

## COMPUTING CREDIT VALUATION ADJUSTMENT FOR BERMUDAN OPTIONS WITH WRONG WAY RISK

QIAN FENG<sup>\*,†</sup> and CORNELIS W. OOSTERLEE<sup>\*,†,§</sup>

*\*Scientific Computing Group*

*Centrum Wiskunde & Informatica, Science Park 123  
1098 XG, Amsterdam, Netherlands*

*†Applied Mathematics (DIAM), Technische Universiteit Delft  
Mekelweg 4, 2628 CD, Delft, Netherlands*

*†qian@cw.i.nl*

*§c.w.oosterlee@cw.i.nl*

Received 7 October 2016

Revised 1 December 2017

Accepted 5 December 2017

Published 2 January 2018

We study the impact of wrong way risk (WWR) on credit valuation adjustment (CVA) for Bermudan options. WWR is modeled by a dependency between the underlying asset and the intensity of the counterparty's default. Two WWR models are proposed, based on a deterministic function and a CIR-jump (CIRJ) model, respectively. We present a nonnested Monte Carlo approach for computing CVA–VaR and CVA–expected shortfall (ES) for Bermudan options. By varying correlation coefficients, we study the impact of credit quality and WWR on the optimal exercise boundaries and CVA values of Bermudan products. Stress testing is performed.

*Keywords:* Credit Valuation Adjustment (CVA); wrong way risk (WWR); nonnested Monte Carlo; Bermudan options.

### 1. Introduction

Credit valuation adjustment (CVA) is an adjustment to the fair values of over-the-counter (OTC) derivatives to compensate for counterparty credit risk (CCR). In short, CVA measures the price of CCR in the derivatives. During the recent financial crisis, banks suffered significant losses because of the deterioration of the counterparty's creditworthiness rather than from the actual default events. The Basel Committee recognized this risk, and introduced, in BCBS (2011), the Basel III CVA capital charge that banks must hold against the CVA risk. In the advanced approach allowed by Basel III, banks need to compute the VaR and Stressed VaR at 99% confidence levels of the computed distribution of CVA losses.

Gregory (2010) presented a calculation formula for the unilateral CVA against a counterparty, where three components need to be specified: the default probability

of a counterparty, the discounted expected exposure to the counterparty and loss given default (LGD). The calculation is then often based on the assumption that the probability of the counterparty's default is independent of the exposure to this counterparty. However, when this assumption does not hold, there is another risk which may amplify the CVA value.

The so-called wrong way risk (WWR) occurs when the credit quality of the counterparty and the exposure to this counterparty are adversely correlated. In short, when the exposure to a counterparty increases, the counterparty is more likely to default. BCBS (2011) regulates that banks must consider WWR when calculating and hedging the associated CVA risk of the derivatives in their portfolio.

It is challenging to integrate WWR into pricing CVA. In the literature, one way is to model the dependency, or correlation structure, between the intensity of a counterparty's default probability and the exposure, or market variables related to the exposure. A dependency model can be developed as, for example, a Gaussian copula between exposure profile and default event, see Cespedes *et al.* (2010) or by a parametric relation between the intensity of the counterparty default and exposure (or the underlying asset), see Hull & White (2012), or the intensity can also be modeled by a stochastic variable which is correlated to the underlying asset, for example, in Brigo *et al.* (2014). Another direction is presented by Glasserman & Yang (2016), where the WWR–CVA problem is formulated as a limiting problem to determine the “worst-case” CVA of derivatives.

In these approaches, the CVA calculation is based on the exposure profile on a set of simulated paths. Drawbacks are that the computation of the exposure profile is expensive, as it is determined by the number of scenarios and time. The calculation of CVA–VaR and CVA–expected shortfall (ES) requires at least one layer of Monte Carlo sub-simulation under this approach. Moreover, there is an underlying assumption in this formula that the exposure on these paths can be computed the same way as when there were no risk of default event. This is not practical for derivatives with early exercise features, such as Bermudan options. When the option writer's credit is significantly downgraded, option holders with early-exercise rights may more likely adjust their expectations by taking into account the price of the counterparty default and thus change the so-called early-exercise decision. In other words, the impact of CVA should be considered when pricing a Bermudan option.

We will set up a valuation framework of computing CVA for such options. Instead of using the calculation formula by Gregory (2010), one can directly employ the definition of CVA for its calculation, in which CVA is the difference between the default risk-free option value and the market value which takes CCR into account. We will present two algorithms, based on the Fourier transform and Monte Carlo simulation, respectively, that can efficiently compute these two option values under a proposed default probability model. Compared to the regular method based on the exposure profile over a large number of simulated scenarios, the computational cost is lower due to the avoidance of sub-simulations.

Another contribution in this paper is to consider the impact of counterparty's CVA to the early-exercise boundary in the valuation of Bermudan options. In addition to the regular exercise strategy, another exercise strategy may be to extract the price of a counterparty default from the expected payoff when determining the optimal exercising asset price. Breton & Marzouk (2015) presented a similar study on the impact of CVA. In this paper, we will show how the early-exercise boundaries vary with WWR. By changing the correlation, we can analyze the effect of WWR on the early-exercise boundary.

In order to incorporate WWR into the calculations, we set up a parametric relation between the intensity of the counterparty and the underlying asset. The dependency between the underlying stock and the intensity rate is modeled in two ways, i.e. the intensity is either a deterministic function of the log-stock value, or by means of jumps with correlated jump sizes in the processes. The stochastic models are based on the class of affine jump diffusion (AJD) processes, see Duffie *et al.* (2000).

The paper is organized as follows. Section 2 presents the intensity model setting. Section 3 discusses the two WWR models to describe the dependency structure between the underlying and the intensity, in the AJD class. Section 4 shows the CVA calculation of Bermudan options, and two early-exercise strategy. Section 5 presents a procedure of computing CVA ES via–nonnested Monte Carlo simulation. Numerical results in Sec. 6 are used to study the effect of WWR, the impact of the dependency structure on CVA and on the CVA–ES. Details of numerical schemes are presented in the Appendices A–C.

## 2. Intensity Model

Intensity models have been studied for many years. Jarrow & Turnbull (1995), Madan & Unal (1998) used intensity models for pricing the risk of default. Duffie & Singleton (1999) presented intensity models for the valuation of defaultable corporate bonds, where the default-adjusted value of a contingent claim was modeled by discounting at a *default-adjusted* short rate. Bielecki & Rutkowski (2002) explained the use of an enlarged filtration for valuation when dealing with both default and market information, and established a valuation framework for claims subject to credit risk, and Lando (2009) presented another intensity model.

We start with the general setting, assuming that a probability space  $(\Omega, \mathcal{F}, \mathbb{Q})$  is given over the time horizon  $[0, T]$ . The sample space  $\Omega$  is the set of all possible outcomes  $w$ , the sigma algebra  $\mathcal{F}$  is the collection of all events and probability measure  $\mathbb{Q}$  is a function  $\mathbb{Q} : \Omega \rightarrow [0, 1]$ . In this paper, we consider derivatives valuation, and it is thus convenient to choose  $\mathbb{Q}$  as a risk-neutral measure. Under this risk-neutral measure, all securities can be priced as discounted values. We assume the risk-free short rate  $r$  to be constant, and denote the risk-free savings account by  $B_t = \exp(rt)$ .

Let a Markov process  $\{\mathbf{X}_t, t \in [0, T]\}$  on some  $d$ -dimensional space  $\mathcal{D} \subset \mathbb{R}^d$  represent all information of the underlying market factors, such as the underlying

equity, interest rate, credit spread of the counterparty. The filtration is generated by  $\mathbf{X}_t$  and is denoted by  $\mathcal{G}_t = \sigma(\mathbf{X}_s, s \in [0, t])$ . Nonnegative function  $h : \mathbb{R}^d \rightarrow \mathbb{R}$  is the  $\mathcal{G}_t$ -measurable intensity of a jump process  $N_t$ . We take the first jump-to-default time  $\tau$  of the process  $N_t$  as the default time of the counterparty,  $\tau := \inf\{t : \int_0^t h(\mathbf{X}_s) ds \geq E_1\}$ , where  $E_1$  is an exponential random variable with mean 1, independent of  $\mathcal{G}_t$ . Define the jump process as  $N_t = \mathbf{1}(\tau \leq t)$ . The filtration that includes the default information is  $\mathcal{H}_t = \sigma(N_s, s \in [0, t])$ . The variable  $h$  is the state-dependent (pre-default) intensity of the jump time  $\tau$ . We construct an enlarged filtration  $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{H}_t$ , which contains information of market factors and default events.

Denote the survival probability with respect to the filtration  $\mathcal{G}$  by  $G$ ,

$$G_t = \mathbb{E}^{\mathbb{Q}}[\mathbf{1}(\tau > t) | \mathcal{G}_t] = \exp\left(-\int_0^t h(\mathbf{X}_s) ds\right), \quad (2.1)$$

where  $h(\mathbf{X}_t)$  depends on the state variables.

Based on a result by Bielecki & Rutkowski (2002, Corollary 5.1.1, p. 145), one can further show that the value of the defaultable claim is the risk-neutral expectation of the payoff discounted at the *default-adjusted rate*,  $R(\mathbf{X}_s) := r + h(\mathbf{X}_s)$ . Here, we define

$$D_t := \exp\left(\int_0^t R(\mathbf{X}_s) ds\right) = \frac{B_t}{G_t}, \quad (2.2)$$

which includes the counterparty credit quality.

### *Defaultable ZCB and market credit spread*

Assume that for the issued  $T$ -period defaultable bonds, the counterparty pays 1 at expiration time  $T$  in the event of no default, and zero otherwise. We can read the price of the defaultable bonds directly from the market. It is well-known that corporate bonds offer a higher yield because of extra risk (default risk). The difference between the yield on a corporate bond and a government bond is called the *credit spread*. The  $t = 0$  value of a  $T$ -period defaultable zero-coupon-bond (ZCB) is given by

$$p(0, T) = \mathbb{E}^{\mathbb{Q}}\left[\frac{1}{D_T}\right], \quad (2.3)$$

where the intensity can be considered as the credit spread associated to the counterparty.

Suppose that the market credit spread is flat over different tenors, denoted by  $\bar{h}$ , then the market price of this defaultable ZCB is given by

$$p_{\text{market}}(0, T) = \exp(-(r + \bar{h})T), \quad (2.4)$$

which gives the market information of the credit quality of the counterparty.

The market credit spread  $\bar{h}$  varies with the credit rating of a counterparty. It grows when the credit quality of the counterparty decreases. One can hence retrieve

the information of the counterparty credit quality directly from the defaultable bonds issued by the counterparty, or from its credit default swap (CDS). The calibration of a default model can be done by comparing formula (2.3) to the market price of the defaultable bonds (2.4).

### 3. WWR Modeling

We use the AJD class to describe the dependency structure between the underlying asset and the intensity. Affine dynamics are convenient and give rise to efficient computation. The AJD class is defined by an SDE system, see Duffie *et al.* (2000). We present the SDE equations in the AJD class model in Appendix A. We will define two ways to describe the dependency structure of the log-asset price and the intensity based on the AJD models.

#### 3.1. The DF model: A deterministic function

The point of departure in our study is asset dynamics based on geometric Brownian motion (GBM) and jump diffusion, which we correlate with an intensity process. Stochastic volatility for asset prices is not considered here. We have some chosen reference models, not necessarily used in practice, with the purpose of analyzing the impact of WWR under different model assumptions.

With  $X_t = \log(S_t)$  the one-dimensional log-asset process, as a first model, we model the dynamics of the underlying and the intensity before a default event by

$$dX_t = \left( r - \frac{1}{2}\sigma_x^2 \right) dt + \sigma_x dW_t, \quad (3.1)$$

$$h(X_t) = a(t) + bX_t,$$

with  $\sigma_x > 0$  the volatility of the log-asset process,  $W_t$  a standard Brownian motion, coefficient  $b \in \mathbb{R}$  is constant and  $a(t)$  is a deterministic function that satisfies  $a(0) = \bar{h} - bX_0$ . We call this model *the DF model*. Coefficient  $b$  denotes the strength of the dependence between the log-asset and the intensity. When holding put options, WWR occurs when  $b$  is negative, otherwise we deal with right way risk (RWR).

