An Application of Deep-xVA Framework on Rate Products under Hull-White Model

Wenyong Zhang

November 2021

Abstract

We adapt the algorithm in Deep-Xva Solver in Gnoatto et al. (2020) to test the interest swap CVA (CVA_{IRS}). We find that IRS CVA can be decomposed into summation of several $\overline{CVA}^{T_2}_{[t,T_1]}-$type IRS CVA with terminal payoff after the end of default time interval. We explore the CVA_{IRS} with default time period between the start time and the expiration time (first swap time). In the first section, we review the theoretical part of general xVA BSDE and introduce the derivation of them. In the second section, we introduce the interest swap CVA and justify the $\overline{CVA}^{T_2}_{[t,T_1]}-$type IRS CVA we need to explore and make the connection between the first part and second part. The third section is the algorithm design: we do the time discretization and simulation. The last section is the implementation and comparison.

1 BSDE of General xVA

1.1 Basic Setting and Dynamic Replication

The basic setting is as same as what in Biagini F, Gnoatto A, Oliva I (2021). We fix a time horizon $T < \infty$ for the trading activity of two agents named the bank (B) and the counterparty (C). Unless otherwise stated, throughout the paper we assume the bank's perspective and refer to the bank as the hedger. All underlying processes are modeled over a probability space $((\Omega, \mathcal{G}, \mathbb{G}, \mathbb{Q}))$, where $\mathbb{G} = (\mathcal{G}_t)_{t\in[0,T]} \subseteq \mathcal{G}$ is a filtration satisfying the usual assumptions ($\mathcal{G}_0$ is assumed to be trivial). We denote by τ_B and τ_C the time of default of the bank and the counterparty, respectively. Specifically, we assume that $\mathbb{G} = \mathbb{F} \vee \mathbb{H}$, where $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is a reference filtration satisfying the usual assumptions and $\mathbb{H} = \mathbb{H}^B \vee \mathbb{H}^C$ for $\mathcal{H}^j_t = \sigma(H_u \mid u \leq t)$ and $H^j_t := \mathbb{I}_{\{\tau^j \leq t\}}, j \in \{B, C\}$. Trivially speaking, the filtration $\mathbb{F}$ is the natural filtration generated by Brownian motion and filtration $\mathbb{H}$ contain the information of jump process, then the filtration $\mathbb{G}$ is the one consists of both Brownian motion and jump process. We set

$$\tau = \tau^C \wedge \tau^B \wedge T$$

Basic trading assets We now introduce the market model.

- (Risky assets) For $d \geq 1$, we denote by S_i, $i = 1, ..., d$ the ex-dividend price (i.e. the price) of risky assets with associated cumulative dividend processes D^i . All S are assumed to be cadlag $\mathbb{F}$-semimartingales, while the cumulative dividend streams D_i are $\mathbb{F}$-adapted processes of finite variation with $D^i_0 = 0$. In our setting, for simplicity, we don't take divident into consideration. The matrix process $\sigma(t, S_t)$ is invertible at every point in time. The dynamics of risky underlying under $\mathbb{P}$-measure is:

$$\begin{cases} dS_t = \mu(t, S_t)dt + \sigma(t, S_t)\, dW^{\mathbb{P}}_t \\ S_0 = s_0 \in \mathbb{R}^d \end{cases} \tag{1}$$

- (Cash account) We assume the existence of an indexed family of cash accounts (money market accounts) $(B^x)_{x\in\mathcal{I}}$, where the stochastic process $r^x := (r_t^x)_{t\geq 0}$ is bounded and $\mathbb{F}$-adapted for all $x \in \mathcal{I}$. The set of indices $\mathcal{I}$ embodies the type of agreement the counterparties establish in order to mitigate the counterparty credit risk. We will specify the characteristics of the aforementioned indices later on. All cash accounts, with unitary value at time 0, are assumed to be strictly positive continuous processes of finite variation of the form
$$B_t^x := \exp\left\{\int_0^t r_s^x ds\right\}, t \in [0,T] \tag{2}$$

- (Risky bond) We introduce two risky bonds with maturity $T^\star \leq T$ and rate of return $r^j + \lambda^j, j \in \{B, C\}$. issued by the bank and the counterparty, respectively, with dynamics
$$dP_t^j = \left(r_t^j + \lambda_t^j\right) P_t^j dt - P_{t-}^j dH_t^j, j \in \{B, C\} \tag{3}$$

- (Payment stream) The payment stream of a financial contract is represented by an $\mathbb{F}$-adapted cadlag process of finite variation $A = (A_t)_{t\in[0,T]}$. We use the notation $\Delta A_t := A_t - A_{t-}$ for the jumps of A. To include the presence of default events, we define the process $\bar{A} = \left(\bar{A}_t\right)_{t\in[0,T]}$ by setting
$$\bar{A}_t := 1_{\{t<\tau\}} A_t + 1_{\{t\geq\tau\}} A_{\tau-} \tag{4}$$
We need mention that $\bar{A}_t$ is in the filtration $\mathbb{G}$ and A_t is in the filtration $\mathbb{F}$ and this will be distinguished from each other in our discussion in $\mathbb{F}$-BSDE and $\mathbb{G}$-BSDE.

Repo-trading and collateralization In line with the existing literature, we assume that the trading activity on the risky assets is collateralized. This means that borrowing and lending activities related to risky securities are financed via security lending or repo market. Since transactions on the repo market are collateralized by the risky assets, repo rates are lower than unsecured funding rates. Aassuming that all assets are traded via repo markets is not restrictive.

- (Financed by collateralization) In case the transactions are fully collateralized, this translates in the following equality
$$\xi_t^i S_t^i + \psi_t^i B_t^i = 0, i = 1\ldots, d, t \in [0,T] \tag{5}$$
where $B^1, ..., B^d$ are the cash accounts associated to the risky assets $S^1, ..., S^d$. It is worth noting that ξ_t^i may be either positive or negative. Here ξ_t^i is the position of the tradable asset S_t^i and ψ_t^i is the corresponded position of money market account. Therefore, this is a self-financing portfolio. $\xi_t^i > 0$ means that we are in a long position, which has to be financed by collateralization. On the other hand, $\xi_t^i < 0$ implies that the i-th asset is shorted, so that the whole amount of collateral is deposited in the riskless asset. This condition plays an important role in precluding trivial arbitrage opportunities among different cash accounts.

- (Simultaneously borrowing and lending is precluded)
$$\psi_t^{f,l} \psi_t^{f,b} = 0, \text{ for all } t \in [0,T] \tag{6}$$
This means that the trading desk can not borrow from and lend money to the treasury desk at the same time so that there must be at least one position be zero.

- (Variation margin) A collateral is posted between the bank and the counterparty to mitigate counterparty risk. The collateral process $C = (C_t), t \in [0,T]$ is assumed to be $\mathbb{G}$-adapted since when $\tau < T$ the collateral will become zero. We assume that $C_t := f\left(\hat{V}_t\right), t \in [0,T]$, where $f : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function and $\hat{V}_t$ is the clean value process which will be defined later.

- If $C_t > 0$, we say that the bank is the collateral provider. It means that the counterparty measures a positive exposure towards the bank, so it is a potential lender to the bank, hence the bank provides/lends collateral to reduce its exposure.
- If $C_t < 0$, we say that the bank is the collateral taker. It means that the bank measures a positive exposure towards the counterparty, so it is a potential lender to the counterparty, hence the counterparty provides/lends collateral to reduce its exposure. We associate the following interest rates to the collateral account:
  * $r^{c,l}$ with account $B^{c,l}$, representing the rate on the collateral amount received by the bank who posted collateral to the counterparty.
  * $r^{c,b}$ with account $B^{c,b}$, representing the rate on the collateral amount paid by the bank who received collateral from the counterparty.

