Bilateral XVA pricing under stochastic default intensity: PDE modelling and computation

Yuwei Chen and Christina C. Christara
Department of Computer Science
University of Toronto
Toronto, Ontario M5S 2E4, Canada
{ywchen,ccc}@cs.toronto.edu

Abstract

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. We derive a novel partial differential equation (PDE) model for derivative pricing including the adjustment for default risk, assuming that the default risk of one of the counterparties (the buyer) 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 is based on standard finite differences, and a penalty-like iteration for handling the nonlinearity. We also develop and analyze a novel asymptotic approximation formula for the adjusted price of derivatives, resulting in a very efficient, accurate, and convenient for practitioners formula. We present numerical results that indicate stable second order convergence for the 2D PDE solution in terms of the discretization size. We compare the effectiveness of the 2D PDE and asymptotic approximations. We study the effect of various numerical and market parameters to the values of the adjusted prices and to the accuracy of the computed solutions.

AMS subject classification: 65M06, 65M12, 91G20, 91G40, 91G60.

Key words: Partial Differential Equations, Black-Scholes, Crank-Nicolson finite differences, mean reversion CIR process, stochastic default intensity, asymptotic approximation.

1 Introduction

Counterparty default risk [14, 16], is the risk that, one of the counterparties in an agreement fails to fulfill their contractual obligation. Taking into account default risk in derivative pricing, results in adjusted derivative values. Usually, the total valuation adjustment (XVA) includes several different types of adjustments. For example, credit valuation adjustment (CVA) and debt valuation adjustment (DVA) refer to the adjustments due to the buyers and sellers' (counterparty or self-party's) default risks, respectively. Funding valuation adjustment (FVA) makes sure that a seller recovers their average funding costs when trading.

One approach to XVA valuation is via partial differential equation (PDE) formulation. The PDE approach usually gives more accurate valuation, faster convergence for problems in low-dimensions, and more accurate hedging parameters computation. Piterbarg [20] derived the PDE formulation for the price of derivatives, including funding costs and collateral agreements. Burgard and Kjaer [3, 4, 5, 6], extended the PDE to include bilateral default risks. They built a replication portfolio, which included bonds of counterparties, to derive the PDE model for derivatives' values taking into account bilateral risks and funding costs. The PDE arising often involves a nonlinear source term. There are several methods in the literature to numerically solve the XVA PDE problem. In [2], the authors discretize the space variable using a finite element method, apply the characteristic method for timestepping, and employ fixed-point

iteration for the nonlinearity. In [7], we connected the nonlinear problem with a control problem, and developed penalty-like methods to efficiently treat the nonlinearity, for both European and American derivatives.

However, the default intensities in these works are assumed to be constant. In reality, default intensities exhibit stochasticity [8]. 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. Therefore, a multi-stochastic factors' model becomes necessary to reflect these issues. In [9], 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 [19, 1], the authors considered the stochastic short term credit default swap (CDS) spread and resulted in a multi-dimensional in space PDE. In [1], existence and uniqueness of the solution to the nonlinear PDE was proved as well. In our work, 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.

In this paper, 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 the paper 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.

- We present numerical results indicating the second-order convergence of the computed 2D PDE solution, and study the effect of various numerical parameters to its accuracy. We study how the speed of mean reversion affects the quality of the asymptotic approximation. We compare the accuracy of the asymptotic and 2D PDE solutions. Furthermore, we study the effect of various model parameters to the adjusted derivative value and make sure it is consistent for both the 2D PDE and the asymptotic solution.

The outline of the paper is as follows. In Section 2, we derive 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, 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. In Section 4, we present numerical experiments to study the behavior of the computed 2D PDE and asymptotic solutions, in terms of various numerical and model

parameters, and compare the two solution techniques. Section 5 summarizes the conclusions of this work. In the Appendix, we present an analysis of the accuracy of the asymptotic approximation.

2 Formulation

2.1 Assumptions

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

$$dS(t) = \mu(t)S(t)dt + \sigma^S S(t)dW^S(t) \tag{1}$$
$$dP_B = P_B(r_B(t)dt - dJ_B) \tag{2}$$
$$dP_C = P_C(r_C(t)dt - dJ_C) \tag{3}$$
$$dP(t) = r(t)P(t)dt \tag{4}$$

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 the bond rates r_B and r_C are related to the respective default intensities λ_B and λ_C by $r_B = \lambda_B - r$ and $r_C = \lambda_C - r$, respectively.

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

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

$$d\lambda_C(t) = \kappa(t)[\theta(t) - \lambda_C(t)]dt + \sigma^{\lambda_C}\sqrt{\lambda_C(t)}dW^{\lambda_C}(t) \tag{5}$$

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) [8]. 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.

2.3 Formulation of PDE

To formulate the PDE for the XVA pricing problem considering stochastic default intensities, we use dynamic hedging techniques similar to [4]. However, in [4], 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 [12] or the stochastic correlation option pricing model [18].

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:

- $\gamma(t)$ 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 [4], 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

$$\hat{V}(t, S, \lambda_C, 1, 0) = M^+ + R_B M^-, \tag{6}$$

$$\hat{V}(t, S, \lambda_C, 0, 1) = R_C M^+ + M^-, \tag{7}$$

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. The positive and negative values of any asset U are denoted as $U^+ \equiv \max\{U, 0\}$ and $U^- \equiv \min\{U, 0\}$. In this paper, $\hat{V}(t, S, \lambda_C, 0, 0)$ is usually written as $\hat{V}(t, S, \lambda_C)$.

The value of Π is written as

$$-\hat{V} = \Pi = \delta(t)S(t) + \alpha(t)P_B + \beta(t)P_C + \gamma(t)\tilde{V} + D(t). \tag{8}$$

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

$$-d\hat{V} = d\Pi = \delta(t)dS(t) + \alpha(t)dP_B + \beta(t)dP_C + \gamma(t)d\tilde{V} + d\bar{D}(t) \tag{9}$$

where the change in the cash account [1] is

$$d\bar{D}(t) = \delta(t)(-r_R)S(t)dt + \{r(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^+ + r_F(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^-\}dt - r\beta(t)P_C dt. \tag{10}$$

[1] More details about the mechanism of cash account can found in [4]

Then the first term [2] in (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

$$d\hat{V} = \mathcal{M}\hat{V}dt + (\sigma^S)S\frac{\partial \hat{V}}{\partial S}dW^S + (\sigma^{\lambda_C})\sqrt{\lambda_C}\frac{\partial \hat{V}}{\partial \lambda_C}dW^{\lambda_C} + \Delta\hat{V}_B dJ_B + \Delta\hat{V}_C dJ_C, \tag{11}$$

$$d\tilde{V} = \mathcal{M}\tilde{V}dt + (\sigma^S)S\frac{\partial \tilde{V}}{\partial S}dW^S + (\sigma^{\lambda_C})\sqrt{\lambda_C}\frac{\partial \tilde{V}}{\partial \lambda_C}dW^{\lambda_C} + \Delta\tilde{V}_B dJ_B + \Delta\tilde{V}_C dJ_C, \tag{12}$$

where

$$\Delta\hat{V}_B = \hat{V}(t,S,\lambda_C,1,0) - \hat{V}(t,S,\lambda_C), \tag{13}$$

$$\Delta\tilde{V}_B = \tilde{V}(t,S,\lambda_C,1,0) - \tilde{V}(t,S,\lambda_C), \tag{14}$$

$$\Delta\hat{V}_C = \hat{V}(t,S,\lambda_C,0,1) - \hat{V}(t,S,\lambda_C), \tag{15}$$

$$\Delta\tilde{V}_C = \tilde{V}(t,S,\lambda_C,0,1) - \tilde{V}(t,S,\lambda_C), \tag{16}$$

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

$$\mathcal{M}V = \frac{\partial V}{\partial t} + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2 V}{\partial \lambda_C^2} + \rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2 V}{\partial S\partial\lambda_C} + \mu S\frac{\partial V}{\partial S} + \kappa(\theta-\lambda_C)\frac{\partial V}{\partial \lambda_C}. \tag{17}$$