The discounted characteristic function (ChF) of this model is given in Appendix A. Comparing to the market price of the ZCB in (2.4), with coefficient  $b$  fixed, the formula for  $a(t)$  is given by

$$a(t) = \bar{h} - bX_0 - \left( r - \frac{1}{2}\sigma_x^2 \right) bt + \frac{1}{2}b^2\sigma_x^2 t^2. \quad (3.2)$$

The log-asset and the intensity processes can be represented by the following formulas:

$$\begin{aligned} X_t &= X_0 + \left( r - \frac{1}{2}\sigma_x^2 \right) t + \sigma_x W_t, \\ h_t &= \bar{h} + \frac{1}{2}\sigma_x^2 t^2 b^2 + b\sigma_x W_t. \end{aligned} \quad (3.3)$$

It is not trivial to ensure that the intensity is nonnegative under the affine structure. However, we use specific parameter settings to keep the intensity positive at a high probability over a certain time horizon.

From (3.3), it follows that coefficient  $b$  does not affect  $X_t$ . The variance of the intensity is given by  $b^2\sigma_x^2 t$ . The expectation of the intensity at time  $t$  is given by  $\bar{h} + \frac{1}{2}\sigma_x^2 t^2 b^2$ .

The intensity is governed by the process that also controls the log-asset price. It is positive over the time horizon  $[0, T]$  with probability  $(1 - 2\zeta)$ , when the value of coefficient  $b$  lies in the interval,

$$(\Delta^+(\Phi^{-1}(1 - \zeta), T), \Delta^-(\Phi^{-1}(\zeta), T)), \quad (3.4)$$

where  $\Phi$  is the cumulative distribution function (CDF) of the standard normal distribution, and  $\Delta^\pm$  are given by

$$\Delta^\pm(Z, t) = \sigma_x t^{-\frac{3}{2}}(-Z \pm \sqrt{Z^2 - 2t\bar{h}}), \quad (3.5)$$

see Appendix B.

With this result, we can set limits for coefficient  $b$  with certain tolerance regarding the negative values of the intensity.

**Comment:** Compared to the model given by Hull & White (2012), and the models proposed by Ruiz (2014), the DF model has the drawback that the intensity may become negative. The affine structure of the DF model has, of course, advantages for the computations, like the availability of an analytic formula of the defaultable ZCB price. The deterministic function  $a(t)$  enables fitting the term-structure of bonds with different tenors. The financial meaning of the coefficients seems clear. The stock value of the company impacts the credit spread.

### 3.2. CIR model with jumps

As a second model, we use the CIR model with upward jumps to ensure that the intensity stays nonnegative, which we call *the CIR-jump (CIRJ) model*. Assume that  $(X_t, Y_t)$  is governed by

$$\begin{aligned} dX_t &= \left( r - \frac{1}{2}\sigma_x^2 - \bar{\lambda}\bar{\mu} \right) dt + \sigma_x dW_1(t) + dJ_t^1, \\ dY_t &= \gamma(\theta - Y_t)dt + \sigma_y \sqrt{Y_t} dW_2(t) + dJ_t^2, \\ h(Y_t) &= \psi^{\text{CIR}}(t) + Y_t, \end{aligned} \quad (3.6)$$

where  $\gamma$ ,  $\theta$  and  $\sigma_y$  are positive constants, representing the speed of mean reversion, the reverting level and the intensity volatility, respectively. Here, the diffusion terms are not correlated but we assume that  $J_t^1$  and  $J_t^2$  jump simultaneously, and that their intensities<sup>a</sup> are correlated. The resulting model is an affine model.

<sup>a</sup>Notice that  $\bar{\lambda}$  is the intensity of jumps in the process  $(X_t, Y_t)$ , not related to the default time  $\tau$ .

The marginal distribution of the jumps in  $Y_t$  is exponential with mean  $\mu_y^J$ ; conditional on a realization of the jump size in  $Y_t$ , say  $z_y$ , the jump size in  $X_t$  is normally distributed with mean  $\mu_x^J + \rho_J z_y$  and variance  $\sigma_x^J$ . The jump transform function  $\Theta$  is given by Duffie *et al.* (2000),

$$\Theta(c_1, c_2) = \frac{\exp\left(\mu_x^J c_1 + \frac{1}{2}(\sigma_x^J)^2 c_1^2\right)}{1 - \mu_y^J c_2 - \rho_J \mu_y^J c_1}, \quad (3.7)$$

with the risk-neutral jump size  $\bar{\mu} = \Theta(1, 0)$ , in (3.6); the initial values of the process are  $(X_0, Y_0)$ , and the function satisfies  $h(0) = \bar{h}$ .

CIR process  $Y_t$  is nonnegative, and hence the intensity is nonnegative as long as  $\psi^{\text{CIR}}$  is positive along the time horizon. Using the default-adjusted ChF to calibrate to the market price given by (2.4), function  $\psi^{\text{CIR}}$  is given by

$$\psi^{\text{CIR}}(t) = \bar{h} - \frac{2\bar{\lambda}\mu_y^J(e^{d_1 t} - 1)}{\bar{g} + 2\mu_y^J(e^{d_1 t} - 1)} - \frac{2\gamma\theta}{\bar{g}}(e^{d_1 t} - 1) - \frac{4Y_0 d_1^2}{\bar{g}^2} e^{d_1 t}, \quad (3.8)$$

where

$$d_1 = \sqrt{\gamma^2 + 2\sigma_y^2}, \quad \bar{g} = \gamma e^{d_1 t} - \gamma + d_1 + d_1 e^{d_1 t}. \quad (3.9)$$

Jumps thus occur at the same time in these two processes. The financial meaning is that due to unexpected events, the stock and the credit quality are simultaneously affected. By this model, we assume that the stock increases/decreases sharply, while the credit rating of the option writer improves/deteriorates at the same time. With a put option, WWR incurs to option holders when  $\rho_J$  is negative, as the exposure to the counterparty increases.

The mean and variance of the log-asset model are given by

$$\begin{aligned} \text{mean}(X_t) &= X_0 + \left(r - \frac{1}{2}\sigma_x^2\right)t + \bar{\lambda}t(\mu_x^J - \bar{\mu} + \rho_J \mu_y^J), \\ \text{var}(X_t) &= t\sigma_x + t\bar{\lambda}\{(\sigma_x^J)^2 + (\mu_x^J)^2 + 2(\mu_y^J \rho_J)^2 - 2\mu_y^J \mu_x^J \rho_J\}, \\ \text{var}(Y_t) &= \frac{\sigma_y^2 \theta}{2\gamma} + \frac{\bar{\lambda}(\mu_y^J)^2}{\gamma}, \end{aligned} \quad (3.10)$$

where the variances can be obtained by the first and second moments using the nondiscounted ChF, and the long-term variance of  $Y_t$  can be obtained by taking the limit.

From (3.10), we see that the parameters  $\mu_y$  and  $\bar{\lambda}$  control the default risk, and also have an impact on the stock value. Correlation  $\rho_J$  is the key correlation between these two processes, and it influences the impact of the jump term in the stock process. By varying three parameters  $(\mu_y, \bar{\lambda}, \rho_J)$  in the jump terms the default-adjusted and default-free values are influenced, since the underlying process will change.

**Comment:** The use of a jump-diffusion process for the intensity can be found in Duffie & Garleanu (2001), Brigo *et al.* (2014). We have not seen studies on the

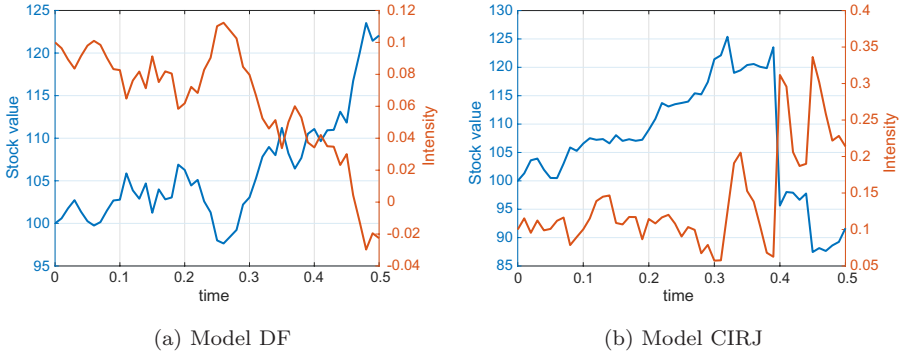


Fig. 1. One path under the risk-neutral measure: the blue lines represent the stock values and the red-dashed lines the value of the  $Y$  process. Plot (a) is by the DF model with  $b = -0.6$ ,  $\sigma_x = 0.2$ ; plot (b) by the CIRJ model  $\sigma_x^{\text{BS}} = 0.2$ ,  $\bar{\lambda} = 5$ ,  $\rho_J = -0.9$ ,  $\mu_y^J = 0.04$ .

simultaneous jumps in the intensity process and the underlying asset process. The CIRJ model considers the correlation between the two factors when there is some sudden change in the market. Even when the jump size correlation coefficient  $\rho_J = 0$ , there may still be correlation between variables  $X_t$  and  $Y_t$ , as the jumps in these two processes happen simultaneously with intensity  $\bar{\lambda}$ . One can show that when  $\rho_J = 0$ ,

$$\mathbb{E}[X_t Y_t] - \mathbb{E}[X] \mathbb{E}[Y] = \frac{\bar{\lambda} \mu_y^J \mu_x^J}{\gamma} (1 - e^{-\gamma t}), \quad (3.11)$$

which implies that  $X_t$  and  $Y_t$  are independent, when at least one of the parameters  $(\bar{\lambda}, \mu_y^J, \mu_x^J)$  equals zero.

### 3.3. Comparison of two models

For these two models, we present one Monte Carlo path in Fig. 1, where we see that for the paths for the CIRJ model high values of intensity do not necessarily imply lower stock values. The dependency structure defined by these two models is different. The impact of WWR on options may thus also be different.

## 4. CVA

In this section, we present the valuation framework of computing CVA for Bermudan options.

### 4.1. Valuation of defaultable Bermudan options

A Bermudan option can be exercised at a set of dates  $\mathcal{T} = \{t_1, \dots, t_M\}$ , where  $t_1 > 0$  is the first and  $t_M$  is the last exercise date. The payoff of immediate exercise  $t$  is given by  $f(X_t)$ , with  $f$  as follows:

$$f(X_t) = \begin{cases} \max(\exp(X_t) - K, 0), & \text{for a call,} \\ \max(K - \exp(X_t), 0), & \text{for a put,} \end{cases} \quad (4.1)$$



where  $K$  is the strike price,  $X_t$  represents for the log-asset variable. Of course,  $f(X_t) \geq 0$  for all  $X_t \in \mathbb{R}$ .

When the claim is defaultable and CCR is present, a specific exercise strategy can be applied. For example, it is may be optimal to exercise the Bermudan option earlier when a default is likely to happen in the near future compared to the default-free case. One exercise policy is to optimize the default-adjusted value, taking explicitly into account the change in the credit quality of a counterparty. When the probability of counterparty default is low, it may be fine to exercise the option by optimizing the default-free value.

We refer to the strategy that *optimizes the default-free value of the option as Policy I*, and the one that *optimizes the default-adjusted value of the option as Policy II*. It is easy to see that the early-exercise boundaries under these two policies will be different. In addition, under each policy, there is a “default-free” value and a “default-adjusted” value of the option related to the associated optimal early-exercise time. The change in the early-exercise boundaries will impact the “default-free” value and the “default-adjusted” value at the same time.

