A Multi-Level Monte Carlo Lagrange-Galerkin method for solving a XVA hybrid model in European options*

Graziana Colonna[a], Ana M. Ferreiro-Ferreiro[a], José A. García-Rodríguez[a], Carlos Vázquez[a,*]

[a]*Department of Mathematics and CITIC, University of A Coruña, Campus Elviña s/n, 15071-A Coruña (Spain)*

Abstract

In this article we consider XVA pricing models for European options that incorporate three stochastic factors, namely, the price of the underlying asset and the intensities of default of the investor and the hedger, with the corresponding stochastic differential equations (SDEs) that govern their dynamics. By using a dynamic hedging methodology, linear and nonlinear partial differential equations (PDEs) have been posed in [15]. In order to avoid the numerical solution PDEs with three spatial variables, we propose a hybrid model based on PDEs in one spatial variable, keeping the two intensities of default as stochastic coefficients of the PDEs satisfying their respective associated SDEs. For the numerical solution, we propose a Multi-Level Monte Carlo Lagrange-Galerkin method, where the PDEs obtained for each sample path of the stochastic coefficients is solved by combining a semi-Lagrangian method for time discretization and a finite element technique for the spatial discretization, adding a fixed point iteration in the nonlinear case. Numerical results allow to compare the different models and the qualitative and quantitative behaviour of the XVA.

Keywords: Counterparty risk, XVA, European options, Multi-Level Monte Carlo, semi-Lagrangian method, finite elements.

1. Introduction

As a consequence of the credit crisis that started around 2008, involving the defaults of big companies, a highly increasing attention has been paid to the analysis and modelling of counterparty risk, both in industry and academia (for example, see the books [9, 13, 18]). Particularly, the presence of counterparty risk arises in the context of derivative contracts, where counterparty risk is related to the possibility that one counterparty of the contract defaults while owing money associated to the derivatives contract or while the mark-to-market value of the derivative is positive for the other part of the contract. Apart

*This work has been partially funded by EU H2020-MSCA-ITN-2018 (ABC-EU-XVA Grant Agreement 813261), Spanish Ministry of Science and Innovation (Grant PID2019-108584RB-I00) and by Galician Government (Grant ED431C 2022/047), both including FEDER financial support. Authors also acknowledge the support received from the Centro de Investigación de Galicia "CITIC", funded by Xunta de Galicia and the European Union (European Regional Development Fund- Galicia 2014-2020 Program), by grant ED431G 2019/01.

*Corresponding author.

Email addresses: graziana.colonna@udc.es (Graziana Colonna), ana.fferreiro@udc.es (Ana M. Ferreiro-Ferreiro), jose.garcia.rodriguez@udc.es (José A. García-Rodríguez), carlos.vazquez.cendon@udc.es (Carlos Vázquez)

from the posted collateral, a derivative is a contractual agreement that might be broken by one of the two involved parties, therefore exposing the other party to risk. For instance, let us consider an over-the-counter (OTC) option sold by a hedger to an investor. If the option expires in-the-money, the hedger owes the intrinsic value to the investor. Counterparty risk is the credit risk that the hedger will not be able to fulfil its obligation to the investor (for example, the hedger might go bankrupt). In general, counterparty risk appears when the counterparty fails to meet the obligations on the trade.

Therefore, the presence of counterparty risk requires some adjustments in the price of the derivative with respect to the case where there is no counterparty risk. For example, when pricing a European option without counterparty risk with the classical Black-Scholes model there is a welknown analytical formula to get the price. However, this formula does not apply when the adjustments related to the presence of counterparty risk must be taken into account.

Many papers and books have proposed and developed techniques for the valuation of derivatives including counterparty risk by means of valuation adjustments. The set of all of these adjustments is usually referred to as total valued adjustment and denoted as XVA (X Valuation Adjustment). Among the adjustments that can be included in XVA we have (the list is not exhaustive):

- CVA: the cost of hedging counterparty credit risk;
- DVA: the adjustment to a derivative price due to the institution's own default risk;
- FVA: the correction made to the derivative price to account for a funding cost/benefit related to counterparty risk;
- KVA: the cost of holding regulatory capital associated to counterparty risk.

In order to model the derivative value including the price of the XVA, three main approaches are considered in the literature: partial differential equations (PDEs), backward stochastic differential equations (BSDEs) and formulations in terms of expectations.

In the PDEs based approach, the spatial dimension of the time dependent PDE is equal to the number of underlying stochastic factors. PDE modelling starts with the seminal article by Burgard and Kjaer work [8], where the authors pose linear and nonlinear PDEs to obtain the value of risky European derivatives traded between two defaultable parties: the hedger and the investor. For this purpose, they use dynamic hedging arguments and appropriate Ito formulas. Moreover, depending on the choice of the mark-to-market value of the derivative at default, either a linear or a nonlinear PDE is obtained. As they consider the price of the derivative as the unique stochastic factor, PDEs with one spatial dimension are obtained. In their approach the intensities of default for both the hedger and the investor are assumed

to be constant, thus entering as constant coefficients in the PDEs

After the work [8], many extension appeared in the literature. For example, in [15] the authors consider stochastic intensities of default both for the investor and the hedger. In this setting, formulations in terms of linear and nonlinear PDEs with three spatial dimensions are obtained by means of suitable hedging arguments and the use of Ito formula for jump diffusion processes. Moreover, by using suitable Feynman-Fac formulas, equivalent formulations in terms expectations are posed in [15] both for the linear and nonlinear cases. Numerical methods for the formulations based on expectations are proposed and numerical examples are discussed in [15], while the solution of the PDEs models is not addressed. More recently, PDEs formulations with two stochastic factors have been mathematically analysed and numerically solved for pricing the XVA associated to European and American options in [1] and [3], respectively. Also the mathematical analysis and numerical solution of one factor models associated to constant intensities of default has been addressed in [2] and [5]. In these previous works, the numerical solution is based on a semi-Lagrangian method for time discretization and finite elements for spatial discretization, while a fixed point iteration is added for the nonlinear PDEs. More recently, in [10] and [11] penalty methods are proposed for pricing the XVA of European and American options.

As in [15], in this article we aim to consider two stochastic intensities of default, so three stochastic underlying factors are involved in the pricing of the XVA of the derivative contract. As the the PDEs models proposed in [15] for this case requires the numerical solution of PDEs with three spatial dimensions, the computational cost would become high when using deterministic numerical methods, such as finite differences or element methods, for example. Alternative formulations based on expectations would require the numerical solution by means of Monte Carlo based techniques, with the additional use of fixed point iterations in the nonlinear case.

Unlike the approach followed in [15] that considers either PDEs with three spatial variables or formulations based on expectations for the case of stochastic intensities of default, we propose an hybrid model that starts from the PDEs in one spatial variable proposed in [8] for the case of constant intensities of default and replaces these constant coefficients by the stochastic processes posed in [15] for the intensities of default. In this way, we obtain PDEs in one spatial dimension with stochastic coefficients, also referred to as random PDEs. For their numerical solution we propose a Multi-Level Monte Carlo finite element technique which exploits the combination of Monte Carlo for the simulation of the paths of the stochastic coefficients and the finite elements solution of the PDEs arising at each path of the stochastic coefficients. More specifically, we use the Multilevel Monte Carlo (MLMC) method introduced by Michael B. Giles in [16] and later extended in [17]. This method relies on repeated random samplings that are taken on different levels of accuracy. First ideas about MLMC come from the works [19, 20], where it was first ap-

plied for the numerical solution of integral equations. It can be considered an *evolution* of the alternative crude Monte Carlo method. Actually, as indicated in [16, 17], MLMC method can significantly decrease the computational cost of crude Monte Carlo methods by taking most samples with low accuracy and corresponding low cost while only very few samples are taken at high precision and corresponding high cost. Note that a PDE with stochastic coefficients has been proposed and solved with a crude Monte Carlo strategy in [4] for the solution of an investment problem in telecommunications networks planning under uncertainty. The combination of MLMC with finite element methods has been analysed in [6] for parabolic PDEs with random coefficients. In [22], a very interesting and different approach has been proposed, where the dimension of the PDE is reduced by using the numerical solution of a one spatial dimension Black-Scholes PDE conditional to the knowledge of the volatility Monte Carlo realization in a Heston model. The method is extended to different options in [21].

