 0.2, \nu = 0.2$						
PDE FDM	5.6814292	5.4947855	3.3273597	3.2201054	1.3971503	1.3532150
PDE extrap	5.6814640	5.4948193	3.3274199	3.2201636	1.3972536	1.3533148
asymptotic	5.6974803	5.5006028	3.3425304	3.2278715	1.4072341	1.3594163
$\kappa = 2, \sigma^{\lambda_C} = 0.2828, \nu = 0.2$						
PDE FDM	5.6681123	5.5723664	3.3159146	3.2606831	1.3905067	1.3677652
PDE extrap	5.6681469	5.5724005	3.3159749	3.2607424	1.3906101	1.3678669
asymptotic	5.6734174	5.5749787	3.3213732	3.2640437	1.3945293	1.3706204
$\kappa = 3, \sigma^{\lambda_C} = 0.3464, \nu = 0.2$						
PDE FDM	5.6607425	5.5964355	3.3094221	3.2722577	1.3866742	1.3713342
PDE extrap	5.6607769	5.5964696	3.3094823	3.2723172	1.3867777	1.3714365
asymptotic	5.6635130	5.5978872	3.3124404	3.2742207	1.3890844	1.3731451

Table 5.18: Values by different approaches for European put option including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 5.11, except that κ and σ^{λ_C} vary as indicated, at several points. The grid size for the PDE solution is $N = 512, M = 256$, and extrapolation takes place between $N = 256, M = 128$ and $N = 512, M = 256$.

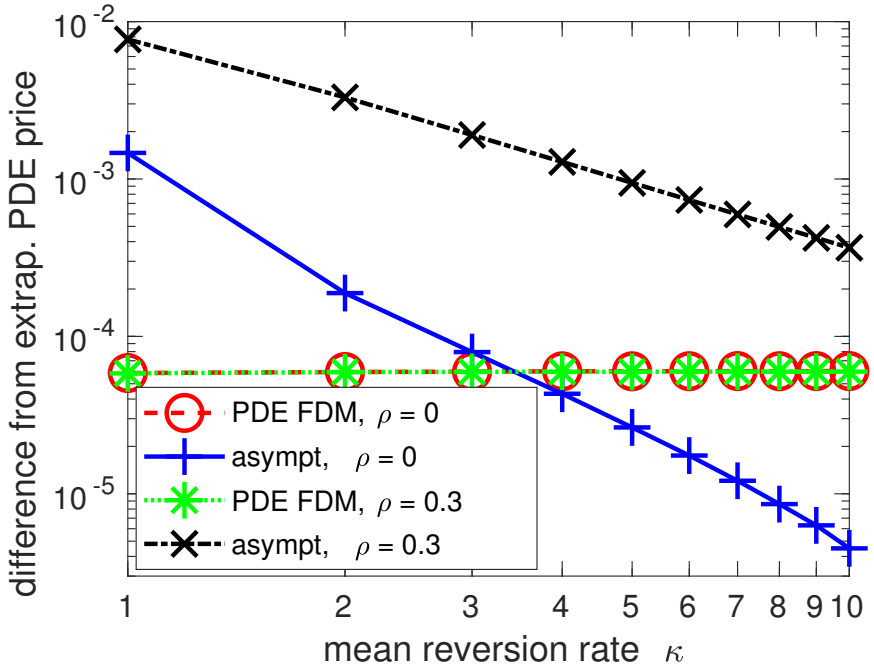

Figure 5.3: Accuracy of different approaches for European put option valuation including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 5.11 except κ and ρ as indicated, and $\sigma^{\lambda_C} = 0.2\sqrt{\kappa}$, versus κ at (15, 0.1).

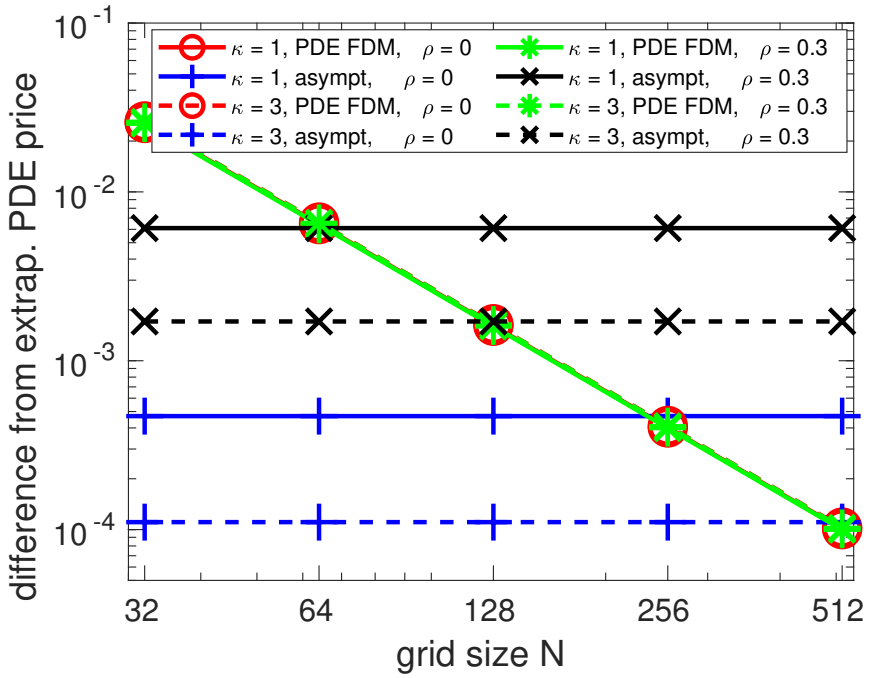

Figure 5.4: Accuracy of different approaches for European put option valuation including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 5.11 except κ and ρ as indicated, and $\sigma^{\lambda_C} = 0.2\sqrt{\kappa}$, versus N at (30, 0.1).

From Figure 5.4, it can be seen that the PDE accuracy is of the same level irrespectively of κ and ρ, while the asymptotic solution accuracy improves with increasing κ, and for $\rho = 0$. Furthermore, in the mean reversion case $\kappa = 3$ and $\rho = 0$, the traditional PDE numerical methods need very fine grids, finer than $N = 512$, to beat the asymptotic approach in accuracy. Likewise, in the mean reversion case $\kappa = 1$ and $\rho = 0.3$, the traditional PDE numerical methods beat the asymptotic approach in accuracy even with $N = 64$.