**Policy I: Optimize the default-free value.** We look for an optimal stopping time  $\tilde{\xi}^I = \tilde{\xi}^I(\mathbf{X}_0, w) \in \mathcal{T}$ , with  $w \in \Omega$ , such that

$$\mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\tilde{\xi}^I})}{B_{\tilde{\xi}^I}} \middle| \mathbf{X}_0 \right] := \sup_{\xi \in \mathcal{T}} \mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\xi})}{B_{\xi}} \middle| \mathbf{X}_0 \right], \quad (4.2)$$

and the corresponding optimal *default-free* expected value obtained with the optimal stopping time  $\tilde{\xi}^I$  reads,

$$V_0^I = \mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\tilde{\xi}^I})}{B_{\tilde{\xi}^I}} \middle| \mathbf{X}_0 \right]. \quad (4.3)$$

The corresponding *default-adjusted* expected value at initial time  $t = 0$ , with stopping time  $\tilde{\xi}^I$ , is then given by

$$U_0^I = \mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\tilde{\xi}^I})}{D_{\tilde{\xi}^I}} \middle| \mathbf{X}_0 \right]. \quad (4.4)$$

**Policy II: Optimize the default-adjusted value.** We look for the optimal stopping time  $\tilde{\xi}^{II} = \tilde{\xi}^{II}(\mathbf{X}_0, w) \in \mathcal{T}$ , with  $w \in \Omega$ , such that

$$\mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\tilde{\xi}^{II}})}{D_{\tilde{\xi}^{II}}} \middle| \mathbf{X}_0 \right] = \sup_{\xi \in \mathcal{T}} \mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\xi})}{D_{\xi}} \middle| \mathbf{X}_0 \right]. \quad (4.5)$$

Using optimal stopping time  $\tilde{\xi}^{II}$ , the corresponding *default-free* expected value is given by

$$V_0^{II} = \mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\tilde{\xi}^{II}})}{B_{\tilde{\xi}^{II}}} \middle| \mathbf{X}_0 \right]. \quad (4.6)$$

We define the corresponding optimal *default-adjusted* expected value by

$$U_0^{\text{II}} = \mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\tilde{\xi}^{\text{II}}})}{D_{\tilde{\xi}^{\text{II}}}} \middle| \mathbf{X}_0 \right]. \quad (4.7)$$

The optimal stopping problem for the default-free value in (4.2) is well-known. As in Andersen & Piterbarg (2010), one can define the Snell envelope for the discounted (at the short rate) exercise values, and the corresponding Hamilton–Jacobi–Bellman equation can be solved by backward recursion. In our setting, we also need to compute the expectation defined in (4.4) under the same exercise policy, which is easy when we know the exercise and continuation regions.

In a similar way, the optimal stopping problem for the default-adjusted value in (4.5) can be solved by constructing sequences of random variables to define the Snell envelope for the exercise values discounted at the intensity-adjusted rate. The difference in the backward recursion procedures for these two problems is to determine the exercise and continuation regions.

Let  $j \in \{\text{I}, \text{II}\}$  be the index of the chosen early-exercise policy. To reduce notation, from now on we write state variable  $\mathbf{X}_{t_m} = \mathbf{X}_m$ , with  $t_m$  the  $m$ th exercise opportunity.

First of all, the default-free and default-adjusted values at expiration date  $t_M = T$ , conditional on  $\mathbf{X}_M$ , are given by the payoff function, i.e.

$$U_M^j = V_M^j = f(X_M), \quad (4.8)$$

where  $X_M$  is the corresponding log-asset value in the state vector  $\mathbf{X}_M$ .

At an exercise date  $t_m < T$ , assuming  $\tau > t_m$ , the expected default-free and default-adjusted payoff of continuing, i.e. holding the option at time  $t_m$ , are given, respectively, by

$$\begin{aligned} c^j(\mathbf{X}_m, t_m, t_{m+1}) &= \mathbb{E}^{\mathbb{Q}} \left[ \frac{B_m}{B_{m+1}} V_{m+1}^j \middle| \mathbf{X}_m \right], \\ g^j(\mathbf{X}_m, t_m, t_{m+1}) &= \mathbb{E}^{\mathbb{Q}} \left[ \frac{D_m}{D_{m+1}} U_{m+1}^j \middle| \mathbf{X}_m \right], \end{aligned} \quad (4.9)$$

where  $c^j$  represents the default-free and  $g^j$  the default-adjusted continuation function under the policy  $j$ , and both are  $L^2$  bounded.

The optimal exercise boundary  $\{\tilde{\mathbf{X}}_m^j\}_{m=1}^M$  using Policy I at  $t_m > 0$  is determined by solving the equation

$$\begin{aligned} \text{Optimal boundary: } & f(\tilde{X}_m^{\text{I}}) - c^{\text{I}}(\tilde{\mathbf{X}}_m^{\text{I}}, t_m, t_{m+1}) = 0, \\ \text{Continuation region: } & \mathcal{C}^{\text{I}}(t_m) = \{\mathbf{X}_m \mid f(X_m) < c^{\text{I}}(\mathbf{X}_m, t_m, t_{m+1})\}, \\ \text{Early-exercise region: } & \mathcal{G}^{\text{I}}(t_m) = \{\mathbf{X}_m \mid f(X_m) \geq c^{\text{I}}(\mathbf{X}_m, t_m, t_{m+1})\}, \end{aligned} \quad (4.10)$$

where the log-asset value  $\tilde{\mathbf{X}}^j$  is the log-asset value in the vector  $\tilde{\mathbf{X}}_m^j$ .

By Policy II, the continuation value accounts for the default risk, being discounted at the default-adjusted value, and one compares this default-adjusted continuation value with the immediate exercise value. Hence, the optimal early exercise boundary, continuation region and early-exercise region using Policy II can be expressed by

$$\begin{aligned} \text{Optimal boundary: } & f(\tilde{X}_m^{\text{II}}) - g^{\text{II}}(\tilde{\mathbf{X}}_m^{\text{II}}, t_m, t_{m+1}) = 0, \\ \text{Continuation region: } & \mathcal{C}^{\text{II}}(t_m) = \{\mathbf{X}_m \mid f(X_m) < g^{\text{II}}(\mathbf{X}_m, t_m, t_{m+1})\}, \\ \text{Early-exercise region: } & \mathcal{G}^{\text{II}}(t_m) = \{\mathbf{X}_m \mid f(X_m) \geq g^{\text{II}}(\mathbf{X}_m, t_m, t_{m+1})\}. \end{aligned} \quad (4.11)$$

The optimal early-exercise time is path-dependent, so are the option values at a future time  $t$ . Given a path  $w$ , the option values, both default-free and default-adjusted values, and the optimal early-exercise time can be obtained via a backward recursion procedure. Under the policy  $j$ , the backward computation can be expressed by,

When  $t_m = T$ ,

$$\begin{aligned} U_M^j(w) &= V_M^j(w) = f(X_M(w)), \\ \xi_M^j(w) &= \begin{cases} t_M, & f(X_M(w)) > 0, \\ +\infty, & \text{otherwise,} \end{cases} \end{aligned} \quad (4.12)$$

where  $\xi_M^j(w)$  is a temporary value for the induction of the optimal early-exercise time  $\tilde{\xi}^j(w)$ .

For  $m = M - 1, M - 2, \dots, 1$ , the option values at time  $t_m$  are computed by

$$\begin{aligned} V_m^j(w) &= \begin{cases} f(X_m(w)), & \mathbf{X}_m(w) \in \mathcal{G}^j(t_m), \\ c^j(\mathbf{X}_m(w), t_m, t_{m+1}), & \mathbf{X}_m(w) \in \mathcal{C}^j(t_m). \end{cases} \\ U_m^j(w) &= \begin{cases} f(X_m(w)), & \mathbf{X}_m(w) \in \mathcal{G}^j(t_m), \\ g^j(\mathbf{X}_m(w), t_m, t_{m+1}), & \mathbf{X}_m(w) \in \mathcal{C}^j(t_m). \end{cases} \\ \xi_m^j(w) &= \begin{cases} t_m, & \mathbf{X}_m(w) \in \mathcal{G}^j(t_m), \\ \xi_{m+1}^j(w), & \mathbf{X}_m(w) \in \mathcal{C}^j(t_m), \end{cases} \end{aligned} \quad (4.13)$$

where  $\xi_m^j(w)$  is a temporary value for the induction of the optimal early-exercise time  $\tilde{\xi}^j(w)$ .

At time  $t_m = 0$ , today's option values and the final optimal early-exercise time  $\tilde{\xi}^j(w)$  is updated as

$$\begin{aligned} V_m^j(w) &= c^j(\mathbf{X}_0(w), t_0, t_1), \\ U_m^j(w) &= g^j(\mathbf{X}_0(w), t_0, t_1), \\ \tilde{\xi}^j(w) &= \xi_1^j(w). \end{aligned} \quad (4.14)$$

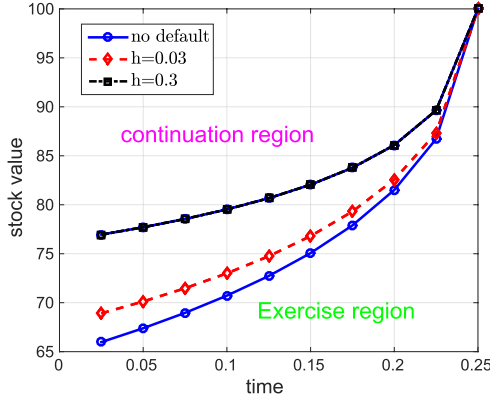


Fig. 2. Optimal exercise boundaries for a Bermudan put with a constant intensity  $h = \{0, -0.03, -0.3\}$ .

As an example for Policy II, consider a Bermudan put option. Assume that the intensity is a constant and plot the optimal early exercise boundaries at each exercise opportunity against time. Figure 2 compares the optimal early exercise boundaries as the credit quality of the counterparty deteriorates from  $h = 0$  to  $h = 0.03$  and finally to  $h = 0.3$ . As the credit spread gets larger, the area of the exercise region expands. It implies that the option holder is more likely to exercise the option in the earlier stage of the contract when the counterparty has a high risk of default.

#### 4.2. Calculation of CVA

When we compute the price of CCR of the contract, one also considers the likeliness of a counterparty default. In this paper, we avoid the discussion about recovery by setting the recovery equal to zero (There are at least two ways of defining recovery, see Duffie & Singleton (1999)). At a future time  $t$ , with the counterparty default time  $\tau < t$ , there is no need for discussion under the zero recovery assumption; otherwise, the time- $t$  value of the corresponding CVA to the option writer, conditional on  $\tau > t$ , is given by

$$\begin{aligned} \text{CVA}_t &= B_t \mathbb{E}^{\mathbb{Q}} \left[ \mathbf{1}(t < \tau < T) \frac{E_\tau}{B_\tau} \middle| \tau > t, \mathbf{X}_t \right], \\ &= \frac{B_t}{G_t} \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T \frac{E_s}{B_s} d(-G_s) \middle| \mathbf{X}_t \right], \end{aligned} \tag{4.15}$$

where  $E_\tau$  is the exposure to the counterparty when default occurs at time  $\tau$  before expiration  $T$ . The second part of the equation is found by Proposition 5.1.1 in Bielecki & Rutkowski (2002, pp. 147).

