Deep xVA under a structural model

Joint work: **Kristoffer Andersson**[1], Athena Picarelli & Alessandro Gnoatto

22nd Winter school on Mathematical Finance,
January 20, 2025

[1] Email address: **kristofferherbert.andersson@univr.it**

Table of contents

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 $\widehat{V}^1, \widehat{V}^2, \ldots, \widehat{V}^P$ 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 $T_1, T_2, \ldots, T_P$.

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

- **CVA/DVA** (Credit/Debit Valuation Adjustment): default risk of the counterparty/bank,
- **ColVA** (Collateral VA): differing 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.

$$X_t^{\mathcal{B}} = 1 + \int_0^t b^{\mathcal{B}}(s, X_s^{\mathcal{B}})\mathrm{d}s + \int_0^s \sigma^{\mathcal{B}}(s, X_s^{\mathcal{B}})\mathrm{d}W_s^{\mathcal{B}}, \tag{1}$$

$$X_t^{\mathcal{C}} = 1 + \int_0^t b^{\mathcal{C}}(s, X_s^{\mathcal{C}})\mathrm{d}s + \int_0^s \sigma^{\mathcal{C}}(s, X_s^{\mathcal{C}})\mathrm{d}W_s^{\mathcal{C}}, \tag{2}$$

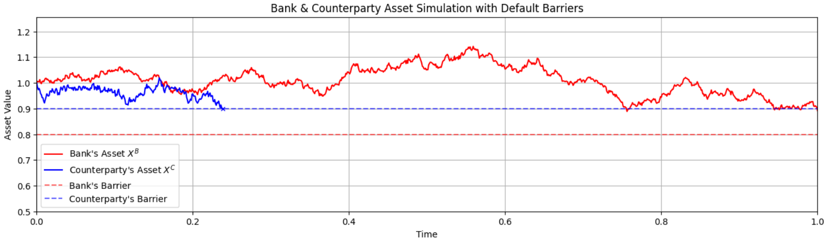
Default times given by the following stopping times

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

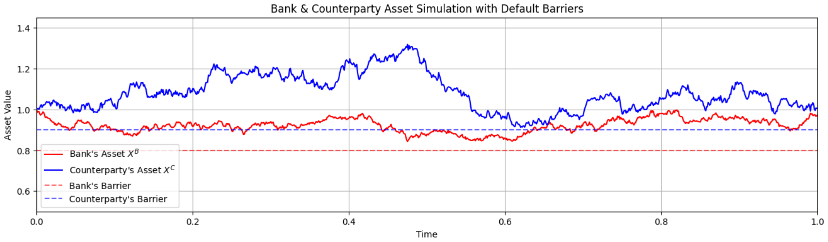
Moreover, define $\tau = \tau^{\mathcal{B}} \wedge \tau^{\mathcal{C}}$ and maturity of netting set given by $\tau \wedge T$.

Visualization of a structural default model

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



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



xVA BSDEs

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

$$-\mathrm{ColVA}_t = \int_{(t,\tau\wedge T)} \left(f^{\mathrm{ColVA}}(s, C_s) - r_s\, \mathrm{ColVA}_s\right) \mathrm{d}s - \int_{(t,\tau\wedge T]} Z_s^{\mathrm{ColVA}} \cdot \mathrm{d}W_s,$$

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

$$\mathrm{CVA}_t = \mathbf{1}_{\{\tau\le T\}}\, \mathbf{1}_{\{\tau=\tau^{\mathcal{C}}\}}\, \mathrm{LGD}^{\mathcal{C}} \left(\widehat{V}_\tau - C_\tau - \mathrm{IM}_\tau^{\mathrm{FC}}\right)^+ - \int_{(t,\tau\wedge T]} r_s\, \mathrm{CVA}_s\, \mathrm{d}s - \int_{(t,\tau\wedge T]} Z_s^{\mathrm{CVA}} \cdot \mathrm{d}W_s,$$