Remark 8 *Taking into account that the computational cost of the 2D numerical PDE solution approach is more than $O(N^2)$, while that of the asymptotic is $O(N)$, it is clear that, for a given problem, the asymptotic solution approach is always more efficient than the 2D numerical PDE one. However, the accuracy of the computed results from the two approaches for a given problem (fixed κ) is not directly comparable, so a fair comparison is hard to carry. It is possibly most important to note that, for a given problem, the 2D numerical PDE solution approach allows the user to adjust the discretization size to match a certain accuracy level, while the asymptotic solution approach does not possess this property. Still, when mild accuracies are satisfactory, researchers and practitioners may find the asymptotic approach attractive.*

5.2.3 Effect of model parameters

We study how the correlation ρ, mean reversion level θ and mean reversion speed κ affect the adjusted values of financial derivatives. We still focus on European call and put options, with parameters in Table 5.11, except when mentioned otherwise.

In Table 5.19 and Figures 5.5 and 5.6, we show the effect of the correlation ρ between spot price S and counter-party default risk λ_C on the values of call and put options computed by solving (3.40). In the put option case, higher ρ leads to higher value of derivatives, while, in the call option case, higher ρ results in lower value derivatives. This effect can be also captured by the asymptotic solution. Recalling the asymptotic solution (3.91), and substituting $\sigma^{\lambda_C} = \frac{\nu}{\sqrt{\epsilon}}$ or $\nu = \sqrt{\epsilon}\sigma^{\lambda_C}$, we get

$$\begin{aligned}
\hat{V}^{\epsilon,1} =& \hat{V}_0 - \sqrt{\epsilon}\tau\rho\sigma^S\nu S(1-R_C)\langle\sqrt{\lambda_C}\rangle\frac{\partial\hat{V}_0^+}{\partial S} \\
&+ \epsilon(1-R_C)(\theta-\lambda_C)\hat{V}_0^+ + \epsilon\tau(1-R_C)^2\frac{\theta\nu^2}{2}\hat{V}_0^+ \\
=& \hat{V}_0 - \epsilon\tau\rho\sigma^S\sigma^{\lambda_C}S(1-R_C)\langle\sqrt{\lambda_C}\rangle\frac{\partial\hat{V}_0^+}{\partial S} \\
&+ \epsilon(1-R_C)(\theta-\lambda_C)\hat{V}_0^+ + \epsilon^2\tau(1-R_C)^2\frac{(\sigma^{\lambda_C})^2}{2}\hat{V}_0^+. \qquad (5.7)
\end{aligned}$$

Recall that, $\hat{V}_0$ is the solution to the Black-Scholes equation (3.43) considering constant counterparty default risk equal to θ. Among the four terms in the right hand side of (5.7), all but the second one are independent of ρ. Considering the second term, in put option case, the sensitivity delta $\frac{\partial \hat{V}_0^+}{\partial S}$ is non-positive, while, in call case, the sensitivity delta $\frac{\partial \hat{V}_0^+}{\partial S}$ is non-negative. This explains the different behaviors of prices in different derivatives.

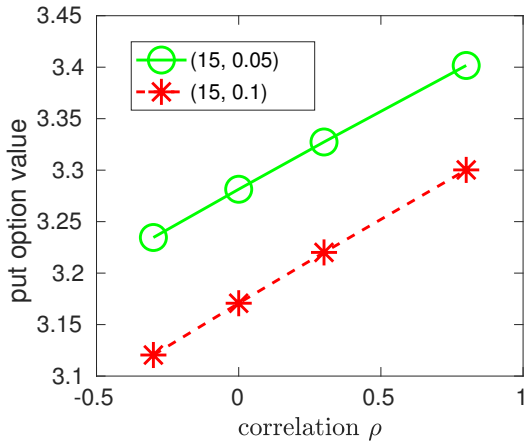

Figure 5.5: Effect of ρ on put option value with XVA, computed by solving (3.40). Other parameters are in Table 5.11.

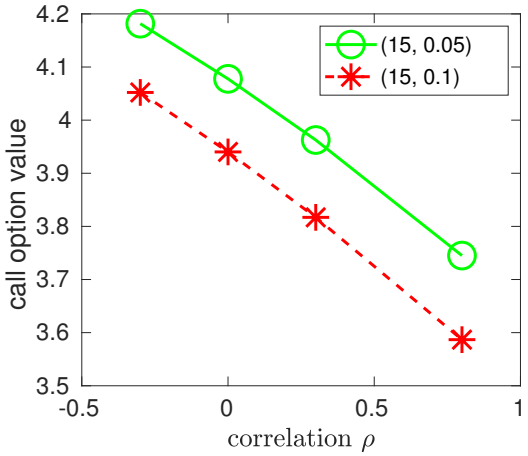

Figure 5.6: Effect of ρ on call option value with XVA, computed by solving (3.40). Other parameters are in Table 5.11.

ρ	put price at (15, 0.05)	put price at (15, 0.10)	call price at (15, 0.10)	call price at (15, 0.10)
-0.3	3.2345962	3.1204577	4.1815355	4.0521265
0	3.2814455	3.1707436	4.0777310	3.9401659
0.3	3.3273597	3.2201054	3.9625865	3.8170009
0.8	3.4016595	3.3002454	3.7450680	3.5867810

Table 5.19: Value comparison for solving (3.40) for European options including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 5.11 with different correlations. The grid size is $N = 512, M = 256$.

Table 5.20 presents results that demonstrate how the long-run mean θ affects the adjusted derivative values. Higher θ results in increased total default probability experienced during derivative life. When the derivative contract is a positive asset to party B, naturally, this results in a lower value of the financial derivatives. From Equations (3.6) and (3.7), higher counterparty default rate means more possibility to receive recovery

value of full contract, especially when this contract is a positive asset to the surviving party. This reduces the value of derivative contract to the surviving party when the counterparty default risk is considered. Figure 5.7 gives a visualization in the case of put option. As we can see in (5.7), the dominant zero-th term value $\hat{V}_0$ decreases as θ decreases, while all other terms increase a bit. These terms (involving ϵ in some positive powers) are dominated by the zero-th order term.

θ	put price at (15, 0.05)	put price at (15, 0.10)	call price at (15, 0.10)	call price at (15, 0.10)
0.01	3.6947761	3.5775433	4.4835585	4.3140849
0.05	3.3273597	3.2201054	3.9625865	3.8170009
0.2	2.2192156	2.1466437	2.5519197	2.4605421

Table 5.20: Value comparison for solving (3.40) for European options including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 5.11 with different mean reversion levels. The grid size is $N = 512, M = 256$.