Combining equations (9)-(17), the following hedging equation is obtained,

$$\begin{aligned}
&-\left\{\left[\frac{\partial \hat{V}}{\partial t} + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2 \hat{V}}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2 \hat{V}}{\partial \lambda_C^2} + \rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2 \hat{V}}{\partial S\partial\lambda_C} + \mu S\frac{\partial \hat{V}}{\partial S} + \kappa(\theta-\lambda_C)\frac{\partial \hat{V}}{\partial \lambda_C}\right]dt\right.\\
&\left. + (\sigma^S)S\frac{\partial \hat{V}}{\partial S}dW^S + (\sigma^{\lambda_C})\sqrt{\lambda_C}\frac{\partial \hat{V}}{\partial \lambda_C}dW^{\lambda_C} + \Delta\hat{V}_B dJ_B + \Delta\hat{V}_C dJ_C\right\}\\
&= \delta(t)\left(\mu S dt + \sigma^S S dW^S\right) + \alpha(t)P_B(r_B dt - dJ_B) + \beta(t)P_C(r_C dt - dJ_C)\\
&+ \gamma(t)\left\{\left[\frac{\partial \tilde{V}}{\partial t} + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2 \tilde{V}}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2 \tilde{V}}{\partial \lambda_C^2} + \rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2 \tilde{V}}{\partial S\partial\lambda_C} + \mu S\frac{\partial \tilde{V}}{\partial S} + \kappa(\theta-\lambda_C)\frac{\partial \tilde{V}}{\partial \lambda_C}\right]dt\right.\\
&\left. + (\sigma^S)S\frac{\partial \tilde{V}}{\partial S}dW^S + (\sigma^{\lambda_C})\sqrt{\lambda_C}\frac{\partial \tilde{V}}{\partial \lambda_C}dW^{\lambda_C} + \Delta\tilde{V}_B dJ_B + \Delta\tilde{V}_C dJ_C\right\}\\
&- \delta(t)r_R S(t)dt + \{r(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^+ + r_F(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^-\}dt - r\beta P_C dt.
\end{aligned} \tag{18}$$

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

In order to remove all the risk factors in (18), the following equations must be satisfied:

$$-(\sigma^{\lambda_C})\sqrt{\lambda_C}\frac{\partial \hat{V}}{\partial \lambda_C}dW^{\lambda_C} = \gamma(t)(\sigma^{\lambda}_C)\sqrt{\lambda_C}\frac{\partial \tilde{V}}{\partial \lambda_C}dW^{\lambda_C}, \tag{19}$$

$$-(\sigma^S)S\frac{\partial \hat{V}}{\partial S}dW^S = \gamma(t)\left((\sigma^S)S\frac{\partial \tilde{V}}{\partial S}dW^S\right) + \delta(t)(\sigma^S)SdW^S, \tag{20}$$

$$-\Delta\hat{V}_B dJ_B = -\alpha(t)P_B dJ_B + \gamma(t)\Delta\tilde{V}_B dJ_B \tag{21}$$

$$-\Delta\hat{V}_C dJ_C = -\beta(t)P_C dJ_C + \gamma(t)\Delta\tilde{V}_C dJ_C. \tag{22}$$

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

$$\gamma(t) = -\frac{\partial \hat{V}}{\partial \lambda_C}\Big/\frac{\partial \tilde{V}}{\partial \lambda_C}, \tag{23}$$

$$\delta(t) = -\gamma(t)\frac{\partial \tilde{V}}{\partial S} - \frac{\partial \hat{V}}{\partial S} = (\frac{\partial \hat{V}}{\partial \lambda_C}\Big/\frac{\partial \tilde{V}}{\partial \lambda_C})\frac{\partial \tilde{V}}{\partial S} - \frac{\partial \hat{V}}{\partial S}, \tag{24}$$

$$\alpha(t) = \frac{1}{P_B}\left[\Delta\hat{V}_B + \gamma(t)\Delta\tilde{V}_B\right] = \frac{1}{P_B}\left[\Delta\hat{V}_B - (\frac{\partial \hat{V}}{\partial \lambda_C}\Big/\frac{\partial \tilde{V}}{\partial \lambda_C})\Delta\tilde{V}_B\right], \tag{25}$$

$$\beta(t) = \frac{1}{P_C}\left[\Delta\hat{V}_C + \gamma(t)\Delta\tilde{V}_C\right] = \frac{1}{P_C}\left[\Delta\hat{V}_C - (\frac{\partial \hat{V}}{\partial \lambda_C}\Big/\frac{\partial \tilde{V}}{\partial \lambda_C})\Delta\tilde{V}_C\right], \tag{26}$$

where

$$\Delta\hat{V}_B = \hat{M}^+ + R_B\hat{M}^- - \hat{V}, \tag{27}$$

$$\Delta\hat{V}_C = R_C\hat{M}^+ + \hat{M}^- - \hat{V}, \tag{28}$$

$$\Delta\tilde{V}_B = \tilde{M}^+ + R_B\tilde{M}^- - \tilde{V}, \tag{29}$$

$$\Delta\tilde{V}_C = R_C\tilde{M}^+ + \tilde{M}^- - \tilde{V}, \tag{30}$$

$\hat{M}$ is close-out value for derivative $\hat{V}$ and $\tilde{M}$ is close-out value for derivative $\tilde{V}$.
Therefore, (18) becomes

$$\begin{aligned}
&-\left[\frac{\partial \hat{V}}{\partial t} + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2 \hat{V}}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2 \hat{V}}{\partial \lambda_C^2} + \rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2 \hat{V}}{\partial S\partial\lambda_C}\right]dt \\
&= \alpha(t)P_B r_B dt + \beta(t)P_C r_C dt + \gamma(t)\left[\frac{\partial \tilde{V}}{\partial t} + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2 \tilde{V}}{\partial S^2} + \frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2 \tilde{V}}{\partial \lambda_C^2} + \rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2 \tilde{V}}{\partial S\partial\lambda_C}\right]dt \\
&- \delta(t)r_R S(t)dt + \{r(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^+ + r_F(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^-\}dt - r\beta P_C dt.
\end{aligned} \tag{31}$$

Rearranging some terms in (31) results in

$$\begin{aligned}
&\alpha(t)P_B r_B dt + \beta(t)P_C r_C dt - r\beta P_C dt \\
&\quad + \{r(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^+ + r_F(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^-\}dt \\
&= \alpha(t)P_B r_B dt + \beta(t)P_C r_C dt - r\beta P_C dt \\
&\quad + r(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})dt + s_F(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^- dt \\
&= \alpha(t)P_B(r_B - r)dt + \beta(t)P_C(r_C - r)dt + r(-\hat{V} - \gamma(t)\tilde{V})dt + s_F(-\hat{V} - \alpha(t)P_B - \gamma(t)\tilde{V})^- dt.
\end{aligned} \tag{32}$$

Taking into account (32), equation (31) becomes

$$
\begin{aligned}
&-\left[\frac{\partial \hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}\right]dt\\
=&\gamma(t)\left[\frac{\partial \tilde{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\tilde{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\tilde{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\tilde{V}}{\partial S\partial\lambda_C}\right]dt\\
&-\delta(t)r_RSdt+\alpha(t)P_B(r_B-r)dt+\beta(t)P_C(r_C-r)dt\\
&+r(-\hat{V}-\gamma(t)\tilde{V})dt+s_F(-\hat{V}-\alpha(t)P_B-\gamma(t)\tilde{V})^-dt.
\end{aligned}
\tag{33}
$$

If we substitute (24)-(26) into (33), we obtain

$$
\begin{aligned}
&-\left[\frac{\partial \hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}-r\hat{V}\right]dt\\
=&\gamma(t)\left[\frac{\partial \tilde{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\tilde{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\tilde{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\tilde{V}}{\partial S\partial\lambda_C}-r\tilde{V}\right]dt\\
&+[\gamma(t)\frac{\partial\tilde{V}}{\partial S}+\frac{\partial\hat{V}}{\partial S}]r_RSdt+(\Delta\hat{V}_B+\gamma(t)\Delta\tilde{V}_B)(r_B-r)dt+(\Delta\hat{V}_C+\gamma(t)\Delta\tilde{V}_C)(r_C-r)dt\\
&+s_F(-\hat{V}-(\Delta\hat{V}_B+\gamma(t)\Delta\tilde{V}_B)-\gamma(t)\tilde{V})^-dt.
\end{aligned}
\tag{34}
$$

Then, if we rearrange (34) and switch the sign on both hand sides, we have

$$
\begin{aligned}
&\left[\frac{\partial \hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}-r\hat{V}+r_RS\frac{\partial\hat{V}}{\partial S}+\lambda_B\Delta\hat{V}_B+\lambda_C\Delta\hat{V}_C\right]\\
=&(-\gamma(t))\left[\frac{\partial \tilde{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\tilde{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\tilde{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\tilde{V}}{\partial S\partial\lambda_C}-r\tilde{V}+r_RS\frac{\partial\tilde{V}}{\partial S}\right.\\
&\left.+\lambda_B\Delta\tilde{V}_B+\lambda_C\Delta\hat{V}_C\right]+s_F((\Delta\hat{V}_B+\tilde{V})+\gamma(t)(\Delta\tilde{V}_B+\tilde{V}))^+.
\end{aligned}
\tag{35}
$$

If we substitute (27)-(30) into (35), we obtain

$$
\begin{aligned}
&\left[\frac{\partial \hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}-r\hat{V}+r_RS\frac{\partial\hat{V}}{\partial S}\right.\\
&\left.-\lambda_B(\hat{M}^+ + R_B\hat{M}^- - \hat{V})-\lambda_C(R_C\hat{M}^+ + \hat{M}^- - \hat{V})\right]\\
=&(-\gamma(t))\left[\frac{\partial \tilde{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\tilde{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\tilde{V}}{\partial\lambda_C^2}+\rho\sigma^S\sigma_C^{\lambda}S\sqrt{\lambda_C}\frac{\partial^2\tilde{V}}{\partial S\partial\lambda_C}-r\tilde{V}+r_RS\frac{\partial\tilde{V}}{\partial S}\right.\\
&\left.-\lambda_B(\tilde{M}^+ + R_B\tilde{M}^- - \tilde{V})-\lambda_C(R_C\tilde{M}^+ + \tilde{M}^- - \tilde{V})\right]\\
&+s_F((\hat{M}^+ + R_B\hat{M}^-)+\gamma(t)(\tilde{M}^+ + R_B\tilde{M}^-))^+.
\end{aligned}
\tag{36}
$$