  But in the case with just CVA and DVA, we can assume the rate related with collateral account is same as market risk free rate.
- if $C_t > 0$, then the bank has lent $\psi_t^c = \psi_t^{c,l} < 0$ units of the collateral account to the counterparty, i.e.
$$\psi_t^{c,l} B_t^{c,l} = -C_t^+, t \in [0,T] \tag{7}$$
- if $C_t < 0$, then the bank has borrowed $\psi_t^c = \psi_t^{c,b} > 0$ units of the collateral account from the counterparty, i.e.
$$\psi_t^{c,b} B_t^{c,b} = C_t^-, t \in [0,T] \tag{8}$$

We now introduce the definition of self-financing strategy in our market model. We recall that we assume the point of view of the bank (i.e. the hedger). A dynamic portfolio, denoted by φ, is given by:

$$\varphi = \left(\xi^1, \dots, \xi^d, \xi^B, \xi^C, \psi^1, \dots, \psi^d, \psi^B, \psi^C, \psi^{f,b}, \psi^{f,l}, \psi^{c,b}, \psi^{c,l}\right)$$

where

1. $\xi^1, ..., \xi^d$ are $\mathbb{G}$-predictable process denoting the number of shares of the risky assets. $S^1, ..., S^d$.
2. ξ^B, ξ^C are $\mathbb{G}$-predictable process, denoting the number of shares of the risky bonds P^B and P^C.
3. $\psi^1, \dots, \psi^d, \psi^B, \psi^C$ are $\mathbb{G}$-predictable process, denoting the number of shares of the repo accounts $B^1, \dots, B^d, B^B, B^C$.
4. $\psi^{f,b}$ is a $\mathbb{G}$-predictable process, denoting the number of shares of the unsecured funding borrowing cash account $B^{f,b}$
5. $\psi^{f,l}$ is a $\mathbb{G}$-predictable process, denoting the number of shares of the unsecured funding lending cash account $B^{f,l}$.
6. $\psi^{c,b}$ is a $\mathbb{G}$-predictable process, denoting the number of shares of the collateral borrowing cash account $B^{c,b}$ for the received cash collateral.
7. $\psi^{c,l}$ is a $\mathbb{G}$-predictable process, denoting the number of shares of the collateral lending cash account $B^{c,l}$ for the posted cash collateral.

The meaning of each entry of the portfolio is given in Biagini F, Gnoatto A, Oliva I (2021). Given a dynamic portfolio, we associate it to xVA. We have the portfolio:

$$\begin{aligned}
dV_t(\varphi) =& \sum_{i=1}^{d} \xi_t^i \left(\mu^i\left(t, S_t\right) dt + \sum_{k=1}^{d} \sigma^{i,k}\left(t, S_t\right) dW_t^{k,\mathbb{P}} \right) \\
&+ \sum_{i=1}^{d} \psi_t^i dB_t^i + \sum_{j \in \{B,C\}} \left(\xi_t^j dP_t^j + \psi_t^j dB_t^j \right) - d\bar{A}_t \\
&+ \psi_t^{f,b} dB_t^{f,b} + \psi_t^{f,l} dB_t^{f,l} - \psi_t^{c,b} dB_t^{c,b} - \psi_t^{c,l} dB_t^{c,l} \\
=& \sum_{i=1}^{d} \xi_t^i \left((\mu^i\left(t, S_t\right) - r_t^i S_t^i) dt + \sum_{k=1}^{d} \sigma^{i,k}\left(t, S_t\right) dW_t^{k,\mathbb{P}} \right) \\
&+ \sum_{j \in \{B,C\}} \left(\xi_t^j dP_t^j - r_t^i \xi_t^j P_t^j dt \right) - d\bar{A}_t \\
&+ \psi_t^{f,b} dB_t^{f,b} + \psi_t^{f,l} dB_t^{f,l} - \psi_t^{c,b} dB_t^{c,b} - \psi_t^{c,l} dB_t^{c,l}
\end{aligned} \tag{9}$$

with the wealth process given by:

$$V_t(\varphi) := \psi_t^{f,b} B_t^{f,b} + \psi_t^{f,l} B_t^{f,l} - \left(\psi_t^{c,b} B_t^{c,b} + \psi_t^{c,l} B_t^{c,l} \right) \tag{10}$$

By the fact that $\sigma(t, S_t)$ is invertible, we have the equivalent martingale measure such that the solution of equation $\sigma(t, S_t)\Theta_t = \mu(t, S_t) - r_t S_t$ w.r.t Θ exists. We have the dynamics of the portfolio in $\mathbb{Q}$ measure:

$$\begin{aligned}
dV_t(\varphi) =& \sum_{i=1}^{d} \xi_t^i B_t^i d\tilde{S}_t^i + \sum_{j \in \{B,C\}} \xi_t^j B_t^j d\tilde{P}_t^j - d\bar{A}_t \\
&+ \psi_t^{f,l} dB_t^{f,l} + \psi_t^{f,b} dB_t^{f,b} - \psi_t^{c,l} dB_t^{c,l} - \psi_t^{c,b} dB_t^{c,b}
\end{aligned} \tag{11}$$

Denote that

$$\begin{aligned}
d\tilde{S}_t^i &= \frac{1}{B_t^i} \left(dS_t^i - r_t^i S_t^i dt \right) = \sum_{k=1}^{d} \frac{\sigma^{i,k}\left(t, S_t\right)}{B_t^i} dW_t^{k,\mathbb{Q}}, i = 1, \ldots, d \\
d\tilde{P}_t^j &= \frac{1}{B_t^j} \left(dP_t^j - r_t^j P_{t-}^j dt \right) = -\tilde{P}_{t-}^j dM_t^{j,\mathbb{Q}}, j \in \{B, C\}
\end{aligned} \tag{12}$$

where $\tilde{S}_t$ is the discounted asset price and $\tilde{P}_t$ is the discounted risky bond price, both of whom are under $\mathbb{Q}$-measure. Moreover,

$$M_t^{j,\mathbb{Q}} := H_t^{j,\mathbb{Q}} - \int_0^{t \wedge \tau^j} \lambda_s^{j,\mathbb{Q}} ds \tag{13}$$

with

$$\lambda_t^{j,\mathbb{Q}} := r_t^f - \lambda_t^{j,\mathbb{P}} - r_t^j, t \in [0, T], j \in \{B, C\} \tag{14}$$

With constraints 5,6,7,8. We have:

$$\psi_t^{f,l} = (V_t(\varphi) - C_t)^+ \left(B_t^{f,l} \right)^{-1} \tag{15}$$

$$\psi_t^{f,b} = -\left(V_t(\varphi) - C_t\right)^- \left(B_t^{f,b} \right)^{-1} \tag{16}$$

$$-\psi_t^{c,l} dB_t^{c,l} = -\psi_t^{c,l} r_t^{c,l} B_t^{c,l} dt = +r_t^{c,l} C_t^+ dt \tag{17}$$

$$-\psi_t^{c,b} dB_t^{c,b} = -\psi_t^{c,b} r_t^{c,b} B_t^{c,b} dt = -r_t^{c,b} C_t^- dt \tag{18}$$

respectively, and we have:

$$\begin{aligned}
dV_t(\varphi) &= \sum_{i=1}^{d} \xi_t^i B_t^i d\tilde{S}_t^i + \sum_{j \in \{B,C\}} \xi_t^j B_t^j d\tilde{P}_t^j - d\bar{A}_t \\
&+ \Big[\left(r_t^{f,l} - r_t \right) (V_t(\varphi) - C_t)^+ - \left(r_t^{f,b} - r_t \right) (V_t(\varphi) - C_t)^- \\
&+ \left(r_t^{c,l} - r_t \right) C_t^+ - \left(r_t^{c,b} - r_t \right) C_t^- + r_t V_t(\varphi) \Big] dt \\
&= \sum_{i,k=1}^{d} \xi_t^i \sigma^{i,k}(t, S_t) dW_t^{k,\mathbb{Q}} - \sum_{j \in \{B,C\}} \xi_t^j B_t^j \tilde{P}_{t-}^j dM_t^{j,\mathbb{Q}} - d\bar{A}_t \\
&+ \Big[\left(r_t^{f,l} - r_t \right) (V_t(\varphi) - C_t)^+ - \left(r_t^{f,b} - r_t \right) (V_t(\varphi) - C_t)^- \\
&+ \left(r_t^{c,l} - r_t \right) C_t^+ - \left(r_t^{c,b} - r_t \right) C_t^- + r_t V_t(\varphi) \Big] dt
\end{aligned} \tag{19}$$

Setting

$$\begin{aligned}
Z_t^k &:= \frac{\sum_{i=1}^{d} \xi_t^i \sigma^{i,k}(t, S_t)}{B_t^i} \\
U_t^j &:= -\xi_t^j \tilde{P}_{t-}^j \\
f(t, V, C) &:= -\Big[\left(r_t^{f,l} - r_t \right) (V_t(\varphi) - C_t)^+ - \left(r_t^{f,b} - r_t \right) (V_t(\varphi) - C_t)^- \\
&\quad + \left(r_t^{c,l} - r_t \right) C_t^+ - \left(r_t^{c,b} - r_t \right) C_t^- \Big].
\end{aligned} \tag{20}$$

1.2 BSDE

Clean Value We first introduce the concept of clean value. A clean price is an ideal value process that would be acceptable between two agents entering a perfectly collateralized transaction with entering the transaction are default-free. We can also say that clean value is the product value without credit risk between the agent and counterparty. For instance, when we consider an XVA of European call option, then the BS option price is the clean value. Therefore the clean value has following property:

- no bid-offer spread in the funding accounts, i.e.,$r^{f,l} = r^{f,b} = r^f$;
- no bid-offer spread in the collateral accounts, i.e., $r^{c,l} = r^{c,b} = r^c$;
- the collateral rate is equal to the fictious market interest rate, i.e., $r^c = r$;
- there is no default, i.e. $\tau^B = \tau^C = \infty$ and risky bonds are excluded from the market;
- perfect collateralization, i.e., $V_t = C_t$, for all $t \in [0, T$, where we use V to denote the value process of a collateralized hedging strategy in the fictious market without default-risk.

