

## Deep xVA under a structural model

Joint work: **Kristoffer Andersson**<sup>1</sup> & Alessandro Gnoatto

12th General AMaMeF Conference  
Università Stagli di Verona,  
June 22, 2025

---

<sup>1</sup>Email address: [kristofferherbert.andersson@univr.it](mailto:kristofferherbert.andersson@univr.it)

# Table of contents

- 1 A structural xVA model
- 2 A system of BSDEs describing the xVAs
- 3 Deep learning based numerical method
- 4 Numerical example

## Financial setup

Consider a portfolio of  $P \in \mathbb{N}^+$  derivatives (options, swaps, swaptions, etc.) traded between a bank ( $\mathcal{B}$ ) and some financial counterparty ( $\mathcal{C}$ ). **Clean values** given by  $\widehat{V}^1, \widehat{V}^2, \dots, \widehat{V}^P$  and the portfolio value  $\widehat{V} = \sum_{j=1}^P \widehat{V}^j$ . Each derivative written on a subset of a  $\mathbb{N}^+ \ni d$ -dimensional underlying assets process,  $\widehat{X}$ . Different maturities  $T_1, T_2, \dots, T_P$ .

In an xVA context, we adjust  $\widehat{V}$  by:

- **CVA/DVA** (Credit/Debit Valuation Adjustment): default risk of the counterparty/bank,
- **CoIVA** (Collateral VA): different lending/borrowing rates for collateral,
- **MVA** (Margin VA): funding costs/benefits for Initial Margin,
- **FVA** (Funding VA): additional funding costs (e.g., hedging market risk).

**Focus of this presentation:** We concentrate on modeling

- (i) defaults (structural default model),
- (ii) initial margin (as a risk measure),
- (iii) a change of measure technique for rare events (particularly defaults).

## A structural default model

In a *structural model* of default, the evolution of each firm's ( $\mathcal{B}/\mathcal{C}$ ) assets is modeled directly, and default is triggered when the asset value falls below a specified boundary.

$$X_t^{\mathcal{B}} = 1 + \int_0^t b^{\mathcal{B}}(s, X_s^{\mathcal{B}}) ds + \int_0^t \sigma^{\mathcal{B}}(s, X_s^{\mathcal{B}}) dW_s^{\mathcal{B}}, \quad (1)$$

$$X_t^{\mathcal{C}} = 1 + \int_0^t b^{\mathcal{C}}(s, X_s^{\mathcal{C}}) ds + \int_0^t \sigma^{\mathcal{C}}(s, X_s^{\mathcal{C}}) dW_s^{\mathcal{C}}, \quad (2)$$

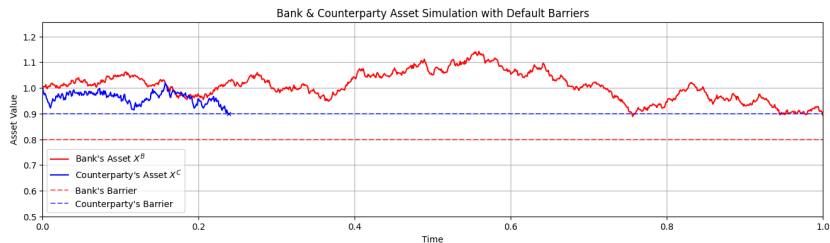
Default times given by the following stopping times

$$\tau^{\mathcal{B}} = \inf\{t \in (0, T]: X_t^{\mathcal{B}} \leq \xi_t^{\mathcal{B}}\}, \quad \tau^{\mathcal{C}} = \inf\{t \in (0, T]: X_t^{\mathcal{C}} \leq \xi_t^{\mathcal{C}}\}. \quad (3)$$

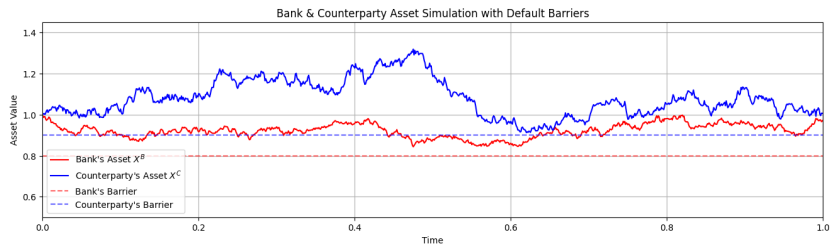
Moreover, define  $\tau = \tau^{\mathcal{B}} \wedge \tau^{\mathcal{C}}$  and maturity of netting set given by  $\tau \wedge T$ .