If we 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, using (23), we get

$$
\begin{aligned}
&(1\Big/\frac{\partial \hat{V}}{\partial \lambda_C})\Big[\frac{\partial \hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial \lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}-r\hat{V}+r_RS\frac{\partial\hat{V}}{\partial S}\\
&+\lambda_B(\hat{M}^+ + R_B\hat{M}^- - \hat{V})+\lambda_C(R_C\hat{M}^+ + \hat{M}^- - \hat{V})-s_F\hat{M}^+\Big]\\
=&(1\Big/\frac{\partial \tilde{V}}{\partial \lambda_C})\Big[\frac{\partial \tilde{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\tilde{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\tilde{V}}{\partial \lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\tilde{V}}{\partial S\partial\lambda_C}-r\tilde{V}+r_RS\frac{\partial\tilde{V}}{\partial S}\\
&+\lambda_B(\tilde{M}^+ + R_B\tilde{M}^- - \tilde{V})+\lambda_C(R_C\tilde{M}^+ + \tilde{M}^- - \tilde{V})-s_F\tilde{M}^+\Big].
\end{aligned}
\tag{37}
$$

Equation (37) holds for any derivatives $\hat{V}$ and $\tilde{V}$. Assuming that both sides of (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

$$
\begin{aligned}
&\frac{\partial \hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial \lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}\\
&\qquad +r_RS\frac{\partial\hat{V}}{\partial S}+\kappa[\theta-\lambda_C]\frac{\partial\hat{V}}{\partial\lambda_C}-(r+\lambda_B+\lambda_C)\hat{V}\\
&\quad =-\lambda_B(\hat{M}^+ + R_B\hat{M}^-)-\lambda_C(R_C\hat{M}^+ + \hat{M}^-)+s_F\hat{M}^+\\
&\quad =(s_F-\lambda_B-R_C\lambda_C)\hat{M}^+ + (-R_B\lambda_B-\lambda_C)\hat{M}^-.
\end{aligned}
\tag{38}
$$

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

$$
\begin{aligned}
&\frac{\partial \hat{V}}{\partial t}+\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial \lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}\\
&\qquad +r_RS\frac{\partial\hat{V}}{\partial S}+\kappa[\theta-\lambda_C]\frac{\partial\hat{V}}{\partial\lambda_C}-r\hat{V}\\
&\quad =(s_F+(1-R_C)\lambda_C)\hat{V}^+ + (1-R_B)\lambda_B\hat{V}^-.
\end{aligned}
\tag{39}
$$

If backward time $\tau=T-t$ is applied, (39) becomes

$$
\frac{\partial \hat{V}}{\partial \tau}=\mathcal{L}\hat{V}+f(\lambda_C,\hat{V}),
\tag{40}
$$

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

$$
\begin{aligned}
\mathcal{L}\hat{V} \equiv&\frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2\hat{V}}{\partial S^2}+\frac{1}{2}(\sigma^{\lambda_C})^2\lambda_C\frac{\partial^2\hat{V}}{\partial \lambda_C^2}+\rho\sigma^S\sigma^{\lambda_C}S\sqrt{\lambda_C}\frac{\partial^2\hat{V}}{\partial S\partial\lambda_C}\\
&+r_RS\frac{\partial\hat{V}}{\partial S}+\kappa[\theta-\lambda_C]\frac{\partial\hat{V}}{\partial\lambda_C}-r\hat{V},
\end{aligned}
\tag{41}
$$

$$
f(\lambda_C,\hat{V}) \equiv -(s_F+(1-R_C)\lambda_C)\hat{V}^+ - (1-R_B)\lambda_B\hat{V}^-.
\tag{42}
$$

The initial (or terminal) condition to (40) is the pay-off 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.

2.4 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:

$$\frac{\partial \hat{V}^c}{\partial \tau} = \frac{1}{2}(\sigma^S)^2 S^2 \frac{\partial^2 \hat{V}^c}{\partial S^2} + r_R S \frac{\partial \hat{V}^c}{\partial S} - r\hat{V}^c - (s_F + (1-R_C)\lambda_C)(\hat{V}^c)^+ - (1-R_B)\lambda_B(\hat{V}^c)^-. \quad (43)$$

PDE (43) is derived in [4] and numerically solved in [7].

3 Numerical solution

PDE (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 (40), special care is still needed to introduce appropriate boundary conditions and to deal with the nonlinearity in the source term.

3.1 Discretizations

In this subsection, we present the discretization of (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}. 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 Section 5. Standard second-order centered finite differences are used for the space discretization of (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 [3], 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 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:

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

- On the $S = 0$ boundary, i.e. on $\{(S, \lambda_C) \in \{S = 0\} \times [0, \lambda_C^{max}]\}$, substitute $S = 0$ into (40). This results in a one-dimensional time-dependent PDE,

$$\frac{\partial \hat{V}}{\partial \tau} = \frac{1}{2}(\sigma^{\lambda_C})^2 \lambda_C \frac{\partial^2 \hat{V}}{\partial \lambda_C^2} + \kappa[\theta - \lambda_C]\frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{44}$$

  PDE (44) is numerically solved and its computed solution used as Dirichlet boundary condition for solving (40). Standard centered differences are used for the discretization of (44). However, in order to numerically solve (44), appropriate boundary conditions are needed. On the corner point $(0, 0)$, we substitute $\lambda_C = 0$ into (44), and get

$$\frac{\partial \hat{V}}{\partial \tau} = \kappa\theta \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{45}$$

  On the corner point $(0, \lambda_C^{max})$, we substitute $\frac{\partial^2 \hat{V}}{\partial \lambda_C^2} = 0$ into (44), and get

$$\frac{\partial \hat{V}}{\partial \tau} = \kappa[\theta - \lambda_C]\frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{46}$$

  Relations (45) and (46) provide the near and far field boundary conditions, respectively, for (44). The first derivative terms of (45) and (46) are discretized by one-sided finite differences (forward or backward, respectively).
  PDE (44) together with boundary conditions (45) and (46) provide approximations to $\hat{V}(\tau, 0, \lambda_C)$ to be used as Dirichlet boundary condition for solving (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

$$\frac{\partial^2 \hat{V}}{\partial S^2} = 0. \tag{47}$$

  Instead of discretizing this condition directly, we substitute $\frac{\partial^2 \hat{V}}{\partial S^2} = 0$ into the PDE (40), and get

$$\frac{\partial \hat{V}}{\partial \tau} = \frac{1}{2}(\sigma^{\lambda_C})^2 \lambda_C \frac{\partial^2 \hat{V}}{\partial \lambda_C^2} + \rho\sigma^S\sigma^{\lambda_C} S\sqrt{\lambda_C}\frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C} + r_R S\frac{\partial \hat{V}}{\partial S} + \kappa[\theta - \lambda_C]\frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{48}$$

  Relation (48) forms the boundary condition at $S = S^{max}$. The first derivative term $\frac{\partial \hat{V}}{\partial S}$ in (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 (48) becomes

$$\frac{\partial \hat{V}}{\partial \tau} = r_R S\frac{\partial \hat{V}}{\partial S} + \kappa\theta \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}), \tag{49}$$

  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 (47), we also impose the linear boundary condition on λ_C

$$\frac{\partial^2 \hat{V}}{\partial \lambda_C^2} = 0. \tag{50}$$

With (47) and (50), PDE (40) becomes

$$\frac{\partial \hat{V}}{\partial \tau} = \rho \sigma^S \sigma^{\lambda_C} S \sqrt{\lambda_C} \frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C} + r_R S \frac{\partial \hat{V}}{\partial S} + \kappa[\theta - \lambda_C] \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{51}$$

For the discretization of $\frac{\partial \hat{V}}{\partial S}$ and $\frac{\partial \hat{V}}{\partial \lambda_C}$ in (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 [17, 22, 15], or correlation [18]. 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 (40), and get

$$\frac{\partial \hat{V}}{\partial \tau} = \frac{1}{2}(\sigma^S)^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + r_R S \frac{\partial \hat{V}}{\partial S} + \kappa\theta \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{52}$$

  This PDE is actually same as the original PDE (40) on $\lambda_C = 0$ boundary. Thus, we do not introduce a new condition, but just apply the PDE itself. In (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 (40), and get

$$\frac{\partial \hat{V}}{\partial \tau} = \frac{1}{2}(\sigma^S)^2 S^2 \frac{\partial^2 \hat{V}}{\partial S^2} + \rho \sigma^S \sigma^{\lambda_C} S \sqrt{\lambda_C} \frac{\partial^2 \hat{V}}{\partial S \partial \lambda_C} + r_R S \frac{\partial \hat{V}}{\partial S} + \kappa[\theta - \lambda_C] \frac{\partial \hat{V}}{\partial \lambda_C} - r\hat{V} + f(\lambda_C, \hat{V}). \tag{53}$$