Table 5.21 and Figure 5.8 demonstrate how the speed of mean reversion κ affects the values of put option. Increasing values of κ, naturally result in prices closer to the Black-Scholes price including XVA with constant default intensity λ_C equal to θ. Again, the asymptotic solution (5.7) expresses this trend well. The terms involving ϵ decrease in absolute value and converge to zero, which makes the asymptotic solution tend to $\hat{V}_0$, the solution of the Black-Scholes equation (3.43) taking default risk into account, with constant default intensity $\lambda_C = \theta$.

point	$\kappa = 1$	$\kappa = 2$	$\kappa = 3$	$\kappa = 4$	$\kappa = 5$
(15, 0.025)	3.3818182	3.3324451	3.3142960	3.3049453	3.2992520
(15, 0.05)	3.3273597	3.3043328	3.2954144	3.2907365	3.2878633
(15, 0.1)	3.2201054	3.2485634	3.2578586	3.2624369	3.2651620
$\hat{V}_0$: 3.27597044					

Table 5.21: Value comparison for solving (3.40) for European put option including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 5.11. The grid size is $N = 512, M = 128$.

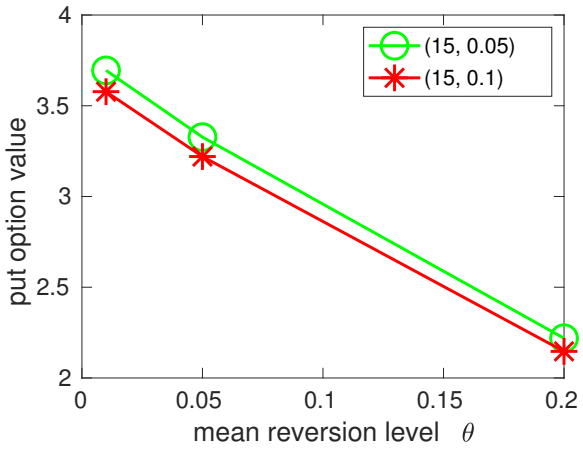

Figure 5.7: Effect of θ on put option value with XVA, computed by solving (3.40). Other parameters are in Table 5.11.

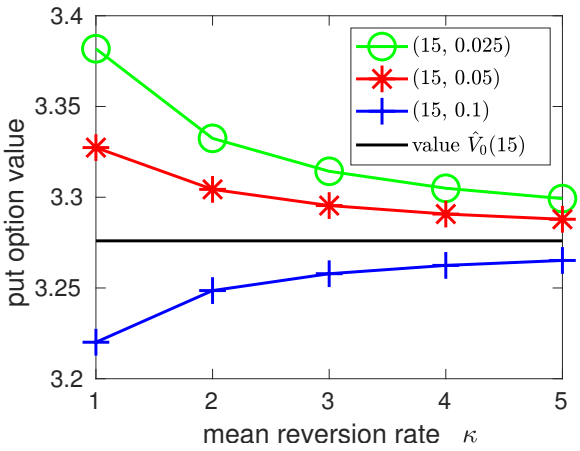

Figure 5.8: Effect of κ on put option value with XVA, computed by solving (3.40). Other parameters are in Table 5.11.

5.3 Numerical results of XVA with stochastic default intensity in American derivatives

In this section, we show results of numerical experiments from applying the two proposed methods (numerical 2D PDE and asymptotic) on American put options. Table 5.22 presents the values of parameters we used in the experiments, which are almost the same as those in Table 5.11, except the maturity time $T = 1$ instead of $T = 5$.

Parameter	Value
Domain of S	$[0, 8K]$
Domain of λ_C	$[0, 6.05]$
Strike Price, K	15
Time to maturity, T	1
Volatility of asset, σ^S	0.4
Volatility of intensity of party C, σ^{λ_C}	0.2
Correlation between S and λ_C, ρ	0.3
Mean reversion level of intensity of party C, θ	0.05
Mean reversion rate of intensity of party C, κ	1
Repo rate minus dividend, $r_R = q - \gamma$	0.015
Interest rate, r	0.03
Default intensity of party B, λ_B	0.02
Recovery rate of party B, R_B	0.4
Recovery rate of party C, R_C	0.3
Funding spread, s_F	$(1 - R_B)\lambda_B$

Table 5.22: Model parameters for bilateral XVA with stochastic default intensity in American put options.

5.3.1 Numerical 2D PDE with double-penalty iterations

The spatial domain of (S, λ_C), $[0, S_{max}] \times [0, \lambda_C^{max}]$, is discretized into $N \times M$ subintervals. We choose the truncated boundaries, $S_{max} = 8K$ and $\lambda_C^{max} = 6.05$, when $\sigma^{\lambda_C} = 0.2$, as suggested in Section 5.2.1. More generally, $\lambda_C^{max} = \theta + 30\sigma^{\lambda_C}$. In Section 5.2.1, the effect of truncated boundaries is studied in details for European options. We do not expect different behavior for American options, as far as truncated boundaries are concerned. We also choose to use nonuniform gridpoints on both S dimension and λ_C dimension. The gridpoints on S are concentrated around the strike price K, while the gridpoints on λ_C are concentrated towards 0. The formulae of the nonuniform mappings from uniform grids on these two spatial dimensions can be found in Section 5.2.1.

The spatial derivatives are discretized by standard second-order centered differences, except the first derivatives in the boundary conditions, which are discretized by first-order forward or backward differences. The number of timesteps in denoted by N_t, and $\Delta t = T/N_t$, as defined in Section 4.2.1. As explained in Section 4.2.1, the timestepping scheme is Crank-Nicolson-Rannacher. Algorithm 5 is used at each timestep. In all tables in this section, "iter tot" and "iter avg" mean total and average (per timestep) number

of iterations. The tolerance *tol* of double-penalty iteration is set to 10^{-7}, thus $p = 10^7$.

We present results from pricing the XVA of American put options with stochastic counterparty default intensity, and with the parameter settings in Table 5.22. For American put options, the XVA with stochastic counterparty default intensity does not have an analytical solution. The error at one resolution is estimated by the difference from the previous (coarser) resolution. In Table 5.23, we show the results at-the-money with different resolutions. We notice that the average number of double-penalty iterations is around 2, and varies very little with grid size. The numerical results do not exhibit any instability, and the order of convergence is 2, which is the same as the theoretically expected convergence. In Table 5.24, we also list numerical results for several spot prices and different levels of default risk. In this problem with $K = 15$, asset price $S = 7.5$ is expected to be in the exercise region, while asset prices $S = 15$ and $S = 30$ are expected to be in the hold region. From Tables 5.23 and 5.24, we notice that when the underlying asset prices $S = 15$ and $S = 30$, the order of convergence is approximately 2 as expected. From Table 5.24, at $S = 7.5$, the numerical solution converges quickly to 7.5, and this is because the errors in the exercise region are related to $\frac{1}{p} = 10^{-7}$ and less affected by the discretization size. Therefore, for high resolutions, the errors are beyond the seven digits.