We then have the dynamics of clean value $\hat{V}_t$:

$$\begin{cases} -d\hat{V}_t(\varphi) = dA_t - r_t \hat{V}_t(\varphi) dt - \sum_{k=1}^{d} Z_t^k dW_t^{k,\mathbb{Q}} \\ \hat{V}_T(\varphi) = 0 \end{cases} \tag{21}$$

For simplicity, we restrict ourselves to Europan-type contracts and write $A_t = \mathbb{I}_{\{t=T\}} g(S_T)$, for a family of Lipschitz functions g Then, equation 21 reads:

$$\begin{cases} -\mathrm{d}\widehat{V}_t = -r_t \widehat{V}_t \ \mathrm{d}t - \sum_{k=1}^{d} \widehat{Z}_t^k \ \mathrm{d}W_t^{k,\mathbb{Q}} \\ \widehat{V}_T = g(S_T) \end{cases} \tag{22}$$

On the time period $\tau > t$, the close condition at τ has three circumstance:

- No default, which is associated with the clean value.
- Default of bank, which is associated with the case $\tau^C > \tau^B$.
- Default of counterparty, which is associated with the case $\tau^C < \tau^B$.

Then the $\mathbb{G}$-BSDE for the portfolio's dynamics under $\mathbb{Q}$ has the form

$$\begin{cases} -\mathrm{d}V_t = \mathrm{d}\bar{A}_t + (f(t, V_t, C_t) - r_t V_t)\,\mathrm{d}t - \sum_{k=1}^{d} Z_t^k \,\mathrm{d}W_t^{k,\mathbb{Q}} - \sum_{j\in\{B,C\}} U_t^j \,\mathrm{d}M_t^{j,\mathbb{Q}} \\ V_\tau = \theta_\tau(\widehat{V}, C), \quad \text{with} \\ \theta_\tau(\widehat{V}, C) := \widehat{V}_\tau + \mathbf{1}_{\{\tau^C<\tau^B\}}\left(1 - R^C\right)\left(\widehat{V}_\tau - C_{\tau-}\right)^- - \mathbf{1}_{\{\tau^B<\tau^C\}}\left(1 - R^B\right)\left(\widehat{V}_\tau - C_{\tau-}\right)^+ \end{cases} \tag{23}$$

where R^B, R^C are two positive constants representing the recovery rate of the bank and the counterparty. We mention again that $\bar{A}_t$ is in the filtration $\mathbb{G}$ and A_t is in the filtration $\mathbb{F}$ and this will be distinguished from each other in our discussion in $\mathbb{F}$-BSDE and $\mathbb{G}$-BSDE. Then V_t can be represented as

$$V_t = B_t^r \mathbb{E}^{\mathbb{Q}}\left[\int_{(t,\tau\wedge T]} \frac{\mathrm{d}\bar{A}_u}{B_u^r} + \int_t^{\tau\wedge T} \frac{f(u, V, C)}{B_u^r}\,\mathrm{d}u + \mathbf{1}_{\{\tau\le T\}} \frac{\theta_\tau(\widehat{V}, C)}{B_\tau^r} \mid \mathcal{G}_t\right] \tag{24}$$

With the fact that $XVA_t = \hat{V}_t - V_t$. We have $XVA_\tau = -\theta_\tau + \hat{V}_\tau$. By equation 23 and 21. We have the $\mathbb{G}$-BSDE of pre-default XVA_t on $\{t < \tau\}$.

$$dXVA_t = -(f(t, \hat{V}_t - XVA_t, C_t) + r_t XVA_t)dt - \sum_{k=1}^{d} \tilde{Z}_t^k \,\mathrm{d}W_t^{k,\mathbb{Q}} - \sum_{j\in\{B,C\}} \tilde{U}_t^j \,\mathrm{d}M_t^{j,\mathbb{Q}} \tag{25}$$

where $\tilde{Z}_t^k = \hat{Z}_t^k - Z_t^k$ and $\tilde{U}_t^j = -U_t^j$. Then we construct a process $\overline{XVA}_t$ such that it satisfy the $\mathbb{F}$-BSDE:

$$\begin{cases} -d\overline{XVA}_t = \bar{f}\left(\hat{V}_t - \overline{XVA}_t\right)dt - \sum_{k=1}^{d} \bar{Z}_t^k dW_t^{k,\mathbb{Q}} \\ \overline{XVA}_T = 0 \end{cases} \tag{26}$$

where

$$\bar{f}\left(\hat{V}_t - \overline{XVA}_t\right) := -f\left(t, \hat{V} - \overline{XVA}_t, C\right) - \left(r_t + \lambda_t^{C,\mathbb{Q}} + \lambda_t^{B,\mathbb{Q}}\right)\overline{XVA}_t - \lambda_t^{C,\mathbb{Q}}\theta_t^C + \lambda_t^{B,\mathbb{Q}}\theta_t^B \tag{27}$$

and

$$\theta_t^C := \left(1 - R^C\right)\left(\hat{V}_t - C_t\right)^-$$
$$\theta_t^B := \left(1 - R^B\right)\left(\hat{V}_t - C_t\right)^+$$

define

$$X_t := \overline{XVA}_t J_t + \mathbf{1}_{\{\tau\le t\}}\vartheta_\tau, t \in [0, \tau\wedge T] \tag{28}$$

where $J_t := 1_{\{t<\tau\}}$ and $\vartheta_t := -\theta_t + \hat{V}_t, t \in [0, T]$. With the fact that

$$\sum_{j\in\{B,C\}} \tilde{U}_t^j dM_t^{j,\mathbb{Q}} = -\left(\left(\vartheta_t - \overline{XVA}_t\right)dJ_t + \lambda_t^{C,\mathbb{Q}}\left(-\theta_t^C - \overline{XVA}_t\right)dt + \lambda_t^{B,\mathbb{Q}}\left(\theta_t^B - \overline{XVA}_t\right)dt\right. \tag{29}$$

the martingale property of the right hand side has been proved in Crepey et al. (2014) in Lemma 5.2.9. Then by the multiply law of difference, we have:

$$\begin{aligned} dX_t &= d\left(\overline{XVA}_t J_t\right) + d\left(1_{\{\tau<T\}}\vartheta_\tau\right) \\ &= d\overline{XVA}_{t\wedge\tau} + \overline{XVA}_t dJ_t - \vartheta_t dJ_t = -\bar{f}\left(\hat{V}_t - \overline{XVA}_t\right)dt - \sum_{k=1}^{d} \tilde{Z}_t^k \,\mathrm{d}W_t^{k,\mathbb{Q}} + \overline{XVA}_t dJ_t - \vartheta_t dJ_t \end{aligned} \tag{30}$$

and

$$
dXVA_t = -\bar{f}\left(\hat{V}_t - XVA_t\right)dt - \sum_{k=1}^{d} \tilde{Z}_t^k \, \mathrm{d}W_t^{k,\mathbb{Q}} + XVA_t dJ_t - \vartheta_t dJ_t \tag{31}
$$