As discussed, the CVA value of the option under these two policies can be computed by taking the difference between the “default-free” and the “default-adjusted” values. For options with early-exercise feature, future CVA values are

path-dependent. Given a path  $w$ , the time  $t$  value of CVA of the Bermudan options using Policy  $j$ , where  $j \in \{\text{I}, \text{II}\}$ , are defined by

$$\text{CVA}_t^j(w) = \begin{cases} V_t^j(w) - U_t^j(w), & t < \xi^j(w), \\ 0, & t \geq \xi^j(w). \end{cases} \quad (4.16)$$

In this section, we will show that using the CVA definition in (4.16) is equivalent to the calculation formula in (4.15), when the exposure to the counterparty at a future time  $t$  prior to the default event on  $w \in \Omega$  is defined as the replacement costs, given by

$$E_t^j(w) = \begin{cases} V_t^j(w), & t < \xi^j(w), \\ 0, & t \geq \xi^j(w), \end{cases} \quad (4.17)$$

where  $E^j(w, t)$  is the exposure under Policy  $j$  at time  $t$  on path  $w$ , the optimal stopping time  $\xi^j(w)$  is determined by Policy I (4.2) or Policy II (4.5) and  $V^j$  is the expected value of continuing the option at time  $t$  using Policy  $j$ , i.e.

$$V_t^j(w) = B_t \mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\xi^j(w)})}{B_{\xi^j(w)}} \middle| \mathbf{X}_t(w) \right], \quad t < \xi^j(w). \quad (4.18)$$

Given a path  $w$ , at time  $t \geq \xi^j(w)$ , it does not make sense to compute CVA as the option has been exercised, so we mark the associated CVA value as zero. Otherwise, at time  $t \leq \xi^j(w)$ , there is the following relation:

$$\begin{aligned} & V_t^j(w) - U_t^j(w) \\ &= B_t \mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\xi^j(w)})}{B_{\xi^j(w)}} \middle| \mathbf{X}_t(w) \right] - D_t \mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\xi^j(w)})}{D_{\xi^j(w)}} \middle| \mathbf{X}_t(w) \right] \\ &= \frac{B_t}{G_t} \mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\xi^j(w)})}{B_{\xi^j(w)}} (G_t - G_{\xi^j(w)}) \middle| \mathbf{X}_t(w) \right] \\ &= \frac{B_t}{G_t} \mathbb{E}^{\mathbb{Q}} \left[ \int_t^{\xi^j(w)} \underbrace{\frac{1}{B_s} B_s \mathbb{E}^{\mathbb{Q}} \left[ \frac{f(X_{\xi^j(w)})}{B_{\xi^j(w)}} \middle| \mathbf{X}_s \right]}_{E_s^j(w)} d(-G_s) \middle| \mathbf{X}_t(w) \right] \\ &= \text{CVA}_t^j(w). \end{aligned} \quad (4.19)$$

Hence, Eq. (4.15) is equivalent to Eq. (4.16) for computing the CVA of Bermudan options.

## 5. Calculation of CVA–ES

The following is stated by BCBS (2015)[Annex1]: “...CVA risk is defined as the risk of losses arising from changing CVA values in response to changes in counterparty credit spread and market risk factors that drive market prices of derivative transactions...”. The CVA–ES is an important measure of the CVA risk.

The CVA–ES reflects the expected value of the losses in the worst 2.5% of the scenarios under the real-world probability measure  $\mathbb{P}$ . The formula for the CVA–ES at time  $t$  is as follows:

$$\text{CVA–ES}(t) = \mathbb{E}^{\mathbb{P}}[L_t | L_t > \text{VaR}_{L_t}(97.5\%)], \quad (5.1)$$

where the CVA loss  $L_t := \text{CVA}_t - \text{CVA}_0$ ;  $\text{VaR}_{L_t}(97.5\%)$  is the 97.5% quantile of the loss distribution  $L_t$ ;  $\text{CVA}_t$  represents the CVA value at time  $t$ .

The calculation of the CVA–ES can be summarized as in Brigo *et al.* (2013).

- First generate scenarios under the real-world probability measure  $\mathbb{P}$  at a set of discrete time points over the time horizon  $[0, T]$ .
- Then, for each real-world scenario, the CVA values are calculated at a set of discrete future time points until expiration.
- Subsequently, one determines the empirical real-world ( $\mathbb{P}$ ) CVA distribution, selects a quantile at a confidence level (say 97.5%) and computes the CVA–VaR, denoted by  $\text{VaR}_{L_t}$ , which is the 97.5% quantile of the CVA loss ( $\text{CVA}_t - \text{CVA}_0$ ).

Calculation of the CVA–VaR and the CVA–ES require different probability measures: the value of CVA is computed under the risk-neutral probability measure  $\mathbb{Q}$  while the CVA–ES is obtained under the real-world probability measure  $\mathbb{P}$ .

We present a nonnested Monte Carlo procedure for computing CVA for Bermudan products, given the formula (4.16), based on backward recursion.

- Generate  $N$  paths at time points  $0 < t_1 < t_2 < \dots < t_M = T$  under the real-world probability measure  $\mathbb{P}$ ; Denote the realized value of the market variable by  $\{x_{m,n}\}$ ,  $n = 1, \dots, N$ ,  $m = 1, \dots, M$ .
- For each path at time  $t_m$ ,  $m = M - 1, M - 2, \dots, 1$ ,
  - Compute the default-free value of continuing the option by the function  $c(x_{m,n}, t_m, t_{m+1})$ , the default-adjusted value of continuing the option by the function  $g(x_{m,n}, t_m, t_{m+1})$ , and the received payoff of exercising the option  $f(x_{m,n})$
  - At the  $n$ th path exercise takes place, when  $c(x_{m,n}, t_m, t_{m+1}) < f(x_{m,n})$  under Policy I, and  $g(x_{m,n}, t_m, t_{m+1}) < f(x_{m,n})$  under Policy II.
  - The default-free and default-adjusted values of the option at the  $n$ th path are given by

$$\begin{aligned}
 v_{m,n} &= \begin{cases} f(x_{m,n}), & \text{if exercised,} \\ c(x_{m,n}, t_m, t_{m+1}), & \text{otherwise,} \end{cases} \\
 u_{m,n} &= \begin{cases} f(x_{m,n}), & \text{if exercised,} \\ g(x_{m,n}, t_m, t_{m+1}), & \text{otherwise,} \end{cases} \\
 \text{cva}_{m,n} &= \begin{cases} 0, & \text{if exercised,} \\ v_{m,n} - u_{m,n}, & \text{otherwise.} \end{cases}
 \end{aligned} \quad (5.2)$$

- When the option at the  $n$ th path at time  $t_m$  is exercised, update the CVA values to zero along the  $n$ th path from  $t_{m+1}$  to  $t_M$ , i.e.  $cva_{m+1,n} = cva_{m+2,n} = \dots = cva_{M,n} = 0$ .
- At time  $t = 0$ , all paths share the same initial value  $x_0$ , hence

$$v_0 = c(x_0, t_m, t_{m+1}), \quad u_0 = g(x_0, t_m, t_{m+1}), \quad cva_0 = v_0 - u_0. \quad (5.3)$$

Following this procedure, as long as the functions  $c$  and  $g$  are known at each time point  $t_m$ , the CVA values along each path at each future time point can be computed without nested simulation. The approximation of the function  $c$  and  $g$  will be done via two numerical schemes in this paper.

We will employ two approaches of estimating the functions  $c$  and  $g$ : The COS method and the stochastic grid bundling method (SGBM). The former is based on Fourier transformation, and the latter is a combination of regression, path bundling and Monte Carlo simulation. Employing these two approaches for efficient computation of exposure profile on both real-world and risk-neutral scenarios, without sub-simulation, can be found in Feng *et al.* (2016).

Details of these two numerical schemes are presented in Appendix C. In the COS method, the approximation of the functions  $c$  and  $g$  is based on backward recursion; In the SGBM approach, the estimation is done based on a set of risk-neutral scenarios.

## 6. Numerical Results

We will study the change in the optimal exercise boundary and CVA values of Bermudan products due to the change in the early-exercise strategy. Further, we employ stress testing to analyze the role played by the WWR parameters. A numerical example of computing CVA–ES is also presented based on the DF model.

### 6.1. Impact of WWR on Bermudan products

In order to compare the proposed WWR models, we assume that the parameters of the models are calibrated to the same market prices of derivatives, as follows:

- The market indicator of the stock value is the at-the-money (ATM) implied volatility  $\sigma_x^{\text{BS}}$ .
  - For the DF model, the WWR coefficient  $b$  only affects the intensity value. The default-free value of a European option is not influenced at all. Hence, we have that  $\sigma_x^{\text{DF}} = \sigma_x^{\text{BS}}$ .
  - For the CIRJ model, jumps add extra volatility into the stock process. To have a fair comparison, we first fix other parameters, and calibrate the log-asset volatility by  $\sigma_x^{\text{CIRJ}}$  by

$$\sigma_x^{\text{CIRJ}} = \underset{\sigma_x > 0}{\operatorname{argmin}} \left( \text{BS-formula}(\sigma_x^{\text{BS}}) - V_{\text{CIRJ}}^{\text{European}}(\sigma_x) \right). \quad (6.1)$$

- The market indicator of counterparty’s credit quality is the credit spread  $\bar{h}$ . The value  $\bar{h}$  is an input parameter in both models. For the DF model, see Eq. (3.3), and for the CIRJ model, see Eq. (3.8).

In this way, we can ensure that two models can be compared under the same market condition, implied by the values of the market indicators  $\{\sigma_x^{\text{BS}}, \bar{h}\}$ . We then vary the WWR related parameters to analyze the impact of these parameters on CVA.

In this section, we consider a Bermudan option that can be exercised  $M = 10$  times prior to and upon expiration  $T$ . The other parameters are chosen as  $S_0 = 100$ ,  $K = 100$ ,  $r = 0.004$ ,  $\bar{h} = 0.1$ ,  $T = 0.5$ ,  $\sigma_x^{\text{BS}} = 0.2$ .

### Model DF

The DF model may be a basic model for the intensity, but it provides a useful insight in the connection between the underlying stock and the intensity. We choose a 90% nonnegativity intensity interval for coefficient  $b$ , which varies as  $\sigma_x$  changes. The interval is larger when the volatility is smaller.

Figure 3 presents the optimal exercise boundaries for a Bermudan put option, where we vary  $b = \{0, -0.2, 0.6\}$ . Under Policy I variations in  $b$  do not have an impact on the optimal exercise boundary, and under Policy II, when taking CCR into account, the optimal exercise boundaries of a put option increase. By comparing the optimal exercise boundary for  $b = -0.2$  with that for  $b = -0.6$ , we see that by WWR the optimal exercise boundaries are raised.

Table 1 compares the default-free, default-adjusted and CVA values using the two policies. Both the COS method and the SGBM numerical schemes are employed to ensure the accuracy of the results. Using Policy I, the default-free value is not impacted by variation in  $b$ , but the default-adjusted value gets smaller as  $b$  decreases. Hence, the CVA value increases under Policy I when WWR increases: by varying

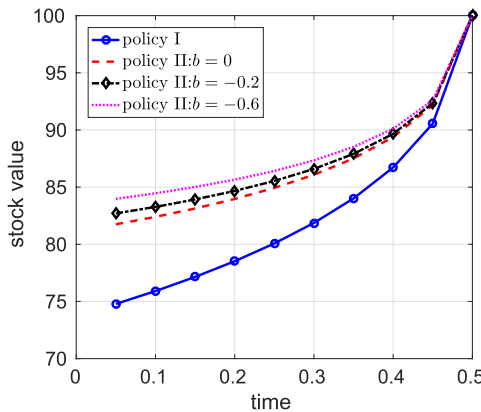


Fig. 3. Model DF: Optimal exercise boundaries for a Bermudan put using the two exercise policies with  $b = \{0, -0.2, -0.6\}$ .



Table 1. Model DF: Default-free and default-adjusted values and CVA for a Bermudan put exercised by two policies, respectively, at time 0.

DF model		Policy I			Policy II		
$b$	Method	$V^I$	$U^I$	$CVA^I$	$V^{II}$	$U^{II}$	$CVA^{II}$
0	COS	5.541	5.317	0.224	5.531	5.330	0.201
	SGBM	5.540	5.317	0.223	5.530	5.330	0.201
	std	0.002	0.001	0.001	0.002	0.001	0.000
-0.2	COS	5.541	5.292	0.248	5.527	5.313	0.214
	SGBM	5.540	5.291	0.249	5.526	5.312	0.214
	std	0.001	0.001	0.000	0.001	0.001	0.000
-0.6	COS	5.541	5.243	0.298	5.518	5.281	0.237
	SGBM	5.541	5.243	0.298	5.518	5.281	0.237
	std	0.001	0.001	0.001	0.001	0.001	0.000

$b$ , from zero to  $-0.2$  and  $-0.6$ , CVA is increased by 0.02 and 0.07, respectively. Policy II reduces the CVA value, and the impact of WWR on CVA using Policy II is smaller: By varying  $b$  from zero to  $-0.2$  and  $-0.6$ , CVA values using Policy II increase by 0.01 and 0.04, respectively.

Under the dependency structure described by the DF model, the CVA value is higher when WWR increases for Bermudan options under either policy.

### Model CIRJ

The dependency in the CIRJ model is in the jump term with different correlation  $\rho_J$  for the jump intensities. Here, the intensity process directly influences the underlying asset process in the CIRJ model. Changing parameters  $\{\bar{\lambda}, \mu_y^J, \rho_J\}$  will affect the underlying log-asset process. Hence, for each combination of selected values for  $\{\bar{\lambda}, \mu_y^J, \rho_J\}$ , we will compute the corresponding  $\sigma_x^{CIRJ}$  fixing  $\sigma_x^{BS} = 0.2$  by solving the root-finding problem (6.1), using the bisection method.