We are also interested in the behavior of the free boundaries for different counterparty default intensities λ_C. In Table 5.25, we show the location and the convergence of the free boundaries when $\lambda_C = 0.0615927$ and $\lambda_C = 0.1364300$ for the double-penalty method for several grid sizes. These λ_C points are chosen because they are grid points of all λ_C-grids from the coarsest to the finest. We could have picked other arbitrary λ_C points, but in such case, we need to apply interpolation to the computed values on the grid points to obtain the values at the arbitrary λ_C points, a procedure that may increase the errors. From Table 5.25, we notice that although the order of convergence of the free boundaries of the double-penalty method is not very stable, the changes from coarse to fine grids generally decrease and go down to levels of about 10^{-3} or 10^{-4}. In Table 5.26, we show the values of the free boundaries for several different λ_C. It is clear that the free boundary increases as the λ_C value increases, which means that the American derivative will be exercised for a larger range of S values.

We also investigate how the free boundary location changes with varying κ. Here, we assume that, as κ changes, σ^{λ_C} also changes, so that ν remains constant, i.e. $\sigma^{\lambda_C} = \nu\sqrt{\kappa}$. This scaling is the same as the one assumed in Section 4.3, and helps obtain results comparable to those obtained from the asymptotic approximation. Figure 5.9 plots the values of the free boundaries, versus counterparty default intensity λ_C, with various

κ. From Table 5.26 and Figure 5.9, we can see that the free boundary increases with increasing default intensities λ_C. Usually, large counterparty default intensity results in valuation reduction in adjusted price. With the "push-up" effect of the payoff constraints from American put options, the free boundary should move to far-side of S-dimension. We also notice, from Figure 5.9, the free boundaries are less sensitive with respect to κ around $\lambda_C = 0.05$, which is the long-run mean of mean-reversion process. When λ_C is smaller, the free boundaries increase with increasing κ, while, when λ_C is large, the free boundaries decrease with increasing κ.

N	M	N_t	iter tot	iter avg	$\hat{V}$ value for put option		
					value	diff in $\hat{V}$	order
16	8	10	19	1.90	2.1254085	–	–
32	16	18	33	1.83	2.1619127	3.65e-02	–
64	32	34	72	2.12	2.1714040	9.49e-03	1.94
128	64	66	149	2.26	2.1741303	2.73e-03	1.80
256	128	130	310	2.38	2.1748854	7.55e-04	1.85
512	256	258	678	2.63	2.1750101	1.25e-04	2.60
Richardson extrapolated value: 2.1750516							

Table 5.23: Results from solving (4.9) for American put option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 5 with the parameters in Table 5.22 when S is at-the-money ($S = K = 15$) and $\lambda_C = \theta$. Nonuniform grids are used.

N, M, N_t	(7.5, 0.025)	(7.5, 0.05)	(7.5, 0.1)	(30, 0.01)	(30, 0.05)	(30, 0.01)
128, 66, 68	7.4999991	7.4999954	7.5000017	0.1224851	0.1213862	0.1191487
256,128,130	7.5000007	7.4999984	7.5000000	0.1227464	0.1216452	0.1194031
512,256,258	7.5000000	7.5000000	7.5000000	0.1227959	0.1216946	0.1194522
order	—-	—-	—-	2.40	2.39	2.39

Table 5.24: Results at various points (S, λ_C) from solving (4.9) for American put option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 5 with the parameters in Table 5.22. Nonuniform grids are used.

N	FB at $\lambda_C = 0.0615927$			FB at $\lambda_C = 0.1364300$		
	value	diff	order	value	diff	order
16	8.2288945	–	–	8.3357731	–	–
32	8.4959278	2.67e-01	–	8.8004742	4.65e-01	–
64	8.4409673	-5.50e-02	2.28	8.8252269	2.48e-02	4.23
128	8.4474892	6.52e-03	3.08	8.8955801	7.04e-02	-1.51
256	8.4528315	5.34e-03	0.29	8.9078471	1.23e-02	2.52
512	8.4526494	-1.82e-04	4.87	8.9092269	1.38e-03	3.15

Table 5.25: Free boundary locations and orders of convergence for various counterparty default intensity λ_C-points from solving (4.9) for American put option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 5 with the parameters in Table 5.22. Nonuniform grids are used. Algorithm 6 is used for the calculation of free boundaries.

λ_C	0.01	0.025	0.05	0.1	0.2
free boundary	8.0062692	8.1547595	8.3671579	8.7067881	9.2033891

Table 5.26: Free boundary locations for various counterparty default intensity λ_C-points from solving (4.9) for American put option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 5 with the parameters in Table 5.22. Nonuniform grids are used and $N = 512$. Algorithm 6 is used for the calculation of free boundaries.

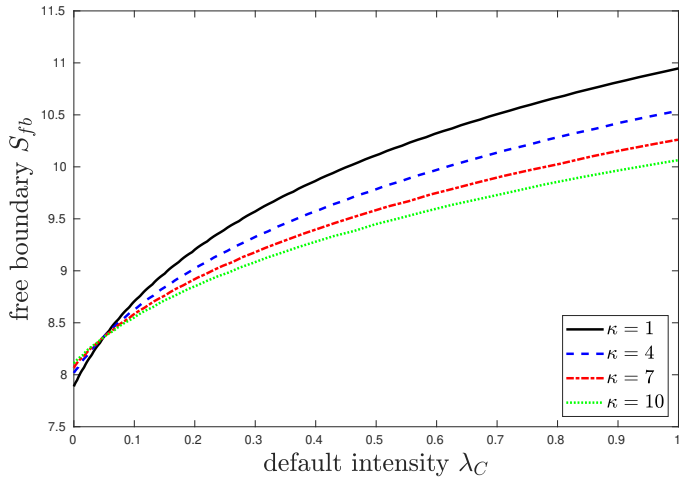

Figure 5.9: Free boundary locations versus the counterparty default intensity λ_C, with various mean-reversion speeds κ from solving (4.9) for American put option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 5 with the parameters in Table 5.22. Nonuniform grids are used and $N = 512$. Algorithm 6 is used for the calculation of free boundaries.