The plan of the article is as follows. In the Section 2 we introduce the proposed hybrid model. In Section 3, the numerical methods for solving the the hybrid model are described. In Section 4, some numerical examples are presented and discussed. In the last section we present some conclusions.

2. Hybrid model for stochastic intensities of default

As in [8] and [15] , we consider two defaultable counterparties: the hedger (H) and the investor (I). As pointed out in the previous introduction, the main objective of this work is to propose a hybrid model with three stochastic factors, which is governed by PDEs with two coefficients that are stochastic factors. This approach avoids the alternative consideration of PDEs with three spatial variables, the numerical solution of which is more computationally demanding. Thus, in this section we pose PDEs models with one spatial dimension and two stochastic coefficients.

For this purpose, we start by recalling the model proposed in [8] with just one stochastic factor: the price of the underlying asset. Thus, we consider a portfolio with four traded assets:

- P^R: default risk-free, zero-coupon bond, with yield r^R;
- P: default risk-free, zero-coupon bond, with yield equal to the risk free rate r;
- P^H: default risky, with constant R^H recovery rate, zero-coupon bond issued by H, with yield r^H;
- P^I: default risky, with constant R^I recovery rate, zero-coupon bond issued by I, with yield r^I;
- S: underlying asset with no default risk.

The two risky bonds P^H and P^I pay one currency unit at the expiry date T if the issuing party has not defaulted, otherwise their recovery rates are R^H and R^I, respectively. Unlike [8], we prefer to start

considering recovery rates in these risky bonds. Under the real probability measure $\mathbb{P}$, the prices of the assets P^H, P^I, P^R, P and S satisfy the following stochastic differential equations (SDEs), respectively:

$$\begin{aligned} dP_t^H &= r^H(t)P_t^H dt - (1-R^H)P_t^H dJ_t^H, \\ dP_t^I &= r^I(t)P_t^I dt - (1-R^I)P_t^I dJ_t^I, \\ dP_t^R &= r^R(t)P_t^R dt, \\ dP_t &= r(t)P_t dt, \\ dS_t &= \mu(t)S_t dt + \sigma(t)S_t dW_t, \end{aligned}$$

where $r^I(t) > 0$, $r^H(t) > 0$, $r^R(t) > 0$, $r(t) > 0$, $\mu(t)$ and $\sigma(t) > 0$ denote deterministic functions of the time t. More precisely, r^H is the yield on recovery-less bond of hedger H, r^I is the yield on recovery-less bond of investor H, r^R is the rate paid for the underlying asset in a repurchase agreement and r is the risk-free rate, while μ and σ are the drift and the volatility of the underlying asset, the dynamics of which is therefore assumed to follow a Geometric Brownian Motion with time dependent parameters. Moreover, W_t represents a standard Brownian motion process, while J_t^H and J_t^I represent two independent Poisson processes that incorporate jumps.

Throughout this article we refer to the value of the risky derivative as the value of the derivative when counterparty risk is taken into account and is given by the process $\hat{V}_t = \hat{V}(t, S_t, J_t^H, J_t^I)$ for a function $\hat{V}$. Analogously, we refer to the value of the risk-free derivative as its value when there is no counterparty risk (both parties of the contract cannot default). In this case the risk-free value is denoted by $V_t = V(t, S_t)$ for a given function V. In the case of European vanilla call or put options, the function V can be computed using a Black-Scholes formula, while in other cases V satisfies the corresponding Black-Sholes PDE model.

Another relevant issue in the modelling of pricing problems including counterparty risk concerns to the definition of the mark-to-market value (also know as close-out value) of the derivative at default of either the counterparty or the hedger. This mark-to-market value is given by $M(t, S_t)$, where the function M can be defined as $M(t, S) = \hat{V}(t, S, 0, 0)$ or $M(t, S) = V(t, S)$.

In terms of the previous mark-to-market choice and depending on which part defaults first, the value of the risky derivative at default is defined as follows:

- if the investor I defaults first, then

$$\hat{V}(t, S, 1, 0) = M^+(t, S) + R^I M^-(t, S); \tag{1}$$

- if the hedger H defaults first, then

$$\hat{V}(t, S, 0, 1) = M^-(t, S) + R^H M^+(t, S); \tag{2}$$

where $R^I \in [0,1]$ (respectively $R^H \in [0,1]$) represents the recovery rate on the derivative mark-to-market value in case the party I (respectively H) defaults first.

A self-financing portfolio that covers all the underlying risk factors of the model is built in [8]. More precisely, at time t the hedger sets up a portfolio Π_t with $\delta(t)$ units of the asset with price S_t, $\alpha^H(t)$ units of the bond with price P_t^H, $\alpha^I(t)$ units of the bond with price P_t^I, and $\beta(t)$ units of cash. After imposing the self-financing condition to the portfolio, its value at time t hedges out the value of the derivative contract to the hedger, i.e. $\hat{V}_t + \Pi_t = 0$. Thus, we have

$$-\hat{V}_t = \Pi_t = \delta(t)\, S_t + \alpha^H(t)\, P_t^H + \alpha^I(t)\, P_t^I + \beta(t).$$

Next, by using Ito lemma for jump-diffusion processes we can compute the change in the risky derivative price and select the appropriate weights in the portfolio so that all risks can be removed and apply the arbitrage-free argument to the resulting risk-free portfolio (see [23], for example). Finally, depending on the choice of the mark-to-market, the following two PDEs are obtained :

- Non Linear PDE (when $M(t,S) = \hat{V}(t,S,0,0)$):

$$\begin{cases} \partial_t \hat{V} + \mathcal{A}\hat{V} - r\hat{V} = (1-R^H)\lambda^H(\hat{V})^- + (1-R^I)\lambda^I(\hat{V})^+ + s^F(\hat{V})^+, \\ U(T,S) = H(S). \end{cases} \tag{3}$$

- Linear PDE (when $M(t,S) = V(t,S)$):

$$\begin{cases} \partial_t \hat{V} + \mathcal{A}\hat{V} - (r + \lambda^H + \lambda^I)\hat{V} = -(R^H\lambda^H + \lambda^I)\hat{V}^- - (R^I\lambda^I + \lambda^H)\hat{V}^+ + s^F(\hat{V})^+, \\ \hat{V}(T,S) = H(S). \end{cases} \tag{4}$$

where $\mathcal{A}[\cdot] = \dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2}{\partial S^2}[\cdot] + r_R S \dfrac{\partial}{\partial S}[\cdot]$.

Note that our aim is the computation of the total value adjustment, $U = \hat{V} - V$, which is the difference between the price of the risky and risk-free derivative. Taking into account that V satisfies the classical Black-Scholes equation of a European vanilla option, the consideration of this classical equation jointy with the respective equations (3) and (4) leads to the following PDEs for the XVA value function U.