# Visualization of a structural default model

Counterparty defaults,  $\tau = \tau^C$ , and portfolio maturity  $\tau^C$ :



No default,  $\tau > T$ , and portfolio maturity  $T$ :



## xVA BSDEs

$$\widehat{V}_t^j = - \int_{(t,T]} dA_s^j - \int_{(t,T]} r_s \widehat{V}_s^j ds - \int_{(t,T]} \widehat{Z}_s^{\mathcal{I}j} \cdot d\widehat{W}_s^{\mathcal{I}j}, \quad j \in \{1, \dots, P\}, \quad \widehat{V} = \sum_{j=1}^P \widehat{V}^j,$$

$$-\text{ColVA}_t = \int_{(t,\tau \wedge T]} (f^{\text{ColVA}}(s, C_s) - r_s \text{ColVA}_s) ds - \int_{(t,\tau \wedge T]} Z_s^{\text{ColVA}} \cdot dW_s,$$

$$\text{IM}_t^{\text{FC}} = \text{VaR}_\alpha((\widehat{V}_{t+\text{MPR}_t} - \widehat{V}_t)^+ | \widehat{\mathcal{F}}_t), \quad \text{IM}_t^{\text{TC}} = -\text{VaR}_{1-\alpha}((\widehat{V}_{t+\text{MPR}_t} - \widehat{V}_t)^- | \widehat{\mathcal{F}}_t),$$

$$\text{CVA}_t = \mathbf{1}_{\{\tau \leq T\}} \mathbf{1}_{\{\tau = \tau^{\mathcal{C}}\}} \text{LGD}^{\mathcal{C}}(\widehat{V}_\tau - C_\tau - \text{IM}_\tau^{\text{FC}})^+ - \int_{(t,\tau \wedge T]} r_s \text{CVA}_s ds - \int_{(t,\tau \wedge T]} Z_s^{\text{CVA}} \cdot dW_s,$$

$$\text{DVA}_t = \mathbf{1}_{\{\tau \leq T\}} \mathbf{1}_{\{\tau = \tau^{\mathcal{B}}\}} \text{LGD}^{\mathcal{B}}(\widehat{V}_\tau - C_\tau - \text{IM}_\tau^{\text{TC}})^- - \int_{(t,\tau \wedge T]} r_s \text{DVA}_s ds - \int_{(t,\tau \wedge T]} Z_s^{\text{DVA}} \cdot dW_s,$$

$$-\text{MVA}_t = \int_{(t,\tau \wedge T]} (f^{\text{MVA}}(s, \text{IM}_s^{\text{FC}}, \text{IM}_s^{\text{TC}}) - r_s \text{MVA}_s) ds - \int_{(t,\tau \wedge T]} Z_s^{\text{MVA}} \cdot dW_s,$$

$$\mathbf{xVA}_t = (\text{ColVA}_t, \text{CVA}_t, \text{DVA}_t, \text{MVA}_t, \text{FVA}_t),$$

$$-\text{FVA}_t = \int_{(t,\tau \wedge T]} (f^{\text{FVA}}(s, \widehat{V}_s, C_s, \mathbf{xVA}_s, \text{IM}_s^{\text{TC}}) - r_s \text{FVA}_s) ds - \int_{(t,\tau \wedge T]} Z_s^{\text{FVA}} \cdot dW_s.$$

## General form of the BSDEs

Forward SDE given by  $X := \text{Concat}(\hat{X}, X^B, X^C)$ , where  $\hat{X}$  is the  $d$ -dimensional underlying assets process.

Each xVA BSDE is on the form:

$$\begin{cases} X_t = x_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) \odot dW_s, \\ Y_t = 1_{\{\tau \leq T\}} \chi_\tau - \int_{(t, \tau \wedge T]} d\Lambda_s + \int_{(t, \tau \wedge T]} (f_s^Y - r_s Y_s) ds - \int_{(t, \tau \wedge T]} Z_s \cdot dW_s, \end{cases} \quad (4)$$

**Non-standard features:**

- ❶ The process  $(f_t^Y)_{t \in \{0, \tau \wedge T\}}$  could be reformulated as a Markovian function of  $X$ , and  $\mathcal{L}[\hat{V}]$  (the IM). It is actually an **anticipated McKean–Vlasov BSDE**.
- ❷ We have a random terminal time given by  $\tau \wedge T$ , a **BSDE with stopping time**.

