Efficient XVA computation under local Lévy models

Anastasia Borovykh*, Andrea Pascucci†, and Cornelis W. Oosterlee‡

Abstract. Various valuation adjustments, or XVAs, can be written in terms of non-linear PIDEs equivalent to FBSDEs. In this paper we develop a Fourier-based method for solving FBSDEs in order to efficiently and accurately price Bermudan derivatives, including options and swaptions, with XVA under the flexible dynamics of a local Lévy model: this framework includes a local volatility function and a local jump measure. Due to the unavailability of the characteristic function for such processes, we use an asymptotic approximation based on the adjoint formulation of the problem.

Key words. Fast Fourier Transform, CVA, XVA, BSDE, characteristic function

AMS subject classifications. 35R09, 65C30, 91B70, 60E10

1. Introduction.

After the financial crisis in 2007, it was recognized that Counterparty Credit Risk (CCR) poses a substantial risk for financial institutions. In 2010 in the Basel III framework an additional capital charge requirement, called Credit Valuation Adjustment (CVA), was introduced to cover the risk of losses on a counterparty default event for over-the-counter (OTC) uncollateralized derivatives. The CVA is the expected loss arising from a default by the counterparty and can be defined as the difference between the risky value and the current risk-free value of a derivatives contract. CVA is calculated and hedged in the same way as derivatives by many banks, therefore having efficient ways of calculating the value and the Greeks of these adjustments is important.

One common way of pricing CVA is to use the concept of expected exposure, defined as the mean of the exposure distribution at a future date. Calculating these exposures typically involve computationally time-consuming Monte Carlo procedures, like nested Monte Carlo schemes or the more efficient least squares Monte Carlo method (LSM)([19]). Recently the Stochastic Grid Bundling method (SGBM) was introduced as an improvement of the standard LSM ([15]). This method was extended to pricing CVA for Bermudan options in [10]. Another recently introduced alternative is the so-called finite-differences Monte Carlo method (FDMC), see [7]. The FDMC method uses the scenario generation from the Monte Carlo method combined with finite-difference option valuation.

Besides CVA, many other valuation adjustments, collectively called XVA, have been introduced

*Dipartimento di Matematica, Università di Bologna, Bologna, Italy. (anastasia.borovykh2@unibo.it)

†Dipartimento di Matematica, Università di Bologna, Bologna, Italy. (andrea.pascucci@unibo.it)

‡Centrum Wiskunde & Informatica, Amsterdam, The Netherlands and Delft University of Technology, Delft, The Netherlands. (c.w.oosterlee@cwi.nl)

in option pricing in the recent years, causing a change in the way derivatives contracts are priced. For instance, a companies own credit risk is taken into account with a debt value adjustment (DVA). The DVA is the expected gain that will be experienced by the bank in the event that the bank defaults on its portfolio of derivatives with a counterparty. To reduce the credit risk in a derivatives contract, the parties can include a credit support annex (CSA), requiring one or both of the parties to post collateral. Valuation of derivatives under CSA was first done in [23]. A margin valuation adjustment (MVA) arises when the parties are required to post an initial margin. In this case the cost of posting the initial margin to the counterparty over the length of the contract is known as MVA. Funding value adjustments (FVA) can be interpreted as a funding cost or benefit associated to the hedge of market risk of an uncollateralized transaction through a collateralized market. While there is still a debate going on about whether to include or exclude this adjustment, see [14], [13] and [5] for an in-depth overview of the arguments, most dealers now seem to indeed take into account the FVA. The capital value adjustment (KVA) refers to the cost of funding the additional capital that is required for derivative trades. This capital acts as a buffer against unexpected losses and thus, as argued in [12], has to be included in derivative pricing.

For pricing in the presence of XVA, one needs to redefine the pricing partial differential equation (PDE) by constructing a hedging portfolio with cashflows that are consistent with the additional funding requirements. This has been done for unilateral CCR in [23], bilateral CCR and XVA in [2] and extended to stochastic rates in [17]. This results in a non-linear PDE.

Non-linear PDEs can be solved with e.g. finite-difference methods or the LSM for solving the corresponsing backward stochastic differential equation (BSDE). In [24] an efficient forward simulation algorithm that gives the solution of the non-linear PDE as an optimum over solutions of related but linear PDEs is introduced, with the computational cost being of the same order as one forward Monte Carlo simulation. The downside of these numerical methods is the computational time that is required to reach an accurate solution. An efficient alternative might be to use Fourier methods for solving the (non-)linear PDE or related BSDE, such as the COS method, as was introduced in [8], extended to Bermudan options in [9] and to BSDEs in [25]. In certain cases the efficiency of these methods is further increased due the ability to the use the fast Fourier transform (FFT).

In this paper we consider an exponential Lévy-type model with a state-dependent jump measure and propose an efficient Fourier-based method to solve for Bermudan derivatives, including options and swaptions, with XVA. We derive, in the presence of jumps, a non-linear partial integro-differential equation (PIDE) and its corresponding BSDE for an OTC derivative between the bank B and its counterparty C in the presence of CCR, bilateral collateralization, MVA, FVA and KVA. We extend the Fourier-based method known as the BCOS method, developed in [25], to solve the

BSDE under Lévy models with non-constant coefficients. As this method requires the knowledge of the characteristic function of the forward process, which, in the case of the Lévy process with variable coefficients, is not known, we will use an approximation of the characteristic function obtained by the adjoint expansion method developed in [21], [20] and extended to the defaultable Lévy process with a state-dependent jump measure in [1]. Compared to other state-of-the-art methods for calculating XVAs, like Monte Carlo methods and PDE solvers, our method is both more efficient and multipurpose. Furthermore we propose an alternative Fourier-based method for explicitly pricing the CVA term in case of unilateral CCR for Bermudan derivatives under the local Lévy model. The advantage of this method is that is allows us to use the FFT, resulting in a fast and efficient calculation. The Greeks, used for hedging CVA, can be computed at almost no additional cost.

The rest of the paper is structured as follows. In Section 2 we introduce the Lévy models with non-constant coefficients. In Section 3 we derive the non-linear PIDE and corresponding BSDE for pricing contracts under XVA. In Section 4 we propose the Fourier-based method for solving this BSDE and in Section 5.1 this method is extended to pricing Bermudan contracts. In Section 5.2 an alternative FFT-based method for pricing and hedging the CVA term is proposed and Section 6 presents numerical examples validating the accuracy and efficiency of the proposed methods.

2. The model.

We consider a defaultable asset S_t whose risk-neutral dynamics are given by

$$
\begin{aligned}
S_t &= \mathbb{1}_{\{t<\zeta\}} e^{X_t},\\
dX_t &= \mu(t,X_t)dt + \sigma(t,X_t)dW_t + \int_{\mathbb{R}} q d\tilde{N}_t(t,X_{t-},dq),\\
d\tilde{N}_t(t,X_{t-},dq) &= dN_t(t,X_{t-},dq) - a(t,X_{t-})\nu(t,dq)dt, \qquad (1)\\
\zeta &= \inf\{t\geq 0 : \int_0^t \gamma(s,X_s)ds \geq \varepsilon\},
\end{aligned}
$$

where $d\tilde{N}_t(t,X_{t-},dq)$ is a compensated random measure with state-dependent Lévy measure

$$\nu(t,X_{t-},dq) = a(t,X_{t-})\nu(dq).$$

The default time ζ of S_t is defined in a canonical way as the first arrival time of a doubly stochastic Poisson process with local intensity function $\gamma(t,x)\geq 0$, and $\varepsilon \sim \mathrm{Exp}(1)$ and is independent of X_t. This way of modeling default is also considered in a diffusive setting in [4] and for exponential Lévy models in [3]. Thus our model includes a local volatility function, a local jump measure, and a default probability which is dependent on the underlying. We define the filtration of the market observer to be $\mathcal{G} = \mathcal{F}^X \vee \mathcal{F}^D$, where $\mathcal{F}^X$ is the filtration generated by X and $\mathcal{F}_t^D := \sigma(\{\zeta \leq u\}, u \leq t)$, for $t\geq 0$, is the filtration of the default. Using this definition of default, the probability of default is

$$\text{PD}(t) := \mathbb{P}(\zeta \le t) = 1 - e^{-\int_0^t \gamma(s,x)ds}. \tag{2}$$

We assume furthermore

$$\int_{\mathbb{R}} e^{|q|} a(t,x)\nu(dq) < \infty.$$

If we were to impose that the discounted asset price $\tilde{S}_t := e^{-rt}S_t$ is a $\mathcal{G}$-martingale under the risk-neutral measure, we get the following restriction on the drift coefficient:

$$\mu(t,x) = \gamma(t,x) + r - \frac{\sigma^2(t,x)}{2} - a(t,x)\int_{\mathbb{R}} \nu(dq)(e^q - 1 - q),$$

with r being the risk-free (collateralized) rate. In the whole of the paper we assume deterministic, constant interest rates, while the derivations can easily be extended to time-dependent rates. The integro-differential operator of the process is given by (see e.g. [22])

$$\begin{aligned} Lu(t,x) =& \partial_t u(t,x) + \mu(t,x)\partial_x u(t,x) - \gamma(t,x)u(t,x) + \frac{\sigma^2(t,x)}{2}\partial_{xx}u(t,x) \\ &+ a(t,x)\int_{\mathbb{R}} \nu(dq)(u(t,x+q) - u(t,x) - q\partial_x u(t,x)). \end{aligned}$$

3. XVA computation.

Consider the bank B and its counterparty C, both of whom might default. Assume the dynamics of the underlying as in (1) with $\gamma(t,x) = 0$. Define $\hat{u}(t,x)$ to be the value to the bank of the (default risky) portfolio with valuation adjustments referred to as XVA and $u(t,x)$ to be the risk-free value. Note that the difference between these two values,

$$\text{TVA} := \hat{u}(t,x) - u(t,x),$$

is called the total valuation adjustment and in our setting this consists of

$$\text{TVA} = \text{CVA} + \text{DVA} + \text{KVA} + \text{MVA} + \text{FVA}. \tag{3}$$

The risk-free value $u(t,x)$ solves a linear PIDE:

$$\begin{aligned} Lu(t,x) &= ru(t,x), \\ u(T,x) &= \phi(x), \end{aligned} \tag{4}$$

where L is given in (2) with $\gamma(t,x) = 0$. Assuming the dynamics in (1), this linear PIDE can be solved with the methods presented in [1].