  In (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 intensity calls or puts cases, λ_C acts as discounting rate (exponential decay rate) – see formula (5.1) in [7]; 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 [23, 17, 22, 15], when considering mean reversion stochastic asset volatility dynamics. Condition (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 [22], 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 put cases. The conditions considered are either homogeneous Dirichlet or Neumann (which linear boundary conditions cover) or non-homogeneous Dirichlet (which are usually problem/payoff-dependent). More specifically, the authors of [22], derive a non-homogeneous Dirichlet boundary condition for the far-side of the stochastic volatility variable in the put option case. To do this, they use homogeneous Neumann *and* linear boundary condition to degenerate the Heston operator to vanilla Black-Scholes operator, which, for the put, has a closed-form solution. We believe that this technique is useful when Dirichlet conditions are needed, but is problem/payoff-dependent. For the XVA problem with all payoffs except the call and put, a one-dimensional XVA (nonlinear) PDE needs to be numerically solved to obtain a Dirichlet condition on $\lambda_C = \lambda_C^{max}$. In order to keep the boundary condition more general, we use the linear boundary condition (53).

Taking into account that the Dirichlet conditions are computed at $S = 0$, in advance of the main simulation of (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.

3.3 Nonlinear iteration method

In this subsection, we present an iterative method for handling the nonlinearity in (40). We refer to it as *discrete penalty-like iteration*, or, simply, *penalty iteration*, motivated by the similarly named method in [10] 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 [7], for the one-dimensional XVA problem. This paper extends the ideas in [7] 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

$$[P(v)]_{i,i} \equiv \begin{cases} -\lambda_B(1 - R_B) & \text{if } v_i < 0 \\ -(\lambda_C)_{i_2}(1 - R_C) - s_F & \text{if } v_i \geq 0. \end{cases} \tag{54}$$

Thus, in contrast with the method in [7], 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

$$f(\hat{v}) = P(\hat{v})\hat{v}, \tag{55}$$

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

$$[\mathbb{I} - \vartheta\Delta\tau^j(A + P(\hat{v}^j))]\hat{v}^j = (\mathbb{I} + (1-\vartheta)\Delta\tau^j A)\hat{v}^{j-1} + (1-\vartheta)\Delta\tau^j P(\hat{v}^{j-1})\hat{v}^{j-1}. \tag{56}$$

Algorithm 1 Discrete penalty iteration for (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 $g^j = (\mathbb{I} + (1-\vartheta)\Delta\tau^j A)\hat{v}^{j-1} + (1-\vartheta)\Delta\tau^j P(\hat{v}^{j-1})\hat{v}^{j-1}$.
Initialize $\hat{v}^{j,0} = \hat{v}^{j-1}$, and $P^{j,0} = P(\hat{v}^{j,0})$
for $k = 1, \ldots, maxit$ **do**
  Solve $[\mathbb{I} - \vartheta\Delta\tau^j(A + P^{j,k-1})]\hat{v}^{j,k} = g^j$
  Compute $P^{j,k} = P(\hat{v}^{j,k})$
  if stopping criterion satisfied **then**
    Break
  end if
end for
Set $\hat{v}^j = \hat{v}^{j,k}$

The proposed discrete penalty-like iteration for (40) is described in Algorithm 1.

The stopping criterion in Algorithm 1 is

$$(P^{j,k} = P^{j,k-1}) \text{ or } (\max_i \frac{|\hat{v}_i^{j,k} - \hat{v}_i^{j,k-1}|}{\max(1, |\hat{v}_i^{j,k}|)} \leq tol). \tag{57}$$

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 (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

$$[P(v)]_{i_2,i_2} \equiv \begin{cases} -\lambda_B(1 - R_B) & \text{if } v_{i_2} < 0 \\ -(\lambda_C)_{i_2}(1 - R_C) - s_F & \text{if } v_{i_2} \geq 0, \end{cases} \tag{58}$$

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

4 Asymptotic solution to XVA problem

Solving the time-dependent two-dimensional PDE derived in Section 2 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 literature, especially when some stochastic variables evolve according to some mean-reverting stochastic processes, [13, 21, 18, 11]. In this section, we derive such an asymptotic approximation for the price of derivatives considering mean-reverting stochastic default intensity.

4.1 More about mean-reverting CIR

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

$$d\lambda_C(t) = \kappa[\theta - \lambda_C(t)]dt + \sigma^{\lambda_C}\sqrt{\lambda_C(t)}dW^{\lambda_C}(t). \tag{59}$$

Although one cannot derive a closed-form solution for (59), the conditional distribution is a non-central chi-square distribution with

$$\mathbb{E}[\lambda_C(t))] = e^{-\kappa t}\lambda_C(0) + \theta(1 - e^{-\kappa t}), \tag{60}$$

$$\text{Var}[\lambda_C(t)] = \frac{(\sigma^{\lambda_C})^2}{\kappa}\lambda_C(0)(e^{-\kappa t} - e^{-2\kappa t}) + \frac{\theta(\sigma^{\lambda_C})^2}{2\kappa}(1 - e^{-\kappa t})^2. \tag{61}$$

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 [21]

$$\Phi(\lambda) = \frac{e^{-\lambda/\beta}\lambda^{\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}}. \tag{62}$$

Following [13, 18], 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 (40), which $\hat{V}(\tau, S, \lambda_C)$ satisfies becomes

$$(\frac{1}{\epsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 + \mathcal{L}_2)\hat{V} = 0, \tag{63}$$

where

$$\mathcal{L}_0 \equiv \frac{1}{2}\nu^2\lambda_C\frac{\partial^2}{\partial\lambda_C^2} + (\theta - \lambda_C)\frac{\partial}{\partial\lambda_C}, \tag{64}$$

$$\mathcal{L}_1 \equiv \rho\sigma^S\nu S\sqrt{\lambda_C}\frac{\partial^2}{\partial S\partial\lambda_C}, \tag{65}$$

$$\mathcal{L}_2 \equiv (-\frac{\partial}{\partial\tau}) + \frac{1}{2}(\sigma^S)^2S^2\frac{\partial^2}{\partial S^2} + r_R S\frac{\partial}{\partial S} - r\mathcal{I} + f(\lambda_C, \hat{V}), \tag{66}$$

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

4.2 Asymptotic approximation for zero correlation

When $\rho = 0$, the partial differential equation (63), which $\hat{V}(\tau, S, \lambda_C)$ satisfies, becomes

$$(\frac{1}{\epsilon}\mathcal{L}_0 + \mathcal{L}_2)\hat{V} = 0. \tag{67}$$

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

$$\hat{V}^\epsilon = \hat{V}_0 + \epsilon\hat{V}_1 + \epsilon^2\hat{V}_2 + \dots \tag{68}$$

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 (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 (68) into (67) and equating the lower order terms to zero, we have:

$$\mathcal{O}(\frac{1}{\epsilon}) : \mathcal{L}_0\hat{V}_0 = 0 \tag{69}$$

$$\mathcal{O}(1) : \mathcal{L}_0\hat{V}_1 + \mathcal{L}_2\hat{V}_0 = 0 \tag{70}$$

$$\mathcal{O}(\epsilon) : \mathcal{L}_0\hat{V}_2 + \mathcal{L}_2\hat{V}_1 = 0. \tag{71}$$

Equation (69) implies that $\hat{V}_0$ is independent of λ_C, i.e. $\hat{V}_0 = \hat{V}_0(\tau, S)$. Equation (69) is a Poisson equation with respect to the operator $\mathcal{L}_0$ in the variable λ_C, which implies the centering condition

$$\langle \mathcal{L}_2\hat{V}_0\rangle = 0. \tag{72}$$

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 (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 [4]. 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 (nonlinear) parabolic PDE needs to be numerically solved, see, for example, [7].