5.3.2 Comparison of asymptotic and numerical 2D PDE approximations

In order to numerically investigate the accuracy of the asymptotic approximation and compare to the finite difference PDE solutions in American type financial derivatives, we show the American put options including XVA with mean-reversion counterparty default intensity, with the parameters of Table 5.22, except that we vary the speed of mean-reversion parameter, κ. We present Table 5.27, in which κ is varying from 1 to 5. This table can also help us numerically analyze how the asymptotic solution accuracy is affected by the speed of mean-reversion parameter κ. In Table 5.27, we keep $\nu = \sigma^{\lambda_C}\sqrt{\epsilon}$ constant, where $\kappa = \frac{1}{\sqrt{\epsilon}}$, and σ^{λ_C} is varying with varying κ. The exact solutions to these problems are not known. Therefore, we compute the numerical approximations by the 2D PDE approach and asymptotic approximations. Also, we compute a highly accurate numerical approximation extrapolated from the two finest grids of the 2D PDE approximations using Richardson extrapolation. The results by Richardson extrapolation are considered as the most accurate approximations. In Table 5.27, we notice, for points in the exercise region, both the 2D PDE solution (with high resolution) and the asymptotic give exact results within tolerance $\mathcal{O}(\frac{1}{p})$. In this case, Richardson extrapolation deviates from the exact results, as extrapolation brings wrong digits in the numerical 2D PDE solution beyond the seventh digit to the left. For these points, extrapolation with respect to the grid size does not make sense.

	(7.5, 0.05)	(7.5, 0.1)	(15, 0.05)	(15, 0.1)	(30, 0.05)	(30, 0.1)
$\kappa = 1, \sigma^{\lambda_C} = 0.2, \nu = 0.2$						
PDE FDM	7.5000000	7.5000000	2.1750101	2.1403166	0.1216946	0.1194522
PDE extrap	7.5000001	7.5000000	2.1750516	2.1403572	0.1217110	0.1194685
asymptotic	7.5000000	7.5000000	2.1947775	2.1187456	0.1236463	0.1194102
$\kappa = 2, \sigma^{\lambda_C} = 0.28, \nu = 0.2$						
PDE FDM	7.5000000	7.5000000	2.1756124	2.1497291	0.1217586	0.1202117
PDE extrap	7.5000001	7.5000000	2.1756538	2.1497695	0.1217750	0.1202280
asymptotic	7.5000000	7.5000000	2.1879855	2.1499696	0.1228682	0.1207502
$\kappa = 3, \sigma^{\lambda_C} = 0.35, \nu = 0.2$						
PDE FDM	7.5000000	7.5000000	2.1758164	2.1556302	0.1217499	0.1205987
PDE extrap	7.5000001	7.5000000	2.1758578	2.1556707	0.1217664	0.1206150
asymptotic	7.5000000	7.5000000	2.1850350	2.1596910	0.1225268	0.1211148
$\kappa = 4, \sigma^{\lambda_C} = 0.4, \nu = 0.2$						
PDE FDM	7.5000000	7.5000000	2.1758642	2.1595322	0.1217224	0.1208164
PDE extrap	7.5000001	7.5000000	2.1759056	2.1595729	0.1217389	0.1208328
asymptotic	7.5000000	7.5000000	2.1832931	2.1642851	0.1223242	0.1212652
$\kappa = 5, \sigma^{\lambda_C} = 0.45, \nu = 0.2$						
PDE FDM	7.5000000	7.5000000	2.1758420	2.1622297	0.1216910	0.1209478
PDE extrap	7.5000001	7.5000000	2.1758835	2.1622708	0.1217075	0.1209643
asymptotic	7.5000000	7.5000000	2.1821117	2.1669053	0.1221864	0.1213391

Table 5.27: Values of $\hat{V}$ at several points (S, λ_C) and by different approaches for American put option including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 5.22, except that κ varies as indicated, and $\sigma^{\lambda_C} = 0.2\sqrt{\kappa}$. The grid size for the PDE solution is $N = 512, M = 256$, and extrapolation takes place between $N = 256, M = 128$ and $N = 512, M = 256$.

In Figure 5.10, we plot the "errors" of the PDE and the asymptotic solutions versus κ. The errors are approximated by the difference between the respective approximations and the PDE extrapolated solution. From Figure 5.10, we can see the errors of the asymptotic solutions are decreasing with increasing κ at some constant order with few fluctuations, while the errors of the PDE solution are κ-independent. The latter errors are mainly affected by the S-points of evaluation, and are larger on the strike K than far away. Numerically, we observe that the order of convergence of asymptotic solution with respect to κ^{-1} is approximately 0.75. We believe that the reason why this order is lower

than in the European case, is that the free boundary approximation is of low order.

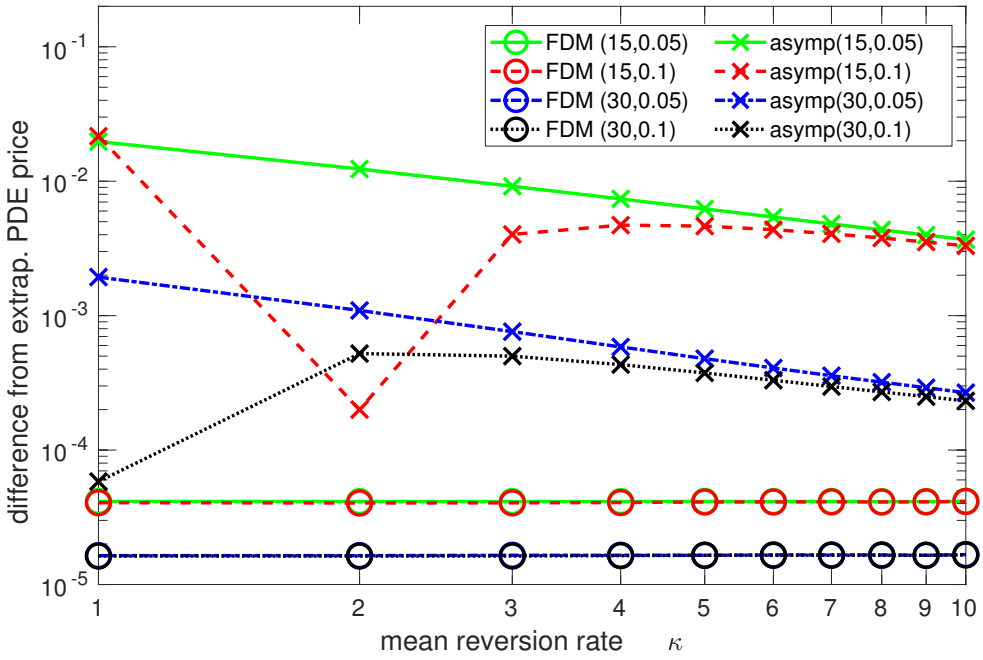

Figure 5.10: Accuracy comparison of different approaches for American put option valuation including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 5.22 except varying κ as indicated, and $\sigma^{\lambda_C} = 0.2\sqrt{\kappa}$, versus κ.