3.1. Derivative pricing under CCR and bilateral CSA agreements. In [2], the authors derive an extension to the Black-Scholes PDE in the presence of a bilateral counterparty risk in a jump-to-default model with the underlying being a diffusion, using replication arguments that include the funding costs. In [17] this derivation is extended to a multivariate diffusion setting with stochastic rates in the presence of CCR, assuming that both parties B and C are subject to default. To mitigate the CCR, both parties exchange collateral consisting of the initial margin and the variation margin. The parties are obliged to hold regulatory capital, the cost of which is the KVA and face the costs of funding uncollateralized positions, known as FVA. Both [2] and [17] extend the approach of [23], in which unilateral collateralization was considered. We extend their approach to derive the value of $\hat{u}(t, x)$ when the underlying follows the jump-diffusion defined in (1). We assume a one-dimensional underlying diffusion and consider all rates to be deterministic and, for ease of notation, constant. As it is unrealistic to assume that market participants can freely borrow and lend at a single risk-free interest rate, we specify different rates, defined in 3.1, for different types of lending.

Rate	Definition
r	the risk-free rate
r_R	the rate received on funding secured by the underlying asset
r_D	the dividend rate in case the stock pays dividends
r_F	the rate received on unsecured funding
r_B	the yield on a bond of the bank B
r_C	the yield on the bond of the counterparty C
λ_B	$\lambda_B := r_B - r$
λ_C	$\lambda_C := r_C - r$
λ_F	$\lambda_F := r_F - r$
R_B	the recovery rate of the bank
R_C	the recovery rate of the counterparty

Table 3.1

Definitions of the rates used throughout this chapter.

Assume that the parties B and C enter into a derivatives contract on the spot asset that pays the bank B the amount $\phi(X_t)$ at maturity T. The value of this derivative to the bank at time t is denoted by $\hat{u}(t, x, J^B, J^C)$ and depends on the value of the underlying X and the default states J^B and J^C of the bank B and counterparty C.

The cashflows are viewed from the perspective of the bank B. At the default time of either the counterparty or the bank, the value of the derivative to the bank $\hat{u}(t, x)$ is determined with a mark-to-market rule M, which may be equal to either the derivative value $\hat{u}(t,x,0,0)$ prior to default or the risk-free derivative value $u(t,x)$, depending on the specifications in the ISDA master agreement. Denote by τ^B and τ^C the random default times of the bank and the counterparty respectively. Define I^{TC} to be the initial margin posted by the bank to the counterparty, I^{FC} the initial margin posted by the counterparty to the bank and $I^V(t)$ to be the variation margin on which a rate r_I is paid or received. The initial margin is constant throughout the duration of the contract and $K(t)$ is the regulatory capital on which a rate of r_K is paid/received. We will use the notation $x^+ = \max(x,0)$ and $x^- = \min(x,0)$. In a situation in which the counterparty defaults, the bank is already in the possession of $I^V + I^{FC}$. If the outstanding value $M - (I^V + I^{FC})$ is negative, the bank has to pay the full amount $(M - I^V - I^{FC})^-$, while if the contract has a positive value to the bank, it will recover only $R_C(M - I^V - I^{FC})^+$. Using a similar argument in case the bank defaults, we find the following boundary conditions:

$$\theta^B := u(t,x,1,0) = I^V - I^{TC} + (M - I^V + I^{TC})^+ + R^B(M - I^V + I^{TC})^-,$$
$$\theta^C := u(t,x,0,1) = I^V + I^{FC} + R^C(M - I^V - I^{FC})^+ + (M - I^V - I^{FC})^-,$$

so that the portfolio value at default is given by

$$\theta_\tau = 1_{\tau^C < \tau^B}\theta^C_\tau + 1_{\tau^B < \tau^C}\theta^B_\tau,$$

with $\tau = \min(\tau^B, \tau^C)$. Further we introduce the default risky, zero-recovery bonds (ZCBs) P^B and P^C with respective maturities T^B and T^C and face value one if the issuer has not defaulted, and zero otherwise. The dynamics of P^B and P^C are given by

$$dP^B_t = r_B P^B_t dt - P^B_{t-} dJ^B_t,$$
$$dP^C_t = r_C P^C_t dt - P^C_{t-} dJ^C_t,$$

where $J^B_t = 1_{\tau^B \le t}$ and $J^C_t = 1_{\tau^C \le t}$. Both counting processes J^B, J^C are two independent point processes that jump from zero to one on default of B and C with intensities γ^B and γ^C, respectively.

We construct a hedging portolio consisting of the shorted derivative, Δ units of X, g units of cash, α_C units of P^C and α_B units of P^B:

$$\Pi(t) = -\hat{u}(t,x) + \Delta(t)X_t + \alpha_B(t)P^B_t + \alpha_C(t)P^C_t + g(t).$$

The shares position provides a dividend income of $r_D\Delta(t)X_t dt$ and requires a financing cost of $r_R\Delta(t)X_t dt$. The seller will short the counterparty bond through a repurchase agreement and incur the financing costs of $-r\alpha_C(t)P^C_t$, assuming no haircut. The cashflows from the collateralization follow from the rate r_{TC} received and r_{FC} paid on the initial margin and the rate r_I paid or received on the collateral, depending on whether $I^V > 0$ and the bank receives collateral or $I^V < 0$ and the bank pays collateral repectively. From holding the regulatory capital we incur a cost of $r_K K(t)$. Finally, the rates r and r_F are respectively received or paid on the surplus cash in the account: $-\hat{u}(t,x) - I^V(t) + I^{TC} - \alpha_B(t)P_t^B$. Thus, the change in the cash account is given by

$$\begin{aligned} dg(t) =& [(r_D - r_R)\Delta(t)X_t - r\alpha_C(t)P_t^C + r_{TC}I_{TC} - r_{FC}I_{FC} - r_I I^V(t) - r_K K(t) \\ & + r(-\hat{u}(t,x) - I^V(t) + I_{TC} - \alpha_B(t)P_t^B) + \lambda_F(-\hat{u}(t,x) - I^V(t) + I_{TC} - \alpha_B(t)P_t^B)^-]dt. \end{aligned}$$

Assuming the portfolio is self-financing we have

$$\begin{aligned} d\Pi(t) =& - d\hat{u}(t,x) + \Delta(t)dX_t + \alpha_B(t)dP_t^B + \alpha_C(t)dP_t^C + dg(t) \\ =& - d\hat{u}(t,x) + \Delta(t)\mu(t,x)dt + \Delta(t)\sigma(t,x)dW_t + \Delta(t)\int_{\mathbb{R}} q d\tilde{N}_t(t, X_{t-}, dq) \\ & + \alpha_B(t)dP_t^B + \alpha_C(t)dP_t^C + dg(t). \end{aligned}$$

Applying Itô's Lemma to $\hat{u}(t,x)$ gives us:

$$\begin{aligned} d\hat{u}(t,x) =& L\hat{u}(t,x)dt + \sigma(t,x)\partial_x\hat{u}(t,x)dW_t + \int_{\mathbb{R}}(\hat{u}(t,x+q) - \hat{u}(t,x))d\tilde{N}(t,x,dq) \\ & - (\theta^B - \hat{u}(t,x))dJ_t^B - (\theta^C - \hat{u}(t,x))dJ_t^C. \end{aligned}$$

Thus, we find,

$$\begin{aligned} d\Pi =& - L\hat{u}(t,x)dt - \sigma(t,x)\partial_x\hat{u}(t,x)dW_t - \int_{\mathbb{R}}(\hat{u}(t,x+q) - \hat{u}(t,x))d\tilde{N}(t,X_{t-},dq) \\ & + (\theta^B - \hat{u}(t,x))dJ_t^B + (\theta^C - \hat{u}(t,x))dJ_t^C \\ & + \Delta(t)\sigma(t,x)dW_t + \Delta(t)\int_{\mathbb{R}} q d\tilde{N}_t(t,X_{t-},dq) - \alpha^B(t)P_{t-}^B dJ_t^B - \alpha^C(t)P_{t-}^C dJ_t^C \\ & + [\Delta(t)(\mu(t,x) + (r_D - r_R)x) + \alpha^B(t)\lambda_B P_t^B + \alpha^C(t)\lambda_C P_t^C \\ & + (r_{TC} + r)I^{TC} - r_{FC}I^{FC} - (r_I + r)I^V(t) - r_K K(t) + r\hat{u}(t,x) \\ & + \lambda_F(-\hat{u}(t,x) - I^V(t) + I^{TC} - \alpha^B(t)P_t^B)^-]dt. \end{aligned}$$

By choosing

$$\Delta = \partial_x u(t,x), \quad \alpha_B = -\frac{\theta^B - \hat{u}(t,x)}{P_B}, \quad \alpha_C = -\frac{\theta^C - \hat{u}(t,x)}{P_C},$$

we hedge the Brownian motion and jump-to-default risk in the hedging portfolio, i.e.,

$$\begin{aligned} d\Pi =& - L\hat{u}(t,x)dt - \int_{\mathbb{R}}(\hat{u}(t,x+q) - \hat{u}(t,x))d\tilde{N}(t,X_{t-},dq) + \partial_x\hat{u}(t,x)\int_{\mathbb{R}} q d\tilde{N}_t(t,X_{t-},dq) \\ & + [\partial_x\hat{u}(t,x)(\mu(t,x) + (r_D - r_R)x) - (\theta^B - \hat{u}(t,x))\lambda_B - (\theta^C - \hat{u}(t,x))\lambda_C \\ & + (r_{TC} + r)I^{TC} - r_{FC}I^{FC} - (r_I + r)I^V(t) - r_K K(t) + r\hat{u}(t,x) \\ & + \lambda_F(\theta^B - I^V(t) + I^{TC})^-]dt. \end{aligned}$$