- If $M(t,S) = \hat{V}(t,S,0,0)$, then U satisfies the nonlinear PDE problem:

$$\begin{cases} \partial_t U + \mathcal{A}U - rU = (1-R^H)\lambda^H(V+U)^- + (1-R^I)\lambda^I(V+U)^+ + s^F(V+U)^+, \\ U(T,S) = 0. \end{cases} \tag{5}$$

- If $M(t,S) = V(t,S)$) then U satisfies the linear PDE problem:

$$\begin{cases} \partial_t U + \mathcal{A}U - (r + \lambda^H + \lambda^I)U = (1-R^H)\lambda^H(V)^- + (1-R^I)\lambda^I(V)^+ + s^F(V)^+, \\ U(T,S) = 0. \end{cases} \tag{6}$$

where

$$\lambda^H = \frac{h^H}{1-R^H}, \quad \lambda^I = \frac{H^I}{1-R^I} \tag{7}$$

provides the expression of the constant intensities of default of the hedger and investor, respectively, with $h^H = r^H - r$ and $h^I = r^I - r$ being the respective constant spreads of the bonds P_I and P_H. Moreover, r^F is the hedger funding rate for borrowed cash on hedger's derivatives replication cash account and $s^F = r^F - r$ is the funding spread.

The previous PDEs models correspond to the case of constant intensities of default, as there is only one stochastic factor S_t and therefore the spatial dimension of the governing PDEs is equal to one. In [5] the authors solved the problems (5) and (6) when considering constant intensities of default λ^I and λ^H.

As previously indicated, the main objective of this work is to consider stochastic intensities of default, instead of constant intensities. For this reason, we pose a hybrid model, we assume that the intensities of default of the investor and the hedger (λ_I and λ_H) follow their corresponding stochastic dynamics.

For this purpose, let us consider that when assuming that the short CDS spreads h_t^I and h_t^H are stochastic, the following SDEs govern their respective dynamics:

$$dh_t^I = \mu^I(t, h_t^I)dt + \sigma^I(t, h_t^I)dW_t^I, \tag{8}$$

$$dh_t^H = \mu^H(t, h_t^H)dt + \sigma^H(t, h_t^H)dW_t^H, \tag{9}$$

where $\mu^I(t, h_t^I)$ and $\mu^H(t, h_t^H)$ are the drifts of the respective processes under the real probability measure $\mathbb{P}$, while $\sigma^I(t, h_t^I)$ and $\sigma^H(t, h_t^H)$ represent their corresponding volatilities. Moreover, W_t^I and W_t^H are standard Brownian motions under the real world measure $\mathbb{P}$.

More precisely, we consider Geometric Brownian motions for the stochastic processes h_t^I and h_t^H and choose the following expressions for their respective drift and volatility terms:

$$\mu^I(t, h_t^I) = -\frac{k^I}{1-R^I}h_t^I, \quad \sigma^I(t, h_t) = \sigma^I h_t^I, \tag{10}$$

$$\mu^H(t, h_t^H) = -\frac{k^H}{1-R^H}h_t^H, \quad \sigma^H(t, h_t^H) = \sigma^H h_t^H, \tag{11}$$

where k^I, σ^I, k^H and σ^H are constant. Note that in the stochastic spreads setting we will assume that there is no correlation between the two processes defined by equations (8) and (9), as well as between the stochastic spreads and the price of the underlying asset. Therefore, the identities (7) for the constant case turn into the following relations between the stochastic processes of intensities of default and spreads:

$$\lambda_t^H = \frac{h_t^H}{1-R^H}, \quad \lambda_t^I = \frac{h_t^I}{1-R^I}. \tag{12}$$

Next, applying Itô formula in (12) and taking into account the dynamics for the stochastic spreads defined by SDEs (8) and (9), we can obtain the SDEs for the dynamics of the stochastic processes λ_t^I and λ_t^H:

$$d\lambda_t^I = -\frac{k^I}{1-R^I}\lambda_t^I dt + \sigma_h^I \lambda_t^I dW_t^I, \tag{13}$$

$$d\lambda_t^H = -\frac{k^H}{1-R^H}\lambda_t^H dt + \sigma_h^H \lambda_t^H dW_t^H. \tag{14}$$

Once the SDEs satisfied by the stochastic processes λ_t^I and λ_t^H have been specified in (13) and (14), the hybrid model we propose consists in either the linear PDEs problem defined by (6) or the one given by (5), where the coefficients λ^I and λ^H are stochastic and satisfy (13) and (14).

So, the proposed hybrid models can be framed as formulations in terms of PDEs with stochastic coefficients, also known as random PDEs. Note that for each path of the stochastic coefficients in the PDEs we obtain a path of the solution, as the solutions of these hybrid formulations are stochastic processes.

3. Numerical methods for the hybrid model

In this section, we describe the proposed numerical methodologies to solve the hybrid model that has been posed in the previous section. This hybrid model is formulated in terms of (linear or nonlinear) PDEs with a couple of coefficients that are stochastic processes (namely, the intensities of default) that satisfy the SDEs (13) and (14). Therefore, for each path of the intensities of default we obtain PDEs with time dependent coefficients. So, the idea is to simulate a large enough number of paths of the stochastic intensities of default, obtain the corresponding solutions for each path and compute the expectation of this solutions as the XVA price.

Basically, the numerical techniques combine two methods: the Multi-Level Monte Carlo method introduced in [16, 17], in order to deal with the stochastic factors λ^I and λ^H appearing in the PDEs (5) and (6), jointly with a Lagrange-Galerkin method proposed in [5] to solve the PDEs with one spatial dimension that arise in the case of constant intensities of default. In next sections we briefly describe both techniques.

3.1. Lagrange-Galerkin method for solving PDE models

In this section we assume that the paths of the stochastic intensities of default that satisfy their corresponding SDEs have been obtained or approximated by suitable stochastic numerical methods. Therefore, next step consists of the numerical solution of the linear and nonlinear PDEs for each path of the stochastic coefficients. Among the different possibilities, we propose a Lagrange-Galerkin method which combines a first order semi-Lagrangian (also referred to as characteristics) method for the time discretization with a

piecewise linear Lagrange finite element method for the discretization in the spatial variable. This combination has been introduced [5] for European options with constant intensities of default and later used in [1, 3] for European and American options, respectively, with one stochastic intensity of default (i.e., two spatial dimensions in the PDEs). Although a detailed description of the proposed method can be found in [5], in next paragraphs we briefly summarize the main steps by taking into account an updated notation for the paths of the intensities of default.

As we will need to compute the numerical solution of PDEs in Monte Carlo paths for different meshes, we show here how to compute the numerical solution U_p of the PDE for a fixed path p. Moreover, for simplicity, only the linear PDE case is explained. Note that the nonlinear case involves an additional fixed point iteration in the numerical solution with the sign nonlinearity of the right hand side of the PDE (5) evaluated at the previous step.

Taking into the previous remarks, for each path p, the PDE (6) is rewritten with the updated notation in the equivalent divergence form, also applying a change of the time variable from the physical time t to the time to maturity $\tau = T - t$:

$$\begin{cases} \dfrac{\partial U_p}{\partial \tau} - \dfrac{\partial}{\partial S}\left(\dfrac{\sigma^2}{2}S^2\dfrac{\partial U_p}{\partial S}\right) + (\sigma^2 - r_R)S\dfrac{\partial U_p}{\partial S} - (r + \lambda_p^H + \lambda_p^I)U_p \\ \qquad\qquad = (1-R^H)\lambda_p^H(V)^- + (1-R^I)\lambda_p^I(V)^+ + s^F(V)^+, \\ U_p(0,S) = 0, \end{cases} \tag{15}$$

where λ_p^I and λ_p^H denote the (time dependent) paths p of the intensities of default and U_p denotes the associated path p of the XVA stochastic process. Next, we introduce the material derivative of the function U_p, as follows:

$$\frac{DU_p}{D\tau}(\tau, S) = \frac{dU_p}{d\tau}(\tau, S(\tau)) = \frac{\partial U_p}{\partial \tau}(\tau, S) + \frac{\partial U_p}{\partial S}(\tau, S)\frac{dS}{d\tau}(\tau), \tag{16}$$

where $S = S(\tau)$. Note that the material derivative of U_p represents the derivative along the integral paths (characteristics curves) associated to the vector field $v(S) = (\sigma^2 - r^R)$. In terms of the material derivative, equation (15) can be written in the following way:

$$\frac{DU_p}{D\tau} - \frac{\sigma^2}{2}\frac{\partial}{\partial S}\left(S^2\frac{\partial U_p}{\partial S}\right) + (r + \lambda_p^H + \lambda_p^I)U_p = (1-R^H)\lambda_p^H(V)^- + (1-R^I)\lambda_p^I(V)^+ + s^F(V)^+. \tag{17}$$