In Table 5.28, we also compare the free boundary locations computed by the numerical 2D PDE solutions and the asymptotic approach at specific λ_C points. As mentioned in the previous subsection, the free boundary location changes very little with respect to κ when $\lambda_C = 0.05$. Considering the 2D PDE approach, when λ_C is less than the mean, the free boundaries are increasing with κ, while when λ_C is larger than the mean, the free boundaries are decreasing with κ. Furthermore, in the case that λ_C is away from the mean, the differences of free boundaries between the numerical PDE and asymptotic approximations become smaller as κ increases, because of the mean-reversion effects.

κ	1	2	3	4	5
$\lambda_C = 0.01$					
PDE FDM	8.0062692	8.0483576	8.0728779	8.1015927	8.1206571
asymptotic	8.3318264	8.3322154	8.3323877	8.3324904	8.3325605
difference	3.26e-01	2.84e-01	2.601e-01	2.31e-01	2.12e-01
$\lambda_C = 0.05$					
PDE FDM	8.3671579	8.3708289	8.3713726	8.3709081	8.3700482
asymptotic	8.3318264	8.3322154	8.3323877	8.3324904	8.3325605
difference	3.53e-02	3.86e-02	3.90e-02	3.84e-02	3.75e-02
$\lambda_C = 0.1$					
PDE FDM	8.7067881	8.6742030	8.6456813	8.6274888	8.6120856
asymptotic	8.3318264	8.3322154	8.3323877	8.3324904	8.3325605
difference	3.75e-01	3.42e-01	3.13e-01	2.95e-01	2.80e-01

Table 5.28: Comparison of free boundary locations given by different approaches at various λ_C points for American put option including bilateral XVA with the parameters in Table 5.22, except that κ varies as indicated, and $\sigma^{\lambda_C} = 0.2\sqrt{\kappa}$. The grid size for the PDE solution is $N = 512, M = 256$. Algorithms 6 and 7 are used for the calculation of free boundaries for the (2D) PDE FDM and the asymptotic methods, respectively.

5.3.3 Comparison of American and European type XVA

In order to compare the American and European put options with XVA, we solve the respective problem for European case, considering the same parameter settings, as shown in Table 5.22, and produce results for European options to be compared to the results in Tables 5.23 and 5.24 for American options. The results for European options are presented in Tables 5.29 and 5.30. Comparing the results of Tables 5.23 and 5.27, we easily see that the average number of iterations in each timestep is about 1.1 to 1.2 for European options, which is smaller than the ones for American options. We expect smaller number of iterations for European derivatives, since the American case has more nonlinearity. Also, for American derivatives, the average of number of iterations depends a little more on problem size than for the European case. Furthermore, from Tables 5.23, 5.24, 5.29 and 5.30, we can directly observe that the American derivative including XVA has slightly larger values than the European one, which is expected from the nature of American derivatives.

N	M	N_t	iter tot	iter avg	$\hat{V}$ value for put option		
					value	diff in $\hat{V}$	order
16	8	10	12	1.20	2.0790884	–	–
32	16	18	22	1.22	2.1114901	3.24e-02	–
64	32	34	40	1.18	2.1196343	8.14e-03	1.99
128	64	66	77	1.17	2.1218013	2.17e-03	1.91
256	128	130	150	1.15	2.1223410	5.40e-04	2.01
512	256	258	297	1.15	2.1224747	1.34e-04	2.01
Richardson extrapolated value: 2.1750516							

Table 5.29: Results for European put option including bilateral XVA with stochastic default intensity on counterparty using penalty-like algorithm in Section 3.2 with the parameters in Table 5.22 when S is at-the-money ($S = K = 15$) and $\lambda_C = \theta$. Nonuniform grids are used.

N, M, N_t	(7.5, 0.025)	(7.5, 0.05)	(7.5, 0.1)	(30, 0.01)	(30, 0.05)	(30, 0.01)
128, 66, 68	6.9954401	6.9205456	6.7723182	0.1215826	0.1204373	0.1180971
256,128,130	6.9956602	6.9207611	6.7725231	0.1218132	0.1206653	0.1183202
512,256,258	6.9957152	6.9208149	6.7725743	0.1218705	0.1207220	0.1183757
order	2.00	2.00	2.00	2.01	2.01	2.01

Table 5.30: Results at various points (S, λ_C) for European put option including bilateral XVA with stochastic default intensity on counterparty using penalty-like algorithm in Section 3.2 with the parameters in Table 5.22 at various points. Nonuniform grids are used.

Chapter 6

Conclusions and future works

6.1 Summary and conclusions of research

In this thesis, we explored the XVA pricing problem by PDE approaches and the related computation issues. After the introduction in Chapter 1, Chapter 2 includes the analysis of iterations to deal with the nonlinearity taking place in the XVA pricing PDE. The third and forth chapters of this thesis focus on developing computational methods for the XVA pricing problem with stochastic counterparty default intensity.

In Chapter 2, we formulated and studied iterative methods for the nonlinearity of the PDE problem arising when taking into account the (possibly bilateral) credit risk in valuing financial derivatives. It is shown that simple fixed-point iteration methods for the European XVA PDE problem require 3-4 iterations per timestep, while the proposed discrete penalty iteration converges in less than 1.1 (average) iterations per timestep. For the American XVA PDE problem the proposed double-penalty iteration converges in less than 1.15 (average) iterations per timestep. The main conclusion of this work is that the proposed penalty iteration methods are powerful (nearly optimal) techniques for handling nonlinearities in the context of financial derivative XVA pricing by a PDE model. The penalty methods proposed can be extended to multi-asset financial derivatives pricing, and to stochastic default intensities XVA pricing, once the associated PDE problem is derived. Besides the analysis of different iterative methods for the nonlinearities, regarding the boundary conditions, another conclusion is that the linear far-side condition and its proposed discretization is appropriate for all European and American contingent claims considered.

In Chapter 3, we studied the bilateral XVA pricing of financial derivatives assuming stochastic counterparty default risk, with a focus on numerical computation issues. We developed the corresponding 2D time-dependent PDE, and two approaches to numeri-

cally approximate the options values including XVA. The first approach is to approximate the solution of the 2D time-dependent PDE by numerical PDE techniques. We developed appropriate boundary conditions, used finite differences for space discretization, Crank-Nicolson-Rannacher timestepping, and the penalty iteration method in Chapter 2 to handle the non-linearity of the source term. The second approach is asymptotic approximation, based on singular perturbation theory [40, 26], assuming the stochastic default intensity exhibits fast mean reversion. The asymptotic approximation approach uses the solution to the 1D XVA pricing PDE with constant counterparty default intensity equal to the mean reversion level, and a few correction (expansion) terms that we developed based on the 2D XVA PDE.