Notice that we are in an incomplete market, as it is not possible to choose $\Delta(t)$ such that the portfolio is risk-free (due to the presence of the state-dependent jumps). Following standard arguments, see e.g. [11] and [6], we assume that an investor holds a diversified portfolio of several hedging portfolios and that the jumps for the different portfolios are uncorrelated. The variance of this 'portfolio of portfolios' will then be small and the *expected* return on the portfolio is given by

$$\mathbb{E}[d\Pi] = 0.$$

The assumption of the jump risk being diversifiable is valid if the jump parameters are adjusted to contain the so-called market price of risk, as can be done by e.g. fitting them from the market. We find the pricing PIDE to be

$$L\hat{u}(t,x) = f(t,x,\hat{u}(t,x),\partial_x\hat{u}(t,x)), \tag{5}$$

where we have defined

$$\begin{aligned} f(t,x,\hat{u}(t,x),\partial_x\hat{u}(t,x)) =& \partial_x\hat{u}(t,x)(\mu(t,x)+(r_D-r_R)x) - (\theta^B(t)-\hat{u}(t,x))\lambda_B \\ & -(\theta^C(t)-\hat{u}(t,x))\lambda_C + (r_{TC}+r)I^{TC} - r_{FC}I^{FC} - (r_I+r)I^V(t) \\ & - r_K K(t) + r\hat{u}(t,x) + \lambda_F(\theta^B - I^V(t) + I^{TC})^-, \end{aligned}$$

and used

$$\mathbb{E}\left[\int_{\mathbb{R}} (\hat{u}(t,x+q)-\hat{u}(t,x)-q\partial_x\hat{u}(t,x))d\tilde{N}(t,X_{t-},q)\right] = 0,$$

due to the jump measure being compensated.

3.2. BSDE representation. In this section we will cast the PIDE in (5) in the form of a Backward Stochastic Differential Equation. We begin by recalling the non-linear Feynman-Kac theorem in the presence of jumps, see e.g. [16].

Theorem 1 (Non-linear Feynman-Kac Theorem). *Consider X_t as in (1) and the BSDE*

$$\begin{aligned} Y_t = \phi(X_T) &+ \int_t^T f\left(s,X_s,Y_s,Z_s,a(s,X_{s-})\int_{\mathbb{R}} V_s(q)\delta(s,q)\nu(dq)\right)ds - \int_t^T Z_s dW_s \\ &- \int_t^T \int_{\mathbb{R}} V_s(q)d\tilde{N}_s(s,X_s,q), \end{aligned} \tag{6}$$

where $\delta(t,q)$ is a non-negative function such that $\int_{\mathbb{R}} |\delta(s,q)|^2\nu(dq) < \infty$, T is the time horizon, f is the generator and ϕ is the terminal condition. The functions μ, σ, a and the generator f are assumed to be uniformly Lipschitz continuous in the space variables, for all $t \in [0,T]$. Consider the *non-linear PIDE*

$$(7)\qquad \begin{cases} Lu(t,x) = f(t,x,u(t,x),\partial_x u(t,x)\sigma(t,x),a(t,x)\int_{\mathbb{R}}(u(t,x+q)-u(t,x))\delta(t,q)\nu(dq)), \\ u(T,x) = \psi(x). \end{cases}$$

If the PIDE in (7) has a solution $u(t,x) \in C^{1,2}$, the solution (Y_t, Z_t, V_t) of the FBSDE in (6) can be represented as

$$\begin{aligned} Y_s^{t,x} &= u(s, X_s^{t,x}), \\ Z_s^{t,x} &= \partial_x u(s, X_s^{t,x})\sigma(s, X_s^{t,x}), \\ V_s^{t,x}(q) &= u(s, X_s^{t,x}+q) - u(s, X_s^{t,x}), \qquad q \in \mathbb{R}, \end{aligned}$$

for all $s \in [t,T]$, where Y is a continuous, real-valued and adapted processes and where Z and V are continuous, real-valued and predictable processes.

In our case, the BSDE corresponding to the PIDE in (5) is given by

$$(8)\qquad Y_t = \phi(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s - \int_t^T \int_{\mathbb{R}} V_s(q) d\tilde{N}(s, X_s, dq),$$

where we have defined the driver function to be

$$\begin{aligned} f(t,x,y,z) =& z\sigma(t,x)^{-1}(\mu(t,x) + (r_D - r_R)x) - \lambda_B(\theta^B - y) - \lambda_C(\theta^C - y) \\ &+ (r_{TC} + r)I^{TC} - r_{FC}I^{FC} - (r_I + r)I^V(t) - r_K K(t) + ry \\ &+ \lambda_F(\theta^B - I^V(t) + I^{TC})^-. \end{aligned}$$

3.2.1. Close-out value $M = \hat{u}(t,x)$. We derive, for completion, the driver function in the scenario in which the close-out value has a mark-to-market rule M equal to $\hat{u}$, the risky portfolio value. Then the driver function has the following form

$$\begin{aligned} f(t,x,y,z) =& z\sigma(t,x)^{-1}(\mu(t,x) + (r_D - r_R)x) - r_K K(t) \\ &+ (r_{TC} + r_B)I^{TC} - (r_{FC} + \lambda_C)I^{FC} - (r_I + r_B + \lambda_C)I^V(t) \\ &+ (r_B + \lambda_C)y - \lambda_B((y - I^V(t) + I^{TC})^+ + R^B(y - I^V(t) + I^{TC})^-) \\ &- \lambda_C(R^C(y - I^V(t) - I^{FC})^+ + (y - I^V(t) - I^{FC})^-) \\ &- \lambda_F(y - I^V(t) + I^{TC})^-, \end{aligned}$$

where we have used $(y - I^V(t) + I^{TC})^+ + R_B(y - I^V(t) + I^{TC})^+ = (y - I^V(t) + I^{TC})^-$.

3.2.2. Close-out value $M = u(t,x)$. We also consider the case of the close-out value being equal to u, the risk-free portflio value. This convention is most often used in the industry. In this case the driver function becomes

$$\begin{aligned} f(t,x,y,z) =& z\sigma(t,x)^{-1}(\mu(t,x)+(r_D-r_R)x)+(r_B+\lambda_C)y \\ &-r_K K(t)-(r_{TC}+r_B)I^{TC}-(r_{FC}+\lambda_C)I^{FC}-(r_I+r_B+\lambda_C)I^V(t) \\ &-\lambda_B((u-I^V(t)+I^{TC})^+ + R^B(u-I^V(t)+I^{TC})^-) \\ &-\lambda_C(R^C(u-I^V(t)-I^{FC})^+ + (u-I^V(t)-I^{FC})^-) \\ &-\lambda_F(u-I^V(t)+I^{TC})^-, \end{aligned}$$

where $u(t,x)$ is the solution to the linear PIDE given in (4) so that the driver function is linear in y. This results in a linear PIDE which can be solved with the method in [1], without the use of BSDEs.

3.2.3. A simplified driver function. Following [12], one can derive that the KVA is a function of trade properties (i.e. maturity, strike) and/or the exposure at default, which in turn is a function of the portfolio value, so that the cost of holding the capital can be rewritten as

$$r_K K(t) = r_K c_1 \hat{u}(t,x),$$

with c_1 being a function of the trade properties. The collateral is paid when the portfolio has a negative value, and received when the collateral has a positive value. Assuming the collateral is a multiple of the portfolio value we have

$$I^V(t) = c_2 \hat{u}(t,x),$$

where c_2 is some constant. Then, the driver function is simply a function of the portfolio value and its first derivative.

Remark 2. *Note that in the case of 'no collateralization' or 'perfect collateralization', the driver function reduces to $f(t,\hat{u}(t,x)) = r_u(t)\max(\hat{u}(t,x),0)$, for a function r_u here left unspecified. In this case the BSDE is similar to the one considered in [24].*

4. Solving FBSDEs. In this section we extend the BCOS method from [25] to solving FBSDEs under local Lévy models with variable coefficients and jumps. The conditional expectations resulting from the discretization of the FBSDE are approximated using the COS method. This requires the characteristic function, which we approximate using the Adjoint Expansion Method of [21] and [1].

4.1. Discretization of the BSDE. Consider the forward process X_t as in (1) and the BSDE Y_t as in (8). Define a partition $0 = t_0 < t_1 < ... < t_N = T$ of $[0, T]$ with a fixed time step $\Delta t = t_{n+1} - t_n$, for $n = N-1, ...0$. Rewriting the set of FBSDEs we find,

$$X_{n+1} = X_n + \int_{t_n}^{t_{n+1}} \mu(s, X_s)ds + \int_{t_n}^{t_{n+1}} \sigma(s, X_s)dW_s + \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}} q d\tilde{N}_s(s, X_{s-}, dq),$$

$$Y_n = Y_{n+1} + \int_{t_n}^{t_{n+1}} f(s, X_s, Y_s, Z_s)\, ds - \int_{t_n}^{t_{n+1}} Z_s dW_s - \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}} V_s(q) d\tilde{N}_s(s, X_{s-}, dq). \tag{9}$$

One can obtain an approximation of the process Y_t by taking conditional expectations with respect to the underlying filtration $\mathcal{G}_n$, using the independence of W_t and $\tilde{N}_t(t, X_{t-}, dq)$ and by approximating the integrals that appear with a theta method, as first done in [26] and extended to BSDEs with jumps in [25]:

$$Y_n \approx \mathbb{E}_n[Y_{n+1}] + \Delta t \theta_1 f(t_n, X_n, Y_n, Z_n) + \Delta t(1-\theta_1)\mathbb{E}_n[f(t_{n+1}, X_{n+1}, Y_{n+1}, Z_{n+1})].$$