Next, for the time discretization of equation (17) we consider an upwind approximation of the material derivative along the characteristics curves at the time discretization nodes. More precisely, for a natural number $N > 0$, a uniform mesh with the time step $\Delta\tau = T/N_T$ and the nodes $\tau^n = n\Delta\tau$, $n = 0, 1, \ldots, N_T$ are considered. For each node τ^n, the following finite differences approximation is chosen:

$$\frac{DU_p}{D\tau}(\tau^{n+1}, .) \approx \frac{U_p^{n+1} - U_p^n \circ \chi^n}{\Delta\tau}, \tag{18}$$

where $\chi^n(S)$ represents the position at time τ^n of a point moving through the characteristics curve χ of the velocity field that passes at point S at time τ^{n+1}. Therefore, $\chi^n(S)$ can be obtained from the solution $\chi(\tau^{n+1}, S; .)$ of the following Cauchy problem:

$$\begin{cases} \dfrac{d\chi}{d\tau}(\tau^{n+1}, S; \tau) = (\sigma^2 - r_R)\chi(\tau^{n+1}, S; \tau), \\ \chi(\tau^{n+1}, S; \tau^{n+1}) = S, \end{cases} .$$

Actually, the analytical solution of (3.1) implies that $\chi^n(S) = \chi(\tau^{n+1}, S; \tau^n) = \exp((r_R - \sigma^2)\Delta\tau$, for $n = 0, 1, \ldots, N_T - 1$.

Next, replacing the approximation (18) in (17), we pose the semidiscretized problem in time at step n:

$$\begin{aligned} \frac{U_p^{n+1} - U_p^n \circ \chi^n}{\Delta\tau} - \frac{\sigma^2}{2}\frac{\partial}{\partial S}\left(S^2 \frac{\partial U_p^{n+1}}{\partial S}\right) + (r + \lambda_{p,n+1}^H + \lambda_{p,n+1}^I)U_p^{n+1} = \\ (1 - R^H)\lambda_{p,n+1}^H (V_p^{n+1})^- + (1 - R^I)\lambda_{p,n+1}^I (V_p^{n+1})^+ + s^F (V_p^{n+1})^+, \end{aligned} \tag{19}$$

where $\lambda_{p,n+1}^H$ and $\lambda_{p,n+1}^I$ are the simulated values at time t^{n+1} on the path p of the intensities of default.

For the spatial discretization, we consider a large enough value S_∞ to define the bounded computational domain $\Omega = [0, S_\infty]$ and use a piecewise linear Lagrange finite element method as in [5]. For this purpose, for a fixed natural number $N_S > 0$, we consider a uniform mesh of the domain $\Omega = [0, S_\infty]$, with nodes $S_i = i\Delta S, i = 1, \ldots . N_S + 1$, where the mesh step is $\Delta S = S_\infty / N_S$. Associated to this uniform mesh, a piecewise linear Lagrange finite element discretization is considered. Thus, the fully discretized problem aims to find $U_p^{n+1} \in W_k$, such that:

$$\begin{aligned} & (1 + (r + \lambda_{p,n+1}^H + \lambda_{p,n+1}^I)\Delta\tau) \int_0^{S_\infty} U_p^{n+1} \phi \, dS - \frac{\Delta\tau\, \sigma^2}{2} \int_0^{S_\infty} S^2 \frac{\partial U_p^{n+1}}{\partial S}\frac{\partial \phi}{\partial S} \, dS \\ = & \int_0^{S_\infty} (U_p^n \circ \chi^n)(S)\phi \, dS \\ & -\Delta\tau \int_0^{S_\infty} ((1 - R^H)\lambda_{p,n+1}^H (V_p^{n+1})^- + (1 - R^I)\lambda_{i,n+1}^I (V_p^{n+1})^+ + s^F (V_p^{n+1})^+)\phi \, dS, \end{aligned} \tag{20}$$

for all $\phi \in W_{k,0}$, where

$$W_k = \{\phi : (0, S_\infty) \to \mathbb{R} | \phi \in \mathcal{C}(0, S_\infty), \phi_{|[S_i, S_{i+1}]} \in \mathcal{P}_1\}, \tag{21}$$

$$W_{k,0} = \{\phi \in W_k | \phi(0) = 0, \phi(S_\infty) = 0\}, \tag{22}$$

with $\mathcal{P}_1$ being the space of polynomials of degree less or equal than one. Note that a variational formulation is implicitly posed before the discretization with the finite element method (see [5] for the details).

3.2. Discretization of the paths of intensities of default

In order to solve the hybrid model proposed in Section 2, we must address the solution of PDEs with stochastic coefficients. More precisely, for each path the values of the intensities of default at the time

discretization nodes t^n are needed for the numerical solution of the PDE. For this purpose, the paths of the intensities of default are simulated by solving their respective SDEs (13) and (14).

For the numerical solution of SDEs (13) and (14), an Euler-Maruyama scheme, with the same time step as the time discretization of the PDE, is applied. In the case of a crude Monte Carlo method, starting from $\lambda^H_{p,0} = \lambda^H_0$ and $\lambda^I_{p,0} = \lambda^I_0$ given, we approximate the path p, with $p = 1, \ldots, N_P$, as follows:

$$\lambda^H_{p,n+1} = \lambda^H_{p,n} - \frac{k^I}{1 - R^H}\lambda^H_{p,n}\Delta\tau + \sigma^H\lambda^H_{p,n}\Delta W^H_{p,n}, \tag{23}$$

$$\lambda^I_{p,n+1} = \lambda^I_{p,n} - \frac{k^I}{1 - R^I}\lambda^I_{p,n}\Delta\tau + \sigma^I\lambda^I_{p,n}\Delta W^I_{p,n}, \tag{24}$$

for $n = 0, 1, \ldots, N_T - 1$, and $\Delta W^H_{p,n}$ and $\Delta W^I_{p,n}$ being the Brownian increments of path p sampled from two independent normal distributions $\mathcal{N}(0, \sqrt{\Delta\tau})$.

Note that in the Multi-Level Monte Carlo method a more smart sampling procedure is addressed and notation becomes a bit more complex.

3.3. Monte Carlo combined with Lagrange-Galerkin method

As we will compare it with the proposed Multi-Level Monte Carlo, in this section we describe the use of the crude Monte Carlo (MC) method that involves the following three steps:

1. *Sample*: We approximate N_P sample paths of the stochastic intensities of default λ^I and λ^H by using the previously described Euler-Maruyama scheme according to expressions (23)-(24).

2. *Solve*: For each approximated path of the stochastic intensities of default λ^I and λ^H, we solve numerically the linear PDE (6) or the nonlinear one (5) by means of the proposed Lagrange-Galerkin method. We denote the approximated solutions at the mesh nodes at the path p, by $U^{i,n}_p = U_p(t^n, S_i)$.