$$\mathrm{DVA}_t = \mathbf{1}_{\{\tau\le T\}}\, \mathbf{1}_{\{\tau=\tau^{\mathcal{B}}\}}\, \mathrm{LGD}^{\mathcal{B}} \left(\widehat{V}_\tau - C_\tau - \mathrm{IM}_\tau^{\mathrm{TC}}\right)^- - \int_{(t,\tau\wedge T]} r_s\, \mathrm{DVA}_s\, \mathrm{d}s - \int_{(t,\tau\wedge T]} Z_s^{\mathrm{DVA}} \cdot \mathrm{d}W_s,$$

$$-\mathrm{MVA}_t = \int_{(t,\tau\wedge T]} \left(f^{\mathrm{MVA}}(s, \mathrm{IM}_s^{\mathrm{FC}}, \mathrm{IM}_s^{\mathrm{TC}}) - r_s\, \mathrm{MVA}_s\right) \mathrm{d}s - \int_{(t,\tau\wedge T]} Z_s^{\mathrm{MVA}} \cdot \mathrm{d}W_s,$$

$$\mathbf{XVA}_t = (\mathrm{ColVA}_t, \mathrm{CVA}_t, \mathrm{DVA}_t, \mathrm{MVA}_t, \mathrm{FVA}_t),$$

$$-\mathrm{FVA}_t = \int_{(t,\tau\wedge T]} \left(f^{\mathrm{FVA}}(s, \widehat{V}_s, C_s, \mathbf{XVA}_s, \mathrm{IM}_s^{\mathrm{TC}}) - r_s\, \mathrm{FVA}_s\right) \mathrm{d}s - \int_{(t,\tau\wedge T]} Z_s^{\mathrm{FVA}} \cdot \mathrm{d}W_s.$$

General form of the BSDEs

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

Each xVA BSDE is on the form:

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

Non-standard features:

1. The process $(f_t^Y)_{t \in \{0, \tau \wedge T\}}$ could be reformulated as a Markovian function of X, $\mathcal{L}[\widehat{V}]$ (the IM). It is actually an **anticipated McKean–Vlasov BSDE**.
2. 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 $\langle q \oslash \sigma, Z \rangle$ in the BSDE driver:

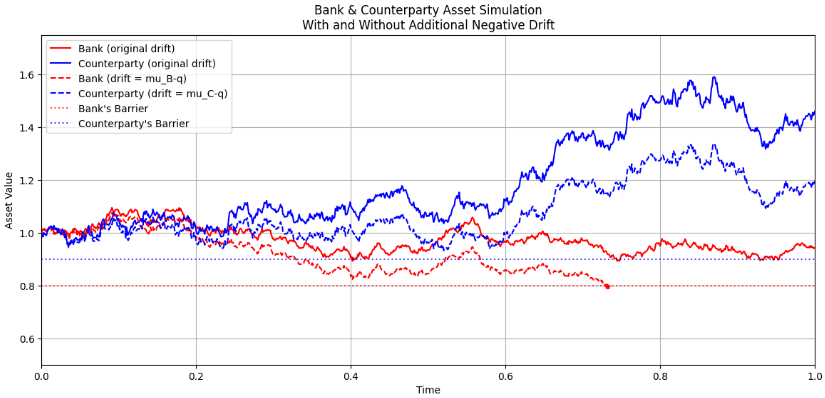
$$\begin{cases} X_t^q = x_0 + \int_0^t \big(b(s, X_s^q) - q(s, X_s^q)\big)\mathrm{d}s + \int_0^t \sigma(s, X_s^q) \odot \mathrm{d}W_s, \\ Y_t^q = \mathbb{I}_{\{\tau^q \le T\}} \chi^q_{\tau^q} - \int_{(t, \tau^q \wedge T]} \mathrm{d}\Lambda_s^q \\ \qquad + \int_{(t, \tau^q \wedge T]} \big(f_s^{Y,q} - r_s Y_s^q - \langle q(s, X_s^q) \oslash \sigma(s, X_s^q), Z_s^q \rangle\big)\,\mathrm{d}s - \int_{(t, \tau^q \wedge T]} Z_s^q \cdot \mathrm{d}W_s. \end{cases} \tag{5}$$

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, $\tau^q = \tau^{\mathcal{B},q}$, and portfolio maturity $\tau^{\mathcal{B},q}$