## BSDE under a new measure

**Problem:** Default events are rare and is not well captured by the neural network unless the training data is huge.

**Solution:** Use a change of measure technique to increase the default probabilities.

---

Subtract  $q$  from the drift in the forward SDE and compensate by  $\langle q \odot \sigma, Z \rangle$  in the BSDE driver:

$$\begin{cases} X_t^q = x_0 + \int_0^t (b(s, X_s^q) - q(s, X_s^q)) ds + \int_0^t \sigma(s, X_s^q) \odot dW_s, \\ Y_t^q = \mathbb{I}_{\{\tau^q \leq T\}} X_{\tau^q}^q - \int_{(t, \tau^q \wedge T]} d\Lambda_s^q \\ \quad + \int_{(t, \tau^q \wedge T]} (f_s^{Y, q} - r_s Y_s^q - \langle q(s, X_s^q) \odot \sigma(s, X_s^q), Z_s^q \rangle) ds - \int_{(t, \tau^q \wedge T]} Z_s^q \cdot dW_s. \end{cases} \quad (5)$$

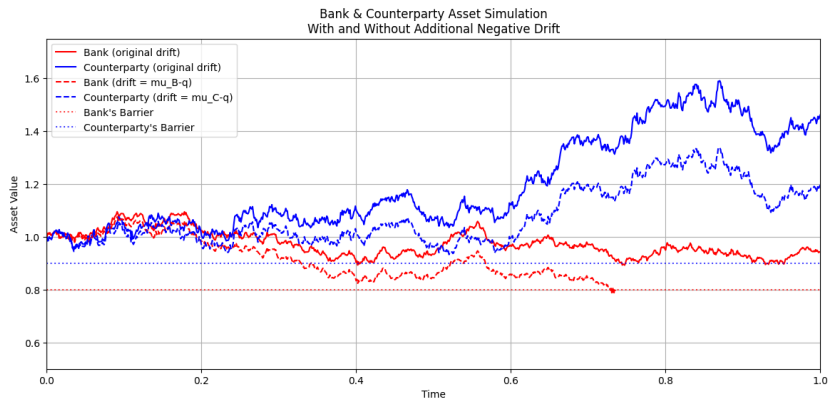
For any  $q$  (regular enough), it holds that:

- The initial condition,  $Y_0^q \equiv Y_0$ ,
- The associate PDE, with solution  $u$ , has the stochastic representation  $Y_t^q = u(t, X_t^q)$ , i.e., for any  $q$ , the associate PDE is the same.

# Visualization of the structural default model under a measure change

**Under original measure:** No default,  $\tau > T$ , and portfolio maturity  $T$

**Under new measure:** Counterparty defaults,  $\tau^q = \tau^{B,q}$ , and portfolio maturity  $\tau^{B,q}$



## Equivalent formulation as a variational problem

The deep BSDE method relies on the reformulation of the FBSDE into the variational problem:

$$\left\{ \begin{array}{l} \inf_{y_0, (Z_t^q)_{t \in [0, T]}} \mathbb{E} \left[ \left| \mathbb{I}_{\{\tau^q \leq T\}} \chi_{\tau^q}^q - \Delta \Lambda_{\tau^q \wedge T}^q - Y_{\tau^q \wedge T}^q \right|^2 \right], \quad \text{where,} \\ X_t^q = x_0 + \int_0^t (b(s, X_s^q) - q(s, X_s^q)) ds + \int_0^t \sigma(s, X_s^q) \odot dW_s, \\ Y_t^q = y_0 + \int_{(0, t]} d\Lambda_s^q - \int_{(0, t]} (f_s^{Y, q} - r_s Y_s^q - \langle q(s, X_s^q) \otimes \sigma(s, X_s^q), Z_s^q \rangle) ds \\ \quad + \int_{(0, t]} Z_s^q \cdot dW_s. \end{array} \right. \quad (6)$$

### Motivation:

- A solution to (5) (the FBSDE) solves (6),
- By well-posedness of the FBSDE, this solution is unique.