3. *Expected value*: As we are interested in the XVA value at time to maturity $\tau = T$ (which corresponds to time $t = 0$) and $T = \tau_{N_T}$ in the time discretization mesh, the expectation of the stochastic process U at the spatial mesh node S_i and the time mesh node τ_{N_T} is estimated by the sample mean (average) of the previously computed approximated solutions U^{i,N_T}_p, for $p = 1 \ldots N_P$, i.e.:

$$\mathbb{E}[U^{i,N_T}] = \frac{1}{N_P}\sum_{p=1}^{N_P} U^{i,N_T}_p. \tag{25}$$

3.4. Multi-Level Monte Carlo combined with Lagrange-Galerkin method

As previously explained, in this work we aim to solve a PDE with stochastic coefficients by combining the numerical the solution of a PDE for the given sample paths of the stochastic process with Monte Carlo techniques to compute the expected value of the solution. As the use of a crude Monte Carlo technique

could be very expensive from a computational point of view, we propose the use of the Multi-Level Monte Carlo method (MLMC), introduced and described in detail in [16, 17] and can be also understood as a variance reduction technique. As it is shown and proved in [16, 17], the MLMC is used to evaluate an expected value with reduced computational costs respect to the use Monte Carlo method. One of the main ideas of MLMC comes from the sampling of different approximation of the quantity whose expectation we want to estimate. Also it borrows some ideas from multigrid methods for solving PDEs. A detailed mathematical analysis of MLMC combined with finite elements for parabolic PDEs with stochastic coefficientes can be found in [6].

Note that MLMC has already been used for solving elliptic PDEs with random coefficients in [12], for example. In this setting, one of the main ideas behind MLMC-FEM methods is to simultaneously draw appropriate Monte Carlo samples on a hierarchy of nested spatial grids. As in our case, we are dealing with parabolic PDEs, we will consider an appropriate sampling of the random coefficients for the corresponding nested spatial and time meshes, which uses more sample paths for coarser meshes and decreases the number of paths as soon as meshes become finer.

More precisely, we consider a set of N_L nested uniform spatial meshes of the computational domain $[0, S_\infty]$. For $l = 0, \ldots, N_L$, let Δx_l be the step size of the spatial mesh at the level l of refinement, with Δx_0 being the step size of the mesh at level 0. We denote by N_S^l the number of nodes at the spatial mesh of level l. We choose $\Delta x_l = 2^{-l}\Delta x_0$. Concerning the nested meshes in the time discretization, if $\Delta\tau_0$ is the time step of the mesh at level 0, the time step for a mesh at level l is chosen as $\Delta\tau_l = 4^{-l}\Delta\tau_0$. Moreover, we denote by N_T^l the number of nodes in the time mesh of level l.

At level l, each sample path of each intensity of default requires the sampling of N_T^l values, so that we have a sample at each discretization time. More precisely, in MLMC setting, starting from $\lambda_{p,0}^H$ and $\lambda_{p,0}^H$ given, at level l we approximate the path p as follows:

$$\lambda_{p,n+1,l}^H = \lambda_{p,n,l}^H - \frac{k^I}{1-R^H}\lambda_{p,n,l}^H\Delta^l\tau + \sigma^H\lambda_{p,n,l}^H\Delta W_{p,n,l}^H, \tag{26}$$

$$\lambda_{p,n+1,l}^I = \lambda_{p,n,l}^I - \frac{k^I}{1-R^I}\lambda_{p,n,l}^I\Delta^l\tau + \sigma^I\lambda_{p,n,l}^I\Delta W_{p,n,l}^I, \tag{27}$$

for $n = 0, 1, \ldots, N_T^l$, $p = 1, \ldots, N_P^l$ with $\Delta W_{p,n,l}^H$ and $\Delta W_{p,n,l}^I$ being the Brownian increments of path p sampled from two independent normal distributions $\mathcal{N}(0, \sqrt{\Delta\tau_l})$. By using these samples, we first solve the PDE with the semi-Lagrangian method for the spatial and time meshes corresponding to level l. Moreover, assuming that $N_T^l = 4N_T^{l-1}$, for each of the previous sample paths at the level l, we can extrapolate the N_T^{l-1} samples of the intensities corresponding to the immediate coarser nested time mesh and use them to solve the PDE in the mesh of level $l-1$.

If we denote the number of sample paths at level l by N_P^l, then we can estimate the expectation of the difference between the results obtained with meshes of levels l and $l-1$, that is:

$$\mathbb{E}[U^{i,N_T^l,l} - U^{i,N_T^{l-1},l-1}] = \frac{1}{N_P^l}\sum_{p=1}^{N_P^l}(U_p^{i,N_T^l,l} - U_p^{i,N_T^{l-1},l-1})\,. \tag{28}$$

As we are using the same sample paths due to the extrapolation procedure, we reduce the variance of the estimator of (28) in the MLMC method for a fixed computational cost.

Starting from $l=1$, we can repeat the process and the estimation of the expectation of the XVA at $\tau = T$ is given by

$$\mathbb{E}[U^{i,N_T^L,L}] = \mathbb{E}[U^{i,N_T^0,0}] + \sum_{l=1}^{N_L}\mathbb{E}[U^{i,N_T^l,l} - U^{i,N_T^{l-1},l-1}]\,. \tag{29}$$

Previously, at level $l=0$ we start solving the problem in the coarsest meshes in space and time on the different paths from the samples at that level to obtain $\mathbb{E}[U^{i,N_T^0,0}]$, which is given by:

$$\mathbb{E}[U^{i,N_T^0,0}] = \frac{1}{N_P^0}\sum_{p=1}^{N_P^0}U_p^{i,N_T^0,0}\,. \tag{30}$$

Note that a different number of paths N_P^l is used at each level, so that a larger number of paths is used in the coarser meshes, while using a lower number of paths for the finer meshes.

In view of the previous description, the main steps of the MLMC Lagrange Galerkin algorithm are the following:

1. *Start with l=0:*
   (a) *Sample:* Simulate N_P^0 paths of λ_0^I and λ_0^H with N_T^0 time steps,
   (b) *Solve:* For each path, solve the nonlinear PDE (5) or the linear PDE (6) with the spatial and time meshes at level $l=0$.
   (c) *Expected value:* Obtain the expectation associated to level $l=0$.
2. *For each* $l = 1, \ldots, N_L$:
   (a) *Sample:* Simulate N_P^l sample paths of λ_l^I and λ_l^H with N_T^l time steps.
   (b) *Solve:* For the paths λ_l^I and λ_l^H, the PDEs (5) or (6) are solved. Also solve the same PDE for the extrapolated samples λ_{l-1}^I and λ_{l-1}^H obtained from the previous samples of λ_l^I and λ_l^H.
   (c) *Expected values:* Compute the expectation in formula (28).
3. *XVA price at final time* $\tau = T$*, (i.e.,* $t = 0$*)*: Compute the expectation given by expression (29).

4. Numerical results

In this section we aim to illustrate the performance of the here proposed hybrid model and its numerical solution by means of Multilevel Monte Carlo technique when combined with a Lagrange-Galerkin method to solve the PDEs arising for each sample path of the stochastic coefficients. For this purpose, we consider the example of a European put option, the payoff of which is given by:

$$H(S) = \max(K - S, 0). \tag{31}$$

For the put option we assume that the maturity is $T = 0.5$ and the strike is $K = 2$. Moreover, the initial time is set to zero.

Besides the previous data about the put option, Table 1 contains the values of the other financial parameters that have been considered to obtain the numerical results.