Let $\Delta W_s := W_s - W_n$ for $t_n \le s \le t_{n+1}$. Multiplying both sides of equation (9) by ΔW_{n+1}, taking conditional expectations and applying the theta-method gives

$$\begin{aligned} Z_n \approx & -\theta_2^{-1}(1-\theta_2)\mathbb{E}_n[Z_{n+1}] + \frac{1}{\Delta t}\theta_2^{-1}\mathbb{E}_n[Y_{n+1}\Delta W_{n+1}] \\ & + \theta_2^{-1}(1-\theta_2)\mathbb{E}_n[f(t_{n+1}, X_{n+1}, Y_{n+1}, Z_{n+1})\Delta W_{n+1}]. \end{aligned}$$

Since in our scheme the terminal values are functions of time t and the Markov process X, it is easily seen that there exist deterministic functions $y(t_n, x)$ and $z(t_n, x)$ so that

$$Y_n = y(t_n, X_n), \quad Z_n = z(t_n, X_n).$$

The functions $y(t_n, x)$ and $z(t_n, x)$ are obtained in a backward manner using the following scheme

$$y(t_N, x) = \phi(x), \quad z(t_N, x) = \partial_x \phi(x)\sigma(t_N, x),$$

$$\text{for } n = N-1, ..., 0:$$

$$y(t_n, x) = \mathbb{E}_n[y(t_{n+1}, X_{n+1})] + \Delta t \theta_1 f(t_n, x) + \Delta t(1-\theta_1)\mathbb{E}_n[f(t_{n+1}, X_{n+1})], \tag{10}$$

$$\begin{aligned} z(t_n, x) = & -\frac{1-\theta_2}{\theta_2}\mathbb{E}_n[z(t_{n+1}, X_{n+1})] + \frac{1}{\Delta t}\theta_2^{-1}\mathbb{E}_n[y(t_{n+1}, X_{n+1})\Delta W_{n+1}] \\ & + \frac{1-\theta_2}{\theta_2}\mathbb{E}_n[f(t_{n+1}, X_{n+1})\Delta W_{n+1}], \end{aligned} \tag{11}$$

where we have simplified notations with

$$f(t, X_t) := f(t, X_t, y(t, X_t), z(t, X_t)).$$

In the case $\theta_1 > 0$ we obtain an implicit dependence on $y(t_n, x)$ in (10) and we use P Picard iterations starting with initial guess $\mathbb{E}_n[y(t_{n+1}, X_{n+1})]$ to determine $y(t_n, x)$. Note that due to the independence of the driver function on $V_s(q)$, we choose not to calculate $V_n(q) = v(t_n, X_n, q)$ in the interation. This simplifies the computation and reduces the computational time.

4.2. The characteristic function. Is it well-known (see, for instance, [18, Section 2.2]) that the price V of a European option with maturity T and payoff $\Phi(S_T)$ is given by

$$V_t = \mathbb{1}_{\{\zeta>t\}} e^{-r(T-t)} \mathbb{E}\left[e^{-\int_t^T \gamma(s,X_s)ds} \phi(X_T) | X_t \right], \quad t \le T,$$

in the measure corresponding to the dynamics in (1) and where we have defined $\phi(x) := \Phi(e^x)$. Thus, in order to compute the price of an option, we must evaluate functions of the form

$$v(t,x) := \mathbb{E}\left[e^{-\int_t^T \gamma(s,X_s)ds} \phi(X_T) | X_t = x \right]. \tag{12}$$

Under standard assumptions, by the Feynman-Kac theorem, v can be expressed as the classical solution of the following Cauchy problem

$$\begin{cases} Lv(t,x) = 0, & t \in [0,T[, \; x \in \mathbb{R}, \\ v(T,x) = \phi(x), & x \in \mathbb{R}, \end{cases}$$

with L as in (2).

The function v in (12) can be represented as an integral with respect to the transition distribution of the defaultable log-price process $\log S_t$:

$$v(t,x) = \int_{\mathbb{R}} \phi(y) \Gamma(t,x;T,dy).$$

Here we notice explicitly that $\Gamma(t,x;T,dy)$ is not necessarily a standard probability measure because its integral over $\mathbb{R}$ can be strictly less than one; nevertheless, with a slight abuse of notation, we say that its Fourier transform

$$\hat{\Gamma}(t,x;T,\xi) := \mathcal{F}(\Gamma(t,x;T,\cdot))(\xi) := \int_{\mathbb{R}} e^{i\xi y} \Gamma(t,x;T,dy), \qquad \xi \in \mathbb{R},$$

is the characteristic function of $\log S$. Following [21] and [1] we expand the state-dependent coefficients

$$s(t,x) := \frac{\sigma^2(t,x)}{2}, \qquad \mu(t,x), \qquad \gamma(t,x), \qquad a(t,x),$$

around some point $\bar{x}$. The coefficients $s(t,x)$, $\gamma(t,x)$ and $a(t,x)$ are assumed to be continuously differentiable with respect to x up to order $n \in \mathbb{N}$.

Introduce the nth-order approximation of L in (2):

$$L_n = L_0 + \sum_{k=1}^{n} \Big((x-\bar{x})^k \mu_k(t) + (x-\bar{x})^k s_k(t) \partial_{xx} - (x-\bar{x})^k \gamma_k(t) + \int_{\mathbb{R}} (x-\bar{x})^k a_k(t) \nu(dq)(e^{q\partial_x} - 1 - q\partial_x) \Big),$$

where

$$L_0 = \partial_t + \mu_0(t)\partial_x + s_0(t)\partial_{xx} - \gamma_0(t) + \int_{\mathbb{R}} a_0(t)\nu(dq)(e^{q\partial_x} - 1 - q\partial_x),$$

and

$$s_k = \frac{\partial_x^k s(\cdot,\bar{x})}{k!}, \qquad \gamma_k = \frac{\partial_x^k \gamma(\cdot,\bar{x})}{k!}, \qquad \mu_k(dq) = \frac{\partial_x^k \mu(\cdot,\bar{x})}{k!}, \qquad a_k = \frac{\partial_x^k a(\cdot,\bar{x})}{k!} \qquad k \geq 0.$$

The basepoint $\bar{x}$ is a constant parameter which can be chosen freely. In general the simplest choice is $\bar{x} = x$ (the value of the underlying at initial time t).

Assume for a moment that L_0 has a fundamental solution $G^0(t,x;T,y)$ that is defined as the solution of the Cauchy problem

$$\begin{cases} L_0 G^0(t,x;T,y) = 0 & t \in [0,T[,\ x \in \mathbb{R}, \\ G^0(T,\cdot;T,y) = \delta_y. \end{cases}$$

In this case we define the nth-order approximation of Γ as

$$\Gamma^{(n)}(t,x;T,y) = \sum_{k=0}^{n} G^k(t,x;T,y),$$

where, for any $k \geq 1$ and (T,y), $G^k(\cdot,\cdot;T,y)$ is defined recursively through the following Cauchy problem

$$\begin{cases} L_0 G^k(t,x;T,y) = -\sum_{h=1}^{k} (L_h - L_{h-1}) G^{k-h}(t,x;T,y) & t \in [0,T[,\ x \in \mathbb{R}, \\ G^k(T,x;T,y) = 0, & x \in \mathbb{R}. \end{cases}$$

Notice that

$$\begin{aligned} L_k - L_{k-1} = &(x-\bar{x})^k \mu_h(t)\partial_x + (x-\bar{x})^k s_k(t)\partial_{xx} - (x-\bar{x})^k \gamma_k(t) \\ &+ \int_{\mathbb{R}} (x-\bar{x})^k a_k(t)\nu(dq)(e^{q\partial_x} - 1 - q\partial_x). \end{aligned}$$

Correspondingly, the nth-order approximation of the characteristic function $\hat{\Gamma}$ is defined to be

$$\hat{\Gamma}^{(n)}(t,x;T,\xi) = \sum_{k=0}^{n} \mathcal{F}\left(G^k(t,x;T,\cdot)\right)(\xi) := \sum_{k=0}^{n} \hat{G}^k(t,x;T,\xi), \qquad \xi \in \mathbb{R}.$$

Now, by transforming the simplified Cauchy problems into adjoint problems and solving these in the Fourier space we find

$$\hat{G}^0(t,x;T,\xi) = e^{i\xi x} e^{\int_t^T \psi(s,\xi)ds},$$

$$\hat{G}^k(t,x;T,\xi) = -\int_t^T e^{\int_s^T \psi(\tau,\xi)d\tau} \mathcal{F}\left(\sum_{h=1}^{k} \left(\tilde{L}_h^{(s,\cdot)}(s) - \tilde{L}_{h-1}^{(s,\cdot)}(s)\right) G^{k-h}(t,x;s,\cdot)\right)(\xi)ds,$$

with

$$\psi(t,\xi) = i\xi\mu_0(t) + s_0(t)\xi^2 + \int_{\mathbb{R}} a_0\nu(t,dq)(e^{iz\xi} - 1 - iz\xi),$$

$$\begin{aligned}\tilde{L}_h^{(t,y)}(t) - \tilde{L}_{h-1}^{(t,y)}(t) = {} & \mu_h(t)h(y-\bar{x})^{h-1} + \mu_h(t)(y-\bar{x})^h\partial_y - \gamma_h(t)(y-\bar{x})^h \\ & + s_h(t)h(h-1)(y-\bar{x})^{h-2} + s_h(t)(y-\bar{x})^{h-1}\left(2h\partial_y + (y-\bar{x})\partial_{yy}\right) \\ & + \int_{\mathbb{R}} a_h(t)\bar{\nu}(dq)\left((y+q-\bar{x})^h e^{q\partial_y} - (y-\bar{x})^h - q\left(h(y-\bar{x})^{h-1} - (y-\bar{x})^h\partial_y\right)\right),\end{aligned}$$

where $\bar{\nu}(dq) = \nu(-dq)$.