## Variational problem for multiple FBSDEs

We want to solve the FBSDE under multiple different measures generated by shifting the drift by  $q_1, q_2, \dots, q_K$

$$\left\{ \begin{array}{l} \inf_{y_0, (Z_t^{q_1}, \dots, Z_t^{q_K})_{t \in [0, T]}} \sum_{i=1}^K \mathbb{E} \left[ \left| \mathbb{I}_{\{\tau^{q_i} \leq T\}} \chi_{\tau^{q_i}}^{q_i} - \Delta \Lambda_{\tau^{q_i} \wedge T}^{q_i} - Y_{\tau^{q_i} \wedge T}^{q_i} \right|^2 \right], \\ \text{where for } i \in \{1, 2, \dots, K\} : \\ X_t^{q_i} = x_0 + \int_0^t (b(s, X_s^{q_i}) - q_i(s, X_s^{q_i})) ds + \int_0^t \sigma(s, X_s^{q_i}) \odot dW_s, \\ Y_t^{q_i} = y_0 + \int_{(0, t]} d\Lambda_s^{q_i} - \int_{(0, t]} (f_s^{Y, q_i} - r_s Y_s^{q_i} - \langle q_i(s, X_s^{q_i}) \oslash \sigma(s, X_s^{q_i}), Z_s^{q_i} \rangle) ds \\ \quad + \int_{(0, t]} Z_s^{q_i} \cdot dW_s. \end{array} \right.$$

- All BSDEs have the same initial condition, hence only one  $y_0^*$ ,
- Functional form  $Z$  is the same for all  $q_i$ , i.e.,  $Z_t^{q_i} = Z(t, X_t^{q_i})$ ,
- Solving multiple BSDEs makes default modeling more accurate. Can also increase robustness.

## Semi-discrete variational problem

Equidistant time grid  $\pi := \{0 = t_0, t_1, \dots, t_N = T\}$ , with  $h = t_{n+1} - t_n$  and Brownian increments  $\Delta W_n = W_{n+1} - W_n$ . The discretized stopping times are given by  $n_{\tau}^i \mathcal{B} = \inf\{n \in \mathbb{N}^+ : [X_n^{\pi, q_i}]_{d+1} \leq \xi_{t_n}^1\}$ ,  $n_{\tau}^i \mathcal{C} = \inf\{n \in \mathbb{N}^+ : [X_n^{\pi, q_i}]_{d+2} \leq \xi_{t_n}^2\}$ .

**Time discrete formulation:**

$$\left\{ \begin{array}{l} \mathcal{J}(\mathbf{y}_0, \mathcal{Z}) := \inf_{\mathbf{y}_0, \mathcal{Z}} \sum_{i=1}^K \mathbb{E} \left[ \left| \mathbb{I}_{\{n_{\tau}^i \leq N\}} \chi_{n_{\tau}^i}^{\pi, q_i} - \Delta \Lambda_{n_{\tau}^i \wedge N}^{\pi, q_i} - Y_{n_{\tau}^i \wedge N}^{\pi} \right|^2 \right], \text{ s.t. for } i \in \{1, \dots, K\} : \\ \chi_n^{\pi, q_i} = x_0 + \sum_{k=0}^{n-1} [b(t_k, X_k^{\pi, q_i}) - q_i(t_k, X_k^{\pi, q_i})] h + \sum_{k=0}^{n-1} \sigma(t_k, X_k^{\pi, q_i}) \odot \Delta W_k, \\ Y_n^{\pi, q_i} = \mathbf{y}_0 + \sum_{k=0}^{n-1} \Delta \Lambda_k^{\pi, q_i} + \sum_{k=0}^{n-1} Z_k^{\pi, q_i} \cdot \Delta W_k \\ \quad - \sum_{k=0}^{n-1} (f_k^{\pi, Y, q_i} - r_k^{\pi} Y_k^{\pi, q_i} - \langle q_i(t_k, X_k^{\pi, q_i}) \otimes \sigma(t_k, X_k^{\pi, q_i}), Z_k^{\pi, q_i} \rangle) h \\ Z_n^{\pi, q_i} = \mathcal{Z}(t_n, X_n^{\pi, q_i}), \end{array} \right. \quad (7)$$

Represent  $\mathbf{y}_0$  with a trainable parameter and represent  $\mathcal{Z}$  with a neural network.  
Optimize subject to objective functional  $\mathcal{J}$ .

- ① Stopping times approximated ✓,
- ② Still no approximation of initial margin (McKean–Vlasov components) ✗. Computed through neural network based quantile regression (see paper for details).

## High-level picture of full algorithm

Deeper layers depend on approximations obtained from higher layers.

- 1 Approximation of the clean value BSDE,
- 2 Approximation of the mapping generating the risk measure value at risk,
- 3 Approximation of the CoIVA, CVA, DVA and MVA BSDEs,
- 4 Approximation of the FVA BSDE,

In this presentation we compare our approximations with reference solutions obtained from (extremely expensive) nested Monte–Carlo sampling.