Parameter	**Definition**	**Value**
r	Risk-free rate	0.04
σ	Volatility of the underlying asset S	0.3
λ_0^H	Hedger intensity of default at time 0	0.04
R^H	Hedger recovery rate	0.4
λ_0^I	Investor intensity of default at time 0	0.04
R^I	Investor recovery rate	0.3
r^R	Rate associated to a repurchase agreement	0.06
σ^H	Volatility of λ^H	0.2
σ^I	Volatility of λ^I	0.2
k^H	constant in the drift term of λ^H	0.1/0.3/0.5/0.7
k^I	constant in the drift term of λ^I	0.1/0.3/0.5/0.7

Table 1: Values of the parameters used in the numerical tests

Throughout this section we consider the bounded computational domain that corresponds to $S_\infty = 20$ for the involved PDEs. In order to compute the reference solution, we use a MLMC technique with $N_L = 6$ levels, with $N_S^6 = 2049$ nodes in the finer spatial finite element mesh, while $N_T^6 = 16384$ in the finer time mesh and considering $N_P^6 = 26$ simulations at the finer level.

In the computation of the numerical solution with MLMC, the number of nodes in the spatial finite element meshes for the different levels is $\{N_S^l,\, l = 0, \ldots, 5\} = \{33,\, 65,\, 127,\, 257,\, 513,\, 1025\}$, while in the time meshes is $\{N_T^l,\, l = 0, \ldots, 5\} = \{4,\, 16,\, 64,\, 256,\, 1024,\, 4096\}$. Moreover, the number of paths

computed at the different levels is $\{N_p^l,\, l = 0, \ldots, 5\} = \{1664, 832, 416, 208, 104, 52\}$.

In the MC technique we have used the different spatial and times meshes of the different MLMC levels and $N_P = 1664$ simulations.

In Figure 1, we show the error in MC and MLMC methods versus the execution time of the code. More precisely, in the MLMC case, the error between the previously described reference solution and the ones obtained at the different levels $l = 1, \ldots, 5$ is shown. In the case of MC, we consider the difference between the reference solution and the ones obtained with the meshes at different MLMC levels and $N_S = 1664$ simulation. As we can see in Figure 1, the error provided by the MLMC is smaller than the MC error for a given computational cost, although we have used a more reduced number of samples for the MLMC. This justifies the choice of MLMC for our algorithm to solve the XVA computation with the hybrid model.

All algorithms have been implemented from scratch in C++. When using MLMC with $N_L = 5$, the code takes only 90 seconds to compute the final result, on a laptop provided with an Intel i7 processor (4 Physical CPUs, 8 Virtual) and 16GB of Memory.

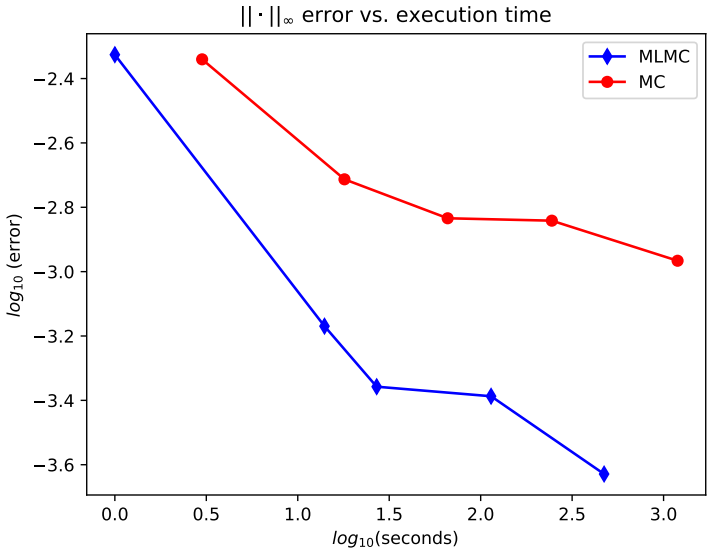

Figure 1: Error MLMC VS Error MC, with respect to execution times.

.

Next, we present the comparison of the results obtained with the hybrid model that incorporates stochastic intensities of default with the ones obtained when both intensities are assumed to be constant. The case of constant intensities of default has been numerically addressed in [5] for the case of European options, while the mathematical analysis of the model has been developed in [2]. The comparison has

been addressed for both the linear and the nonlinear PDEs.

More precisely, we start with a set of examples concerning the linear PDE case. In Figure 2 we consider null intensity of default of the investor (i.e., $\lambda_I = 0$), while the hedger can have constant or stochastic intensity of default. In the stochastic case, we assume an initial intensity of default λ_0^h equal to the constant one $\lambda^H = 0.04$. Note that the more negative value of the XVA corresponds to the constant case, while the XVA becomes less negative when the drift parameter k^H of the intensity of default becomes higher. Next, in Figure 3 we just consider $\lambda^I = 0.04$ instead of $\lambda^I = 0$ as in Figure 2. We observe the same behaviour when increasing the drift parameter k^H, although the respective XVA values are more negative. Next, in Figure 4 we consider the comparison between both λ^I and λ^H stochastic and both constant. Again, the initial values of the stochastic case coincide with the respective constant values. Note that the consideration of stochastic λ^I instead of constant makes the XVA values less negative, as we can see from the comparison of Figures 3 and 4.

In summary, for the linear PDE case, we have distinguished three different cases:

- λ^H is constant and λ^I is null, compared to the results of the same PDE, with λ^H is stochastic ($k^H = 0.1, 0.3, 0.5, 0.7$) and λ^I null.
- λ^H and λ^I are constant, compared to the results of the same PDE, with λ^H stochastic ($k^H = 0.1, 0.3, 0.5, 0.7$) and λ^I constant.
- λ^H and λ^I are constant, compared to the results of the same PDE, with λ^H and λ^I both stochastic ($k^H = k^I = 0.1, 0.3, 0.5, 0.7$) .

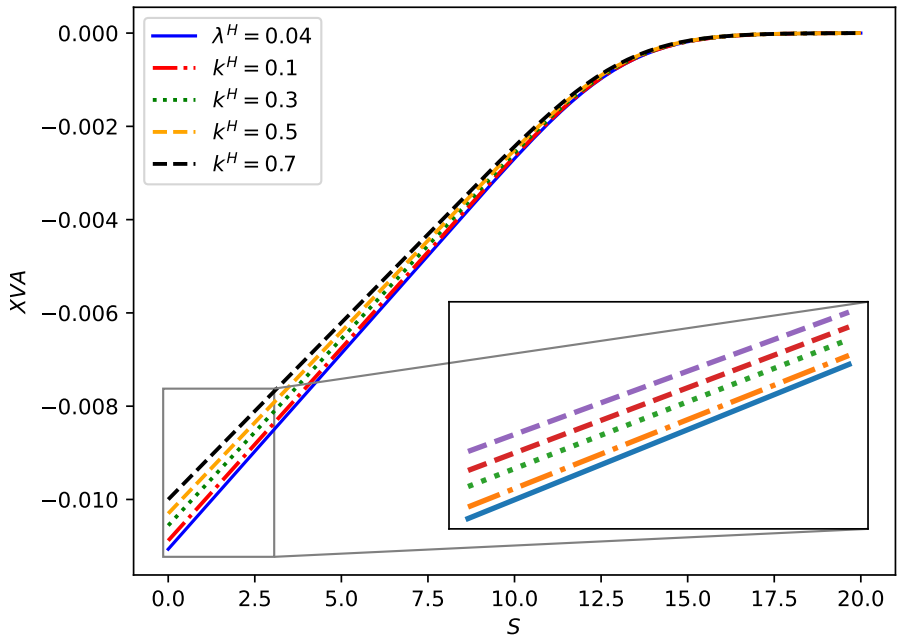

Figure 2: XVA with linear PDE: λ^I null and λ^H stochastic ($\lambda_0^H = 0.04$) versus $\lambda^H = 0.04$ constant.