Given  $\bar{\lambda} = 5$ ,  $\mu_y^J = 0.04$ , we compare the optimal exercise boundaries when varying the jump size correlation  $\rho_J = \{0, -0.9\}$  in Fig. 4. In this two-dimensional setting, the optimal exercise boundary at each exercise date is a surface, and we choose to plot the stock value along the time horizon of this optimal exercise surface at a fixed point of the  $Y$ -grid. As discussed, varying  $\rho_J$  has an immediate effect on the log-asset process, and the optimal exercise boundaries for  $\rho_J = \{0, -0.9\}$  using Policy I are different. The exercise region (the area below the curve) always increases when  $\rho_J$  gets negative, for each policy. In the CIRJ model, when the jump in the intensity process is upward, the jump in the log-asset process tends to be downward when  $\rho_J$  is negative (notice that we choose  $\mu_x^J = 0$ ).

Results in Table 2 show that the impact of  $\rho_J$  on the CVA values of a Bermudan put is small. The CVA values even become smaller when  $\rho_J$  is negative. Under this model, the CVA of a Bermudan put in the presence of WWR need not be higher than CVA without WWR.

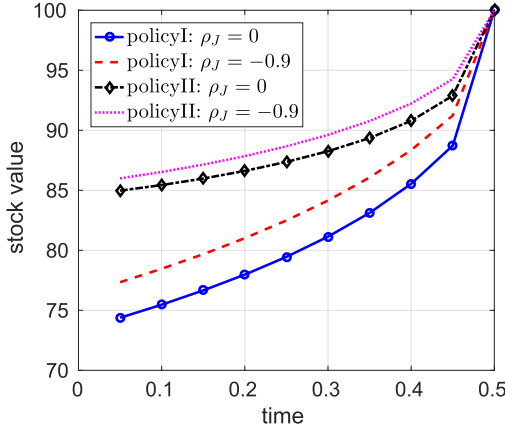


Fig. 4. Model CIRJ: Optimal exercise boundary along the time horizon for a Bermudan put,  $M = 10$ . The stock value along the time horizon, fixing the value on the  $Y$ -grid,  $\bar{w}_q = 0.0979$ . Parameters  $\bar{\lambda} = 5$ ,  $\mu_y^J = 0.04$ ,  $\sigma_x^{\text{BS}} = 0.2$ .

Table 2. Model CIRJ: Initial values of a Bermudan put. The volatility in the BS formula is  $\sigma_x^{\text{BS}} = 0.2$ ;  $\mu_y^J = 0.04$  and  $\bar{\lambda} = 5$ .

CIRJ model		Policy I			Policy II		
$\rho_J$	Method	$V^I$	$U^I$	$\text{CVA}^I$	$V^{II}$	$U^{II}$	$\text{CVA}^{II}$
0	COS	5.539	5.308	0.230	5.526	5.327	0.199
	SGBM	5.541	5.312	0.229	5.526	5.335	0.191
	std	0.002	0.002	0.001	0.002	0.002	0.001
-0.9	COS	5.544	5.322	0.222	5.530	5.346	0.184
	SGBM	5.551	5.335	0.217	5.537	5.361	0.177
	std	0.001	0.002	0.001	0.001	0.002	0.002

## 6.2. Stress testing

The effect of WWR on options can be analyzed by a CVA stress test, for example, by varying the dependency coefficients. In this section, we use the European option instead of the Bermudan products to study the impact of a varying dependency structure.

As mentioned, the CVA is governed by three components: the credit quality, the underlying asset and their dependency structure, i.e. by WWR or RWR. In order to isolate the impact of WWR (RWR) from the other components, we define a *WWR ratio*:  $\frac{\text{CVA}_W}{\text{CVA}_I}$ , where only the correlation/dependency coefficients (such as the coefficients  $b$ ,  $\rho_J$  or  $\bar{\rho}$ ) are varied, while keeping other parameters fixed.  $\text{CVA}_W$  is the CVA value when the intensity and the underlying asset are “wrongly” correlated in the worst case, and  $\text{CVA}_I$  is the CVA value when the underlying asset and intensity are independent.

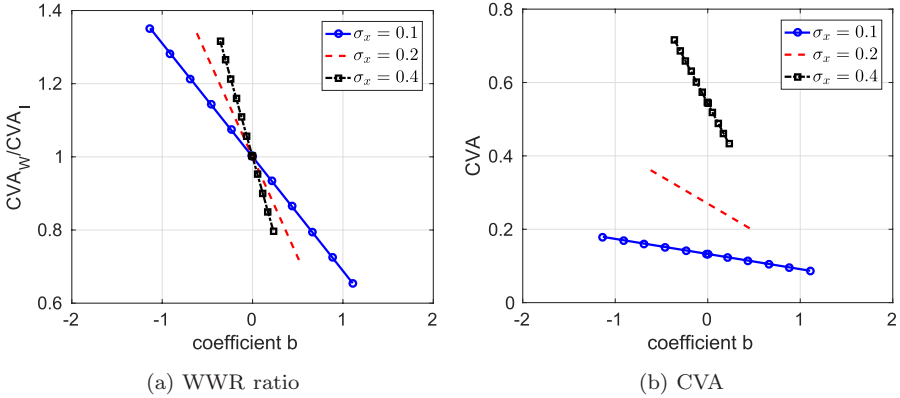


Fig. 5. Model DF: CVA test for a European put varying coefficient,  $b \in [-1.134, 1.108]$  when  $\sigma_x = 0.1$ ,  $b \in [-0.617, 0.514]$  when  $\sigma_x = 0.2$ ,  $b \in [-0.358, 0.231]$  when  $\sigma_x = 0.4$ . The blue line represents  $\sigma_x = 0.1$ , the red-dashed line  $\sigma_x = 0.2$  and the black-squares line  $\sigma_x = 0.4$ .

### Model DF

We present the WWR ratio in plot (a) and the CVA value in plot (b) of Fig. 5. CVA increases as the value of coefficient  $b$  gets smaller.

Figure 5(a) shows that, given a fixed value of coefficient  $b$ , the impact of WWR on the CVA value of the associated option is larger for higher  $\sigma_x$ . From Eq. (3.3), it is clear that the volatility of the intensity is related to  $|b|\sigma_x$ . The intensity model's volatility thus plays an important role in the WWR. Figure 5(b) shows that  $\sigma_x$  is the determining factor for the CVA value. In the DF model, the volatility of the asset does not only impact the absolute value of CVA, but the strength of the WWR as well.

### Model CIRJ

Figure 6 presents the WWR and CVA values for a European put by varying the parameters. When  $\rho_J = 0$ , variations in the other parameters do not have any impact on the WWR or CVA values; in Fig. 6(a), it is shown that with  $\rho_J \neq 0$ , the WWR increases when the jump parameters  $\mu_y^J$  and  $\bar{\lambda}_J$  increase and  $\sigma_x^{BS}$  decreases. It appears to be the contribution of the correlated jump terms to the total variance of the log-asset process that determines the WWR. Figure 6 shows that  $\sigma_x^{BS}$ , which represents the level of the total variance in the log-asset process, is an important factor for the CVA values.

### 6.3. Calculation of CVA-ES

As an example, we compute the CVA-ES for a Bermudan put under the DF model via the nonnested Monte Carlo procedure. It is well-known that the real-world drift

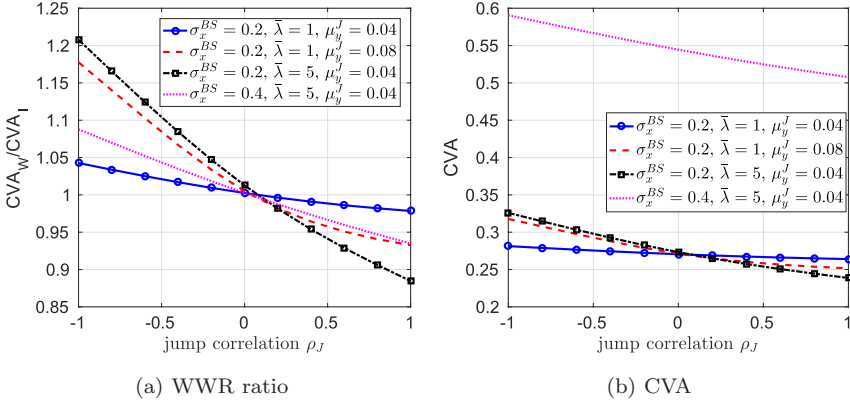


Fig. 6. Model CIRJ: CVA test for a European put varying  $\{\sigma_x^{BS}, \bar{\lambda}, \mu_y^J, \rho_J\}$ .  $S_0 = K = 100, T = 0.5, r = 0.004, \bar{h} = 0.1, Y_0 = 0.02, \gamma = 5$ , see Eq. (3.6),  $\theta = 0.02, \sigma_y = 0.1, \mu_x^J = 0, \sigma_x^J = 0.05$ .

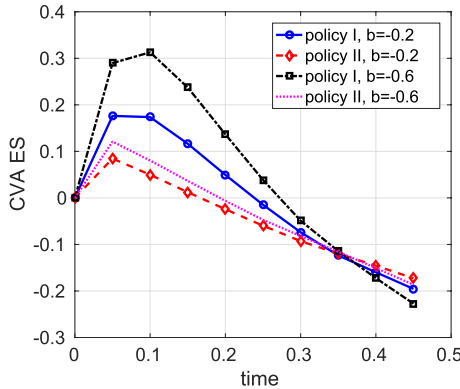


Fig. 7. The CVA-ES for a Bermudan put under the DF model;  $S_0 = 100, K = 100, r = 0.004, \sigma_x^Q = 0.2$ , real-world log-return  $\mu^P = 0.006$ , real-world volatility  $\sigma_x^P = 0.2, \bar{h} = 0.1, T = 0.5$ , and coefficient  $b = \{-0.2, -0.6\}$ .

and the risk-neutral drift are different. In this example, we show an example by setting that the real-world log-return  $\mu^P = 0.006 > r = 0.004$ .

Figure 7 shows the CVA-ES from time zero until the option's expiration by the COS method. We compare the CVA-ES values for  $b = \{-0.2, -0.6\}$  using the two policies. It shows that WWR has a significant impact on the CVA-ES values, as compared to the time-zero CVA values. By Policy II, the CVA-ES values decrease significantly.

## 7. Conclusion

We studied the effect of WWR on the European and Bermudan options by means of two dependency structures. The results showed that the effect of WWR for

European and Bermudan options are different. Different exercise policies can be followed when CCR is present, which may impact the corresponding CVA values of Bermudan options. The way to define the dependency structure also played a role regarding the WWR. The two types of dependency structure we considered, gave different results for the optimal exercise boundaries of the Bermudan options and the corresponding option values. Increasing WWR always increased CVA in the case of European options, but for some dependency structures, the CVA of a Bermudan option may decrease when WWR incurs. In addition to the correlation coefficients, the volatility of the intensity plays an important role regarding the WWR. When computing the CVA as the difference between the default-free and default-adjusted values, the CVA-ES can be computed without sub-simulations by means of the developed algorithms.

## Appendix A. AJD Class

The AJD class is defined by the following SDE system, see Duffie *et al.* (2000):

$$d\mathbf{X}_t = \mu(\mathbf{X}_t)dt + \sigma(\mathbf{X}_t)d\mathbf{W}_t + d\mathbf{J}_t, \quad (\text{A.1})$$

where  $\mathbf{W}_t$  is a standard Brownian motion in  $\mathbb{R}^d$ ;  $\mu : \mathcal{D} \rightarrow \mathbb{R}^d$ ,  $\sigma : \mathcal{D} \rightarrow \mathbb{R}^{d \times d}$  and  $\mathbf{J}_t \in \mathbb{R}^d$  is a pure jump process independent of  $\mathbf{W}_t$ , whose jump sizes are governed by a probability distribution  $\nu$  on  $\mathbb{R}^d$ . More specifically, the jump times of  $\mathbf{J}_t$  are the jump times of a Poisson process with jump-arrival rate  $\bar{\lambda}$ , and the  $\mathbf{J}_t$  jump sizes are governed by probability  $\nu$ . For complex-valued numbers  $\mathbf{c} \in \mathbb{C}^d$ , we define  $\Theta(\mathbf{c}) = \int_{\mathbb{R}^d} \exp(\mathbf{c} \cdot \mathbf{z}) d\nu(\mathbf{z})$ , where the integral should be well-defined. This *jump transform function*  $\Theta$  determines the jump-size distribution, see Duffie *et al.* (2000). We can choose jump distribution  $\nu$  so that the corresponding jump transform  $\Theta$  is known explicitly.