- Model the underlying processes with a 7-dimensional Geometric Brownian Motion,
- A portfolio of  $P = 33$  European basket options,
- In total, 93 dimensional problem,

$$\sigma = \begin{pmatrix} 0.2 \\ 0.25 \\ 0.25 \\ 0.25 \\ 0.3 \\ 0.2 \\ 0.3 \end{pmatrix}, \quad \rho = \begin{pmatrix} 1.0 & 0.9 & 0.2 & 0.5 & 0.1 & 0.1 & 0.2 \\ 0.9 & 1.0 & 0.4 & 0.3 & 0.2 & 0.3 & 0.2 \\ 0.2 & 0.4 & 1.0 & 0.2 & 0.75 & 0.15 & 0.25 \\ 0.5 & 0.3 & 0.2 & 1.0 & 0.35 & 0.05 & 0.15 \\ 0.1 & 0.2 & 0.75 & 0.35 & 1.0 & 0.15 & 0.05 \\ 0.1 & 0.3 & 0.15 & 0.05 & 0.15 & 1.0 & 0.25 \\ 0.2 & 0.2 & 0.25 & 0.15 & 0.05 & 0.25 & 1.0 \end{pmatrix}.$$

$$T^P = \{1, 1, 1, 0.8, 0.8, 0.6, 0.6, 0.4, 0.4, 0.2, 0.2, 1, 1, 1, 0.7, 0.7, 0.5, 0.5, 0.3, 0.3, 0.1, 0.1, 1, 0.7, 0.7, 0.5, 0.5, 0.3, 1, 0.8, 0.8, 0.6, 0.6\},$$

$$K^P = \{1.05, 1.1, 1.05, 1.05, 0.7, 0.7, 0.75, 0.75, 1, 0.9, 0.8, 0.9, 1.1, 1.05, 0.85, 0.9, 0.9, 1.05, 1, 1, 0.9, 0.95, 1.05, 0.7, 0.7, 0.75, 0.75, 1, 0.9, 0.8, 0.9, 1.1, 1.05\}, \text{ and}$$

$$\begin{aligned} \mathcal{I} &= \{\mathcal{I}_1, \mathcal{I}_2, \dots, \mathcal{I}_{33}\} \\ &= \{[1, 2, 3, 4, 5], [2, 3, 4, 5], [1, 3, 4, 5], [1, 2, 4, 5], [1, 2, 3, 5], \\ &\quad [1, 2, 3, 4], [1, 2, 3], [1, 2, 4], [1, 2, 5], [1, 3, 4], \\ &\quad [1, 3, 5], [2, 3, 4], [2, 3, 5], [3, 4, 5], [1, 2], \\ &\quad [2, 3], [1, 3], [1], [2], [3], [4], [5], \\ &\quad [1, 2], [2, 3], [1, 3], [1], [1, 2, 3, 4, 5], [1, 2, 3, 4, 5], \\ &\quad [1, 2, 3, 4, 5], [2, 3, 4, 5], [1, 2, 3], [1, 2], [2, 3]\}. \end{aligned}$$

## Clean values

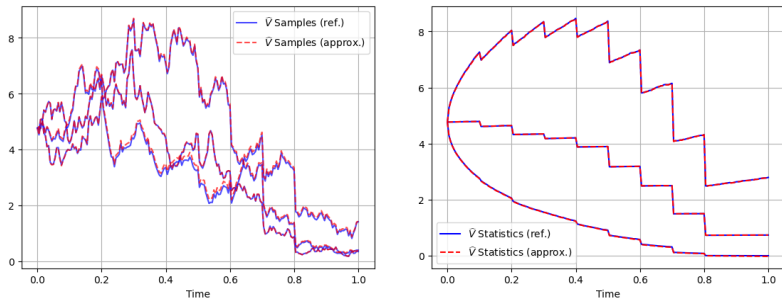


**Figure:** Approximate portfolio values compared with their analytical counterparts. **Left:** Three representative samples. **Right:** Empirical mean, 99th and 1st percentiles.