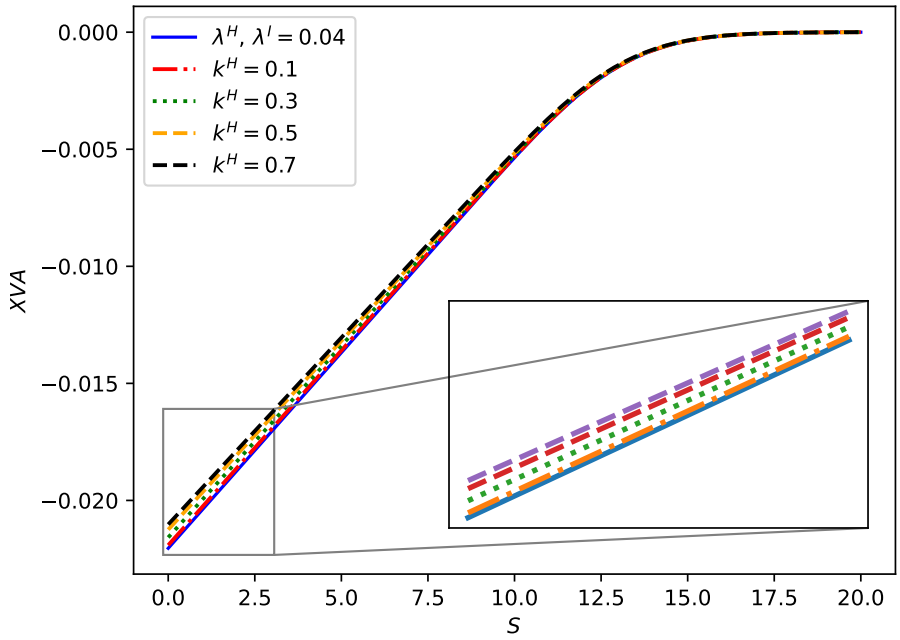

Figure 3: XVA with linear PDE: λ^H stochastic ($\lambda_0^H = 0.04$) and $\lambda^I = 0.04$ constant versus $\lambda^H = 0.04$ and $\lambda^I = 0.04$ constant.

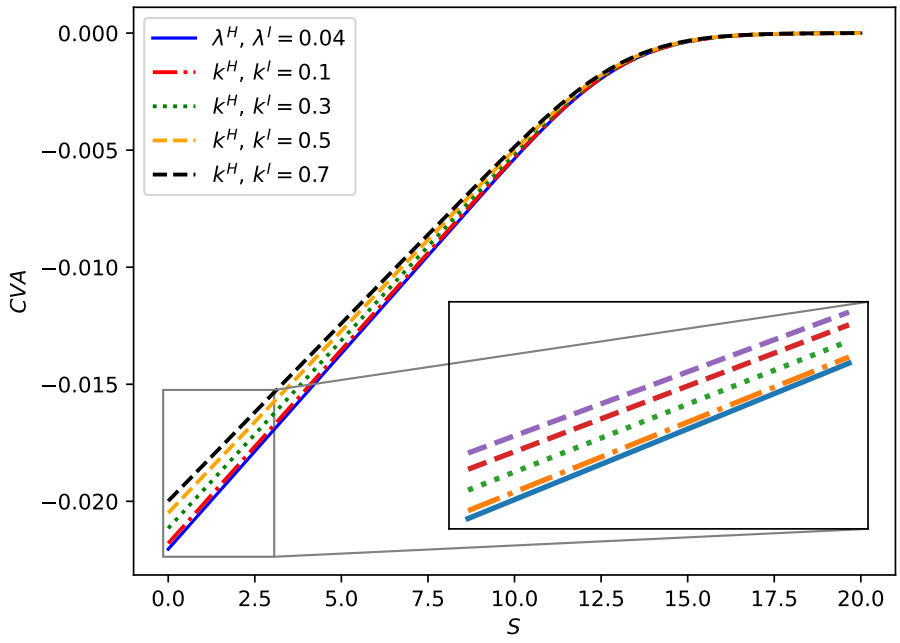

Figure 4: XVA with linear PDE: λ^H and λ^I stochastic ($\lambda_0^H = 0.04, \lambda_0^I = 0.04$) versus $\lambda^H = 0.04$ and $\lambda^I = 0.04$ constant.

Next, in Figures 5, 6 and 7, we display the result for the analogous examples when non linear PDE is considered. The behavior observed in each example of the nonlinear PDE case is similar to the corresponding one of the linear case. Also note that the consideration or not of stochastic intensities usually has more impact than the choice of a linear or nonlinear PDE (which is related to the choice of the mark-to market value).

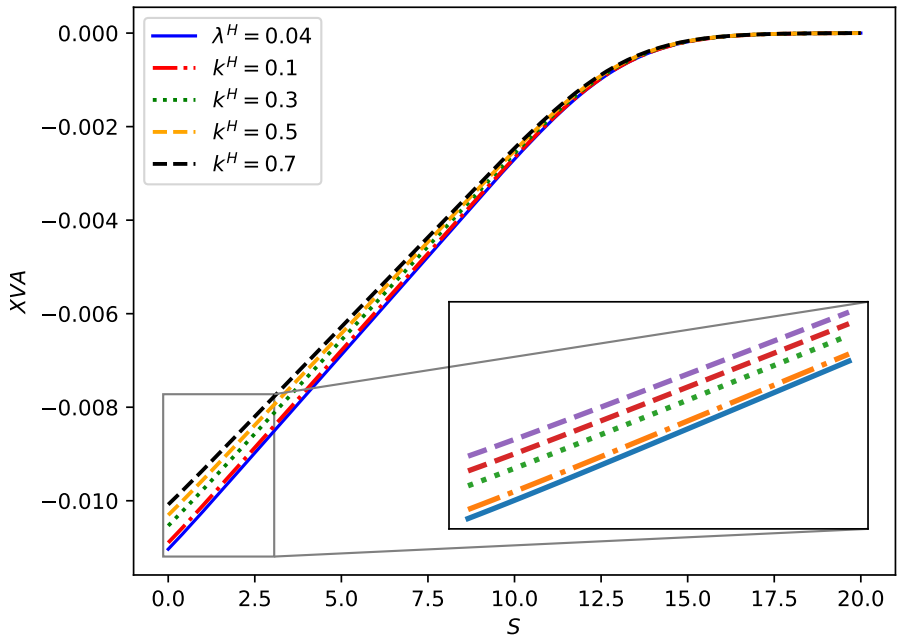

Figure 5: XVA with nonlinear PDE: λ^I null and λ^H stochastic ($\lambda_0^H = 0.04$) versus $\lambda^H = 0.04$ constant.

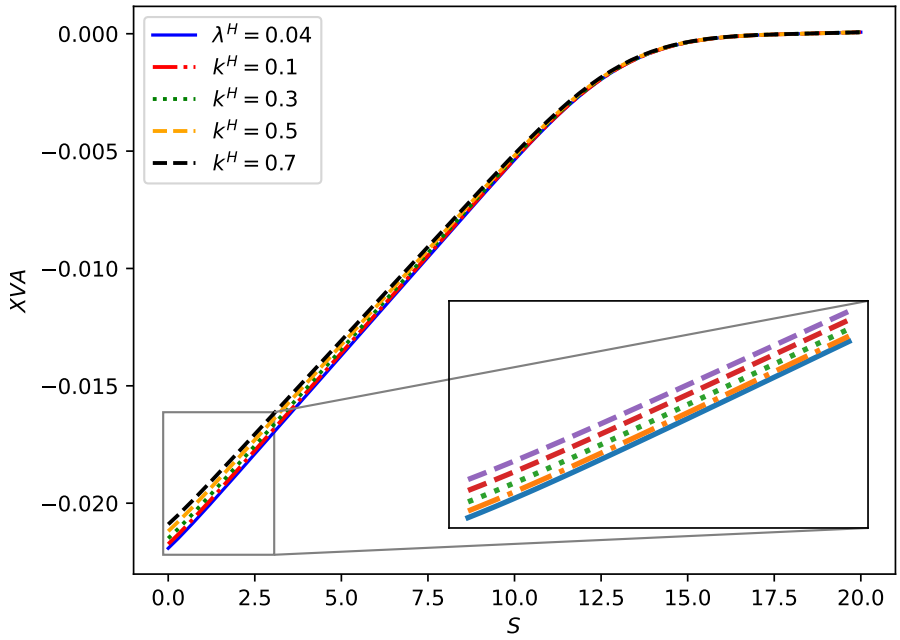

Figure 6: XVA with nonlinear PDE: λ^H stochastic ($\lambda_0^H = 0.04$) and $\lambda^I = 0.04$ constant versus $\lambda^H = 0.04$ and $\lambda^I = 0.04$ constant.