Given a fixed affine discount rate  $\delta : \mathcal{D} \rightarrow \mathbb{R}$ , the affinity is satisfied when the coefficients are defined by

$$\begin{aligned} \mu(\mathbf{X}_t) &= K_0 + K_1 \mathbf{X}_t, & K_0 \in \mathbb{R}^d, K_1 \in \mathbb{R}^{d \times d}, \\ (\sigma(\mathbf{X}_t)\sigma(\mathbf{X}_t)^T)_{i,j} &= (H_0) + (H_1)_{i,j} \mathbf{X}_t, & H_0 \in \mathbb{R}^{d \times d}, H_1 \in \mathbb{R}^{d \times d \times d} \\ \delta(\mathbf{X}_t) &= r_0 + r_1 \mathbf{X}_t, & r_0 \in \mathbb{R}, r_1 \in \mathbb{R}^d. \end{aligned} \quad (\text{A.2})$$

The discounted (at rate  $\delta$ ) ChF is defined by

$$\phi_\delta(\mathbf{u}, t, T, \mathbf{X}_t) = \mathbb{E}^\mathbb{Q} \left[ \exp \left( - \int_t^T \delta(\mathbf{X}_s) ds \right) \exp(i\mathbf{u}\mathbf{X}_T) \middle| \mathbf{X}_t \right]. \quad (\text{A.3})$$

With the short rate, i.e.  $\delta(\mathbf{x}) = r$ , we denote the discounted ChF by  $\phi_r$ ; when discounted at the intensity-adjusted rate,  $\delta(\mathbf{x}) = R(\mathbf{x})$ , we denote the intensity-adjusted discounted ChF by  $\phi_R$ . Duffie *et al.* (2000) showed that the discounted ChF for process  $\mathbf{X}_t$  is given by the following formula:

$$\phi_\delta(\mathbf{u}, t, T, \mathbf{X}_t) = \exp(\alpha(T - t) + \mathbf{X}_t \boldsymbol{\beta}(T - t)), \quad (\text{A.4})$$

where the coefficients  $\alpha$  and  $\beta$  satisfy the following ODEs:

$$\begin{aligned}\frac{d\beta(s)}{ds} &= -r_1 + K_1^T \beta(s) + \frac{1}{2} \beta(s)^T H_1 \beta(s), \\ \frac{d\alpha(s)}{ds} &= -r_0 + K_0 \beta(s) + \frac{1}{2} \beta(s)^T H_0 \beta(s) + \bar{\lambda}[\Theta(\beta(s)) - 1],\end{aligned}\tag{A.5}$$

with  $s = T - t$  and initial conditions  $\beta(0) = i\mathbf{u}$  and  $\alpha(0) = 0$ .

The discounted ChF is convenient for computation. In the COS method (see Appendix C), for example, we use the discounted ChF to recover the corresponding discounted density function; while in the SGBM, we need the discounted ChF to derive the discounted moments. We will present the details in Appendix C.

### The DF model

The default-adjusted discounted ChF of the DF model is given by

$$\phi_R(u, t, T, X_t) = \exp\left(-\int_t^T a(x)dx + \alpha^R(s) + \beta^R(s)X_t\right),\tag{A.6}$$

with  $s = T - t$ , and the coefficients given by

$$\begin{aligned}\alpha^R(s) &= \left[-r + iu\left(r - \frac{1}{2}\sigma_x^2\right) - \frac{1}{2}\sigma_x^2 u^2\right] s \\ &\quad \times \frac{b}{2} \left[iu\sigma_x^2 - \left(r - \frac{1}{2}\sigma_x^2\right)\right] s^2 + \frac{1}{6}\sigma_x^2 s^3, \\ \beta^R(s) &= iu - bs,\end{aligned}\tag{A.7}$$

The corresponding short rate discounted ChF of the DF model is given by

$$\phi_r(X_t, t, T) = \exp(\alpha^r(s) + \beta^r(s)X_t),\tag{A.8}$$

where the coefficients are given by

$$\alpha^r(s) = \left[-r + iu\left(r - \frac{1}{2}\sigma_x^2\right) - \frac{1}{2}\sigma_x^2 u^2\right] s, \quad \beta^r(s) = iu.\tag{A.9}$$

### The CIRJ model

The default-adjusted discounted ChF of the CIRJ model is given by

$$\phi_R\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, t, T, \begin{bmatrix} X_t \\ Y_t \end{bmatrix}\right) = \exp\left(-\int_t^T \psi^{\text{CIR}}(x)dx + \alpha^R(s) + \beta_1^R(s)X_t + \beta_2^R(s)Y_t\right),\tag{A.10}$$

with  $s = T - t$  and where the coefficients are given by

$$\begin{aligned}\alpha^R(s) &= \left[-r - \bar{\lambda} + iu_1\left(r - \frac{1}{2}\sigma_x^2 - \bar{\lambda}\bar{\mu}\right) - \frac{1}{2}\sigma_x^2 u_1^2\right] s \\ &\quad + \frac{\gamma\theta}{\sigma_y^2} \left[(\gamma + d_1)s - 2 \log\left(\frac{d_2 e^{d_1 s} - 1}{d_2 - 1}\right)\right]\end{aligned}$$

$$\begin{aligned}
 & + \bar{\lambda} \bar{G} \left[ \frac{2\sigma_y^2 \mu_y^J}{g_1 g_2} \log \left( \frac{g_1 + d_2 g_2 e^{d_1 s}}{g_1 + d_2 g_2} \right) - \frac{\sigma_y^2 s}{g_1} \right], \\
 \beta_1^R(s) & = iu_1, \quad \beta_2^R(s) = \frac{\gamma}{\sigma_y^2} + \frac{d_1}{\sigma_y^2} \frac{1 + d_2 e^{d_1 s}}{1 - d_2 e^{d_1 s}},
 \end{aligned} \tag{A.11}$$

and

$$\begin{aligned}
 d_1 & = \sqrt{\gamma^2 + 2\sigma_y^2}, \quad d_2 = \frac{i u_2 \sigma_y^2 - \gamma - d_1}{i u_2 \sigma_y^2 - \gamma + d_1}, \\
 \bar{G} & = \exp \left( i u_1 \mu_x^J - \frac{1}{2} (u_1 \sigma_x^J)^2 \right), \\
 g_1 & = -(1 - i u_1 \rho_J \mu_y^J) \sigma_y^2 + (d_1 + \gamma) \mu_y^J, \\
 g_2 & = (1 - i u_1 \rho_J \mu_y^J) \sigma_y^2 + (d_1 - \gamma) \mu_y^J.
 \end{aligned} \tag{A.12}$$

The short rate discounted ChF of the CIRJ model is given by

$$\phi_r \left( \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, t, T, \begin{bmatrix} X_t \\ Y_t \end{bmatrix} \right) = \exp(\alpha^r(s) + \beta_1^r(s) X_t + \beta_2^r(s) Y_t), \tag{A.13}$$

where the coefficients are given by

$$\begin{aligned}
 \alpha^r(s) & = \left[ -r - \bar{\lambda} + i u_1 \left( r - \frac{1}{2} \sigma_x^2 - \bar{\lambda} \bar{\mu} \right) - \frac{1}{2} \sigma_x^2 u_1^2 \right] s \\
 & + \frac{2\gamma\theta}{\sigma_y^2} \left[ \gamma s - \log \left( \frac{\tilde{c}_0 e^{\gamma s} - i u_2 \sigma_y^2}{\tilde{c}_0 - i u_2 \sigma_y^2} \right) \right] \\
 & + \bar{\lambda} \bar{G} \left[ \frac{\sigma_y^2 s}{\tilde{c}_1} + \log \left( \frac{i u_2 \tilde{c}_1 - \tilde{c}_0 \tilde{c}_2 e^{\gamma s}}{i u_2 \tilde{c}_1 - \tilde{c}_0 \tilde{c}_2} \right) \cdot \frac{(\tilde{c}_1 - \tilde{c}_2) \sigma_y^2}{\gamma \tilde{c}_1 \tilde{c}_2} \right], \\
 \beta_1^r(s) & = iu_1, \quad \beta_2^r(s) = \frac{2i u_2 \gamma}{i u_2 \sigma_y^2 + e^{\gamma s} (-i u_2 \sigma_y^2 + 2\gamma)},
 \end{aligned} \tag{A.14}$$

where

$$\tilde{c}_1 = 1 - i u_1 \rho_J \mu_y^J - 2\gamma \mu_y^J, \quad \tilde{c}_2 = 1 - i u_1 \rho_J \mu_y^J, \quad \tilde{c}_0 = i u_2 \sigma_y^2 - 2\gamma. \tag{A.15}$$

## Appendix B. The Range of the Coefficient $b$

The intensity  $h_t$  can be regarded as a function of coefficient  $b$ , given by

$$h_t = \left( \frac{1}{2} \sigma_x^2 t^2 \right) b^2 + (\sigma_x \sqrt{t} Z) b + \bar{h}, \tag{B.1}$$

where  $Z \sim \mathcal{N}(0, 1)$  and  $t \in [0, T]$ .

The value of random variable  $Z$  is between  $\Phi^{-1}(\zeta)$  and  $\Phi^{-1}(1 - \zeta)$  at a probability  $(1 - 2\zeta)$ , where  $\zeta < 0.5$ . With  $z_0 := \Phi^{-1}(1 - \zeta) > 0$ , interval  $[-z_0, z_0]$  is the confidence interval of  $Z$  at level  $((1 - 2\zeta))$ .

By fixing  $Z$  and  $t$ , the expression in (B.1) becomes a quadratic function in  $b$ , and its roots are given by the solution to the function,

$$\Delta^\pm(Z, t) = \sigma_x^{-1} t^{-\frac{3}{2}} (-Z^2 \pm \sqrt{Z^2 - 2t\bar{h}}). \quad (\text{B.2})$$

The two roots are determined by  $Z$  and  $t$ ; when fixing  $t$ , we find,

- When  $Z^2 < 2t\bar{h}$ , i.e. when  $Z \in (-\sqrt{2t\bar{h}}, \sqrt{2t\bar{h}})$ , the expression in (B.1) has two complex-valued roots and intensity  $h$  is thus always positive;
- When  $Z^2 \geq 2t\bar{h}$ , the expression in (B.1) has two real-valued roots; it is positive when  $b < \Delta^-(Z, t)$  or  $b > \Delta^+(Z, t)$ .

We look for values of  $b$  so that  $h_t$  is nonnegative for  $Z \in [-z_0, z_0]$ , where  $z_0 > \sqrt{2t\bar{h}} > 0$ . The derivatives of the two roots,

$$\frac{\partial \Delta^\pm(Z, t)}{\partial Z} = \sigma_x^{-1} t^{-\frac{3}{2}} \left( -1 \pm \frac{Z}{\sqrt{Z^2 - 2t\bar{h}}} \right), \quad (\text{B.3})$$

where  $|\frac{Z}{\sqrt{Z^2 - 2t\bar{h}}}| < 1$ , and hence we have,

- When  $Z \in [-z_0, -\sqrt{2t\bar{h}}]$ ,

$$\frac{\partial \Delta^-(Z, t)}{\partial Z} > 0, \quad \frac{\partial \Delta^+(Z, t)}{\partial Z} < 0, \quad (\text{B.4})$$

from which we see that, on the interval  $Z \in [-z_0, -\sqrt{2t\bar{h}}]$ , the root  $\Delta^-$  is monotonically increasing and the root  $\Delta^+$  is monotonically decreasing, and  $h_t > 0$  for  $Z \in [-z_0, -\sqrt{2t\bar{h}}]$  when the value of the coefficient  $b$  satisfies

$$b \in (-\infty, \Delta^-(-z_0, t)) \cup (\Delta^+(-\sqrt{2t\bar{h}}, t), +\infty). \quad (\text{B.5})$$

- When  $Z \in [\sqrt{2t\bar{h}}, z_0]$ , we have

$$\frac{\partial \Delta^-(Z, t)}{\partial Z} < 0, \quad \frac{\partial \Delta^+(Z, t)}{\partial Z} > 0, \quad (\text{B.6})$$

so, on the interval  $Z \in [\sqrt{2t\bar{h}}, z_0]$ ,  $\Delta^-$  is monotonically decreasing and  $\Delta^+$  is monotonically increasing.