The numerical experiments indicate that the numerical 2D PDE approximation converges with stable second order, so that extrapolation can be used as well. Thus, the numerical PDE method can gives us high precision (up to machine precision minus the conditioning of the problem) for any particular problem (set of parameters). The asymptotic approximation agrees quite well with the numerical 2D PDE approximation, and converges with increasing mean reversion rate κ at rate of at least 1.5, thus its precision for a particular problem (set of parameters) is limited. However, the computational cost of the asymptotic method is substantially lower than that of the numerical 2D PDE method, thus, for reasonable accuracy requirements, the asymptotic method is an attractive alternative to the numerical 2D PDE method, and particularly handy for practitioners, due to its simplicity.

Chapter 4 of the thesis formulates and studies, from a perspective of computation issues, the American type bilateral XVA pricing of financial derivatives, especially American put options, assuming stochastic counterparty default intensity for the counterparty. We formulated a 2D time-dependent linear complementarity problem PDE, reformulated it into a 2D time-dependent PDE with multiple nonlinear source terms, and developed two approaches to numerically approximate the option's value. For the numerical PDE approximation, we applied appropriate boundary conditions, used finite differences scheme for the spatial discretization, Crank-Nicolson-Rannacher timestepping for the time discretization, and double-penalty iteration to handle the double nonlinearity (one from XVA and another one from the American constraints) in the source terms. For the asymptotic approximation, we extended the asymptotic approximation formulation of XVA in European derivatives in Chapter 3 to American derivatives, as well as developed the asymptotic approximation of the free boundary location. The asymptotic approximated price needs the solution to the 1D American XVA PDE with constant counterparty default intensity, and a few more corrections terms. The asymptotic ap-

proximated free boundary needs the free boundary given by the same 1D American XVA PDE with constant counterparty default intensity, plus one more correction. The numerical experiments show that both techniques agree quite well with reasonable accuracy.

6.2 Future work

Several possible extensions to the research work in this thesis are listed below.

6.2.1 Multi-dimensional problems and neural networks

In reality, the XVA pricing problem often involves multiple underlying assets, which results in multiple spatial dimensions. The traditional numerical methods often suffer from the curse of dimensionality. A possible extension of this research work could be the investigation of modern computational methods, such as neural networks approximation in CVA or XVA pricing, especially when facing high-dimensional (more than 5 spatial dimensions) computational problems. In [43], authors proposed an algorithm called "Deep Galerkin Method" to solve high-dimensional PDEs by approximating the solution with a deep neural network which the objective function is based on the residual of the differential operator, initial condition, and boundary conditions. Their algorithm is mesh-free and trained on batches of randomly sampled time and spatial points. In [32], authors proposed an simulation based deep learning method to solve the high-dimensional PDE. In their algorithm, they need to reformulate the high-dimensional parabolic PDEs into the corresponding BSDEs. In [27], authors applied the ideas from [32] to solve the XVA problem which can potentially handle a large number of risk factors, which is called the "deep xVA solver". There are also some other Monte Carlo based deep learning methods that researchers use to derivative pricing problem or PDEs solving problem with the help of Feynman-Kac theorem. For example, authors in [20] directly train a neural network to approximate the derivatives valuation function by generating lot of paths from the factor dynamics.

6.2.2 Analysis of asymptotic approximation in American XVA problem

For American type XVA problem, it is interesting to study the error of the asymptotic formula of the free boundary approximation, as the range of the untrusted area around approximated free boundary is important in the XVA pricing problem. It is also in-

teresting to improve the accuracy of the asymptotic approximation of free boundary to the same accuracy level of the asymptotic approximation of price, which is at least $\mathcal{O}(\epsilon^{3/2})$. This may need one or two more correction terms to our existing asymptotic approximation.

Bibliography

[1] A. Agarwal, S. Juneja, and R. Sircar, *American options under stochastic volatility: control variates, maturity randomization & multiscale asymptotics*, Quantitative Finance, 16 (2016), pp. 17–30.

[2] I. Arregui, B. Salvador, D. Ševčovič, and C. Vázquez, *Total value adjustment for European options with two stochastic factors: Mathematical model, analysis and numerical simulation*, Computers and Mathematics with Applications, 76 (2018), pp. 725–740.

[3] I. Arregui, B. Salvador, and C. Vázquez, *CVA computing by PDE models*, in NAA16: Numerical Analysis and Its Applications, I. Dimov et al, ed., LNCS 10187, Springer, 2017, pp. 15–24.

[4] I. Arregui, B. Salvador, and C. Vázquez, *PDE models and numerical methods for total value adjustment in European and American options with counterparty risk*, Applied Mathematics and Computation, 308 (2017), pp. 31–53.

[5] ——, *A Monte Carlo approach to American options pricing including counterparty risk*, International Journal of Computer Mathematics, 96 (2019), pp. 2157–2176.

[6] M. Bichuch, A. Capponi, and S. Sturm, *Arbitrage-free XVA*, Mathematical Finance, 28 (2018), pp. 582–620.

[7] D. Brigo and A. Capponi, *Bilateral counterparty risk valuation with stochastic dynamical models and application to credit default swaps*, https://arxiv.org/pdf/0812.3705, (2008).

[8] ——, *Bilateral counterparty risk with application to CDSs*, Risk, 23 (2010), pp. 85–90.

[9] D. Brigo, A. Capponi, and A. Pallavicini, *Arbitrage-free bilateral counterparty risk valuation under collateralization and application to credit default swaps*,

Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics, 24 (2014), pp. 125–146.

[10] C. BURGARD AND M. KJAER, *In the balance*, Risk, 24 (2011), pp. 72–75.

[11] C. BURGARD AND M. KJAER, *Partial differential equation representations of derivatives with bilateral counterparty risk and funding costs*, The Journal of Credit Risk, 7 (2011), pp. 75–93.

[12] C. BURGARD AND M. KJAER, *The FVA debate: In theory and practice*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2157634, (2012).

[13] ——, *Funding strategies, funding costs*, Risk, 26 (2013), pp. 82–87.

[14] Y. CHEN AND C. C. CHRISTARA, *Penalty methods for bilateral XVA pricing in European and American contingent claims by a PDE model*, Journal of Computational Finance, 24 (2021), pp. 41–70.