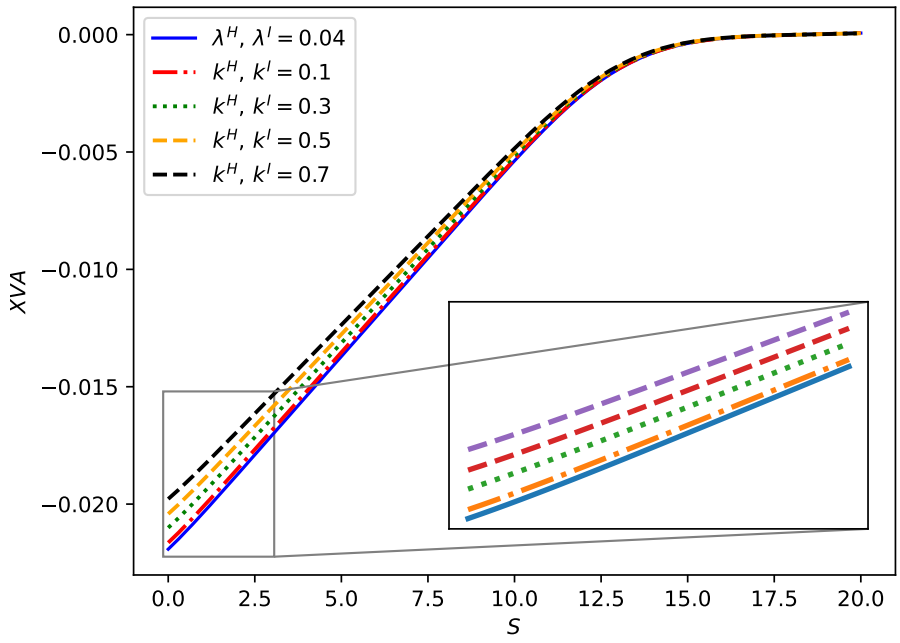

Figure 7: XVA with nonlinear PDE: λ^H and λ^I stochastic ($\lambda_0^H = 0.04, \lambda_0^I = 0.04$) versus $\lambda^H = 0.04$ and $\lambda^I = 0.04$ constant.

5. Conclusions

In this article we propose XVA pricing models for European options that incorporate three stochastic factors, namely, the price of the underlying asset and the intensities of default of the investor and the hedger. The model is formulated in terms of nonlinear or linear PDEs with stochastic coefficients and one spatial dimension. The spatial variable represents the asset price, while the stochastic coefficients are the intensities of default, whose dynamics are prescribed by means of appropriate stochastic differential equations.

For the numerical solution, we propose a Multi-Level Monte Carlo Lagrange-Galerkin method, where the PDEs obtained for each sample of the stochastic coefficients are solved by combining a semi-Lagrangian method for time discretization with a finite elements technique for the spatial discretization (Lagrange-Galerkin mehod). Unlike the more classical Monte-Carlo method for solving PDEs with stochastic coefficients, MLMC provides better results by considering nested meshes in space and time with a larger number of sample paths of the stochastic coefficients for the coarser meshes than for the finer ones. Numerical results illustrate the advantage of MLMC with respect to MC in terms of accuracy and com-

putational cost. Moreover, the numerical results allow to compare the XVA values when models with constant intensities or stochastic ones are consedered.

References

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

[2] I. Arregui, B. Salvador, D. Ševčovič, C. Vázquez, Mathematical analysis of a nonlinear PDE model for European options with counterparty risk, *Comptes Rendus Mathématique*, 357 (2019), 252-257.

[3] I. Arregui, B. Salvador, D. Ševčovič, C. Vázquez, PDE models for American options with counterparty risk and two stochastic factors: mathematical analysis and numerical solution, *Computers & Mathematics with Applications*, 79 (2020), 1525-1542.

[4] I. Arregui, B. Salvador, C. Vázquez, A numerical strategy for telecommunications networks capacity planning under demand and price uncertainty, *Journal of Computational and Applied Mathematics*, 318 (2017) 491-503.

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

[6] A. Barth, A. Lang, Ch. Schwab. Multilevel Monte Carlo method for parabolic stochastic partial differential equations, *BIT Numerical Mathematics*, 53 (2013) 3–27.

[7] P. Boyle, Options: A Monte Carlo approach, *Journal of Financial Economics*, 4:323 (38) 1977.

[8] C. Burgard, M. Kjaer. *PDE representations of derivatives with bilateral counterparty risk and funding costs*, *Journal of Credit Risk*, 7 (2011) 75–93.

[9] D. Brigo, M. Morini, A. Pallavicini. *Counterparty Credit Risk, Collateral and Funding*. Wiley Finance, 2013.

[10] Y. Chen and C. Christara. Penalty methods for bilateral XVA pricing in Eurpean and American contingent claims by PDE model, *Journal of Computational Finance*, 24 (2021) 41–70.

[11] Y. Chen and C. Christara. Bilateral XVA pricing under stochastic default intensity: PDE modelling and computation, *to appear in Applied Numerical Mathematics in 2023*.

[12] K.A. Cliffe, M.B. Giles, R. Scheichl, A.L. Teckentrup, Multilevel Monte Carlo Methods and Applications to Elliptic PDEs with Random Coefficients, *Computing and Visualization in Science*, 14(2011) 3-15.

[13] S. Crepey, T. Bieleckii. *Counterparty risk and funding: a tale of two puzzles*. Chapman and Hall-CRC Press, 2014.

[14] D. Farahany, K. Jackson, S. Jaimungal, Mixing LSMC and PDE Methods to Price Bermudan Options, *SIAM Journal Financial Mathematics*, 11 (2020) 201–239.

[15] L. M. García. *CVA, FVA (and DVA?) with stochastic spreads. A feasible replication approach under realistic assumptions*. MPRA Paper n 44568, February 2013.

[16] M. Giles. Multilevel Monte Carlo path simulation, *Operations Research*, 56 (3) (2008) 607–617.

[17] M. Giles. Multilevel Monte Carlo methods, *Acta Numerica*, 24 (2015) 259–328.

[18] J. Gregory. *Counterparty credit risk and credit value adjustment*. Wiley Finance, 2012.

[19] S. Heinrich. Monte Carlo complexity of global solution of integral equations. *Journal of Complexity*, 14 (1998) 151–75.

[20] S. Heinrich. Multilevel Monte Carlo methods. In: *Lecture Notes in Computer Science*, Vol. 2179, 58-67, Springer, 2001.

[21] T. Lipp, G. Loeper, O. Pironneau. Mixing Monte-Carlo and partial differential equations for pricing options, *Chinese Annals of Mathematics Series B*, 34B (2) (2013) 255–276.

[22] G. Loeper, O. Pironneau. A mixed PDE/Monte Carlo Method for stochastic volatility models, *Comptes Rendues Academie des Sciences de Paris, Séries I: Mathématiques*, 347 (2009) 559–563.

[23] B. Salvador *Modelling, mathematical analysis and numerical simulation of problems related to conterparty risk and CVA*. Ph.D. Thesis, Universidade da Coruña, 2018.