Then,  $h_t > 0$  for  $Z \in [\sqrt{2t\bar{h}}, z_0]$  when the value of  $b$  satisfies

$$b \in (-\infty, \Delta^-(\sqrt{2t\bar{h}}, t)] \cup (\Delta^+(z_0, t), +\infty). \quad (\text{B.7})$$

The intersection of Set (B.5) and Set (B.7), gives us the set of  $b$ -values for which  $h_t > 0$  for  $Z \in [-z_0, z_0]$ , given by

$$b \in (\Delta^+(z_0, t), \Delta^-(-z_0, t)). \quad (\text{B.8})$$

Next, we will show that the set of the  $b$ -value that satisfies  $h_t > 0$ , for  $Z \in [-z_0, z_0]$  over the time horizon  $t \in [0, T]$ , is given by

$$b \in (\Delta^+(z_0, T), \Delta^-(-z_0, T)), \quad (\text{B.9})$$

when  $z_0 > \frac{8}{3}\bar{h}$ .



Fixing  $Z$ , and taking the derivative of the two roots with respect to the time  $t$ , gives

$$\frac{\partial \Delta^\pm(Z, t)}{\partial t} = \frac{\sigma_x^{-1} t^{-\frac{5}{2}}}{2\sqrt{Z^2 - 2t\bar{h}}} (Z \pm \sqrt{Z^2 - 2t\bar{h}})(Z \pm 2\sqrt{Z^2 - 2t\bar{h}}). \quad (\text{B.10})$$

From (B.10),  $\frac{\partial \Delta^-(-z_0, t)}{\partial t} < 0$  and  $\frac{\partial \Delta^+(z_0, t)}{\partial t} > 0$  when  $z_0 > \sqrt{\frac{8}{3}t\bar{h}}$ . We can choose  $\zeta \leq \Phi(-\frac{8}{3}t\bar{h})$  to satisfy this condition. Then,  $\Delta^+(z_0, t)$  is increasing monotonically with respect to time  $t$ , and  $\Delta^-(-z_0, t)$  is decreasing monotonically with respect to time  $t$ . Hence, the bounded interval for the coefficient  $b$ , given by (B.8), satisfies  $h_t > 0$  for the time period  $[0, T]$  at a probability  $(1 - 2\zeta)$ .

### Appendix C. Numerical Schemes

In this appendix, we give details of the COS method and the SGBM for the default-adjusted option valuation. First of all, the algorithm for the one-dimensional problem for the DF model can be developed by using the one-dimensional Fourier-cosine expansion introduced in Fang & Oosterlee (2009) to recover the needed discounted probability density function (PDF).

#### The COS method

The discount term,  $\Gamma = \exp(-\int_t^T \delta(\mathbf{X}_s) ds) \in \mathbb{R}$ , is an  $\mathcal{F}_T$ -measurable random variable. Let  $p(\mathbf{y}, \Gamma | \mathbf{X}_t)$  represent the joint PDF for  $\mathbf{X}_T$  and  $\Gamma$ , conditional on  $\mathbf{X}_t$ . Define the PDF discounted at  $\delta$  by

$$\hat{p}_\delta(\mathbf{y}, t, T | \mathbf{X}_t) := \int_{\mathbb{R}} \Gamma \cdot p(\mathbf{y}, \Gamma | \mathbf{X}_t) d\Gamma. \quad (\text{C.1})$$

The discounted PDF  $\hat{p}_\delta$  is computed by the inverse Fourier transform of the discounted ChF. Hence, it can be recovered by a Fourier cosine expansion on a specific interval as in Fang & Oosterlee (2009), Ruijter & Oosterlee (2012). The formulas for the discounted PDF in a two-dimensional setting according to Ruijter & Oosterlee (2012) are given by

$$\hat{p}_\delta(\mathbf{y}, t, T | \mathbf{X}_t) \approx \sum_{k_1=0}^{N_1} \sum_{k_2=0}^{N_2} \mathcal{P}_{k_1, k_2}^\delta(\mathbf{X}_t, t, T) \mathcal{A}_{k_1, k_2}(\mathbf{y}), \quad (\text{C.2})$$

where the symbol  $\sum'$  indicates that the first term of the summation is weighted by  $\frac{1}{2}$ ;  $\mathbf{y} = [y_1, y_2]^T$ ,  $\mathbf{X}_t = [x_1, x_2]^T$ , and the coefficients are given by

$$\begin{aligned} \mathcal{A}_{k_1, k_2}(\mathbf{y}) &= \cos\left(k_1 \pi \frac{y_1 - a_1}{b_1 - a_1}\right) \cos\left(k_2 \pi \frac{y_2 - a_2}{b_2 - a_2}\right), \\ \mathcal{P}_{k_1, k_2}^\delta(\mathbf{X}_t, t, T) &= \frac{1}{2} \left( \chi_{k_1, k_2}^+(\mathbf{X}_t, t, T) + \chi_{k_1, k_2}^-(\mathbf{X}_t, t, T) \right), \end{aligned}$$

$$\begin{aligned} \chi_{k_1, k_2}^{\pm}(\mathbf{X}_t, t, T) &= \frac{2}{b_1 - a_1} \frac{2}{b_2 - a_2} \operatorname{Re} \left\{ \phi_{\delta} \left( \left[ \begin{array}{c} \frac{k_1 \pi}{b_1 - a_1} \\ \pm \frac{k_2 \pi}{b_2 - a_2} \end{array} \right], t, T, \mathbf{X}_t \right) \right. \\ &\quad \left. \times \exp \left( -\frac{ik_1 a_1 \pi}{b_1 - a_1} \mp \frac{ik_2 a_2 \pi}{b_2 - a_2} \right) \right\}, \end{aligned} \tag{C.3}$$

with  $k_1 = 0, \dots, N_1 - 1$ ,  $k_2 = 0, \dots, N_2 - 1$ , and  $\operatorname{Re}$  represents the real part of the argument; the truncated interval  $[a_1, b_1] \times [a_2, b_2]$  is given by

$$\begin{aligned} a_1 &= \mathbb{E}^{\mathbb{Q}}[X_T] - L_1 \sqrt{\operatorname{var}(X_T)}, & b_1 &= \mathbb{E}^{\mathbb{Q}}[X_T] + L_1 \sqrt{\operatorname{var}(X_T)}, \\ a_2 &= \mathbb{E}^{\mathbb{Q}}[Y_T] - L_2 \sqrt{\operatorname{var}(Y_T)}, & b_2 &= \mathbb{E}^{\mathbb{Q}}[Y_T] + L_2 \sqrt{\operatorname{var}(Y_T)}, \end{aligned} \tag{C.4}$$

where  $L_1$  and  $L_2$  are constants that define the length of the interval.

Notice that we require two discounted ChFs, one discounted by the short rate, another by the intensity-adjusted rate. We use the notation  $\mathcal{P}_{k_1, k_2}^r$  for the coefficients of the PDF discounted at the short rate and  $\mathcal{P}_{k_1, k_2}^R$  when discounting at the intensity-adjusted rate.

We present the backward recursion steps when pricing default-adjusted and default-free Bermudan puts, by either using Policy I or II. The difference is in the way optimal exercise boundaries are determined. We use index  $j = \{\text{Policy I, II}\}$  to refer to the corresponding values under these two policies.

At an exercise time  $t_m < T$ , the expected default-adjusted value of continuing the option is given, respectively, by

$$\begin{aligned} c^j(\mathbf{X}_m, t_m, t_{m+1}) &\approx \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \mathcal{P}_{k_1, k_2}^r(\mathbf{X}_m, t_m, t_{m+1}) \mathcal{V}_{k_1, k_2}^j(t_{m+1}), \\ g^j(\mathbf{X}_m, t_m, t_{m+1}) &\approx \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \mathcal{P}_{k_1, k_2}^R(\mathbf{X}_m, t_m, t_{m+1}) \mathcal{U}_{k_1, k_2}^j(t_{m+1}), \end{aligned} \tag{C.5}$$

where the coefficients read

$$\begin{aligned} \mathcal{V}_{k_1, k_2}^j(t_{m+1}) &\approx \int_{a_1}^{b_1} \int_{a_2}^{b_2} V^j(\mathbf{y}, t_{m+1}) \mathcal{A}_{k_1, k_2}(\mathbf{y}) d\mathbf{y}, \\ \mathcal{U}_{k_1, k_2}^j(t_{m+1}) &\approx \int_{a_1}^{b_1} \int_{a_2}^{b_2} U^j(\mathbf{y}, t_{m+1}) \mathcal{A}_{k_1, k_2}(\mathbf{y}) d\mathbf{y}. \end{aligned} \tag{C.6}$$

At time  $t_M = T$ ,  $V^j = U^j = f(\cdot)$ ; for a put option, the immediate exercise payoff is larger than zero when the stock price at time  $T$  is lower than strike  $K$ . We

have

$$\mathcal{V}_{k_1, k_2}^j(t_M) = \mathcal{U}_{k_1, k_2}^j(t_M) = \int_{a_1}^{\log(K)} f(y_1) \int_{a_2}^{b_2} \mathcal{A}_{k_1, k_2}(\mathbf{y}) dy_2 dy_1. \quad (\text{C.7})$$

The optimal exercise boundary at each exercise date is represented by a two-dimensional surface. We define a grid along the intensity direction, denoted by  $\{a_2 = w_0 < w_2 < \dots < w_Q = b_2\}$ , with  $Q$  grid points. On each interval  $[w_{q-1}, w_q]$ ,  $q = 1, \dots, Q$ , we approximate the optimal exercise log-stock values by a constant  $x_{m,q}^j$ , using the following expression:

$$\begin{aligned} \text{Policy I: } f(x_{m,q}^I) - c^j \left( \begin{bmatrix} x_{m,q}^{\text{II}} \\ \bar{w}_q \end{bmatrix}, t_m, t_{m+1} \right) &= 0, \\ \text{Policy II: } f(x_{m,q}^{\text{II}}) - g^j \left( \begin{bmatrix} x_{m,q}^{\text{II}} \\ \bar{w}_q \end{bmatrix}, t_m, t_{m+1} \right) &= 0, \end{aligned} \quad (\text{C.8})$$

where  $\bar{w}_q = \frac{1}{2}(w_{q-1} + w_q)$ .

For a put, the exercise region  $\mathcal{G}_q^j(t_m, j) = [a_1, x_{m,q}^j] \times [w_q, w_{q+1}]$ , and the continuation region  $\mathcal{C}_q^j(t_m) = [x_{m,q}^j, b_1] \times [w_q, w_{q+1}]$  for each  $q = 1, \dots, Q$ . Hence, the coefficients at time  $t_m$  are given by

$$\begin{aligned} \mathcal{V}_{k_1, k_2}^j(t_m) &= \sum_{q=1}^Q \left( \iint_{\mathcal{C}_q^j(t_m)} c^j(\mathbf{y}, t_m, t_{m+1}) \mathcal{A}_{k_1, k_2}(\mathbf{y}) d\mathbf{y} \right. \\ &\quad \left. + \iint_{\mathcal{G}_q^j(t_m)} f(y_1) \mathcal{A}_{k_1, k_2}(\mathbf{y}) d\mathbf{y} \right) \\ \mathcal{U}_{k_1, k_2}^j(t_m) &= \sum_{q=1}^Q \left( \iint_{\mathcal{C}_q^j(t_m)} g^j(\mathbf{y}, t_m, t_{m+1}) \mathcal{A}_{k_1, k_2}(\mathbf{y}) d\mathbf{y} \right. \\ &\quad \left. + \iint_{\mathcal{G}_q^j(t_m)} f(y_1) \mathcal{A}_{k_1, k_2}(\mathbf{y}) d\mathbf{y} \right), \end{aligned} \quad (\text{C.9})$$

where  $c^j$  and  $g^j$  is given by (C.5).