Remark 3. *After some algebraic manipulations it can be shown, see [1], that the characteristic function approximation of order n is a function of the form*

$$\hat{\Gamma}^{(n)}(t,x;T,\xi) := e^{i\xi x}\sum_{k=0}^{n}(x-\bar{x})^k g_{n,k}(t,T,\xi), \tag{13}$$

where the coefficients $g_{n,k}$, with $0 \le k \le n$, depend only on t, T and ξ, but not on x. The approximation formula can thus always be split into a sum of products of functions depending only on ξ and functions that are linear combinations of $(x-\bar{x})^m e^{i\xi x}$, $m \in \mathbb{N}_0$.

4.3. The COS formulae. The conditional expectations are approximated using the COS method, which was developed in [9] and applied to FBSDEs with jumps in [25]. The conditional expectations arising in the equations (10)-(11)are all of the form $\mathbb{E}_n[h(t_{n+1},X_{n+1})]$ or $\mathbb{E}_n[h(t_{n+1},X_{n+1})\Delta W_{n+1}]$. The COS formula for the first conditional expectation reads

$$\mathbb{E}_n^x[h(t_{n+1},X_{n+1})] \approx {\sum_{j=0}^{J-1}}' H_j(t_{n+1})\mathrm{Re}\left(\hat{\Gamma}\left(t_n,x;t_{n+1},\frac{j\pi}{b-a}\right)\exp\left(ij\pi\frac{-a}{b-a}\right)\right),$$

where $\sum'$ denotes an ordinary summation with the first term weighted by one-half, $J > 0$ is the number of Fourier-cosine coefficients we use, $H_j(t_{n+1})$ denotes the jth Fourier-cosine coefficients of the function $h(t_{n+1},x)$ and $\hat{\Gamma}\left(t_n,x;t_{n+1},\xi\right)$ is the conditional characteristic function of the process X_{n+1} given $X_n = x$. For the second conditional expectation, using integration by parts, we obtain

$$\begin{aligned}&\mathbb{E}_n^x[h(t_{n+1},X_{n+1})\Delta W_n] \\ &\quad\approx \Delta t\sigma(t_n,x){\sum_{j=0}^{J-1}}' H_j(t_{n+1})\mathrm{Re}\left(i\frac{j\pi}{b-a}\hat{\Gamma}\left(t_n,x;t_{n+1},\frac{j\pi}{b-a}\right)\exp\left(ij\pi\frac{-a}{b-a}\right)\right).\end{aligned}$$

See [25] for the full derivations.

Remark 4. *Note that these formulas are obtained by using an Euler approximation of the forward process and using the 2nd-order approximation of the characteristic function of the actual process. We have found this to be more exact than using the characteristic function of the Euler process, which is equivalent to using just the 0th-order approximation of the characteristic function.*

Finally we need to approximate the Fourier-cosine coefficients $H_j(t_{n+1})$ of h at time points t_n, where $n = 0, ..., N$. The Fourier-cosine coefficient of h at time t_{n+1} is defined by

$$H_j(t_{n+1}) = \frac{2}{b-a}\int_a^b h(t_{n+1}, x)\cos\left(j\pi\frac{x-a}{b-a}\right)dx.$$

Due to the structure of the approximated characteristic function of the local Lévy process, see (13), the coefficients of the functions $z(t_{n+1}, x)$ and the explicit part of $y(t_{n+1}, x)$ can be computed using a FFT algorithm, as we do in Appendix A, because of the matrix in (20) being of a certain form. In order to determine $F_j(t_{n+1})$, the Fourier-Cosine coefficient of the function

$$f\left(t_{n+1}, x, y(t_{n+1}, x), z(t_{n+1}, x)\right),$$

due to the intricate dependence on the functions z and y we choose to approximate the integral in F_j with a discrete Fourier-Cosine transform (DCT). For the DCT we compute the integrand, and thus the functions $z(t_{n+1}, x)$ and $y(t_{n+1}, x)$, on an equidistant x-grid. Note that in this case we can easily approximate *all* Fourier-Cosine coefficients with a DCT (instead of the FFT). If we take J grid points defined by $x_i := a + (i + \frac{1}{2})\frac{b-a}{J}$ and $\Delta x = \frac{b-a}{J}$ we find using the mid-point integration rule the approximation

$$H_j(t_{n+1}) \approx \frac{2}{J}\sum_{i=0}^{J-1}{}' h(t_{n+1}, x_i)\cos\left(j\pi\frac{2i+1}{2G}\right),$$

which can be calculated using a DCT algorithm, with the computational time being $O(J\log J)$.

Remark 5. *We define the truncation range* $[a, b]$ *as follows:*

$$[a, b] := \left[c_1 - L\sqrt{c_2 + \sqrt{c_4}}, c_1 + L\sqrt{c_2 + \sqrt{c_4}}\right],$$

where c_n *is the nth cumulant of log-price process* $\log S$, *as proposed in [8]. The cumulants are calculated using the 0th-order approximation of the characteristic function.*

5. XVA computation for Bermudan derivatives. The method in Section 4 allows us to compute the XVA as in (3), consisting of CVA, DVA, MVA, KVA and FVA. In this section, we apply this method to computing Bermudan derivative values with XVA. For the CVA component in the XVA we develop an alternative method, which due to the ability to use the FFT results in a particularly efficient computation.

5.1. XVA computation. Consider an OTC derivative contract between the bank B and the counterparty C with a Bermudan-type exercise possibility: there is a finite set of so-called exercise moments $\{t_1, ..., t_M\}$ prior to the maturity, with $0 \leq t_1 < t_2 < \cdots < t_M = T$. The payoff from the point-of-view of bank B is given by $\Phi(t_m, X_{t_m})$. Denote $\hat{u}(t,x)$ to be the risky Bermudan option value and $c(t,x)$ the so-called continuation value. By the dynamic programming approach, the value for a Bermudan derivative with XVA and M exercise dates $t_1, ..., t_M$ can be expressed by a backward recursion as

$$\hat{u}(t_M, x) = \Phi(t_M, x),$$

and the continuation value solves the non-linear PIDE defined in (5)

$$\begin{cases} \begin{cases} Lc(t,x) = f(t,x,c(t,x),\partial_x c(t,x)), & t \in [t_{m-1}, t_m[\\ c(t_m, x) = \hat{u}(t_m, x) \end{cases} \\ \hat{u}(t_{m-1}, x) = \max\{\Phi(t_{m-1}, x), c(t_{m-1}, x)\}, \quad m \in \{2, \dots, M\}. \end{cases}$$

The derivative value is set to be $\hat{u}(t,x) = c(t,x)$ for $t \in]t_{m-1}, t_m[$, and, if $t_1 > 0$, also for $t \in [0, t_1[$. The payoff function might take on various forms:

1. (Portfolio) Following [24], we can consider X_t to the process of a portfolio which can take on both positive and negative values. Then, when exercised at time t_m, bank B receives the portfolio and $\Phi(t_m, x) = x$.
2. (Bermudan option) In case the Bermudan contract is an option, the option value to the bank can not have a negative value for the bank. At the same time, in case of default of the bank itself, the counterparty loses nothing. In this case the framework simplifies to one with unilateral collateralization and default risk and the payoff at time t_m, if exercised, is given by $\Phi(t_m, x) = (K - e^x)^+$ for a put and $\Phi(t_m, x) = (e^x - K)^+$ for a call with K being the strike price.
3. (Bermudan swaptions) A Bermudan swaption is an option in which the holder, bank B, has the right to exercise and enter into an underlying swap with fixed end date t_{M+1}. If the swaption is exercised at time t_m the underlying swap starts with payment dates $\mathcal{T}_m = \{t_{m+1}, ..., t_{M+1}\}$. Working under the forward measure corresponding to the last reset date t_M, the payoff function is given by

$$\Phi(t_m, x) = N^S \left(\sum_{k=m}^{M} \frac{P(t_m, t_{k+1}, x)}{P(t_m, t_M)} \Delta t \right) \max(c_p(S(t_m, \mathcal{T}_m, x) - K), 0),$$

where N^S is the notional, $c_p = 1$ for a payer swaption and $c_p = -1$ for a receiver swaption, $P(t_m, t_k, x)$ is the price of a ZCB conditional on $X_{t_m} = x$ and $S(t_m, \mathcal{T}_m, x)$ is the forward swap rate given by

$$S(t_m, \mathcal{T}_m, x) = \left(1 - \frac{P(t_m, t_{m+1}, x)}{P(t_m, t_M, x)}\right) / \left(\sum_{k=m}^{M} \frac{P(t_m, t_{k+1}, x)}{P(t_m, t_M, x)} \Delta t\right).$$

