

XVA analysis from the balance sheet

CLAUDIO ALBANESE[†], STÉPHANE CRÉPEY^{*||‡}, RODNEY HOSKINSON[§] and
BOUAZZA SAADEDDINE^{‡¶}

[†]Global Valuation Ltd, London, UK

[‡]LaMME, Paris-Saclay/Univ Evry, Evry, France

[§]ANZ Banking Group, Singapore, Singapore

[¶]Quantitative Research GMD/GMT Credit Agricole CIB, Paris, France

(Received 28 October 2019; accepted 21 August 2020; published online 16 October 2020)

XVAs denote various counterparty risk related valuation adjustments that are applied to financial derivatives since the 2007–2009 crisis. We root a cost-of-capital XVA strategy in a balance sheet perspective which is key to identifying the economic meaning of the XVA terms. Our approach is first detailed in a static setup that is solved explicitly. It is then plugged into the dynamic and trade incremental context of a real derivative banking portfolio. The corresponding cost-of-capital XVA strategy ensures for bank shareholders a submartingale equity process corresponding to a target hurdle rate on their capital at risk, consistently between and throughout deals. Set on a forward/backward SDE formulation, this strategy can be solved efficiently using GPU computing combined with deep learning regression methods in a whole bank balance sheet context. A numerical case study emphasizes the workability and added value of the ensuing pathwise XVA computations.

Keywords: Counterparty risk; Balance sheet of a bank; Market incompleteness; Wealth transfer; X-valuation adjustment (XVA); Deep learning; Quantile regression

Mathematics Subject Classification: 91B25, 91B26, 91B30, 91G20, 91G40, 62G08, 68Q32

JEL Classification: D52, G13, G24, G28, G33, M41

1. Introduction

XVAs, with X as C for credit, D for debt, F for funding, M for margin, or K for capital, are post-2007–2009 crisis valuation adjustments for financial derivatives. In broad terms to be detailed later in the paper (cf. table 1 in Section 2), CVA is what the bank expects to lose due to counterparty defaults in the future; DVA (irrelevant for pricing but material to bank creditors as we will see) is what the bank expects to gain due to its own default; FVA is the expected cost for the bank of having to raise variation margin (re-hypothecable collateral); MVA is the expected cost for the bank of having to raise initial margin (segregated collateral); KVA is the expected cost for the bank of having to remunerate its shareholders through dividends for their capital at risk.

XVAs deeply affect the derivative pricing task by making it global, nonlinear, and entity dependent. However, before these technical implications, the fundamental point is to understand what really deserves to be priced and what

does not, by rooting the pricing approach in a corresponding collateralization, accounting, and dividend policy of the bank.

Coming after several papers on the valuation of defaultable assets in the 90's, such as Duffie and Huang (1996), Bielecki and Rutkowski (2002, equation (14.25) p. 448) obtained the formula

$$\text{CVA} - \text{DVA} \quad (1)$$

for the valuation of bilateral counterparty risk on a swap, assuming risk-free funding. This formula, rediscovered and generalized by others since the 2008–2009 financial crisis (cf. e.g. Brigo and Capponi 2010), is symmetrical, i.e. it is the negative of the analogous quantity considered from the point of view of the counterparty, consistent with the law of one price and the Modigliani and Miller (1958) theorem.

Around 2010, the materiality of the DVA windfall benefit of a bank at its own default time became the topic of intense debates in the quant and academic communities. At least, it seemed reasonable to admit that, if the own default risk of the bank was accounted for in the modelling, in the form of a DVA benefit, then the cost of funding (FVA) implication of this risk should be included as well, leading to the modified formula (CVA – DVA + FVA). See for instance Burgard

*Corresponding author. Email: stephane.crepey@lpsm.paris

||Current address: UFR de Mathématiques & LPSM, Université de Paris, Paris 75013, France

Table 1. Main financial acronyms and place where they are introduced conceptually and/or specified mathematically in the paper, as relevant.