From the existence and uniqueness of the solution of $\mathbb{G}$-BSDE, we can find that $X_t = XVA_t = \overline{XVA}_{t\wedge\tau}$. Using tower property to do the integral under τ^B and τ^C in the definition (3.9) in Biagini F, Gnoatto A, Oliva I (2021). We can have the representation in Corollary 3.17 in that paper, this is also an alternative method to get the FBSDE 26. An example to prove the correctness of CVA under filteration $\mathcal{F}_t$ is that

$$
\begin{aligned}
\overline{CVA}_t &= B_t^r \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[1_{\{\tau^C<\tau^B\}}\left(1-R^C\right)\frac{1}{B_\tau^r}\left(\hat{V}_\tau(\varphi)-C_{\tau-}\right)^-\right] \mid \mathcal{F}_t\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\int_t^T\left(1-R^C\right)\left(\hat{V}_u(\varphi)-C_u\right)^-\frac{B_t^r}{B_u^r}\frac{\lambda_u^{C,\mathbb{Q}}}{\lambda_u^{B,\mathbb{Q}}+\lambda_u^{C,\mathbb{Q}}}(\lambda_u^{B,\mathbb{Q}}+\lambda_u^{C,\mathbb{Q}})\cdot e^{-\int_t^u(\lambda_s^{B,\mathbb{Q}}+\lambda_s^{C,\mathbb{Q}})ds}du \mid \mathcal{F}_t\right] \\
&= B_t^{\tilde{r}} \mathbb{E}^{\mathbb{Q}}\left[\left(1-R^C\right)\int_t^T\frac{1}{B_u^{\tilde{r}}}\left(\hat{V}_u(\varphi)-C_u\right)^-\lambda_u^{C,\mathbb{Q}}du \mid \mathcal{F}_t\right]
\end{aligned} \tag{32}
$$

where $\tilde{r} := r + \lambda^{C,\mathbb{Q}} + \lambda^{B,\mathbb{Q}}$. Then $\overline{DVA}_t$, $\overline{FVA}_t$ and $\overline{ColVA}_t$Can also be calculated as this,

$$
\begin{aligned}
\overline{DVA}_t &:= B_t^{\tilde{r}}\mathbb{E}^{\mathbb{Q}}\left[\left(1-R^B\right)\int_t^T\frac{1}{B_u^{\tilde{r}}}\left(\hat{V}_u(\varphi)-C_u\right)^+\lambda_u^{B,\mathbb{Q}}du \mid \mathcal{F}_t\right] \\
\overline{FVA}_t &:= B_t^{\tilde{r}}\mathbb{E}^{\mathbb{Q}}\left[\int_t^T\frac{\left(r_u^{f,l}-r_u\right)\left(V_u(\varphi)-C_u\right)^+-\left(r_u^{f,b}-r_u\right)\left(V_u(\varphi)-C_u\right)^-}{B_u^{\tilde{r}}}du \mid \mathcal{F}_t\right] \\
\overline{\mathrm{ColVA}}_t &:= B_t^{\tilde{r}}\mathbb{E}^{\mathbb{Q}}\left[\int_t^T\frac{\left(r_u^{c,l}-r_u\right)C_u^+-\left(r_u^{c,b}-r_u\right)C_u^-}{B_u^{\tilde{r}}}du \mid \mathcal{F}_t\right]
\end{aligned} \tag{33}
$$

finally, we have the driver:

$$
\begin{aligned}
\bar{f}\left(t,\widehat{V}_t,\overline{\mathrm{XVA}}_t\right) :=& -\left(1-R^C\right)\left(\widehat{V}_t-C_t\right)^-\lambda_t^{C,\mathbb{Q}} \\
&+\left(1-R^B\right)\left(\widehat{V}_t-C_t\right)^+\lambda_t^{B,\mathbb{Q}} \\
&+\left(r_t^{f,l}-r_t\right)\left(\widehat{V}_t-\overline{\mathrm{XVA}}_t-C_t\right)^+-\left(r_t^{f,b}-r_t\right)\left(\widehat{V}_t-\overline{\mathrm{XVA}}_t-C_t\right)^- \\
&+\left(r_t^{c,l}-r_t\right)C_t^+-\left(r_t^{c,b}-r_t\right)C_t^--\left(r_t+\lambda_t^{C,\mathbb{Q}}+\lambda_t^{B,\mathbb{Q}}\right)\overline{\mathrm{XVA}}_t
\end{aligned} \tag{34}
$$

With the fact that $\mathbb{F}$-BSDE has just the $\sum_{k=1}^{d}\overline{Z}_t^k \, \mathrm{d}W_t^{k,\mathbb{Q}}$ term, we have the BSDE 26.

2 Product Description

2.1 Interest Rate Swap Revisited

let us consider the pricing of the payer's swaps, which take LIBOR, as the reference rate for the floating leg. Let the fixed rate be K and the number of interest-rate payments be n. Denote $P(t,T)$ as the price of zero coupon bond with maturity of T at time t and ΔT_j be the time length between the i-th swap time and (i+1)-th swap time. Then we denote the swap rate as:

$$
R_{m,n}(t) = \frac{P(t,T_m)-P(t,T_n)}{\sum_{j=m}^{n-1}\Delta T_j P(t,T_{j+1})} \tag{35}
$$

Then the value of the swap with expiration date T_m and swap maturity T_n at time t can be derived as follows:

$$\begin{aligned}\operatorname{Swap}\left(T;T_m,T_n,K\right) &= P\left(T,T_m\right)-P\left(T,T_n\right)-\sum_{j=m}^{n-1}\Delta T_j K P\left(T,T_{j+1}\right)\\ &= \left(\sum_{j=m}^{n-1}\Delta T_j P\left(T,T_{j+1}\right)\right)\left(R_{m,n}(T)-K\right)\\ &= A_{m,n}(T)\left(R_{m,n}(T)-K\right)\end{aligned}\tag{36}$$

In the formula, the term $A_{m,n}(T)=\left(\sum_{j=m}^{n-1}\Delta T_j P\left(T,T_{j+1}\right)\right)$ is the annuity from T_m to T_n at time T.

2.2 Interest Rate Swap CVA

Denote the clean value of a derivative is $V(t,T)$ where T is the maturity date of the derivative (option). We have definition of the price of CVA with the default time period $[T_1,T]$:

$$\mathrm{CVA}(t,T_1,T)=\mathbb{E}^{\mathbb{Q}}\left[(1-\mathrm{R}^{\mathrm{C}})\mathbb{I}_{[T_1\le\tau\le T]}\cdot\max(V(\tau,T),0)D(t,\tau)\right]$$

and $D(t,\tau)=\exp\left\{-\int_t^\tau r(s)ds\right\}$ is the value of the money market account at time τ starting with a unit value at time t. R^C stands for the (constant or stochastic) recovery rate of the derivative. So the CVA of the derivative value is the expectation of the irrecoverable part of the exposure in the risk-neutral world. Denote hazard rate as h, $S(t)=Q(\tau\ge t)=1-Q(\tau\le t)$ is the survival function of the counterparty. The expectation may be rewritten into the following form:

$$\begin{aligned}\mathrm{CVA}(t,T_1,T) &= \mathbb{E}^{\mathbb{Q}}\left[(1-\mathrm{R}^{\mathrm{C}})\mathbb{I}_{[T_1\le\tau\le T]}\cdot\max(V(\tau,T),0)D(t,\tau)\right]\\ &= -\mathbb{E}^{\mathbb{Q}}\left[\int_{T_1}^{T}(1-\mathrm{R}^{\mathrm{C}})\cdot D(t,u)V(u,T)^+\mathrm{d}S(u)\right]\\ &= \mathbb{E}^{\mathbb{Q}}\left[\int_{T_1}^{T}(1-\mathrm{R}^{\mathrm{C}})\cdot D(t,u)V(u,T)^+h\cdot e^{-hu}\mathrm{d}u\right]\end{aligned}$$

In our experiment we set $R^C=0$. The CVA of an interest rate swap (hereinafter "CVA_{IRS}") with n swap payments ($t=T_0\le T_1\le\cdots\le T_n=T$) can be decomposed to the summation of CVA_{IRS} of each time intervals.

$$\mathrm{CVA}_{IRS}(t,T)=\sum_{i=0}^{n-1}\mathrm{CVA}_{IRS}\left(t,T_i,T_{i+1}\right)$$

where $\mathrm{CVA}_{IRS}\left(t,T_i,T_{i+1}\right)$ is the expected value of the loss if the counterparty defaults between the times T_i and T_{i+1}. More rigorously,for $i=0,...,n-1$, setting LGD $=1$. The detail of this decomposition is proved in appendix section.

2.3 $\overline{CVA}_{[t,T_1]}^{T_2}$−type IRS CVA

The simplest type of IRS CVA is the one whose default time interval is $[t,T_1]$ ($T_1\le T_2$) and the swap beginning at T_2. Since for this type of CVA, we can calculate the terminal payoff at the end of the default time interval, which is very important for us to make the framework into our product. For a $\overline{CVA}_{[t,T_1]}^{T_2}$−type

IRS CVA, the terminal payoff at time T_1 is:

$$\begin{aligned}\operatorname{Swap}\left(T_1; T_2, T, K\right) &= P\left(T_1, T_2\right) - P\left(T_1, T\right) - \sum_{j=m}^{n-1} \Delta T_j K P\left(T_1, T_{j+1}\right) \\ &= \left(\sum_{j=m}^{n-1} \Delta T_j P\left(T_1, T_{j+1}\right)\right)\left(R_{m,n}(T_1) - K\right).\end{aligned}$$

Where $\sum_{j=m}^{n-1} \Delta T_j P\left(T_1, T_{j+1}\right)$ is the annuity discounted to time T_1 and $R_{m,n}(T_1)$ is the swap rate at T_1 which can be calculated by 38. Therefore we can calculate the terminal payoff of the swap at the end of default time interval for each simulation path of the short rate process.