To solve for the continuation value we define a partition with N steps $t_{m-1} = t_{0,m} < t_{1,m} < t_{2,m} < ... < t_{n,m} < ... < t_{N,m} = t_m$ between two exercise dates t_{m-1} and t_m, with fixed time step $\Delta t_n := t_{n+1,m} - t_{n,m}$. Applying the method developed in Section 4, we find the following time iteration for the continuation value and its derivative

$$c(t_{N,m}, x) = \hat{u}(t_m, x), \qquad z(t_{N,m}, x) = \partial_x \hat{u}(t_m, x)\sigma(t_{N,m}, x)$$

for $n = N-1, ..., 0$

$$\begin{aligned} c(t_{n,m}, x) \approx{}& \Delta t_n \theta_1 f(t_{n,m}, x, c(t_{n,m}, x), z(t_{n,m}, x)) \\ &+ \sum_{j=0}^{J-1}{}' \Psi_j(x)(C_j(t_{n+1,m}) + \Delta t_n(1-\theta_1)F_j(t_{n+1,m})), \end{aligned} \tag{14}$$

$$\begin{aligned} z(t_{n,m}, x) \approx{}& \sum_{j=0}^{J-1}{}' -\frac{1-\theta_2}{\theta 2} Z_j(t_{n+1,m})\Psi_j(x) \\ &+ \left(\frac{1}{\Delta t_n \theta_2} C_j(t_{n+1,m}) + \frac{1-\theta_2}{\theta_2} F_j(t_{n+1,m})\right)\sigma(t_{n+1,m}, x)\Delta t_n \bar{\Psi}_j(x) \end{aligned} \tag{15}$$

where we have defined

$$\Psi_j(x) = \mathrm{Re}\left(\hat{\Gamma}\left(t_{n,m}, x; t_{n+1,m}, \frac{j\pi}{b-a}\right)\exp\left(ij\pi\frac{-a}{b-a}\right)\right),$$

$$\bar{\Psi}_j(x) = \mathrm{Re}\left(i\frac{j\pi}{b-a}\hat{\Gamma}\left(t_{n,m}, x; t_{n+1,m}, \frac{j\pi}{b-a}\right)\exp\left(ij\pi\frac{-a}{b-a}\right)\right),$$

and the Fourier-cosine coefficients are given by

$$C_j(t_{n+1,m}) = \frac{2}{b-a}\int_a^b c(t_{n+1,m}, x)\cos\left(j\pi\frac{x-a}{b-a}\right)dx,$$

$$Z_j(t_{n+1,m}) = \frac{2}{b-a}\int_a^b z(t_{n+1,m}, x)\cos\left(j\pi\frac{x-a}{b-a}\right)dx,$$

$$F_j(t_{n+1,m}) = \frac{2}{b-a}\int_a^b f(t_{n+1,m}, x, c(t_{n+1,m}, x), \partial_x c(t_{n+1,m}, x))\cos\left(j\pi\frac{x-a}{b-a}\right)dx.$$

In order to determine the function $c(t_n, x)$, we will perform P Picard iterations. To evaluate the coefficients with a DCT we need to compute the integrand $f(t_{n+1,m}, x, c(t_{n+1,m}, x), z(t_{n+1,m}, x))$ on the equidistant x-grid with x_i, for $i = 0, ..., J-1$. In order to compute this at each time step $t_{n,m}$ we thus need to evaluate $c(t_{n,m}, x)$ and $z(t_{n+1,m}, x)$ on the x-grid with J equidistant points using formula (14)-(15). This matrix-vector product results in a computational time of order $O(J^2)$.

The total algorithm for computing the value of a Bermudan contract with XVA can be summarised as in Algorithm 1 in Figure 5.1. The total computational time for the algorithm is $O(M \cdot N(J^2 + PJ + J \log J + J))$, consisting of the computation for $M \cdot N$ times the computation of the characteristic function on the x-grid, initialization of the Picard method, computation of the P Picard approximations for $c(t_{n,m}, x)$ and computing the Fourier coefficients $F_j(t_n)$ and $C_j(t_n)$.

1. Define the x-grid with J grid points given by $x_i = a + (i + \frac{1}{2})\frac{b-a}{J}$ for $i = 0, ..., J-1$.
2. Calculate the final exercise date values $c(t_{N,M}, x) = \hat{u}(t_M, x)$ and $z(t_{N,M}, x) = \partial_x \hat{u}(t_M, x)\sigma(t_{N,M}, x)$ on the x-grid and compute the terminal coefficients $C_j(t_M)$, $Z_j(t_M)$ and $F_j(t_M)$ using the DCT.
3. Recursively for the exercise dates $m = M-1, ..., 0$ do:
   (a) For time steps $n = N-1, ..., 0$ do:
      i. Compute $c(t_{n,m}, x)$, $z(t_{n,m}, x)$ using formula (14)-(15) and use these to determine $f(t_{n,m}, x, c(t_{n,m}, x), z(t_{n,m}, x))$ on the x-grid.
      ii. Subsequently, use these to determine $F_j(t_{n,m})$, $Z_j(t_{n,m})$ and $C_j(t_{n,m})$ using the DCT.
   (b) Compute the new terminal conditions $c(t_{N,m-1}, x) = \max\{\phi(t_{0,m}, x), c(t_{0,m}, x)\}$ and $z(t_{N,m-1}, x) = \partial_x \max\{\phi(t_{0,m}, x), c(t_{0,m}, x)\}\sigma(t_{N,m-1}, x)$ (either analytically or numerically) and the corresponding Fourier-cosine coefficients.
4. Finally $v(t_0, x_0) = c(t_{0,0}, x_0)$.

Figure 5.1. *Algorithm 1: Bermudan derivative valuation with XVA*

5.2. An alternative for CVA computation. In this section we present an efficient alternative way of calculating the CVA term in (3) in the case of unilateral CCR using a Fourier-based method. Due to the ability of using the FFT this method is considerably faster for computing the CVA than the method presented in Section 5.1. We use the definition of CVA at time t given by

$$\mathrm{CVA}(t) = \hat{u}(t, X_t) - u(t, X_t),$$

where $u(t, X_t)$ is as usual the default-free value of the Bermudan option, while $\hat{u}(t, X_t)$ is the value including default. We consider the model as defined in (1). We will compute $u(t, X_t)$ and $\hat{u}(t, X_t)$ using the COS method and the approximation of the characteristic function (as derived in Section 4.3), without default ($\gamma(t, x) = 0$) and with default respectively. In case of a default the payoff becomes zero. Note that the risky option value $\hat{u}(t, x)$ computed with the characteristic function for a defaultable underlying corresponds exactly to the option value in which the counterparty might default with the probablity of default, $PD(t)$, defined as in (2). Thus, in this case we have unilateral CCR and $\zeta = \tau_C$, the default time of the counterparty.

Using the definition of the defaultable S_t, it is well-known (see, for instance, [18, Section 2.2]) that the risky no-arbitrage value of the Bermudan option on the defaultable asset S_t at time t is

$$\hat{u}\,(t, X_t) = \mathbb{1}_{\{\zeta>t\}} \sup_{\tau\in\mathcal{T}_t} \mathbb{E}\left[e^{-\int_t^\tau (r+\gamma(s,X_s))ds}\phi(\tau, X_\tau)|X_t\right].$$

Remark 6 (Wrong-way risk). *By allowing the dependence of the default intensity on the underlying, a simplified form of wrong-way risk is incorporated into the CVA valuation.*

Note that the option value at time t becomes 0 if default occurs prior to time t. For a Bermudan put option with strike price K, we simply have $\phi(t, x) = (K - x)^+$. By the dynamic programming approach, the option value can be expressed by a backward recursion as

$$\hat{u}(t_M, x) = \mathbb{1}_{\{\zeta>t_M\}} \max(\phi(t_M, x), 0),$$

and

$$c(t, x) = \mathbb{E}\left[e^{\int_t^{t_m} (r+\gamma(s,X_s))ds}\hat{u}(t_m, X_{t_m})|X_t = x\right], \qquad t \in [t_{m-1}, t_m[$$

$$\hat{u}(t_{m-1}, x) = \mathbb{1}_{\{\zeta>t_{m-1}\}} \max\{\phi(t_{m-1}, x), c(t_{m-1}, x)\}, \qquad m \in \{2, \dots, M\}. \tag{16}$$

Thus to find the risky option price $\hat{u}(t, X_t)$ one uses the defaultable asset and in order to get the default-free value $u(t, X_t)$ one uses the default-free asset by setting $\gamma(t, x) = 0$ and the CVA adjustment is calculated as the difference between the two. Both $\hat{u}(t, x)$ and $u(t, x)$ are calculated using the approximated characteristic function and the COS method applied to the continuation value, as is done in [1]. Due to the characteristic function being of the form (13), we are able to use a FFT in the matrix-vector multiplication. For more details, refer to Appendix A.

5.2.1. Hedging CVA. In practice CVA is hedged and thus practitioners require efficient ways to compute the sensitivity of the CVA with respect to the underlying. The widely used bump-and revalue- method, while resulting in precise calculations, might be slow to compute. Using the Fourier-based approach we find the following explicit formulas allowing for an easy computation of the first- and second-order derivatives of the CVA with respect to the underlying:

$$\begin{aligned}
\hat{\Delta} = {} & e^{-r(t_1-t_0)} \sum_{j=0}^{J-1}{}' \mathrm{Re}\left(e^{ij\pi\frac{x-a}{b-a}}\left(\frac{ij\pi}{b-a} g^d_{n,0}\left(t_0,t_1,\frac{j\pi}{b-a}\right) + g^d_{n,1}\left(t_0,t_1,\frac{j\pi}{b-a}\right)\right)\right)\hat{V}^d_j(t_1) \\
& - e^{-r(t_1-t_0)} \sum_{j=0}^{J-1}{}' \mathrm{Re}\left(e^{ij\pi\frac{x-a}{b-a}}\left(\frac{ij\pi}{b-a} g^r_{n,0}\left(t_0,t_1,\frac{j\pi}{b-a}\right) + g^r_{n,1}\left(t_0,t_1,\frac{j\pi}{b-a}\right)\right)\right)\hat{V}^r_j(t_1), \\
\hat{\Gamma} = {} & e^{-r(t_1-t_0)} \sum_{j=0}^{J-1}{}' \mathrm{Re}\Bigg(e^{ij\pi\frac{x-a}{b-a}}\Bigg(-\frac{ij\pi}{b-a} g^d_{n,0}\left(t_0,t_1,\frac{j\pi}{b-a}\right) - g^d_{n,1}\left(t_0,t_1,\frac{j\pi}{b-a}\right) \\
& + 2\frac{ij\pi}{b-a} g^d_{n,1}\left(t_0,t_1,\frac{j\pi}{b-a}\right) + \left(\frac{ij\pi}{b-a}\right)^2 g^d_{n,0}\left(t_0,t_1,\frac{j\pi}{b-a}\right) + 2g^d_{n,2}\left(t_0,t_1,\frac{j\pi}{b-a}\right)\Bigg)\Bigg)\hat{V}^d_j(t_1) \\
& - e^{-r(t_1-t_0)} \sum_{j=0}^{J-1}{}' \mathrm{Re}\Bigg(e^{ij\pi\frac{x-a}{b-a}}\Bigg(-\frac{ij\pi}{b-a} g^r_{n,0}\left(t_0,t_1,\frac{j\pi}{b-a}\right) - g^r_{n,1}\left(t_0,t_1,\frac{j\pi}{b-a}\right) \\
& - 2\frac{ij\pi}{b-a} g^r_{n,1}\left(t_0,t_1,\frac{j\pi}{b-a}\right) + \left(\frac{ij\pi}{b-a}\right)^2 g^r_{n,0}\left(t_0,t_1,\frac{j\pi}{b-a}\right) + 2g^r_{n,2}\left(t_0,t_1,\frac{j\pi}{b-a}\right)\Bigg)\Bigg)\hat{V}_j(t_1)^r,
\end{aligned}$$

where V^d_k and V^r_k are the Fourier-cosine coefficients with the defaultable and default-free characteristic functions terms, $g^d_{n,h}$ and $g^r_{n,h}$, respectively.