Equivalent formulation as a variational problem

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

$$\begin{cases} \inf_{y_0, (Z^q_t)_{t\in[0,T]}} \mathbb{E}\left[\left|\mathbb{I}_{\{\tau^q \le T\}} \chi^q_{\tau^q} - \Delta\Lambda^q_{\tau^q \wedge T} - Y^q_{\tau^q \wedge T}\right|^2\right], \quad \text{where,} \\ X^q_t = x_0 + \int_0^t \left(b(s, X^q_s) - q(s, X^q_s)\right) ds + \int_0^t \sigma(s, X^q_s) \odot dW_s, \\ Y^q_t = y_0 + \int_{(0,t]} d\Lambda^q_s - \int_{(0,t]} \left(f^{Y,q}_s - r_s Y^q_s - \langle q(s, X^q_s) \oslash \sigma(s, X^q_s), Z^q_s \rangle\right) ds \\ \qquad + \int_{(0,t]} Z^q_s \cdot dW_s. \end{cases} \tag{6}$$

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 $q_1, q_2, \dots, q_K$

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

- 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 $n^i_{\tau^{\mathcal{B}}} = \inf\{n \in \mathbb{N}^+ : [X^{\pi,q_i}_n]_{d+1} \le \xi^1_{t_n}\}, \quad n^i_{\tau^{\mathcal{C}}} = \inf\{n \in \mathbb{N}^+ : [X^{\pi,q_i}_n]_{d+2} \le \xi^2_{t_n}\}$.

Time discrete formulation:

$$\begin{cases} \mathcal{J}(y_0, \mathcal{Z}) := \inf\limits_{y_0, \mathcal{Z}} \sum_{i=1}^K \mathbb{E}\left[\left|\mathbb{I}_{\{n^i_\tau \le N\}} \chi^{\pi,q_i}_{n^i_\tau} - \Delta\Lambda^{\pi,q_i}_{n^i_\tau \wedge N} - Y^\pi_{n^i_\tau \wedge N}\right|^2\right], \text{ s.t. for } i \in \{1, \dots, K\}: \\ X^{\pi,q_i}_n = x_0 + \sum_{k=0}^{n-1} \left[b(t_k, X^{\pi,q_i}_k) - q_i(t_k, X^{\pi,q_i}_k)\right] h + \sum_{k=0}^{n-1} \sigma(t_k, X^{\pi,q_i}_k) \odot \Delta W_k, \\ Y^{\pi,q_i}_n = y_0 + \sum_{k=0}^{n-1} \Delta\Lambda^{\pi,q_i}_k + \sum_{k=0}^{n-1} Z^{\pi,q_i}_k \cdot \Delta W_k \\ \qquad - \sum_{k=0}^{n-1} \left(f^{\pi,Y,q_i}_k - r^\pi_k Y^{\pi,q_i}_k - \langle q_i(t_k, X^{\pi,q_i}_k) \oslash \sigma(t_k, X^{\pi,q_i}_k), Z^{\pi,q_i}_k \rangle\right) h \\ Z^{\pi,q_i}_n = \mathcal{Z}(t_n, X^{\pi,q_i}_n), \end{cases} \tag{7}$$

Represent 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) ✗.

Initial margin approximations

Reformulate VaR as an optimization problem:

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

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

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

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

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 ColVA, CVA, DVA and MVA BSDEs,
4. Approximation of the FVA BSDE,
5. (Approximation of the KVA BSDE).

In this presentation we focus on: Clean Values $\widehat{V}$, Initial Margins, $\mathrm{IM}^{\mathrm{TC}}$ and $\mathrm{IM}^{\mathrm{FC}}$, and Margin Valuation Adjustment, MVA.