[15] ——, *Bilateral XVA pricing under stochastic default intensity: PDE modelling and computation*, Applied Numerical Mathematics, (2023). To appear.

[16] C. CHRISTARA AND D. M. DANG, *Adaptive and high-order methods for valuing American options*, J. Comput. Finance, 14 (2011), pp. 73–113.

[17] S. CRÉPEY, *Bilateral counterparty risk under funding constraints, part II: CVA*, Mathematical Finance, 25 (2015), pp. 23–50.

[18] D. DUFFIE AND N. GARLEANU, *Risk and valuation of collateralized debt obligations*, Financial analysts journal, 57 (2001), pp. 41–59.

[19] Y. FENG, *CVA under Bates model with stochastic default intensity*, Journal of Mathematical Finance, 7 (2017), pp. 682–698.

[20] R. FERGUSON AND A. GREEN, *Deeply learning derivatives*, arXiv preprint arXiv:1809.02233, (2018).

[21] P. A. FORSYTH AND G. LABAHN, *Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance*, J. Comput. Finance, 11 (2007), pp. 1–43.

[22] P. A. FORSYTH AND K. R. VETZAL, *Quadratic convergence for valuing American options using a penalty method*, SIAM J. Sci. Comput., 23 (2002), pp. 2095–2122.

[23] J.-P. FOUQUE, M. LORIG, AND R. SIRCAR, *Second order multiscale stochastic volatility asymptotics: stochastic terminal layer analysis and calibration*, Finance and Stochastics, 20 (2016), pp. 543–588.

[24] J.-P. FOUQUE, G. PAPANICOLAOU, AND K. R. SIRCAR, *Derivatives in financial markets with stochastic volatility*, Cambridge University Press, 2000.

[25] ——, *From the implied volatility skew to a robust correction to Black-Scholes American option prices*, International Journal of Theoretical and Applied Finance, 4 (2001), pp. 651–675.

[26] J. P. FOUQUE, G. PAPANICOLAOU, R. SIRCAR, AND K. SOLNA, *Singular perturbations in option pricing*, SIAM J. Appl. Math., 63 (2003), pp. 1648–1665.

[27] A. GNOATTO, A. PICARELLI, AND C. REISINGER, *Deep xVA solver – A neural network based counterparty credit risk management framework*, arXiv preprint arXiv:2005.02633v3, (2021).

[28] A. GREEN, *XVA: Credit, Funding and Capital Valuation Adjustments*, Wiley, 2015.

[29] J. GREGORY, *Counterparty credit risk and credit value adjustment: A continuing challenge for global financial markets*, John Wiley & Sons, 2012.

[30] J. GREGORY AND I. GERMAN, *Closing out DVA?*, Risk, 26 (2013), pp. 96–100.

[31] T. HAENTJENS AND K. J. IN'T HOUT, *ADI schemes for pricing American options under the Heston model*, Applied Mathematical Finance, 22 (2015), pp. 207–237.

[32] J. HAN, A. JENTZEN, AND W. E, *Solving high-dimensional partial differential equations using deep learning*, Proceedings of the National Academy of Sciences, 115 (2018), pp. 8505–8510.

[33] S. HESTON, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Review of Financial Studies, 6 (1993), pp. 327–343.

[34] J. HULL AND A. WHITE, *CVA and wrong-way risk*, Financial Analysts Journal, 68 (2012), pp. 58–69.

[35] ——, *The FVA debate*, Risk, 25 (2012), pp. 83–85.

[36] J. C. HULL, *Options, Futures, and Other Derivatives*, Pearson, 10 ed., 2017.

[37] K. J. IN'T HOUT AND S. FOULON, *ADI finite difference schemes for option pricing in the Heston model with correlation*, International Journal of Numerical Analysis and Modeling, 7 (2010), pp. 303–320.

[38] K. ITO, C. REISINGER, AND Y. ZHANG, *A neural network based policy iteration algorithm with global H^2-superlinear convergence for stochastic games on domains*, Foundations of Computational Mathematics, 21 (2021), pp. 331–374.

[39] R. KANGRO AND R. NICOLAIDES, *Far field boundary conditions for Black-Scholes equations*, SIAM J. Numer. Anal., 38 (2000), pp. 1357–1368.

[40] N. C.-H. LEUNG, C. CHRISTARA, AND D. M. DANG, *Partial differential equation pricing of contingent claims under stochastic correlation*, SIAM J. Sci. Comput., 40 (2018), pp. B1–B31.

[41] G. MUÑOZ AND L. MANUEL, *CVA, FVA (and DVA?) with stochastic spreads. a feasible replication approach under realistic assumptions.*, Tech. Rep. 44568, Munich Personal RePEc Archive, 2013. `https://mpra.ub.uni-muenchen.de/44568/8/MPRA_paper_44568.pdf`.

[42] V. PITERBARG, *Funding beyond discounting: collateral agreements and derivatives pricing*, Risk, 23 (2010), pp. 42–47.

[43] J. SIRIGNANO AND K. SPILIOPOULOS, *DGM: A deep learning algorithm for solving partial differential equations*, Journal of computational physics, 375 (2018), pp. 1339–1364.

[44] S. H. M. TING, *Asymptotic techniques and stochastic volatility in option pricing problems*, PhD thesis, School of Mathematics and Statistics, University of Sydney, Australia, 2012. `http://hdl.handle.net/2123/8751`.

[45] D. WANG, K. SERKH, AND C. CHRISTARA, *A high-order deferred correction method for the solution of free boundary problems using penalty iteration, with an application to American option pricing*, arXiv preprint arXiv:2205.00617v4, (2022).

[46] M. WIDDICKS, P. W. DUCK, A. D. ANDRICOPOULOS, AND D. P. NEWTON, *The Black-Scholes equation revisited: Asymptotic expansions and singular perturbations*, Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics, 15 (2005), pp. 373–391.

[47] P. WILMOTT, S. HOWISON, AND J. DEWYNNE, *The Mathematics of Financial Derivatives*, Cambridge University Press, 1995.

[48] H. WINDCLIFF, P. A. FORSYTH, AND K. R. VETZAL, *Analysis of the stability of the linear boundary condition for the Black-Scholes equation*, J. Comput. Finance, 8 (2004), pp. 65–92.

[49] S.-P. ZHU AND W.-T. CHEN, *A predictor-corrector scheme based on the ADI method for pricing American puts with stochastic volatility*, Computers and Mathematics with Applications, 62 (2011), pp. 1–26.

[50] R. ZVAN, P. A. FORSYTH, AND K. R. VETZAL, *Penalty methods for American options with stochastic volatility*, Journal of Computational and Applied Mathematics, 91 (1998), pp. 199–218.