The computation of summation in (C.9) can be enhanced by using Fast Fourier Transform (FFT) techniques, see Fang & Oosterlee (2009), Ruijter & Oosterlee (2012).

The COS method can also be employed when the ChF function can be approximated. Borovykh *et al.* (2016) provide a pricing method for Bermudan options based on an analytic approximation of the ChF using the COS method under local Lévy models with default.

## SGBM

The SGBM was developed for pricing multi-dimensional vanilla Bermudan options in Jain & Oosterlee (2015). SGBM can, however, be generalized to pricing

default-adjusted Bermudan options in a straightforward way. We give details of SGBM in the two-dimensional case, for the CIRJ model.

The default-free option values are dependent only on the underlying log-stock  $X = \log(S)$  and the path-dependent early-exercise policy, whereas the default-adjusted values are also impacted by the intensity factor  $Y$ . Within the regression step of SGBM, we choose different basis functions for computing default-free and default-adjusted values. For example, we use the polynomial basis  $\{1, X, X^2\}$  for the default-free case, and the polynomial basis  $\{1, X, Y, X^2, XY, Y^2\}$  for the default-adjusted case. Generally, we choose the set of monomials of the variables that matter for the basis functions. The number of monomials of degree  $\tilde{p}$  in  $d$  variables is given by  $\frac{(\tilde{p}+d-1)!}{(d-1)!\tilde{p}!}$ .

Let  $\left\{ \left[ \begin{matrix} \hat{X}_{1,l} \\ \hat{Y}_{1,l} \end{matrix} \right], \dots, \left[ \begin{matrix} \hat{X}_{M,l} \\ \hat{Y}_{M,l} \end{matrix} \right] \right\}_{l=1}^N$  be  $N$  generated scenarios, where the underlying evolves with the chosen risk-neutral models. We will perform the valuation by backward in time recursion, from time  $t_M = T$  to  $t_0 = 0$ . For notational convenience, we do not specify the policy  $j = \{I, II\}$  of the values.

At the final exercise opportunity,  $t_M$ , the option is either expiring or exercised, and the probability of default happening exactly at time  $t_M = T$  is zero. At time  $t_M$ , the default-free and default-adjusted option values are equal i.e.  $v_{M,l} = f(\hat{X}_{M,l})$  and  $u_{M,l} = f(\hat{X}_{M,l})$ .

At an exercise time  $t_m < T$ , prior to expiry, we define  $J$  bundles for bundling the Monte Carlo paths, which we denote by  $\{\mathcal{B}_{m,q}\}_{q=1}^J$ . We subdivide the paths into bundles based on cross-sectional samples of the state variable, so that paths in the same bundle have very similar realized values at time  $t_m$ . For  $q = 1, \dots, J$ , on each Monte Carlo path within the bundle  $\mathcal{B}_{m,q}$ , we approximate the default-free function  $V$  and default-adjusted function  $U$  at time  $t_{m+1}$ , respectively, by a linear combination of the basis functions.

$$\begin{aligned}
 v_{m+1,l} &= V(\hat{X}_{m+1,l}, t_{m+1}) \approx \sum_{0 \leq k \leq \tilde{p}} \eta_k^r(t_m, q) (\hat{X}_{m+1,l})^k, \\
 u_{m+1,l} &= U \left( \begin{bmatrix} \hat{X}_{m+1,l} \\ \hat{Y}_{m+1,l} \end{bmatrix}, t_{m+1} \right) \\
 &\approx \sum_{0 \leq k_1 + k_2 \leq \tilde{p}} \eta_{k_1, k_2}^R(t_m, q) (\hat{X}_{m+1,l})^{k_1} (\hat{Y}_{m+1,l})^{k_2},
 \end{aligned} \tag{C.10}$$

where  $\tilde{p}$  is the degree of the monomials,  $v_{m+1,l}$  and  $u_{m+1,l}$  are the realized default-free and default-adjusted values, respectively, on the  $l$ th path at time  $t_m$ , where  $l \in \mathcal{B}_{m,q}$ , and the coefficients  $\eta_k^r$  and  $\eta_k^R$  are obtained by minimizing least-squares summation errors.

With the obtained coefficients, the default-free or default-adjusted continuation values defined in Equation (4.9) for paths within the bundle  $\mathcal{B}_{m,q}$  at time  $t_m$  can

be approximated by

$$\begin{aligned}\hat{c}_{m,l} &\approx \sum_{k=1}^{B_1} \eta_k^r(t_m, q) \varphi_k^r \left( \begin{bmatrix} \hat{X}_{m,l} \\ \hat{Y}_{m,l} \end{bmatrix}, t_m, t_{m+1} \right), \\ \hat{g}_{m,l} &\approx \sum_{k=1}^{B_2} \eta_{k_1, k_2}^R(t_m, q) \varphi_{k_1, k_2}^R \left( \begin{bmatrix} \hat{X}_{m,l} \\ \hat{Y}_{m,l} \end{bmatrix}, t_m, t_{m+1} \right),\end{aligned}\tag{C.11}$$

where the functions  $\varphi^r$  and  $\varphi^R$  are defined by

$$\begin{aligned}\varphi_k^r(\mathbf{X}_m, t_m, t_{m+1}) &= B(t_m) \mathbb{E}^{\mathbb{Q}} \left[ \frac{X_{m+1}^k}{B(t_{m+1})} \middle| \mathbf{X}_m \right], \\ \varphi_{k_1, k_2}^R(\mathbf{X}_m, t_m, t_{m+1}) &= D(t_m) \mathbb{E}^{\mathbb{Q}} \left[ \frac{X_{m+1}^{k_1} Y_{m+1}^{k_2}}{D(t_{m+1})} \middle| \mathbf{X}_m \right],\end{aligned}\tag{C.12}$$

which represent the conditional expectations discounted at the corresponding rate, conditional on  $\mathbf{X}_m$ .

We refer to  $\varphi^r$  and  $\varphi^R$  as the *discounted moments*. It is straightforward to derive an analytic formula for the discounted moments from the corresponding discounted ChF. By regressing over each bundle  $\mathcal{B}_{m,q}$ , the expected default-free and default-adjusted values of continuing the option on each path are approximated.

The early-exercise decision at each path is determined by

- Policy I:

$$\begin{aligned}v_{m,l} &= \max(f(\hat{X}_{m,l}), \hat{c}_{m,l}), \\ u_{m,l} &= \begin{cases} f(\hat{X}_{m,l}), & f(\hat{X}_{m,l}) > \hat{c}_{m,l}, \\ \hat{g}_{m,l}, & f(\hat{X}_{m,l}) \leq \hat{c}_{m,l}. \end{cases}\end{aligned}\tag{C.13}$$

- Policy II:

$$\begin{aligned}u_{m,l} &= \max(f(\hat{X}_{m,l}), \hat{g}_{m,l}), \\ v_{m,l} &= \begin{cases} f(\hat{X}_{m,l}), & f(\hat{X}_{m,l}) > \hat{g}_{m,l}, \\ \hat{c}_{m,l}, & f(\hat{X}_{m,l}) \leq \hat{g}_{m,l}. \end{cases}\end{aligned}\tag{C.14}$$

We proceed by moving one step backward in time, to  $t_{m-1}$ , where the paths are again subdivided into new bundles (based on the samples at time  $t_{m-1}$ ) and perform the required SGBM computations. The algorithm iterates back to time  $t_0$ , where we do not need bundles, and we can perform the regression over all paths.

The option value obtained is the so-called *SGBM direct estimator*. An important approach is to also compute option values based on another set of MC paths using the obtained coefficients, which gives the *SGBM path estimator*. The direct estimator is often an upper bound while the path estimator often is a lower bound for the corresponding value. These values should be close.

In this paper, we only present results obtained via the SGBM direct estimator. The SGBM path estimator has been computed and it confirmed the convergence of the algorithm.

## Acknowledgments

We thank the CVA team at the ING bank for helpful discussions. Financial support by the Dutch Technology Foundation STW (Project 12214) is gratefully acknowledged.

## References

- L. B. G. Andersen & V. V. Piterbarg (2010) *Interest Rate Modeling*. London: Atlantic Financial Press.
- Basel Committee on Banking Supervision (2011) Basel III: A global regulatory framework for more resilient banks and banking systems — revised version June 2011, Basel: Bank for International Settlements.
- Basel Committee on Banking Supervision (2015) Consultative document: Review of the credit valuation adjustment risk framework. Basel: Bank for International Settlements.
- T. R. Bielecki & M. Rutkowski (2002) *Credit Risk: Modeling, Valuation and Hedging*. Berlin: Springer Science & Business Media.
- A. Borovykh, A. Pascucci, & C. W. Oosterlee (2016) *Pricing Bermudan Options Under Local Lévy Models with Default*. SSRN: 2771632.
- M. Breton & O. Marzouk (2015) *Evaluation of Counterparty Risk for Derivatives with Early Exercise Features*, in Les cahiers du GERAD. Montréal: GERAD — HEC Montréal.
- D. Brigo, A. Capponi & A. Pallavicini (2014) Arbitrage-free bilateral counterparty risk valuation under collateralization and application to credit default swaps, *Mathematical Finance* **24**(1), 125–146.
- D. Brigo, M. Morini & A. Pallavicini (2013) *Counterparty Credit Risk, Collateral and Funding: With Pricing Cases for All Asset Classes*, The Wiley Finance Series. Chichester, West Sussex: John Wiley & Sons,
- J. C. G. Céspedes, J. A. de Juan Herrero, D. Rosen & D. Saunders (2010) Effective modeling of wrong way risk, counterparty credit risk capital, and alpha in Basel II, *Journal of Risk Model Validation* **4**(1), 71.
- D. Duffie & N. Garleanu (2001) Risk and valuation of collateralized debt obligations. *Financial Analysts Journal* **57**(1), 41–59.
- D. Duffie & K. J. Singleton (1999) Modeling term structures of defaultable bonds, *The Review of Financial Studies* **12**(4), 687–720.
- D. Duffie, J. Pan & K. J. Singleton (2000) Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* **68**(6), 1343–1376.
- F. Fang & C. W. Oosterlee (2009) Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions, *Numerische Mathematik* **114**(1), 27–62.
- Q. Feng, S. Jain, P. Karlsson, B. D. Kandhai & C. W. Oosterlee (2016) Efficient computation of exposure profiles on real-world and risk-neutral scenarios for Bermudan swaptions, *The Journal of Computational Finance* **20**(1), 139–172.
- P. Glasserman & L. Yang (2016) Bounding wrong-way risk in CVA calculation, *Mathematical Finance*.

- J. Gregory (2010) *Counterparty Credit Risk: the New Challenge for Global Financial Markets*, Chichester, West Sussex: John Wiley & Sons.
- J. Hull & A. White (2012) CVA and wrong-way risk, *Financial Analysts Journal* **68**(5), 58–69.
- S. Jain & C. W. Oosterlee (2015) The stochastic grid bundling method: Efficient pricing of Bermudan options and their Greeks, *Applied Mathematics and Computation* **269**, 412–431.
- R. A. Jarrow & S. M. Turnbull (1995) Pricing derivatives on financial securities subject to credit risk, *The Journal of Finance* **50**(1), 53–85.
- D. Lando (2009) *Credit Risk Modeling: Theory and Applications*. New Jersey: Princeton University Press.
- D. B. Madan & H. Unal (1998) Pricing the risks of default, *Review of Derivatives Research* **2**(2–3), 121–160.
- M. J. Ruijter & C. W. Oosterlee (2012) Two-dimensional Fourier cosine series expansion method for pricing financial options, *SIAM Journal on Scientific Computing* **34**(5), B642–B671.
- I. Ruiz (2014) Technical note : On wrong way risk, *The Capco Institute Journal of Financial Transformation* **39**(4), 69–71.