<i>Amounts on dedicated cash accounts of the bank:</i>		
CM	Clean margin	Definition 2.1 and Assumption 2.1
RC	Reserve capital	Definition 2.1 and Assumption 2.1
RM	Risk margin	Definition 2.1 and Assumption 2.1
UC	Uninvested capital	Definition 2.1 and Assumption 2.1
<i>Valuations:</i>		
CA	Contra-assets valuation	(3), (16), and (A3)
CL	Contra-liabilities valuation	Definition 2.1 and (18), (35), and (43)
CVA	Credit valuation adjustment	(17), (16), (A4), and (A12)–(A13)
DVA	Debt valuation adjustment	(18) and (17)
FDA	Funding debt adjustment	(18) and (23)
FV	Firm valuation of counterparty risk	(21) and (23)
FVA	Funding valuation adjustment	Remark 2.1, (17), (16), and (A4)
KVA	Capital valuation adjustment	(4), (26), and (A8)
MtM	Mark-to-market	(4) and (15)
MVA	Margin valuation adjustment	Remark 2.1, (33), (A4), and (A14)
XVA	Generic ‘X’ valuation adjustment	First paragraph
<i>Also:</i>		
CR	Capital at risk	(A6)
CET1	Core equity tier I capital	(3) and (40)
EC	Economic capital	Definitions 3.2 and A.1
FTP	Funds transfer price	(43)
SHC	Shareholder capital (or equity)	(3) and (41)
SCR	Shareholder capital at risk	Assumption 2.1 and (25)

and Kjaer (2011, 2013, 2017), Crépey (2015), Brigo and Pallavicini (2014), or Bichuch *et al.* (2018). See also Bielecki and Rutkowski (2015) for an abstract funding framework (without explicit reference to XVAs), generalizing Piterbarg (2010) to a nonlinear setup.

Then Hull and White (2012) objected that the FVA was only the compensator of another windfall benefit of the bank at its own default, corresponding to the non-reimbursement by the bank of its funding debt. Accounting for the corresponding ‘DVA2’ (akin to the FDA in this paper) brings back to the original firm valuation formula:

$$CVA - DVA + FVA - FDA = CVA - DVA,$$

as $FVA = FDA$ (assuming risky funding fairly priced as we will see).

However, their argument implicitly assumes that the bank can perfectly hedge its own default: cf. Burgard and Kjaer (2013, end of Section 3.1) and see section 3.5. As a bank is an intrinsically leveraged entity, this is not the case in practice. One can mention the related corporate finance notion of debt overhang in Myers (1977), by which a project valuable for the firm as a whole may be rejected by shareholders because the project is mainly valuable to bondholders. But, until recently, such considerations were hardly considered in the field of derivative pricing.

The first ones to recast the XVA debate in the perspective of the balance sheet of the bank were (Burgard and Kjaer 2011), to explain that an appropriately hedged derivative position has no impact on the dealer’s funding costs. Also relying on balance sheet models of a dealer bank, Castagna (2014) and Andersen *et al.* (2019) end up with conflicting conclusions, namely that the FVA should, respectively, should not, be included in the valuation of financial derivatives. Adding the KVA, but in a replication framework, Green *et al.* (2014) conclude that both the FVA and the KVA should be included

as add-ons in entry prices and as liabilities in the balance sheet.

1.1. Contents

Our key premise is that counterparty risk entails two distinct but intertwined sources of market incompleteness:

- A bank cannot perfectly hedge counterparty default losses, by lack of sufficiently liquid CDS markets;
- A bank can even less hedge its own jump-to-default exposure, because this would mean selling protection on its own default, which is nonpractical and, under certain jurisdictions, even legally forbidden (see section 2).

We specify the banking XVA metrics that align derivative entry prices to shareholder interest, given this impossibility for a bank to replicate the jump-to-default related cash flows. We develop a cost-of-capital XVA approach consistent with the accounting standards set out in IFRS 4 Phase II (see International Financial Reporting Standards 2013), inspired from the Swiss solvency test and Solvency II insurance regulatory frameworks (see Swiss Federal Office of Private Insurance 2006, Committee of European Insurance and Occupational Pensions Supervisors 2010), which so far has no analogue in the banking domain. Under this approach, the valuation (CL) of the so-called contra-liabilities and the cost of capital (KVA) are sourced from clients at trade inceptions, on top of the $(CVA - DVA)$ complete market valuation of counterparty risk, in order to compensate bank shareholders for wealth transfer and risk on their capital.