Now let's try to find $\hat{V}_1$. Combining (70) and the fact that $\langle \mathcal{L}_2\hat{V}_0\rangle = \langle \mathcal{L}_2\rangle\hat{V}_0 = 0$, we have

$$\mathcal{L}_0\hat{V}_1 = -\mathcal{L}_2\hat{V}_0 = -(\mathcal{L}_2\hat{V}_0 - \langle \mathcal{L}_2\rangle\hat{V}_0) = -(\mathcal{L}_2 - \langle \mathcal{L}_2\rangle)\hat{V}_0, \tag{73}$$

where

$$(\mathcal{L}_2 - \langle \mathcal{L}_2\rangle)\hat{V}_0 = f(\lambda_C, \hat{V}_0) - f(\theta, \hat{V}_0) = (1 - R_C)(\theta - \lambda_C)\hat{V}_0^+. \tag{74}$$

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

$$\hat{V}_1 = -(1 - R_C)[(\lambda_C - \theta) + \tilde{C}(\tau, S)]\hat{V}_0^+. \tag{75}$$

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

$$\hat{V}_1 = -(1 - R_C)(\lambda_C - \theta)\hat{V}_0^+ + C(\tau, S) = (1 - R_C)(\theta - \lambda_C)\hat{V}_0^+ + C(\tau, S), \tag{76}$$

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 (71) implies the centering condition

$$\langle \mathcal{L}_2\hat{V}_1\rangle = 0. \tag{77}$$

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

$$\begin{aligned}\langle \mathcal{L}_2 \rangle C(\tau, S) &= \langle (\mathcal{L}_2 - \langle \mathcal{L}_2 \rangle)(1 - R_C)(\lambda_C - \theta)\hat{V}_0^+ \rangle \\ &= \langle (1 - R_C)^2(\lambda_C - \theta)(\theta - \lambda_C)\hat{V}_0^+ \rangle \\ &= \langle (\lambda_C - \theta)(\theta - \lambda_C) \rangle (1 - R_C)^2 \hat{V}_0^+ \\ &= -(1 - R_C)^2 \frac{\theta \nu^2}{2} \hat{V}_0^+. \end{aligned} \tag{78}$$

Note that $C(\tau, S) = \tau(1 - R_C)^2 \frac{\theta\nu^2}{2} \hat{V}_0^+$ is a solution to (78). Thus, (76) gives

$$\hat{V}_1 = (1 - R_C)(\theta - \lambda_C)\hat{V}_0^+ + \tau(1 - R_C)^2 \frac{\theta \nu^2}{2} \hat{V}_0^+. \tag{79}$$

Therefore, we obtain the approximation

$$\hat{V}^{\epsilon,1} = \hat{V}_0 + \epsilon(1 - R_C)(\theta - \lambda_C)\hat{V}_0^+ + \epsilon\tau(1 - R_C)^2 \frac{\theta \nu^2}{2} \hat{V}_0^+, \tag{80}$$

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

4.3 Asymptotic approximation for general correlation

Recall that the partial differential equation which $\hat{V}(\tau, S, \lambda_C)$ satisfies is (63). Let $\hat{V}^\epsilon$ be a power series expansion of $\hat{V}$ in $\sqrt{\epsilon}$ [4]:

$$\hat{V}^\epsilon = \hat{V}_0 + \sqrt{\epsilon}\hat{V}_{1/2} + \epsilon\hat{V}_1 + \epsilon\sqrt{\epsilon}\hat{V}_{3/2} + \dots \tag{81}$$

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 (40), that is, $\hat{V}_0(0, S, \lambda_C) = h(S)$.
Substituting (81) into (63) and equating the lower order terms to zero, we have:

$$\mathcal{O}(\frac{1}{\epsilon}) : \mathcal{L}_0 \hat{V}_0 = 0 \tag{82}$$

$$\mathcal{O}(\frac{1}{\sqrt{\epsilon}}) : \mathcal{L}_0 \hat{V}_{1/2} + \mathcal{L}_1 \hat{V}_0 = 0 \tag{83}$$

$$\mathcal{O}(1) : \mathcal{L}_0 \hat{V}_1 + \mathcal{L}_1 \hat{V}_{1/2} + \mathcal{L}_2 \hat{V}_0 = 0 \tag{84}$$

$$\mathcal{O}(\sqrt{\epsilon}) : \mathcal{L}_0 \hat{V}_{3/2} + \mathcal{L}_1 \hat{V}_1 + \mathcal{L}_2 \hat{V}_{1/2} = 0. \tag{85}$$

Equation (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 (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 (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

$$\langle \mathcal{L}_2 \hat{V}_0 \rangle = 0. \tag{86}$$

[4] 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.

Similarly as in the case of zero correlation, as $\hat{V}_0$ is independent of λ_C, the centering condition (86) becomes $\langle \mathcal{L}_2 \rangle \hat{V}_0 = 0$. Therefore, $\hat{V}_0$ is the solution to the Black-Scholes equation (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 (84) can be reduced to $\mathcal{L}_0 \hat{V}_1 + \mathcal{L}_2 \hat{V}_0 = 0$, which is exactly same as Equation (70). Hence, the formula for $\hat{V}_1$ is given in the previous subsection, and is

$$\hat{V}_1 = (1 - R_C)(\theta - \lambda_C)\hat{V}_0^+ + \tau(1 - R_C)^2 \frac{\theta \nu^2}{2} \hat{V}_0^+. \tag{87}$$

For equation (85), again, the solvability of this Poisson equation requires

$$\langle \mathcal{L}_1 \hat{V}_1 + \mathcal{L}_2 \hat{V}_{1/2} \rangle = 0, \tag{88}$$

which gives

$$\begin{aligned} \langle \mathcal{L}_2 \rangle \hat{V}_{1/2} = -\langle \mathcal{L}_1 \hat{V}_1 \rangle &= -\langle \rho \sigma^S \nu S \sqrt{\lambda_C} \frac{\partial^2}{\partial S \partial \lambda_C} [(1 - R_C)(\theta - \lambda_C)\hat{V}_0^+ + \tau(1 - R_C)^2 \frac{\theta \nu^2}{2} \hat{V}_0^+] \rangle \\ &= \langle \rho \sigma^S \nu S \sqrt{\lambda_C} (1 - R_C) \frac{\partial \hat{V}_0^+}{\partial S} \rangle = \rho \sigma^S \nu S (1 - R_C) \langle \sqrt{\lambda_C} \rangle \frac{\partial \hat{V}_0^+}{\partial S}. \end{aligned} \tag{89}$$

Because $\langle \mathcal{L}_2 \rangle \hat{V}_0^+ = 0$, we can verify that the solution to (89) is

$$\hat{V}_{1/2}(\tau, S) = -\tau \rho \sigma^S \nu S (1 - R_C) \langle \sqrt{\lambda_C} \rangle \frac{\partial \hat{V}_0^+}{\partial S}. \tag{90}$$

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

$$\hat{V}^{\epsilon,1} = \hat{V}_0 - \sqrt{\epsilon} \tau \rho \sigma^S \nu S (1 - R_C) \langle \sqrt{\lambda_C} \rangle \frac{\partial \hat{V}_0^+}{\partial S} + \epsilon (1 - R_C)(\theta - \lambda_C) \hat{V}_0^+ + \epsilon \tau (1 - R_C)^2 \frac{\theta \nu^2}{2} \hat{V}_0^+. \tag{91}$$

In Appendix A, 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 an upper bound to the variable λ_C.

4.3.1 More details

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

$$\frac{\partial \hat{V}_0^+}{\partial S} = \begin{cases} \frac{\partial \hat{V}_0}{\partial S} & \hat{V}_0 > 0 \\ 0 & \hat{V}_0 \leq 0 \end{cases}. \tag{92}$$

(2) In approximation (91), by the definition of $\langle \cdot \rangle$, the quantity $\langle \sqrt{\lambda_C} \rangle$ can be computed by

$$\langle \sqrt{\lambda_C} \rangle := \int_0^\infty \sqrt{\lambda_C} \Phi(\lambda_C) d\lambda_C, \tag{93}$$

where $\Phi(\lambda_C)$ is the probability density function of stationary distribution given in (62). Numerical quadrature, e.g. MATLAB's `integral`, can be used if an analytical formula cannot be easily obtained.

5 Numerical experiments

We present results of numerical experiments from applying the proposed methods on several financial derivatives. The market and numerical parameters used in the experiments are listed in Table 1. Whenever different parameters are used for the purpose of testing, we state their values explicitly.

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, $q - \gamma$	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 1: Model parameters for bilateral XVA with stochastic default intensity in European derivatives.

5.1 Numerical 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 [17] 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

$$S_i \equiv w(x_i) = (1 + \frac{\sinh(\beta - (x_i/x_N - \alpha))}{\sinh(\beta\alpha)})K \tag{94}$$

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

$$(\lambda_C)_i \equiv u(y_i) = \sinh(y_i). \tag{95}$$

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 Section 3.2. 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 Section 3.1.

Algorithm 1 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}.

5.1.1 Call and put options

We present results from pricing the XVA with stochastic counterparty default intensity of European options with parameter settings in Table 1. 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 2, we show the results at-the-money with different resolutions. In Tables 3 and 4, 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. This holds even if, with the choice of parameters in Table 1, the Feller condition is not satisfied. From Tables 3 and 4, 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 2: Results from solving (40) for European derivatives including bilateral XVA with stochastic default intensities on counterparty using Algorithm 1 with the parameters in Table 1 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	12.424685
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 3: Results from solving (40) for European call option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 1 with the parameters in Table 1 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 4: Results from solving (40) for European put option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 1 with the parameters in Table 1 at various points. Nonuniform grids are used.