6. Numerical experiments.

In this Section we present numerical examples to justify the accuracy of the methods in practice. We compute the XVA with the method presented in Section 5.1 and the CVA in the case of unilateral CCR with the method from Section 5.2, which we show is more efficient for cases in which one only needs to compute the CVA.

The computer used in the experiments has an Intel Core i7 CPU with a 2.2 GHz processor. We use the second-order approximation of the characteristic function. We have found this to be sufficiently accurate by numerical experiments and theoretical error estimates. The formulas for the second-order approximation are simple, making the methods easy to implement.

6.1. A numerical example for XVA.

In this section we check the accuracy of the method from Section 5.1. We will compute the Bermudan option value with XVA using a simplified drivers function $f(t, \hat{u}(t,x)) = -r\max(\hat{u}(t,x), 0)$. Out method is easily extendible to the drivers functions in Section 3.2. Consider X_t to be a portfolio process and the payoff, if exercised at time t_m, to be given by $\Phi(t_m, x) = x$. In this case the value we can receive at every exercise date is the value of the portfolio.

Consider the model in Section 2 without default, with a local jump measure and a local volatility function with CEV-like dynamics and Gaussian jumps defined by

$$\sigma(x) = be^{\beta x}, \tag{17}$$

$$\nu(x, dq) = \lambda e^{\beta x} \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(\frac{-(q-m)^2}{2\delta^2}\right) dq. \tag{18}$$

We assume the following parameters in equations (17)-(18), unless otherwise mentioned: $b = 0.15$, $\beta = -2$, $\lambda = 0.2$, $\delta = 0.2$, $m = -0.2$, $r = 0.1$, $K = 1$ and $X_0 = 0$. In the LSM the number of time steps is taken to be 100 and we simulate 10^5 paths. In the COS method we take $L = 10$, $J = 256$, $\theta_1 = 0.5$ and $N = 10$, $M = 10$, making the total number of time steps $N \cdot M = 100$.

The results of the method compared to a LSM are presented in Table 6.1. These results show that our method is able to solve non-linear PIDEs accurately. The CPU time of the approximating method depends on the number of time steps $M \cdot N$ and is approximately $5 \cdot (N \cdot M)$ ms. The effects of the non-linear part become clear when we compare the option value with and without XVA. The results are presented in Figure 6.1. In Figure 6.2 we present the convergence results for the parameters in the COS approximation. The number of Fourier-cosine terms in the summation is given by $J = 2^d$, $d = 1, ..., 8$, the number of exercise dates is fixed, $M = 10$, and the number of time steps between each exercise date is set at $N = 1, 10$.

maturity T	X_0	MC value with XVA	COS value with XVA
0.5	0	0.03998-0.04051	0.04169
	0.2	0.2326-0.2330	0.23504
	0.4	0.4251-0.4254	0.4265
	0.6	0.6169-0.6171	0.6172
	0.8	0.8077-0.8079	0.8074
	1	1.000-1.000	1.0000
1	0	007703-0.07785	0.07878
	0.2	0.2611-0.2617	0.2660
	0.4	0.4461-0.4465	0.4493
	0.6	0.6288-0.6291	0.6311
	0.8	0.8126-0.8129	0.8120
	1	1.001-1.001	1.000

Table 6.1

A Bermudan put option with XVA (10 exercise dates, expiry $T = 1$) in the CEV-like model for the 2nd-order approximation of the characteristic function, and a LSM.

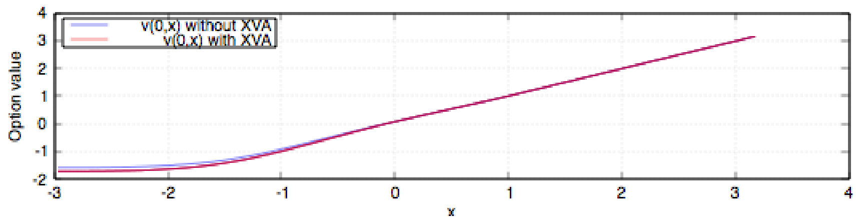

Figure 6.1. *Values for a Bermudan portfolio at time $t = 0$ with and without XVA as a function of x. The payoff function is $\Phi(t_m, x) = x$ and the process is the CEV-like model.*

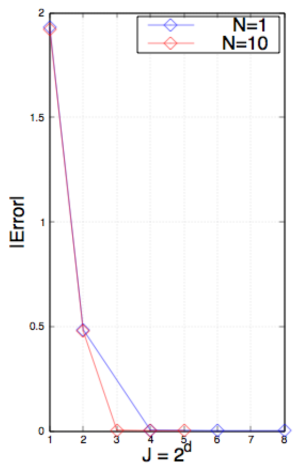

Figure 6.2. *Convergence of the absolute error for a Bermudan portfolio under the CEV-like model with payoff function $\Phi(t_m, x) = x$ for varying N and J.*

6.2. A numerical example for CVA. In this section we validate the accuracy of the method presented in Section 5.2 and compute the CVA in the case of unilateral CCR under the model dynamics given in Section 2 with a local jump measure, a local default function and a local volatility function with CEV-like dynamics and Gaussian jumps defined by defined as in (18) and a local default function $\gamma(x) = ce^{\beta x}$. We assume the same parameters as in 6.2, except $r = 0.05$ and we take $c = 0.1$ in the default function. In the LSM the number of time steps is taken to be 100 and we simulate 10^5 paths. In the COS method we take $L = 10$ and $J = 100$.

The results for the CVA valuation with the FFT-based method and with LSM are presented in Table 6.2. The CPU time of the LSM is at least 5 times the CPU time of the approximating method, which for M exercise dates is approximately $3 \cdot M$ ms, thus more efficient then the computation of the XVA with the method in 5.1. The optimal exercise boundary in Figure 6.3 shows that the exercise region becomes larger when the probability of default increases; this is to be expected: in case of the default probability being greater, the option of exercising early is more valuable and used more often.

maturity T	strike K	MC CVA	COS CVA
0.5	0.6	$4.200 \cdot 10^{-4} - 4.807 \cdot 10^{-4}$	$1.113 \cdot 10^{-4}$
	0.8	0.001525-0.001609	$9.869{\cdot}10^{-4}$
	1	0.01254-0.01273	0.01138
	1.2	0.005908-0.005931	0.005937
	1.4	0.006657-0.06758	0.006898
	1.6	0.007795-0.008008	0.007883
1	0.6	$8.673 \cdot 10^{-4} - 9.574 \cdot 10^{-4}$	$4.463 \cdot 10^{-4}$
	0.8	0.005817-0.006040	0.003535
	1	0.02023-0.02054	0.01882
	1.2	0.01221-0.01222	0.1272
	1.4	0.01378-0.01391	0.01360
	1.6	0.01532-0.01502	0.01554

Table 6.2

CVA for a Bermudan put option (10 exercise dates, expiry $T = 0.5, 1$) in the CEV-like model for the 2nd-order approximation of the characteristic function, and a LSM.

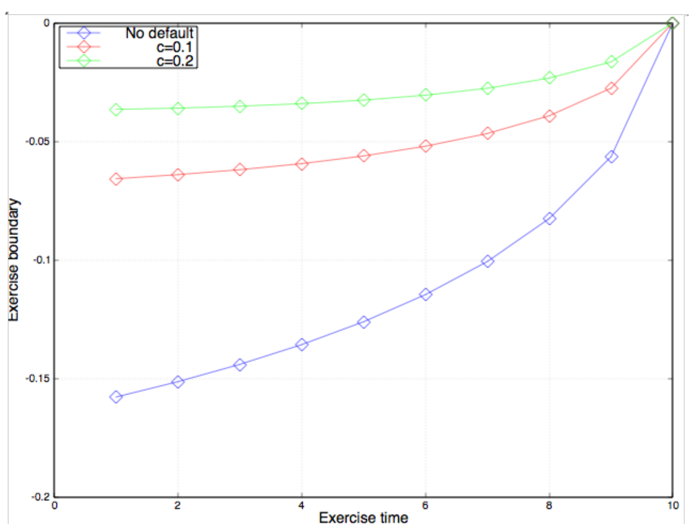

Figure 6.3. *Optimal exercise boundary for a Bermudan put option (10 exercise dates, expiry $T = 1$) in the CEV-like model with varying default $c = 0, 0.1, 0.2$.*

7. Conclusion. In this paper we considered pricing Bermudan derivatives under the presence of XVA, consisting of CVA, DVA, MVA, FVA and KVA. We derived the replicating portfolio with cashflows corresponding to the different rates for different types of lending. This resulted in the PIDE in (5) and its corresponding BSDE (8). We propose to solve the BSDE using a Fourier-cosine method for the resulting conditional expectations and an adjoint expansion method for determining an approximation of the characteristic function of the local Lévy model in (1). This approach is extended to Bermudan option pricing in Section 5.1. In Section 5.2 we present an alternative for computing the CVA term in the case of unilateral collateralization (as is the case when the derivative is an option) without the use of BSDEs. This results in an even more efficient method due to the ability of using the FFT. We verify the accuracy of both methods in Sections 6.1 and 6.2 by comparing it to a LSM and conclude that the method from Section 5.1 is able to price Bermudan options with XVA accurately and the alternative method for CVA computation from Section 5.2 is indeed more efficient than the BSDE method for computing just the CVA term.