## Initial Margin

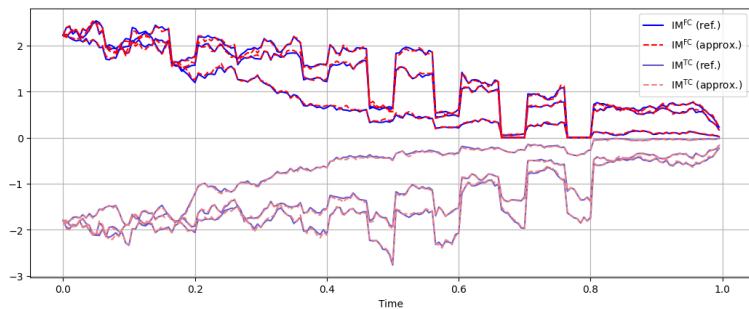


**Figure:** Three representative samples of approximate IM compared with reference solutions obtained by nested Monte-Carlo sampling.

## MVA, CVA, ColVA and FVA

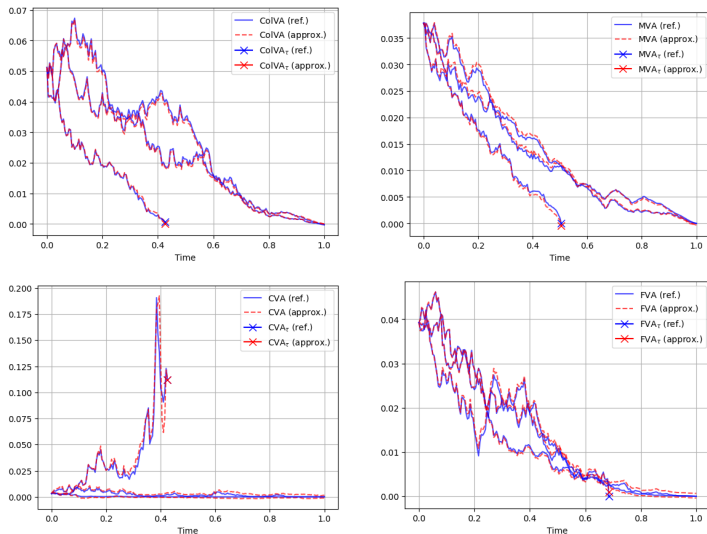


**Figure:** The approximate ColVA, MVA, CVA, and FVA compared with their respective reference solutions.

## Effect of measure change

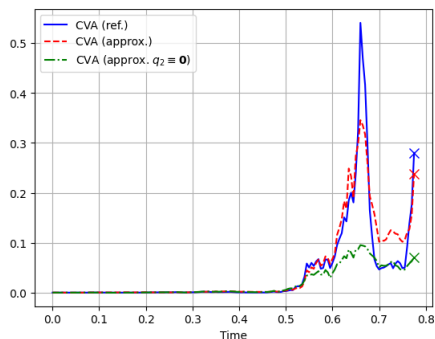
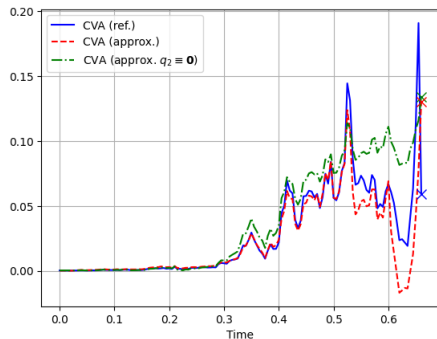


**Figure:** Comparison of the reference solution with our approximations (with and without the measure change technique) for representative scenarios where the counterparty defaults at  $t = 0.66$  (Panel 2) and  $t = 0.775$  (Panel 3).

# Contact Information

## Email address

kristofferherbert.andersson@univr.it

## Related Publications

- Andersson, Kristoffer, and Gnoatto, Alessandro. *Multi-Layer Deep xVA: Structural Credit Models, Measure Changes and Convergence Analysis*. arXiv preprint arXiv:2502.14766 (2025).
- Andersson, Kristoffer, Andersson, Adam, and Cornelis, W. Oosterlee. *The Deep Multi-FBSDE Method: A Robust Deep Learning Method for Coupled FBSDEs*. arXiv preprint arXiv:2503.13193. (2025).
- Andersson, Kristoffer, Andersson, Adam, and Cornelis, W. Oosterlee. "Convergence of a Robust Deep FBSDE Method for Stochastic Control." *SIAM Journal on Scientific Computing* 45, (1), (2023): A226–A255.