The present value of IRS CVA can be decomposed into $\overline{CVA}_{[t,T_1]}^{T_2}$−type IRS CVAs, the proof is given in appendix section.

3 Algorithms

3.1 Dynamics in our problem

3.1.1 BSDE on Clean Value

For simplicity, we restrict ourselves to Europan-type contracts and let $T \leq \tau$, for a family of Lipschitz functions g. Since the underlying tradable asset is the zero coupon bond, we can have it's dynamics by HJM model. Then we have the BSDE of the clean value of the swap under one factor hull white model is:

$$\left\{\begin{array}{l} \mathrm{d}r_t = \kappa(\theta - r_t)\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} \\ \mathrm{d}P_t^\tau = P_t^\tau\left[r_t \mathrm{d}t + \boldsymbol{\Sigma}^{\mathrm{T}}(t,\tau)\,\mathrm{d}W_t^{\mathbb{Q}}\right] \\ -\mathrm{d}\widehat{V}_t = -r_t\widehat{V}_t\mathrm{d}t - \widehat{Z}_t\, \mathrm{d}W_t^{\mathbb{Q}} \\ \widehat{V}_T = g(P_T^\tau) \end{array}\right. \tag{37}$$

Zero coupon bond $P(t,T)$ It is also trivial that our interest rate process is Hull White model (which is also called Vasicek model), we have the analytical formula of its zero coupon bound:

$$\begin{aligned} P(t,T) &= A(t,T)e^{-B(t,T)r_t} \\ B(t,T) &= \frac{1}{\kappa}(1 - e^{-\kappa(T-t)}) \\ A(t,T) &= \exp\left\{\frac{\left(B(t,T) - T + t\right)\left(\kappa^2\theta - \sigma^2/2\right)}{\kappa^2} - \frac{\sigma^2 B^2(t,T)}{4\kappa}\right\} \end{aligned} \tag{38}$$

Since we have analytical formula of $P(t,T)$, we can calculate the value of the swap via 36.

Remark 1: Although we put the bond process into A, we do not exactly simulate the process of P_t^τ. We just need to simulate the short rate under Versicek model and then get the bond price at each time step via 38. The main reason we put the HJM process into A is that the underlying process need to be tradable and this makes sure that our method can be fitted into the framework.

Remark 2: In our paper, for the simplicity, we just let the time of swap be one, that is, we let the swap contract begin at time T and the swap happen at time τ. Then we just need to face the dynamics of the $P(t,\tau)$. In next sections, we denote T_1 as the end of default time and generally $T_1 \leq T$. In our product, without loss of generality, we have $T_1 = T$.

3.1.2 BSDE on Adjustment

In our problem, we set $r_t^f = r_r^c = r_t$ almost every time. The $FVA = ColVA = 0$. Then the $XVA = DVA - CVA$. By setting $\lambda^B = 0$, with process of clean value $\hat{V}_t$, we have the BSDE of CVA as following:

$$\begin{cases} -\mathrm{d}\overline{\mathrm{CVA}}_t = \bar{f}_c\left(t, \widehat{V}_t, \overline{\mathrm{CVA}}_t\right)\mathrm{d}t - \sum_{k=1}^{d} \bar{Z}_t^k \, \mathrm{d}W_t^{k,\mathbb{Q}} \\ \overline{\mathrm{CVA}}_T = 0 \end{cases} \tag{39}$$

Where

$$\begin{aligned} \bar{f}_c\left(t, \widehat{V}_t, \overline{\mathrm{CVA}}_t\right) :=& -\left(1 - R^C\right)\left(\widehat{V}_t - C_t\right)^{-} \lambda^C \\ & - \left(r_t + \lambda^C\right)\overline{\mathrm{CVA}}_t \end{aligned} \tag{40}$$

3.2 Simulation

Since the process of the short rate under $\mathbb{Q}$-measure is:

$$\mathrm{d}r_t = \kappa(\theta - r_t)\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} \tag{41}$$

Then we can simulate the swap value and stochastic discount factor as following steps:

- Construct the discount curve, P_0^τ, any τ, this can be calculated by 38. Where $\frac{P_0^\tau}{B_0} = P_0^\tau$. This can be finished via the Euler scheme.
- we obtain the money market account with analytical solution:

$$\begin{aligned} B_t^{-1} &= \exp\left\{-\int_0^t f(0,u)du - \int_0^t \frac{\sigma^2}{2\kappa^2}\left(1 - e^{-\kappa u}\right)^2 du - \int_0^t \sigma \int_0^u e^{-\kappa(u-s)} dW_s du\right\} \\ &= P(0,t)\exp\left\{-\int_0^t \frac{\sigma^2}{2\kappa^2}\left(1 - e^{-\kappa u}\right)^2 du - \int_0^t \sigma \int_s^t e^{-\kappa(u-s)} du dW_s\right\} \end{aligned}$$

- Simulate a number of Brownian paths and simulate the short rate. Then calculate the P_t^τ under 38.
- Average the payoffs:

$$g(P(T,\tau))$$

  Remark In Hull white model, the integral part can be obtained analytically. We can just sample the Gaussian random number based on the term $\int_0^{T_0} \mathbf{\Sigma}^{\mathrm{T}}(t, T_0)\,\mathrm{d}\mathbf{W}_t^{\mathbb{Q}}$. In Hull White model, $\mathbf{\Sigma}(t, T_i) = -\int_t^T \sigma(t,s)\mathrm{d}s = \frac{\sigma}{\kappa}(1 - e^{-\kappa(T-t)})$

3.3 Time discretization

Given a sample path of interest rate $(r_n)_{n=1,\ldots,N}$ and clean value $(\hat{V}_n)_{n=1,\ldots,N}$, a unifying formula for $\overline{CVA}_{[t,T_1]}^{T_2}-$type IRS CVA and DVA can be written as

$$\mathbb{E}^{\mathbb{Q}}\left[\int_t^{T_1} D(t,u)V(u,T_2,T)^+ h \cdot e^{-hu}\mathrm{d}u|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[\int_t^{T_1} \Phi_u(\hat{V}_u)\mathrm{d}u|\mathcal{F}_t\right]$$

where $D(t_1,t_2) = e^{-\int_{t_1}^{t_2} r_u \mathrm{d}u} = B_{t_1}/B_{t_2}$ and $R^C = R^B = 0$ and $\lambda^B = \lambda^C = h$.

$$\begin{aligned} \Phi_u(v) =& (1 - R^C)\frac{B_\tau}{B_u}(v_u)^+ h \cdot e^{-hu} \quad \text{for CVA} \\ \Phi_u(v) =& (1 - R^B)\frac{B_\tau}{B_u}(v_u)^- h \cdot e^{-hu} \quad \text{for DVA} \end{aligned}$$

Taking $t = t_0 = 0$, we first discretize the time interval $[0, T_1]$ into N uniform grids. Where T_1 is the end time of default interval. The clean value process are also be estimated at each step. Denote the step size as $\eta_n = \Delta t$, the integral can be approximated by

$$\int_0^{T_1} \Phi_u(\hat{V}_u)\mathrm{d}u \approx \sum_{n=0}^{N} \frac{B_{T_1}}{B_n} \cdot \hat{V}_n^+ \cdot he^{-ht_n}\eta_n.$$

the adjustment is the approximated by

$$\frac{1}{P}\sum_{p=1}^{P}\sum_{n=0}^{N} \cdot \left(\frac{B_{T_1}}{B_n} \cdot (\hat{\mathcal{V}}_n^{\xi^*,\rho^*,(p)})^+ \cdot he^{-ht_n}\eta_n \right)$$

Where T_1 denotes the end of default time interval of the $\overline{CVA}_{[t,T_1]}^{T_2}-$type IRS CVA.

Deep CVA Algorithm The following Algorithm is the pseudocode corresponding to our method. Algorithm one is to train the NN for BSDE to simulate the clean value $\hat{V}_n$. Algorithm two is to simulate the value of CVA on the basis of series r_n and $\hat{V}_n$.