REMARK 1 *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 long forwards with constant default intensity are found in [7].*

5.1.2 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 1, except that in Table 5, we set $S_{max} = 4K, 8K, 6K$ and $10K$, and in Table 6, we set $\lambda_C^{max} = 4.05, 6.05$ and 8.05. In Table 5, 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 6, 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 1 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: Results from solving (40) for European call option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 1 with the parameters in Table 1, 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 6: Results from solving (40) for European call option including bilateral XVA with stochastic default intensity on counterparty using Algorithm 1 with the parameters in Table 1, 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 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 default risk with the parameters of Table 1, except that κ varies from 1 to 10. We present two Tables 7 and 8 (κ varying from 1 to 3), one with correlation ρ being zero and one with correlation given in Table 1. 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 Richardson extrapolation, with second order convergence. The numerical results show good agreements among solutions under different approaches.

In Tables 7, 8, and Figure 1 (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.

Comparing the differences in Tables 7 and 8, and looking at Figure 1, 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. def. 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 7: Values by different approaches for European put option including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 1, 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 $N = 512, M = 256$.

	(7.5, 0.05)	(7.5, 0.1)	(15, 0.05)	(15, 0.1)	(30, 0.05)	(30, 0.1)
const. def. 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 8: Values by different approaches for European put option including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 1, 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 $N = 512, M = 256$.

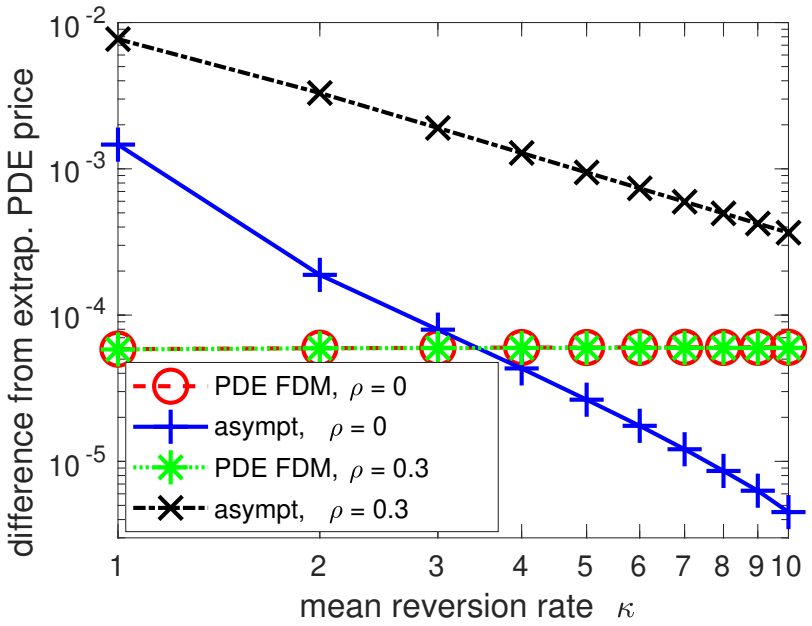

Figure 1: Accuracy of different approaches for European put option valuation including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 1 except κ and ρ as indicated, and $\sigma^{\lambda_C} = 0.2\sqrt{\kappa}$, versus κ at (15, 0.1). The PDE FDM solution is obtained with $N = 512, M = 256$.

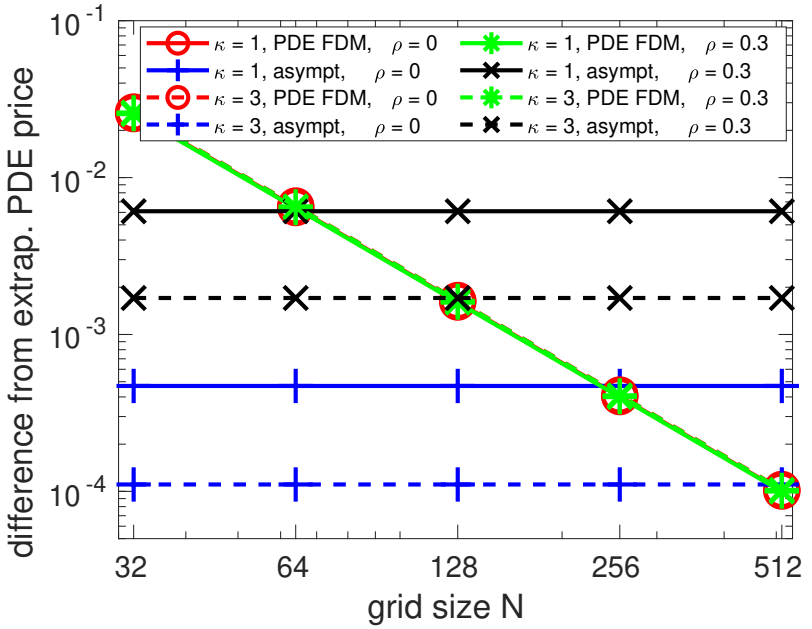

Figure 2: Accuracy of different approaches for European put option valuation including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 1 except κ and ρ as indicated, and $\sigma^{\lambda_C} = 0.2\sqrt{\kappa}$, versus N at (30, 0.1).

From Figure 2, 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. Comparing Figures 1 (point (15,0.1)) and 2 (point (30,0.1)), we note that, for point (15,0.1), $\rho = 0.3$ and all κ's considered, the asymptotic accuracy never reaches the

2D PDE accuracy, while, for point (30,0.1), the asymptotic accuracy may match or exceed the 2D PDE accuracy, depending on the values of κ and ρ considered.

5.3 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 1, except when mentioned otherwise.

In Table 9 and Figures 3 and 4, 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 (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 (91), and substituting $\sigma^{\lambda_C} = \frac{\nu}{\sqrt{\epsilon}}$ or $\nu = \sqrt{\epsilon}\sigma^{\lambda_C}$, we get

$$\begin{aligned}\hat{V}^{\epsilon,1} &= \hat{V}_0 - \sqrt{\epsilon}\tau\rho\sigma^S\nu S(1-R_C)\langle\sqrt{\lambda_C}\rangle\frac{\partial\hat{V}_0^+}{\partial S} + \epsilon(1-R_C)(\theta-\lambda_C)\hat{V}_0^+ + \epsilon\tau(1-R_C)^2\frac{\theta\nu^2}{2}\hat{V}_0^+ \\ &= \hat{V}_0 - \epsilon\tau\rho\sigma^S\sigma^{\lambda_C}S(1-R_C)\langle\sqrt{\lambda_C}\rangle\frac{\partial\hat{V}_0^+}{\partial S} + \epsilon(1-R_C)(\theta-\lambda_C)\hat{V}_0^+ + \epsilon^2\tau(1-R_C)^2\frac{(\sigma^{\lambda_C})^2}{2}\hat{V}_0^+.\end{aligned} \tag{96}$$

Recall that, $\hat{V}_0$ is the solution to the Black-Scholes equation (43) considering constant counterparty default risk equal to θ. Among the four terms in the right hand side of (96), 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.

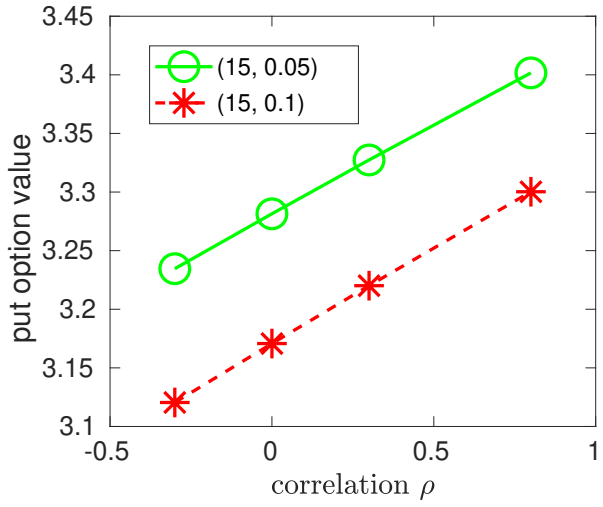

Figure 3: Effect of ρ on put option value with XVA, computed by solving (40). Other parameters are in Table 1

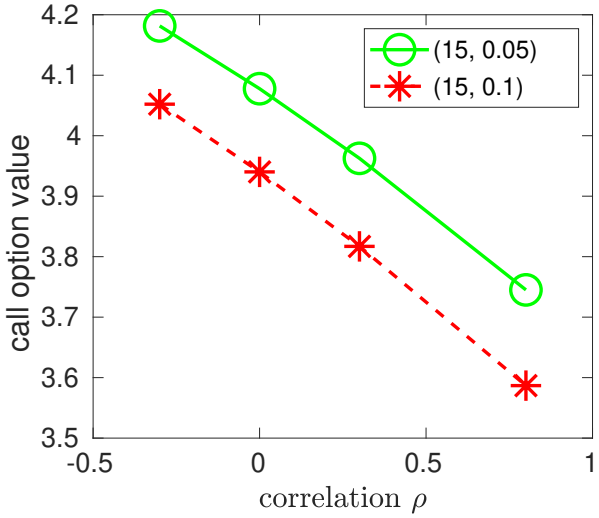

Figure 4: Effect of ρ on call option value with XVA, computed by solving (40). Other parameters are in Table 1

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 9: Value comparison for solving (40) for European options including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 1 with different correlations. The grid size is $N = 512, M = 256$.