# A posteriori error bounds

## Theorem

*Under appropriate assumptions, for every  $\epsilon \in (0, 1/2)$  there is a constant  $C$  such that*

$$\begin{aligned}
 \text{Err}^V(h)^2 &\leq C(h + \epsilon_V), \\
 \text{Err}^{\text{CoIVA}}(h)^2 &\leq C(h^{1/2-\epsilon} + \epsilon_V + \epsilon_{\text{CoIVA}}), \\
 \text{Err}^{\text{MVA}}(h)^2 &\leq C(h^{1/2-\epsilon} + \epsilon_V + \max_n \mathbb{E} |\text{IM}_{t_n} - \text{IM}_n^{\pi, \diamond}|^2 + \epsilon_{\text{MVA}}), \\
 \text{Err}^{\text{CVA}}(h)^2 &\leq C(h^{1/2-\epsilon} + \epsilon_V + \max_n \mathbb{E} |\text{IM}_{t_n} - \text{IM}_n^{\pi, \diamond}|^2 + \epsilon_{\text{CVA}}), \\
 \text{Err}^{\text{DVA}}(h)^2 &\leq C(h^{1/2-\epsilon} + \epsilon_V + \max_n \mathbb{E} |\text{IM}_{t_n} - \text{IM}_n^{\pi, \diamond}|^2 + \epsilon_{\text{DVA}}), \\
 \text{Err}^{\text{FVA}}(h)^2 &\leq C\left(h^{1/2-\epsilon} + \epsilon_V + \max_n \mathbb{E} |\text{IM}_{t_n} - \text{IM}_n^{\pi, \diamond}|^2 \right. \\
 &\quad \left. + \epsilon_{\text{CoIVA}} + \epsilon_{\text{MVA}} + \epsilon_{\text{CVA}} + \epsilon_{\text{DVA}} + \epsilon_{\text{FVA}}\right).
 \end{aligned}$$

## Definitions

$$\text{Err}^V(h)^2 := \sum_{j=1}^P \mathbb{E} \left[ \sup_{t \in [0, T]} |\hat{V}_t^j - \hat{V}_{\phi(t)}^{j, \pi, *}|^2 \right] + \sum_{j=1}^P \int_0^T \mathbb{E} |\hat{Z}_t^j - \hat{Z}_{\phi(t)}^{j, \pi, *}|^2 dt,$$

and for  $(x, y) \in \{(\text{Col}, *), (\text{M}, \diamond), (\text{C}, \diamond), (\text{D}, \diamond), (\text{F}, \bullet)\}$

$$\text{Err}^{x\text{VA}}(h)^2 := \mathbb{E} \left[ \sup_{t \in [0, T]} |x\text{VA}_t - x\text{VA}_{\phi(t)}^{\pi, y}|^2 \right] + \int_0^T \mathbb{E} |Z_t^{x\text{VA}} - Z_{\phi(t)}^{x\text{VA}, \pi, y}|^2 dt,$$

and we define the a posteriori error terms as

$$\begin{aligned} \epsilon_{Vj} &:= \mathbb{E} |\hat{V}_t^j - \hat{V}_{\phi(t)}^{j, \pi, *}|^2, & \epsilon_{\text{ColVA}} &:= \mathbb{E} |\text{ColVA}_{n_\tau \pi \wedge N}^{\pi, \diamond}|^2, & \epsilon_{\text{MVA}} &:= \mathbb{E} |\text{MVA}_{n_\tau \pi \wedge N}^{\pi, \square}|^2, \\ \epsilon_{\text{CVA}} &:= \mathbb{E} |\text{CVA}_{n_\tau \pi \wedge N}^{\pi, \square} - \mathbb{1}_{\{n_\tau \leq N\}} \mathbb{1}_{\{n_\tau = n_{\tau C}\}} \text{LGD}^C(\hat{V}_{n_\tau}^{\pi, *} - C_{n_\tau}^{\pi, *} - \text{IM}_\tau^{\text{FC}, \pi, \diamond})^+|^2, \\ \epsilon_{\text{DVA}} &:= \mathbb{E} \left[ |\text{DVA}_{n_\tau \wedge N}^{\pi, \square} - \mathbb{1}_{\{n_\tau \leq N\}} \mathbb{1}_{\{n_\tau = n_{\tau B}\}} \text{LGD}^B(\hat{V}_{n_\tau}^{\pi, *} - C_{n_\tau}^{\pi, *} - \text{IM}_\tau^{\text{TC}, \pi, \diamond})^-|^2 \right], \\ \epsilon_{\text{FVA}} &:= \mathbb{E} |\text{FVA}_{n_\tau \pi \wedge N}^{\pi, \bullet}|^2. \end{aligned}$$