Algorithm 1: Deep algorithm for exposure sumulation

Input: Parameters: N for the m-th CVA_{IRS}, L. ▷ *N time steps, L paths for inner Monte Carlo loop which is also batch size.*
fix architecture of ANN. ▷ *instrinsically defines the number of parameter R.*
Deep BSDE Solver (N,L):
while the neural network $(\varphi_n^\rho)_{n=1,\dots,N}$ does not converge **do**
  Simulate L paths of short rate $r_n^{(l)}$ at each time step.
  Simulate L paths $(P(t_n,\tau)_n^{(l)})_{n=1,\dots,N}$ with each maturity T_i based on 38 for the forward dynamics.
  Calculate L paths of $(B_{t_n}^{(l)})_{n=1,\dots,N}$ at each time step via 3.2.
  Define the neural network $(\varphi_n^\rho)_{n=1,\dots,N}$.
  Minimize over ξ and ρ by SGD

$$\frac{1}{L}\sum_{l=1}^{L}\left(g(P(t_n,\tau)^{(l)}) - \hat{\mathcal{V}}_n^{\xi,\rho,(l)}\right)^2$$

  subject to

$$\begin{cases} \hat{\mathcal{V}}_{n+1}^{\xi,\rho,(l)} &= \hat{\mathcal{V}}_n^{\xi,\rho,(l)} + r_n^{(l)}\hat{\mathcal{V}}_n^{\xi,\rho,(l)}\Delta t + (\hat{\mathcal{Z}}_n^{\rho,(l)})^T \Delta W_n^{(L)} \\ \hat{\mathcal{V}}_0^{\xi,\rho,(l)} &= \xi \\ \hat{\mathcal{Z}}_n^{\rho,(l)} &= \varphi_n^\rho(P(t_n,\tau)) \end{cases} \tag{42}$$

  Save the optimizer (ξ^*, ρ^*)
end while

---

4 Other Products

For other products such as FX and CDS, the framework is also applicable. We use FX product, CDS as example

Algorithm 2: Deep xVA solver for non-recursive valuation adjustments

1: Apply Algorithm 1.
2: **Input:** Parameters: P. ▷ *P paths for outer Monte Carlo loop.*
3: Simulate, $(\hat{\mathcal{V}}_n^{\xi^*,\rho^*,(p)})_{n=1,...,N}$ by the dynamic 43
4: Compute the adjustment as

$$\frac{1}{P}\sum_{p=1}^{P}\sum_{n=0}^{N}\cdot\left(\frac{B_T}{B_n}\cdot(\hat{\mathcal{V}}_n^{\xi^*,\rho^*,(p)})^+\cdot he^{-ht_n}\eta_n\right)$$

Algorithm 3: Deep xVA solver

1: Apply Algorithm 1.
2: Set parameters: P. ▷ *P paths for outer Monte Carlo loop.*
3: **Deep XVA-BSDE solver** (N,P):
4: **while** the neural network $(\varphi_n^\rho)_{n=1,...,N}$ does not converge **do**
5: Simulate $(\hat{\mathcal{V}}_n^P)_{n=0,...,N}$, $p = 1, ..., P$ for the portfolio value
6: Define the neural networks $\left(\psi_n^\zeta\right)_{n=1,...,N}$.
7: Minimize over γ and ζ by SGD.

$$\frac{1}{P}\sum_{p=1}^{P}\left(\overline{\mathcal{X}}_N^{\gamma,\zeta,(p)}\right)^2$$

8: subject to

$$\begin{cases}\overline{\mathcal{X}}_{n+1}^{\gamma,\zeta,(p)} &=\overline{\mathcal{X}}_n^{\gamma,\zeta,(p)} - \bar{f}\left(t_n,\widehat{\mathcal{V}}_n^{(p)},\overline{\mathcal{X}}_n^{\gamma,\zeta,(p)}\right)\Delta t + \left(\overline{\mathcal{Z}}_n^{\zeta,(p)}\right)^\top \Delta W_n^{(p)}\\ \overline{\mathcal{X}}_0^{\gamma,\zeta,(p)} &=\gamma\\ \overline{\mathcal{Z}}_n^{\zeta,(p)} &=\psi_n^\zeta\left(P(t_n,\tau)\right)\end{cases} \tag{43}$$

9: Save the optimizer (ξ^*,ρ^*)
10: **end while**

4.1 FX Product

Basic setting: We denote r_d as domestic rate r_f as foreign rate. The FX rate is denoted as F and σ_{FX} as the volatility of FX rate. Then we have the dynamics of the rate as:

$$dF_t = (r_d - r_f)\, F_t dt + \sigma_{FX} F_t dW_{FX}^d \tag{44}$$

Here W_{FX}^d is the $\mathbb{Q}$-Brownian motion under domestic measure. Then The BSDE of clean value and CVA is:

$$\begin{cases} dF_t = (r_d - r_f)\, F_t dt + \sigma_{FX} F_t dW_{FX}^d \\ -\mathrm{d}\widehat{V}_t = -r_d \widehat{V}_t \mathrm{d}t - \widehat{Z}_t\ \mathrm{d}W_t^{\mathbb{Q}} \\ \widehat{V}_T = g(F_T) \end{cases} \tag{45}$$

$$\begin{cases} -\mathrm{d}\overline{\mathrm{CVA}}_t = \bar{f}_c\left(t, \widehat{V}_t, \overline{\mathrm{CVA}}_t\right) \mathrm{d}t - \sum_{k=1}^d \bar{Z}_t^k\ \mathrm{d}W_t^{k,\mathbb{Q}} \\ \overline{\mathrm{CVA}}_T = 0 \end{cases} \tag{46}$$

Then Algorithm 1 and 2 can be used to solve 45. Algorithm 3 can be applied on solving 52

4.2 CDS

Basic setting: Denote the process of hazard rate as η. The riskless rate is r. Then assume that

$$\mathrm{d}\eta_t = \kappa(\theta - \eta_t)\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} \tag{47}$$

which is also a Vasicek model. Let s denote the premium paid by the buyer of default protection. Assuming that the premium is paid continuously, the present value of the premium leg of a credit default swap $P(s,T)$ can now be expressed as:

$$P(s,T) = \mathbb{E}\left[s \int_0^T \exp\left(-\int_0^t r + \eta_s ds \right) dt \right] \tag{48}$$

Similarly, the value of the protection leg of a credit default swap $PR(w,T)$ can be expressed as:

$$PR(w,T) = \mathbb{E}\left[w \int_0^T \eta_t \exp\left(-\int_0^t r + \eta_s ds \right) dt \right] \tag{49}$$

The it is suffice to evaluate the CVA of the clean value of 48 and 49. Then this is the product with continuous payoff within the timeline. Then we can deduce that $\mathrm{d}A_t = \exp\left(-\int_0^t \eta_s ds\right) \mathrm{d}t$. Then we have the BSDE of clean value of 48 as:

$$\begin{cases} \mathrm{d}\eta_t = \kappa(\theta - \eta_t)\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} \\ -\mathrm{d}\widehat{V}_t = \exp\left(-\int_0^t \eta_s ds\right) \mathrm{d}t - r\widehat{V}_t \mathrm{d}t - \widehat{Z}_t\ \mathrm{d}W_t^{\mathbb{Q}} \\ \widehat{V}_T = 0 \end{cases} \tag{50}$$

And we have the BSDE of clean value of 49 as:

$$\begin{cases} \mathrm{d}\eta_t = \kappa(\theta - \eta_t)\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} \\ -\mathrm{d}\widehat{V}_t = \eta_t \exp\left(-\int_0^t \eta_s ds\right) \mathrm{d}t - r\widehat{V}_t \mathrm{d}t - \widehat{Z}_t\ \mathrm{d}W_t^{\mathbb{Q}} \\ \widehat{V}_T = 0 \end{cases} \tag{51}$$

The value of $\exp\left(-\int_0^t \eta_s ds\right)$ can be calculated by 3.2. Then we can calculate the CVA by algorithm 2 or algirithm 3 under the BSDE:

$$\begin{cases} -\mathrm{d}\overline{\mathrm{CVA}}_t = \bar{f}_c\left(t, \widehat{V}_t, \overline{\mathrm{CVA}}_t\right) \mathrm{d}t - \sum_{k=1}^d \bar{Z}_t^k\ \mathrm{d}W_t^{k,\mathbb{Q}} \\ \overline{\mathrm{CVA}}_T = 0 \end{cases} \tag{52}$$

5 Experiments

5.1 Introduction

In experiment part we test our model especially the class BSDEsolver. We will describe methods for testing functions and justify how they work. Our product is the forward rate agreement which is also a one-time interest rate swap. First, we train the BSDEsolver to get the simulator of the clean value $\hat{V}_t$. This step is finished throughout algorithm 1. After training the simulator, we test the BSDEsolver. During this process, we generate a batch of validating data (a batch of stochastic interest rate paths) and calculate the analytical clean value and clean value evaluated by the trained BSDE simulator and we calculate the mean absolute error (MAE) between these two clean value and check whether the error is smaller than a threshold. Since the error is smaller than the threshold, we can guarantee that we simulate the right clean value. The next step is to evaluate the value of CVA on the clean value. The value of adjustment can be obtained either by inserting the values in an 'outer' Monte Carlo computation for non-recursive adjustments which is demonstrated in algorithm 2, or applying a second time the Deep BSDE Solver (algorithm 3) in the recursive case. Since the Monte Carlo method is a states-of-the-art method for financial engineering pricing and can also give the confidence interval, we take this method as the test benchmark for checking our price value from algorithm 3.