Acknowledgments. This research is supported by the European Union in the the context of the H2020 EU Marie Curie Initial Training Network project named WAKEUPCALL.

Appendix A. The COS formulae.

Remembering that the expected value $c(t,x)$ in (16) can be rewritten in integral form, we have

$$c(t,x) = e^{-r(t_m - t)} \int_{\mathbb{R}} v(t_m, y)\Gamma(t, x; t_m, dy), \qquad t \in [t_{m-1}, t_m[,$$

where, $v(t_m, y)$ can be either $u(t_m, y)$ or $\hat{u}(t_m, y)$. Then we use the Fourier-cosine expansion to get the approximation:

$$\hat{c}(t,x) = e^{-r(t_m - t)} \sum_{j=0}^{J-1}{}' \operatorname{Re}\left(e^{-ij\pi\frac{a}{b-a}}\hat{\Gamma}\left(t, x; t_m, \frac{j\pi}{b-a}\right)\right) V_j(t_m), \qquad t \in [t_{m-1}, t_m[\tag{19}$$

$$V_j(t_m) = \frac{2}{b-a}\int_a^b \cos\left(j\pi\frac{y-a}{b-a}\right)\max\{\phi(t_m, y), c(t_m, y)\}dy,$$

with $\phi(t,x) = (K - e^x)^+$.

We can recover the coefficients $(V_j(t_m))_{j=0,1,\dots,J-1}$ from $(V_j(t_{m+1}))_{j=0,1,\dots,J-1}$. To this end, we split the integral in the definition of $V_j(t_m)$ into two parts using the early-exercise point x_m^*, which is the point where the continuation value is equal to the payoff, i.e. $c(t_m, x_m^*) = \phi(t_m, x_m^*)$; this point can easily be found by using the Newton method. Thus, we have

$$V_j(t_m) = F_j(t_m, x_m^*) + C_j(t_m, x_m^*), \qquad m = M-1, M-2, \dots, 1,$$

where

$$F_j(t_m, x_m^*) := \frac{2}{b-a}\int_a^{x_m^*} \phi(t_m, y)\cos\left(j\pi\frac{y-a}{b-a}\right)dy,$$

$$C_j(t_m, x_m^*) := \frac{2}{b-a}\int_{x_m^*}^b c(t_m, y)\cos\left(j\pi\frac{y-a}{b-a}\right)dy,$$

and $V_j(t_M) = F_j(t_M, \log K)$.

The coefficients $F_j(t_m, x_m^*)$ can be computed analytically using $x_m^* \leq \log K$, and by inserting the approximation (19) for the continuation value into the formula for $C_j(t_m, x_m^*)$ have the following coefficients $\hat{C}_j$ for $m = M-1, M-2, ..., 1$:

$$\hat{C}_j(t_m, x_m^*) = \frac{2e^{-r(t_{m+1}-t_m)}}{b-a} \cdot {\sum_{k=0}^{N-1}}' V_k(t_{m+1}) \int_{x_m^*}^{b} \mathrm{Re}\left(e^{-ik\pi\frac{a}{b-a}} \hat{\Gamma}\left(t_m, x; t_{m+1}, \frac{k\pi}{b-a}\right)\right) \cos\left(j\pi\frac{x-a}{b-a}\right) dx.$$

From (13) we know that the nth-order approximation of the characteristic function is of the form:

$$\hat{\Gamma}^{(n)}(t_m, x; t_{m+1}, \xi) = e^{i\xi x} \sum_{h=0}^{n} (x-\bar{x})^h g_{n,h}(t_m, t_{m+1}, \xi),$$

where the coefficients $g_{n,h}(t, T, \xi)$, with $0 \leq k \leq n$, depend only on t, T and ξ, but not on x.

Remark 7 (The defaultable and default-free characteristic functions). *To find $u(t,x)$ we use*

$$\hat{\Gamma}^r(t_m, x; t_{m+1}, \xi) := e^{i\xi x} \sum_{h=0}^{n} (x-\bar{x})^h g^r_{n,h}(t_m, t_{m+1}, \xi),$$

the characteristic function with $\gamma(t,x) = 0$. For $\hat{u}(t,x)$ we use

$$\hat{\Gamma}^d(t_m, x; t_{m+1}, \xi) := e^{i\xi x} \sum_{h=0}^{n} (x-\bar{x})^h g^d_{n,h}(t_m, t_{m+1}, \xi),$$

where $\gamma(t,x)$ is chosen to be some specified function.

Using (13) we can write the Fourier coefficients of the continuation value in vectorized form as:

$$\hat{\mathbf{C}}(t_m, x_m^*) = \sum_{h=0}^{n} e^{-r(t_{m+1}-t_m)} \mathrm{Re}\left(\mathbf{V}(t_{m+1}) \mathcal{M}^h(x_m^*, b) \Lambda^h\right),$$

where $\mathbf{V}(t_{m+1})$ is the vector $[V_0(t_{m+1}), ..., V_{J-1}(t_{m+1})]^T$ and $\mathcal{M}^h(x_m^*, b)\Lambda^h$ is a matrix-matrix product with $\mathcal{M}^h$ a matrix with elements $\{M^h_{k,j}\}_{k,j=0}^{J-1}$ defined as

$$M^h_{k,j}(x_m^*, b) := \frac{2}{b-a} \int_{x_m^*}^{b} e^{ij\pi\frac{x-a}{b-a}} (x-\bar{x})^h \cos\left(k\pi\frac{x-a}{b-a}\right) dx, \tag{20}$$

and Λ^h is a diagonal matrix with elements

$$g_{n,h}\Big(t_m, t_{m+1}, \frac{j\pi}{b-a}\Big), \qquad j = 0, \dots, J-1.$$

One can show, see [1], that the resulting matrix $\mathcal{M}^h$ is a sum of a Hankel and Toeplitz matrix and thus the resulting matrix vector product can be calculated using a FFT.

REFERENCES

[1] A. Borovykh, A. Pascucci, and C. W. Oosterlee, *Pricing Bermudan options under local Lévy models with default*, Submitted, (2016).

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

[3] A. Capponi, S. Pagliarani, and T. Vargiolu, *Pricing vulnerable claims in a Lévy-driven model*, Finance Stoch., 18 (2014), pp. 755–789.

[4] P. Carr and V. Linetsky, *A jump to default extended CEV model: an application of Bessel processes*, Finance Stoch., 10 (2006), pp. 303–330.

[5] A. Castagna, *Yes, FVA is a cost for derivatives desks - A note on 'is FVA a cost for derivatives desks?' by Prof. Hull and Prof. White*, working paper, (2012).

[6] R. Cont and P. Tankov, *Financial modelling with jump processes*, vol. 2, Chapman & Hall, 2004.

[7] C. de Graaf, Q. Feng, D. Kandhai, and C. Oosterlee, *Efficient computation of exposure profiles for counterparty credit risk*, International Journal of Theoretical and Applied Finance, 4 (2014).

[8] F. Fang and C. W. Oosterlee, *A novel pricing method for European options based on Fourier-cosine series expansions*, SIAM J. Sci. Comput., 31 (2008/09), pp. 826–848.

[9] ———, *Pricing early-exercise and discrete Barrier options by Fourier-cosine series expansions*, Numer. Math., 114 (2009), pp. 27–62.

[10] Q. Feng and C. Oosterlee, *Monte Carlo calculation of exposure profiles and Greeks for Bermudan and Barrier options under the Heston Hull-White model*, submitted, (2014).

[11] P. Forsyth, *An introduction to computational finance without agonizing pain*, Lecture notes, (2015), pp. 77–80.

[12] A. Green, C. Kenyon, and C. Dennis, *KVA: Capital Valuation Adjustment*, Risk, 12 (2014).

[13] J. Hull and A. White, *The FVA debate*, Risk, 7 (2012), pp. 83–85.

[14] ———, *XVAs: A gap between theory and practice*, Risk, (2016).

[15] S. Jain and C. W. Oosterlee, *The stochastic grid bundling method: Efficient pricing of Bermudan options and their Greeks*, Appl. Math. Comput., 269 (2015), pp. 412–432.

[16] E. Kromer, L. Overbeck, and J. A. Röder, *Path-dependent backward stochastic differential equations with jumps*, preprint, (2015).

[17] A. Lesniewski and A. Richter, *Managing counterparty credit risk via BSDEs*, submitted, (2016).

[18] V. Linetsky, *Pricing equity derivatives subject to bankruptcy*, Math. Finance, 16 (2006), pp. 255–282.

[19] F. Longstaff and E. Schwartz, *Valuing American options by simulation: A simple least-squares approach*, Rev. Financ. Stud., 14 (2001), pp. 113–147.

[20] M. Lorig, S. Pagliarani, and A. Pascucci, *A family of density expansions for Lévy-type processes*, Ann. Appl. Probab., 25 (2015), pp. 235–267.

[21] S. Pagliarani, A. Pascucci, and C. Riga, *Adjoint expansions in local Lévy models*, SIAM J. Financial Math., 4 (2013), pp. 265–296.

[22] A. Pascucci, *PDE and martingale methods in option pricing*, vol. 2 of Bocconi & Springer Series, Springer, Milan; Bocconi University Press, Milan, 2011.

[23] V. Piterbarg, *Funding beyond discounting: Collateral agreements and derivatives pricing*, Risk, 2 (2010), pp. 97 –102.

[24] ———, *A non-linear PDE for XVA by forward Monte Carlo*, Risk, 10 (2015).

[25] M. J. Ruijter and C. W. Oosterlee, *A Fourier-cosine method for an efficient computation of solutions to*

BSDEs, SIAM J. Sci. Comput., 37 (2015), pp. 859–889.

[26] W. Zhao, Y. Li, and G. Zhang, *A generalized θ-scheme for solving backward stochastic differential equations*, DCDS-S, 17 (2012), p. 15851603.