Finally, for  $t \in [0, T]$  and  $n \in \{0, 1, \dots, N-1\}$ , we denote by  $\text{IM}_t = \text{Concat}(\text{IM}_t^{\text{TC}}, \text{IM}_t^{\text{FC}})$  and  $\text{IM}_n^{\pi, \diamond} = \text{Concat}(\text{IM}_n^{\text{TC}, \pi, \diamond}, \text{IM}_n^{\text{FC}, \pi, \diamond})$ , the vector concatenations of the IMs to and from the counterparty and their approximations, respectively.

# Margin Valuation Adjustment - the control process

The control process with and without applying the measure change technique.

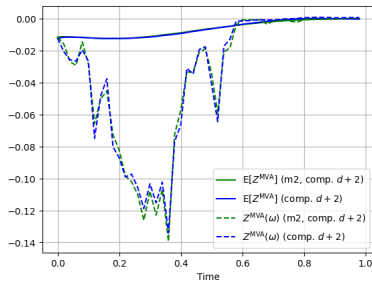
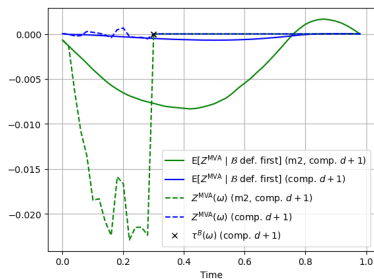


**Figure:** **Left:** When default probability is moderate/low. **Right:** When default probability is unrealistically high.

## Initial margin approximations

Reformulate VaR as an optimization problem:

$$\text{VaR}_\alpha \left( \widehat{V}_{t+\text{MPR}_t} - \widehat{V}_t \mid \widehat{\mathcal{F}}_t \right) = \arg \min_{q \in \mathbb{R}} \mathbb{E} \left[ \varkappa^\alpha(q; \widehat{V}_{t+\text{MPR}_t} - \widehat{V}_t) \mid \widehat{\mathcal{F}}_t \right],$$

where  $\varkappa^\alpha(q; x) = \max(\alpha \cdot (x - q), (\alpha - 1) \cdot (x - q))$ .

$$\left\{ \begin{array}{l} \underset{q_n^+, q_n^-}{\text{minimize}} \quad \mathbb{E} \left[ \varkappa^\alpha(q_n^+; (\widehat{X}_n^\pi, \widehat{V}_n^{\pi,*}); (\widehat{V}_{n+\text{MPR}_n^\pi}^{\pi,*} - \widehat{V}_n^{\pi,*})^+) \right] \\ \quad + \mathbb{E} \left[ \varkappa^\alpha(q_n^-; (\widehat{X}_n^\pi, \widehat{V}_n^{\pi,*}); (\widehat{V}_{n+\text{MPR}_n^\pi}^{\pi,*} - \widehat{V}_n^{\pi,*})^-) \right], \quad \text{where} \\ \widehat{X}_n^\pi = x_0 + \sum_{k=0}^{n-1} \widehat{b}(t_k, \widehat{X}_k^\pi) h + \sum_{k=0}^{n-1} \widehat{\sigma}(t_k, \widehat{X}_k^\pi) \cdot \Delta \widehat{W}_k, \\ \widehat{V}_n^{j,\pi,*} = F_n^{\widehat{V}^j}(\widehat{X}_{0:n-1}^\pi, \widehat{V}_{0:n-1}^{j,\pi,*}, \widehat{Z}_{0:n-1}^{j,\pi,*}, \Delta \widehat{W}_{0:n-1}^{\mathcal{I}_j}), \quad \widehat{V}_0^{j,\pi,*} = \widehat{v}_0^j, \quad j \in \{1, 2, \dots, P\} \\ \widehat{V}^{\pi,*} = \sum_{j=1}^P \widehat{V}^{j,\pi,*}, \quad \widehat{Z}_n^{j,\pi,*} = [\mathcal{Z}^*(t_n, \widehat{X}_n^\pi)]_{i=\mathcal{J}_j}^{\mathcal{J}_j+d_j}, \quad j \in \{1, 2, \dots, P\}. \end{array} \right. \quad (8)$$

Here  $\widehat{V}^{j,\pi,*}$  is approximate clean value of derivative  $j$ . Represent  $(q_n^+, q_n^-)$  and optimize subject to objective function.