5.2 One-time Interest rate swap

We consider an interest rate swap begin at the end of year $T = 1$ and do the swap only once at the end of year $\tau = 2$. This behaviour is equivalent to a forward rate agreement. The short rate model is one-factor Hull White constant model (OU) with three parameters (θ, κ, σ). For the stochastic interest model and product issues, we have the following config setting: For the product, the terminal payoff of the one-time swap is:

Parameter	value
T	1.0
time step	100
τ	2.0
σ	0.002
κ	0.04
θ	1.0
r_0	0.03
K	0.001

Table 1: Parameters of the product and rate model

$$g(P(T,\tau)) = 1 - (1 + (\tau - T)K)P(T,\tau) \tag{53}$$

The analytical clean value of this FRA is

$$\hat{V}_t = P(t,T) - (1 + (\tau - T)K)P(t,\tau) \tag{54}$$

For the case of the swap, we observe that the exposure profile corresponds to the present value of the contract. We first learn the clean value BSDE A and find the learned initial clean value is 0.07791, where the analytical value is 0.07783. The absolute error is 1e-4. Then we do the monte carlo simulation on the expectation representation (algorithm 2) with 100000 paths and get the result 0.007424 with The confidence interval $[0.007383, 0.007466]$ and the price learned via 39 by algorithm 3 is 0.007427 which falls into the associated interval. There exists a problem in the recursive method that the process of evaluating the value of adjustment may induce error from the clean value. Then we calculate the adjustment value again with the analytical value via 54. Then the price is 0.007419 which also falls into the confidence interval. Some associated results will be illustrated in the next subsection.

5.3 Model Testing

In this subsection, we describe the method to test the correctness of BSDEsolver which is the core algorithm in our topic. The unit test has been divided into three separate parts

1. We test the single neural network at each time step. The function name is **FeedForwardSubNet**. For one single neural network, we need to make sure that the output is exactly what we expect. We use two methods to check this: First, we set the bias be zero and initialize the weight randomly and then input zero tensor into the single network. The result is also a zero tensor, which fulfill our expectation. Second, we reduce the dimension to one and cancel the activation function, then the neural map is an affine function. We initialize the weight to one and bias to zero and check whether the output equals to the input. The result fulfill our expectation. Combine these two assertions, we can make sure that our implementation of single network is right.

2. In the next step, we test the whole model associated with the FBSDE (**NonSharedModel**) and make sure the output is exactly the terminal value of the neural BSDE. Since the single neural network at each time step has passed the unit test we can use the same method to test the whole model. We reduce the time step to 2 and dimension of asset to 1, which will make it easier for us to calculate the exact value by hand. Without loss of generality, we set the weight be one and bias to zero. We calculate the output of the whole model under this setting and also calculate the terminal of BSDE by hand. After calculating, we make the comparison and find that the error is smaller than the threshold. Therefore, the model pass the unit test. After this, we also set the asset price and increment of Brownian motion be zero as initialize the weight randomly to double check whether the output of the model be zero. By setting the initial value of BSDE be zero, the test is passed and we can make sure that the implementation of the whole model is right.

3. We test the core method **train** in the class. Before testing train method, we test the two functions: First we test **grad** with some simple differentiable functions to make sure that the gradient which we calculate is right. Second we test the loss function with the dimension be 1 and time step be 2 just as in testing **NonSharedModel**. We compare the value of loss function output by the model and the one calculated by hand. With these two test passed, we test the function **train**. We set the parameter as same as in the product and let the trainer run for 4000 iterations. At the end of training, we generate a batch of validating data and calculate the gradient and the loss value. We check two things: 1.Whether the gradient of model is nearly a zero tensor which is a necessary condition of local optimality. 2.Whether the loss is near to zero, which is also the minimum value of the loss function. Combine with this two condition, we can make sure the the loss function come to a global minimum. After calculating these issues, we find that the 2-norm of the gradient with respect to weights is smaller than 0.000001 and nearly all of the entry of the gradient tensor become zero. The value of loss is smaller than 0.000000001 and does not change for 1000 iterations. We can conclude that the training process comes to global minimum and our implementation of **train** is right.

After the unit test, we do an integration test to test our pipeline. The product we use is described in the last subsection and the test is passed. We list the numerical results to show that the test is passed. We list the numerical results in subsection 5.2.

	clean initial value	CVA initial mean	CVA CI
solver solution	0.07791	0.007427	-
monte carlo	-	0.007424	$[0.007383, 0.007466]$
analytical solution	0.07783	0.007419	-

Table 2: Numerical results

After this, we compute the value of the collateral balance C corresponding to the simulated paths of $\hat{V}_t$, which in turn allows us to compute the post-collateral exposure process $\hat{V}_t - C_t$ that enters the xVA formulas. For illustration, we consider the one-time swap as the example. We introduce a simple example of a collateral agreement where collateral is exchanged between the counterparties at every point in time (a margin call frequency that does not coincide with the simulation time discretization can of course be treated as well). Collateral is exchanged only in case the pre-collateral exposure is above (below) a receiving (posting) threshold which are both set equal to 0.078

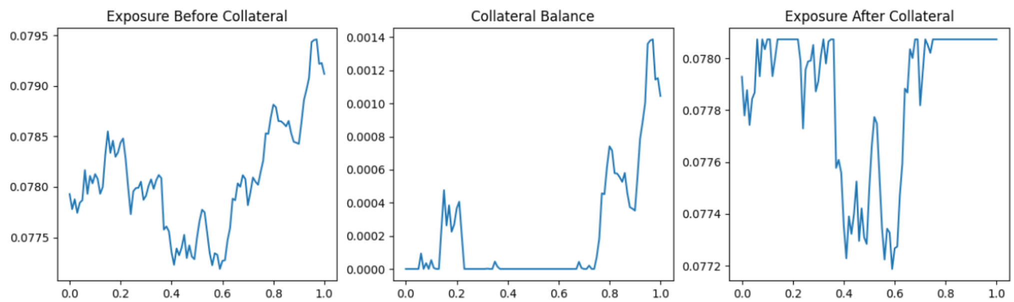

Figure 1: Pathwise simulation of a collateralized exposure. Left: $\hat{V}_t$.Middle: C_t. Right: $\hat{V}_t - C_t$. Posting and receiving threshold are 0.078

6 Conclusions and Extensions

We extent the framework of Deep-XVA solver to rate products and use the price of zero coupon bonds as the assets just as equity in previous work. Then throughout training the network of diffusion at each time step with the input of price of zero coupon bond, we can simulate the clean value of rate product with the terminal value. Given the analytical clean value, we can find that the value of adjustment solved by recursive method falls into the confidence interval of the outer Monte-Carlo method, which can also justify our application make sense on swap type product. For further discussion, one thing is to check the correctness on multiple swap product. This can be achieved with two methods: First, we calculate the price of each CVA with terminal payoff and then add them. Second, we can add the cashflow stream A_t into drift term and directly train the BSDE solver,however, this method should be justified later. Another thing is that we can develop framework to evaluate Greeks of value of adjustment, which is useful for financial industry. Since this is the simplest case of rate related product, we will explore more rate products and moreover do more research on portfolio of equity, fixed income and credit products. In order to solve the problem, another approach is to convert the FBSDE to the semi-linear PDE and use numerical methods to solve it. However traditional methods such as FDM and FEM have dependence on discretization of state space and the traditional deep learning based methods suffers from the same computational issue as classical methods: the optimization problem needs to be solved for every new instance. Then we ca try to use neural operator to solve a class of PDE. Fourier Neural Operator is a candidate since it is the first work that learns the resolution-invariant solution operator and the FFT method in the fourier layer can enhance the efficiency. After the solution of the assets in the portfolio obtained, we can use the trained operator to look for the solution corresponded to the portfolio.