Table 10 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 (6) and (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 gives a visualization in the case of put option. As we can see in (96), 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 10: Value comparison for solving (40) for European options including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 1 with different mean reversion levels. The grid size is $N = 512, M = 256$.

Table 11 and Figure 6 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 (96) 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 (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 11: Value comparison for solving (40) for European put option including bilateral XVA with stochastic default intensity on counterparty with the parameters in Table 1. The grid size is $N = 512, M = 128$.

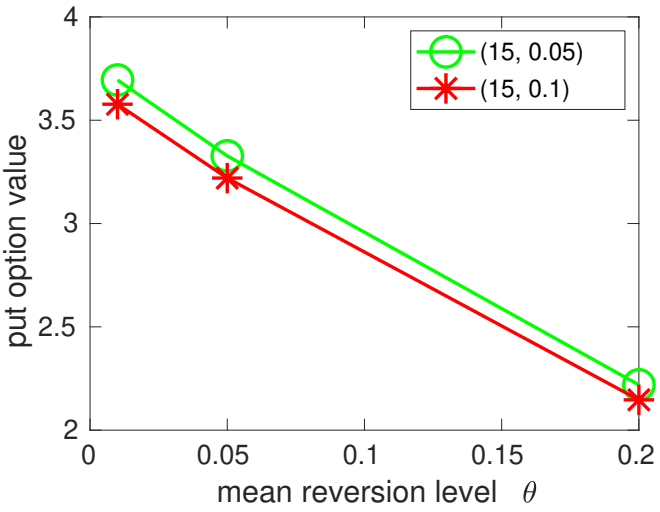

Figure 5: Effect of θ on put option value with XVA, computed by solving (40). Other parameters are in Table 1.

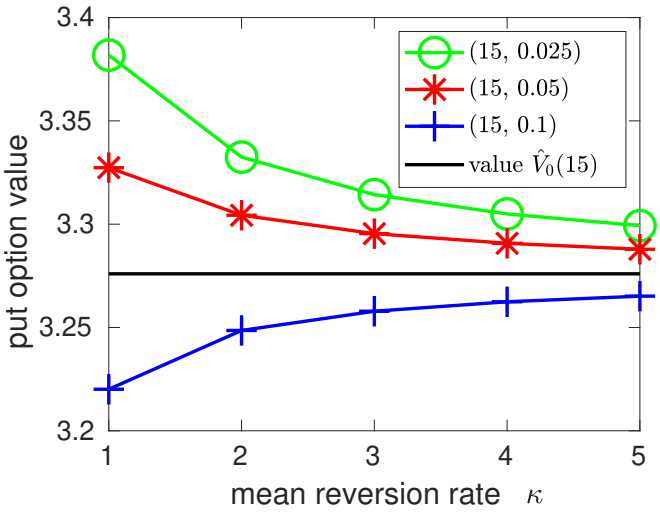

Figure 6: Effect of κ on put option value with XVA, computed by solving (40). Other parameters are in Table 1.

6 Conclusions

In this paper, 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 numerically 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 timestepping, and the penalty iteration method in [7] to handle the non-linearity of the source term. The second approach is asymptotic approximation, based on singular perturbation theory [18, 13], 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, 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.

Appendix A: Accuracy analysis of asymptotic approximation

We consider the case of general correlation. The case of zero correlation is briefly considered at the end of the Appendix. In this analysis, we assume λ_C is bounded from above. In order to analyze the quality

of the approximation $\hat{V}^{\epsilon,1}$ of (91) to $\hat{V}$, we begin by recalling (63), and defining

$$\mathcal{L}^{\epsilon}\hat{V}^{\epsilon} \equiv (\frac{1}{\epsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 + \mathcal{L}_2)\hat{V}^{\epsilon} = 0, \tag{97}$$

$$\mathcal{E} \equiv \hat{V}^{\epsilon} - \hat{V}_0 - \sqrt{\epsilon}\hat{V}_{1/2} - \epsilon\hat{V}_1 - \epsilon\sqrt{\epsilon}\hat{V}_{3/2}, \tag{98}$$

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

$$\mathcal{E}(0, S, \lambda_C) = -\epsilon(1 - R_C)(\theta - \lambda_C)\hat{V}_0^{+}(0, S) - \epsilon\sqrt{\epsilon}\hat{V}_{3/2}(0, S, \lambda_C), \tag{99}$$

since, at $\tau = 0$, $\hat{V}^{\epsilon} = \hat{V}_0$, which is also equal to the pay-off function.
In addition, we have

$$\begin{aligned}
\mathcal{L}^{\epsilon}\mathcal{E} =& (\frac{1}{\epsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 + \mathcal{L}_2)(\hat{V}^{\epsilon} - \hat{V}_0 - \sqrt{\epsilon}\hat{V}_{1/2} - \epsilon\hat{V}_1 - \epsilon\sqrt{\epsilon}\hat{V}_{3/2}) \\
=& -\frac{1}{\epsilon}(\mathcal{L}_0\hat{V}_0) - \frac{1}{\sqrt{\epsilon}}(\mathcal{L}_0\hat{V}_{1/2} + \mathcal{L}_1\hat{V}_0) - (\mathcal{L}_0\hat{V}_1 + \mathcal{L}_1\hat{V}_{1/2} + \mathcal{L}_2\hat{V}_0) - \sqrt{\epsilon}(\mathcal{L}_0\hat{V}_{3/2} + \mathcal{L}_1\hat{V}_1 + \mathcal{L}_2\hat{V}_{1/2}) \\
& - \epsilon(\mathcal{L}_1\hat{V}_{3/2} + \sqrt{\epsilon}\mathcal{L}_2\hat{V}_{3/2}) = -\epsilon(\mathcal{L}_1\hat{V}_{3/2} + \sqrt{\epsilon}\mathcal{L}_2\hat{V}_{3/2}),
\end{aligned} \tag{100}$$

taking also into account (82)-(85). The Feynman-Kac probabilistic representation formula for the solution of (100) is given as

$$\begin{aligned}
\mathcal{E}(\tau, S, \lambda_C) =& -\mathbb{E}^{Q}\left[e^{-r\tau}(\epsilon(1 - R_C)(\theta - \lambda_C)\hat{V}_0^{+} + \epsilon^{3/2}\hat{V}_{3/2} \mid S^{\epsilon}_{T-\tau} = S, (\lambda_C)^{\epsilon}_{T-\tau} = \lambda_C\right] \\
& + \mathbb{E}^{Q}\left[\int_0^{\tau} e^{-r(\tau-s)}(\epsilon\mathcal{L}_1\hat{V}_{3/2} + \epsilon^{3/2}\mathcal{L}_2\hat{V}_{3/2})ds \mid S^{\epsilon}_{T-\tau} = S, (\lambda_C)^{\epsilon}_{T-\tau} = \lambda_C\right] \\
& + \mathbb{E}^{Q}\left[\int_0^{\tau} e^{-r(\tau-s)}f(\lambda_C, \mathcal{E})ds \mid S^{\epsilon}_{T-\tau} = S, (\lambda_C)^{\epsilon}_{T-\tau} = \lambda_C\right].
\end{aligned} \tag{101}$$

By (85), we have

$$\begin{aligned}
\mathcal{L}_0\hat{V}_{3/2} =& -\mathcal{L}_1\hat{V}_1 - \mathcal{L}_2\hat{V}_{1/2} \\
=& \rho\sigma^S\nu S(1 - R_C)(\sqrt{\lambda_C} - \langle\sqrt{\lambda_C}\rangle)\frac{\partial\hat{V}_0^{+}}{\partial S} + f(\theta, \hat{V}_{1/2}) - f(\lambda_C, \hat{V}_{1/2}).
\end{aligned} \tag{102}$$

Hence, following the arguments in [18], $|\hat{V}_{3/2}| \leq C_1|\frac{\partial\hat{V}_0^{+}}{\partial S}|$, where C_1 is a positive constant. Then we have

$$\left|\mathbb{E}^{Q}\left[\epsilon^{3/2}\hat{V}_{3/2} \mid S^{\epsilon}_{T-\tau} = S, (\lambda_C)^{\epsilon}_{T-\tau} = \lambda_C\right]\right| \leq C_2\epsilon^{3/2}, \tag{103}$$

$$\left|\mathbb{E}^{Q}\left[\int_0^{\tau} e^{-r(\tau-s)}\epsilon^{3/2}\mathcal{L}_2\hat{V}_{3/2}ds \mid S^{\epsilon}_{T-\tau} = S, (\lambda_C)^{\epsilon}_{T-\tau} = \lambda_C\right]\right| \leq C_3\epsilon^{3/2}, \tag{104}$$

where C_2 and C_3 are positive constants. We also notice that, since λ_C is bounded from above,