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

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

$T^P = \{1, 1, 1, 0.8, 0.8, 0.6, 0.6, 0.4, 0.4, 0.2, 0.2, 1, 1, 1, 0.7, 0.7, 0.5, 0.5, 0.3, 0.3, 0.1, 0.1, 1, 0.7, 0.7, 0.5, 0.5, 0.3, 1, 0.8, 0.8, 0.6, 0.6\}$,
$K^P = \{1.05, 1.1, 1.05, 1.05, 0.7, 0.7, 0.75, 0.75, 1, 0.9, 0.8, 0.9, 1.1, 1.05, 0.85, 0.9, 0.9, 1.05, 1, 1, 0.9, 0.95, 1.05, 0.7, 0.7, 0.75, 0.75, 1, 0.9, 0.8, 0.9, 1.1, 1.05\}$, and

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

Clean values

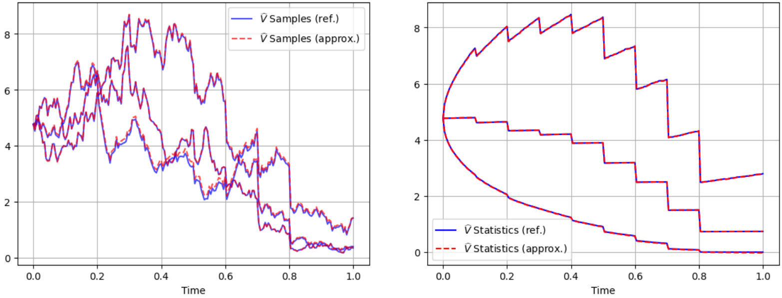


Figure: Approximate portfolio values compared with their analytical counterparts. **Left:** Three representative samples. **Right:** Empirical mean, 99th and 1st percentiles.

Initial Margin

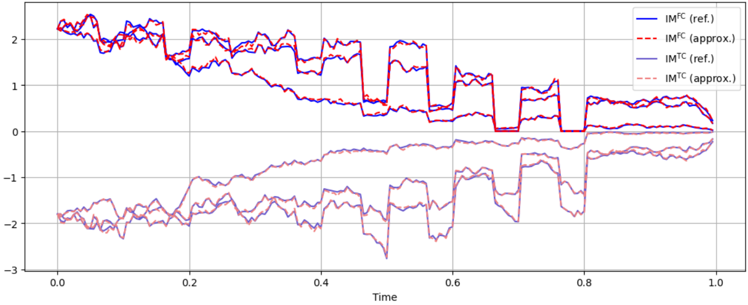


Figure: Three representative samples of approximate IM compared with reference solutions obtained by nested Monte–Carlo sampling.

Margin Valuation Adjustment - the control process

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

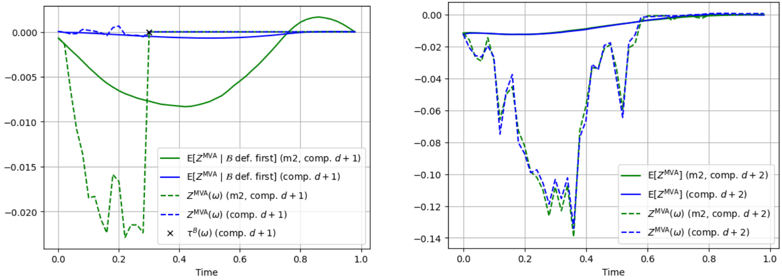


Figure: **Left:** When default probability is moderate/low. **Right:** When default probability is unrealistically high.

Margin Valuation Adjustment

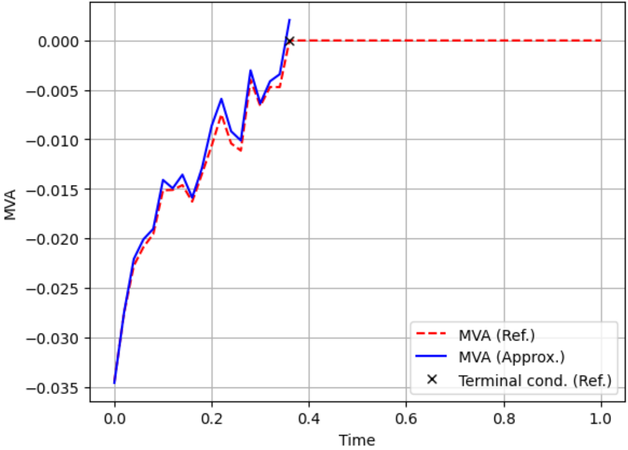


Figure: Dynamic MVA for a rare event of early default with the measure change technique applied.

Email address: **kristofferherbert.andersson@univr.it**