Appendix A Short-Rate and Zero-Coupon Bond Price

Claim I The price of zero coupon bond under Versicek model is:

$$\begin{aligned}
P(t,T) &= A(t,T)e^{-B(t,T)r_t} \\
B(t,T) &= \frac{1}{\kappa}(1-e^{-\kappa(T-t)}) \\
A(t,T) &= \exp\left\{\frac{(B(t,T)-T+t)\left(\kappa^2\theta-\sigma^2/2\right)}{\kappa^2}-\frac{\sigma^2B^2(t,T)}{4\kappa}\right\}
\end{aligned} \tag{55}$$

Proof: We set the price has the form $P(t,T)=A(t,T)e^{-B(t,T)r_t}$. Then by Ito Lemma and the fact that the $P(t,T)/B_t$ is the martingale. We have the drift be zero under any short rate. Then we have the following ODEs:

$$\begin{cases} -A+\kappa AB-AB_t=0 \\ A_t-AB\theta+\frac{1}{2}\sigma^2AB^2=0 \end{cases}$$

Then we have:

$$\begin{cases} B_t=\kappa B-1, B(T)=0 \\ A_t=\kappa AB-\frac{1}{2}\sigma^2AB^2, A(T)=1 \end{cases}$$

Then we can solve $A(t,T)$ and $B(t,T)$. **Claim II** If the dynamics of the short-rate is under Versicek model with parameter (θ,κ,σ). Then the short rate process and it's corresponded HJM model are:

$$\begin{cases} \mathrm{d}r_t=\kappa(\theta-r_t)\mathrm{d}t+\sigma\mathrm{d}W_t^{\mathbb{Q}} \\ \mathrm{d}P(t,\tau)=P(t,\tau)[r_t\mathrm{d}t+\boldsymbol{\Sigma}^{\mathrm{T}}(t,\tau)\,\mathrm{d}W_t^{\mathbb{Q}}] \end{cases}$$

Where $W_t^{\mathbb{Q}}$ is the $\mathbb{Q}$-Brownian motion.

Proof: By Ito's Lemma, we have the dynamics of the bond price $P(t,\tau)$:

$$\mathrm{d}P(t,\tau)=P_t(t,\tau)\mathrm{d}t+P_r(t,\tau)\mathrm{d}r_t+\frac{1}{2}P_{rr}(t,\tau)\mathrm{d}[r,r]_t$$

Since we have the analytical form of the bond price. We plug 38 into the dynamic. We then havethe HJM process whose volatility term: $\int_0^{T_0}\boldsymbol{\Sigma}^{\mathrm{T}}(t,T_0)\,\mathrm{d}\mathbf{W}_t^{\mathbb{Q}}$. In Hull White model, $\boldsymbol{\Sigma}(t,T_i)=-\int_t^T\sigma(t,s)\mathrm{d}s=\frac{\sigma}{\kappa}(1-e^{-\kappa(T-t)})$

Appendix B Decomposing of IRS CVA

Claim I The CVA of an interest rate swap (hereinafter "CVA_{IRS}") with n swap payments $(t=T_0\leq T_1\leq\cdots\leq T_n=T)$ can be decomposed to the summation of CVA_{IRS} of each time intervals.

$$\mathrm{CVA}_{IRS}(t,T)=\sum_{i=0}^{n-1}\mathrm{CVA}_{IRS}\left(t,T_i,T_{i+1}\right)$$

where $\mathrm{CVA}_{IRS}\left(t,T_i,T_{i+1}\right)$ is the expected value of the loss if the counterparty defaults between the times T_i and T_{i+1}. More rigorously,for $i=0,...,n-1$, setting $\mathrm{LGD}=1$.

Proof: By definition, CVA of IRS can be written as:

$$\begin{aligned}
\mathrm{CVA}_{IRS}(t,T) &= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{I}_{[\tau\leq T]}\cdot V(\tau,T)^+D(t,\tau)\right] \\
&= \sum_{i=0}^{n-1}\mathbb{E}^{\mathbb{Q}}\left[\mathbb{I}_{[T_i\leq\tau\leq T_{i+1}]}\cdot V(\tau,T)^+D(t,\tau)\right]
\end{aligned}$$

However, when time goes between T_i and T_{i+1}, we can notice that the swap we are facing is the one begins at T_{i+1} and end at T, since swap from T_0 to T_i has been finished. Then we have the following equation:

$$\mathbb{I}_{[T_i\leq\tau\leq T_{i+1}]}\cdot V(\tau,T)=\mathbb{I}_{[T_i\leq\tau\leq T_{i+1}]}\cdot V(\tau,T_{i+1},T)$$

Therefore we have

$$\begin{aligned}
\mathrm{CVA}_{IRS}(t,T)&=\mathbb{E}^{\mathbb{Q}}\left[\mathbb{I}_{[\tau\leq T]}\cdot V(\tau,T)^{+}D(t,\tau)\right]\\
&=\sum_{i=0}^{n-1}\mathbb{E}^{\mathbb{Q}}\left[\mathbb{I}_{[T_i\leq\tau\leq T_{i+1}]}\cdot V(\tau,T)^{+}D(t,\tau)\right]\\
&=\sum_{i=0}^{n-1}\mathbb{E}^{\mathbb{Q}}\left[\mathbb{I}_{[T_i\leq\tau\leq T_{i+1}]}\cdot V(\tau,T_{i+1},T)^{+}D(t,\tau)\right]\\
&=\sum_{i=0}^{n-1}\mathrm{CVA}_{IRS}\left(t,T_i,T_{i+1}\right)
\end{aligned}$$

Claim II

$$\mathrm{CVA}_{IRS}\left(t,T_i,T_{i+1}\right)=\overline{CVA}^{T_{i+1}}_{[t,T_{i+1}]}-\overline{CVA}^{T_{i+1}}_{[t,T_i]}$$

where $\overline{CVA}^{T_{i+1}}_{[t,T_{i+1}]}$ is the value of the CVA with swap beginning at T_{i+1} and default time $[t,T_{i+1}]$ and $\overline{CVA}^{T_{i+1}}_{[t,T_i]}$ is the value of the CVA with swap beginning at T_{i+1} and default time $[t,T_i]$ (both swaps endding at T).

Proof: We can prove this claim by integral representation of CVA value:

$$\begin{aligned}
\mathrm{CVA}_{IRS}\left(t,T_i,T_{i+1}\right)&=\mathbb{E}^{\mathbb{Q}}\left[\mathbb{I}_{[T_i\leq\tau\leq T_{i+1}]}\cdot V(\tau,T_{i+1},T)^{+}D(t,\tau)\right]\\
&=\mathbb{E}^{\mathbb{Q}}\left[\int_{T_i}^{T_{i+1}}V(u,T_{i+1},T)^{+}D(t,\tau)\mathrm{d}S(u)\right]\\
&=\mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T_{i+1}}V(u,T_{i+1},T)^{+}D(t,\tau)\mathrm{d}S(u)\right]\\
&-\mathbb{E}^{\mathbb{Q}}\left[\int_{t}^{T_i}V(u,T_{i+1},T)^{+}D(t,\tau)\mathrm{d}S(u)\right]\\
&=\overline{CVA}^{T_{i+1}}_{[t,T_{i+1}]}-\overline{CVA}^{T_{i+1}}_{[t,T_i]}
\end{aligned}$$

Where $\overline{CVA}^{T_{i+1}}_{[t,T_{i+1}]}$ swaps starting at T_{i+1}, whose default time interval is $[t,T_{i+1}]$; $\overline{CVA}^{T_{i+1}}_{[t,T_{i+1}]}$ swaps starting at T_{i+1}, whose default time interval is $[t,T_i]$. We can easily find that the $\overline{CVA}^{T_2}_{[t,T_1]}-$type IRS-CVA can be easily priced under Deep-xVA framework.

References

Cerny, J., Witzany, J. 2014. Interest Rate Swap Credit Valuation Adjustment. *Working Paper* https://doi.org/10.3905/jod.2015.23.2.024

Biagini F, Gnoatto A, Oliva I. A unified approach to xVA with CSA discounting and initial margin[J]. Available at SSRN 3394928, 2021.

Gnoatto, A., Picarelli, A., Reisinger, C. 2020. Deep xVA solver – A neural network based counterparty credit risk management framework. *arXiv e-prints* page arXiv:2005.02633

Crépey S, Bielecki T R, Brigo D. Counterparty risk and funding: A tale of two puzzles[M]. Chapman and Hall/CRC, 2014.