$$\begin{aligned}
& \left|\mathbb{E}^{Q}\left[\int_0^{\tau} e^{-r(\tau-s)}\epsilon\mathcal{L}_1\hat{V}_{3/2}ds \mid S^{\epsilon}_{T-\tau} = S, (\lambda_C)^{\epsilon}_{T-\tau} = \lambda_C\right]\right| \\
=& \left|\mathbb{E}^{Q}\left[\int_0^{\tau} e^{-r(\tau-s)}\epsilon\rho\sigma^S\nu S\sqrt{\lambda_C}\frac{\partial^2\hat{V}_{3/2}}{\partial S\partial\lambda_C}ds \mid S^{\epsilon}_{T-\tau} = S, (\lambda_C)^{\epsilon}_{T-\tau} = \lambda_C\right]\right| \\
=& \left|\mathbb{E}^{Q}\left[\int_0^{\tau} e^{-r(\tau-s)}\epsilon^{3/2}\rho\sigma^S S\sigma^{\lambda_C}\sqrt{\lambda_C}\frac{\partial^2\hat{V}_{3/2}}{\partial S\partial\lambda_C}ds \mid S^{\epsilon}_{T-\tau} = S, (\lambda_C)^{\epsilon}_{T-\tau} = \lambda_C\right]\right| \\
\leq& C_4\epsilon^{3/2},
\end{aligned} \tag{105}$$

where C_4 is a positive constant.
With similar arguments as in [11, 18, 21], and taking into account that $\langle \theta - \lambda_C \rangle = 0$, we have, with C_5 and C_6 positive constants, and since λ_C is bounded from above,

$$\begin{aligned}&\left|\mathbb{E}^Q\left[e^{-r\tau}(1-R_C)(\theta-\lambda_C)\hat{V}_0^+ \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right]\right| \\ &\quad\leq C_5\left|\mathbb{E}^Q\left[(\theta-\lambda_C)\mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right]\right| \leq C_5 e^{-C_6\frac{1}{\epsilon}},\end{aligned} \tag{106}$$

which converges exponentially fast as $\epsilon \to 0$.
Now let's consider the last term of the right-hand-side of (101):

$$\begin{aligned}&\mathbb{E}^Q\left[\int_0^\tau e^{-r(\tau-s)}f(\lambda_C,\mathcal{E})ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right] \\ =&\mathbb{E}^Q\left[\int_0^\tau e^{-r(\tau-s)}[-(s_F+(1-R_C)\lambda_C)\mathcal{E}^+-(1-R_B)\lambda_B\mathcal{E}^-]ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right] \\ =&-\mathbb{E}^Q\left[\int_0^\tau e^{-r(\tau-s)}[(s_F+(1-R_C)\lambda_C)\mathcal{E}^++(1-R_B)\lambda_B\mathcal{E}^-]ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\right].\end{aligned} \tag{107}$$

Taking into account (103)-(107), equation (101) results in

$$\begin{aligned}&\left|\mathcal{E}(\tau,S,\lambda_C)+\mathbb{E}^Q\Big[\int_0^\tau e^{-r(\tau-s)}[(s_F+(1-R_C)\lambda_C)\mathcal{E}(\tau,S,\lambda_C)^+\right. \\ &\left.+(1-R_B)\lambda_B\mathcal{E}(\tau,S,\lambda_C)^-]ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\right| \leq C\epsilon^{3/2},\end{aligned} \tag{108}$$

where C is a positive constant.
We consider two cases for the sign of $\mathcal{E}(\tau,S,\lambda_C)$. If $\mathcal{E}(\tau,S,\lambda_C)\leq 0$, equation (108) results in

$$\left|\mathcal{E}(\tau,S,\lambda_C)+\mathbb{E}^Q\Big[\int_0^\tau e^{-r(\tau-s)}(1-R_B)\lambda_B\mathcal{E}(\tau,S,\lambda_C)ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\right| \leq C\epsilon^{3/2}, \tag{109}$$

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

$$\begin{aligned}&\left|\mathbb{E}^Q\Big[\int_0^\tau e^{-r(\tau-s)}(1-R_B)\lambda_B\mathcal{E}(\tau,S,\lambda_C)ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\right| \\ \geq&\left|(1-R_B)\lambda_B e^{-r\tau}\mathbb{E}^Q\Big[\int_0^\tau \mathcal{E}(\tau,S,\lambda_C)ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\right| \\ =&\left|(1-R_B)\lambda_B e^{-r\tau}\mathbb{E}^Q\Big[\mathcal{E}(\tau,S,\lambda_C) \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\right| \\ =&\left|(1-R_B)\lambda_B e^{-r\tau}\tau\mathcal{E}(\tau,S,\lambda_C)\right|.\end{aligned} \tag{110}$$

With (110), relation (109) leads to

$$\left|\mathcal{E}(\tau,S,\lambda_C)+(1-R_B)\lambda_B e^{-r\tau}\tau\mathcal{E}(\tau,S,\lambda_C)\right| \leq C\epsilon^{3/2}, \tag{111}$$

from which we get

$$|1+(1-R_B)\lambda_B e^{-r\tau}\tau||\mathcal{E}(\tau,S,\lambda_C)| \le C\epsilon^{3/2} \implies$$
$$|\mathcal{E}(\tau,S,\lambda_C)| \le \frac{C}{|1+(1-R_B)\lambda_B e^{-r\tau}\tau|}\epsilon^{3/2}. \tag{112}$$

If $\mathcal{E}(\tau,S,\lambda_C)>0$, equation (108) results in

$$\Big|\mathcal{E}(\tau,S,\lambda_C)+\mathbb{E}^Q\Big[\int_0^\tau e^{-r(\tau-s)}[(s_F+(1-R_C)\lambda_C)\mathcal{E}(\tau,S,\lambda_C)^+]ds \\ \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\Big| \le C\epsilon^{3/2}, \tag{113}$$

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

$$\begin{aligned}
&\Big|\mathbb{E}^Q\Big[\int_0^\tau e^{-r(\tau-s)}[(s_F+(1-R_C)\lambda_C)\mathcal{E}(\tau,S,\lambda_C)^+]ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\Big| \\
\ge&\Big|\mathbb{E}^Q\Big[\int_0^\tau e^{-r(\tau)}\delta_C\mathcal{E}(\tau,S,\lambda_C)ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\Big| \\
=&\Big|e^{-r\tau}\delta_C\mathbb{E}^Q\Big[\int_0^\tau \mathcal{E}(\tau,S,\lambda_C)ds \mid S^\epsilon_{T-\tau}=S,(\lambda_C)^\epsilon_{T-\tau}=\lambda_C\Big]\Big| \\
=&\Big|e^{-r\tau}\delta_C\tau\mathcal{E}(\tau,S,\lambda_C)\Big|,
\end{aligned} \tag{114}$$

where $\delta_C\ge 0$ is a constant lower bound of $(s_F+(1-R_C)\lambda_C)$. With (114), relation (113) leads to

$$\Big|\mathcal{E}(\tau,S,\lambda_C)+e^{-r\tau}\delta_C\tau\mathcal{E}(\tau,S,\lambda_C)\Big| \le C\epsilon^{3/2}, \tag{115}$$

from which we get

$$|1+e^{-r\tau}\delta_C\tau||\mathcal{E}(\tau,S,\lambda_C)| \le C\epsilon^{3/2} \implies$$
$$|\mathcal{E}(\tau,S,\lambda_C)| \le \frac{C}{|1+e^{-r\tau}\delta_C\tau|}\epsilon^{3/2}. \tag{116}$$

Combining (112) and (116), 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 (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 (80) is at least of order $\mathcal{O}(\epsilon^2)$.

Acknowledgements

This work is supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Faculty of Arts and Science University of Toronto Tri-Council Bridge Funding Program, and the Ontario Graduate Scholarship (OGS).

References

[1] 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.

[2] 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.

[3] C. BURGARD AND M. KJAER, *In the balance*, Risk, 24 (2011), pp. 72–75.

[4] 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.

[5] C. BURGARD AND M. KJAER, *The FVA debate: In theory and practice*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2157634, (2012).

[6] ——, *Funding strategies, funding costs*, Risk, 26 (2013), pp. 82–87.

[7] 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.

[8] D. DUFFIE AND N. GARLEANU, *Risk and valuation of collateralized debt obligations*, Financial analysts journal, 57 (2001), pp. 41–59.

[9] Y. FENG, *CVA under Bates model with stochastic default intensity*, Journal of Mathematical Finance, 7 (2017), pp. 682–698.

[10] 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.

[11] 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.

[12] J.-P. FOUQUE, G. PAPANICOLAOU, AND K. R. SIRCAR, *Derivatives in financial markets with stochastic volatility*, Cambridge University Press, 2000.

[13] J. P. FOUQUE, G. PAPANICOLAOU, R. SIRCAR, AND K. SOLNA, *Singular perturbations in option pricing*, SIAM J. Appl. Math., 63 (2003), pp. 1648–1665.

[14] A. GREEN, *XVA: Credit, Funding and Capital Valuation Adjustments*, Wiley, 2015.

[15] 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.

[16] J. C. HULL, *Options, Futures, and Other Derivatives*, Pearson, 10 ed., 2017.

[17] 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.

[18] 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.

[19] 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.

[20] V. PITERBARG, *Funding beyond discounting: collateral agreements and derivatives pricing*, Risk, 23 (2010), pp. 42–47.

[21] 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.

[22] 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.

[23] 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.