The cost of the corresponding collateralization, accounting, and dividend policy is, by contrast with the complete market

valuation (CVA – DVA) of counterparty risk,

$$\text{CVA} + \text{FVA} + \text{KVA}, \quad (2)$$

computed unilaterally in a certain sense (even though we do crucially include the default of the bank itself in our modelling), and charged to clients on an incremental run-off basis at every new deal.[†]

All in one, our cost-of-capital XVA strategy makes shareholder equity a submartingale with drift corresponding to a hurdle rate h on shareholder capital at risk, consistently between and throughout deals. Thus we arrive at a sustainable strategy for profits retention, much like in the above-mentioned insurance regulation, but in a consistent continuous-time and banking framework.

Last but not least, our approach can be solved efficiently using GPU computing combined with deep learning regression methods in a whole bank balance sheet context.

1.2. Outline and contributions

Section 2 sets a financial stage where a bank is split across several trading desks and entails different stakeholders. Section 3 develops our cost-of-capital XVA approach in a one-period static setup. Section 4 revisits the approach at the dynamic and trade incremental level. Section 5 is a numerical case study on large, multi-counterparty portfolios of interest rate swaps, based on the continuous-time XVA equations for bilateral trade portfolios recalled in Appendix.

The *main contributions* of the paper are:

- The one-period static XVA model of section 3, with explicit formulas for all the quantities at hand, offering a concrete grasp on the related wealth transfer and risk premium issues;
- Proposition 4.1, which establishes the connections between XVAs and the core equity tier 1 capital of the bank, respectively, bank shareholder equity;
- Proposition 4.2, which establishes that, under the XVA policy represented by the balance conditions (4) between deals and the counterparty risk add-on (43) throughout deals, bank shareholder equity is a submartingale with drift corresponding to a target hurdle rate h on shareholder capital at risk. This perspective solves the puzzle according to which, on the one hand, XVA computations are performed on a run-off portfolio basis, while, on the other hand, they are used for computing pricing add-ons to new deals;
- The XVA deep learning (quantile) regression computational strategy of section 4.4;
- The numerical case study of section 5, which emphasizes the materiality of refined, pathwise XVA computations, as compared to more simplistic XVA approaches.

From a broader point of view, this paper reflects a shift of paradigm regarding the pricing and risk management of financial derivatives, from hedging to balance sheet optimization,

as quantified by relevant XVA metrics. In particular (compare with the last paragraph before section 1.1), our approach implies that the FVA (and also the MVA, see Remark 2.1) should be included as an add-on in entry prices and as a liability in the balance sheet; the KVA should be included as an add-on in entry prices, but not as a liability in the balance sheet.

From a computational point of view, this paper opens the way to second generation XVA GPU implementation. The first generation consisted of nested Monte Carlo implemented by explicit CUDA programming on GPUs (see Albanese *et al.* 2017, Abbas-Turki *et al.* 2018). The second generation takes advantage of GPUs leveraging via pre-coded CUDA/AAD deep learning packages that are used for the XVA embedded regression and quantile regression task. Compared to a regulatory capital-based KVA approach, an economic capital-based KVA approach is then not only conceptually more satisfying but also simpler to implement.

2. Balance sheet and capital structure model of the bank

We consider a dealer bank, which is a market maker involved in bilateral derivative portfolios. For simplicity, we only consider European derivatives. The bank has two kinds of stakeholders, *shareholders* and *bondholders*. The shareholders have the control of the bank and are solely responsible for investment decisions before bank default. The bondholders represent the senior creditors of the bank, who have no decision power until bank default, but are protected by laws, of the *pari-passu* type, forbidding trades that would trigger value away from them to shareholders during the default resolution process of the bank. The bank also has junior creditors, represented in our framework by an *external funder*, who can lend unsecured to the bank and is assumed to suffer an exogenously given loss-given-default in case of default of the bank.

We consider three kinds of business units within the bank (see figure 1 for the corresponding picture of the bank balance sheet and refer to table 1 for a list of the main financial acronyms used in the paper): the *CA desks*, i.e. the CVA desk and the FVA desk (or Treasury) of the bank, in charge of contra-assets, i.e. of counterparty risk and its funding implications for the bank; the *clean desks*, who focus on the market risk of the contracts in their respective business lines; the *management* of the bank, in charge of the dividend release policy of the bank.

Collateral means cash or liquid assets that are posted to guarantee a netted set of transactions against defaults. It comes in two forms: variation margin, which is re-hypothecable, i.e. fungible across netting sets, and initial margin, which is segregated. We assume cash only collateral. Posted collateral is supposed to be remunerated at the risk-free rate (assumed to exist, with overnight index swap rates as a best market proxy).

REMARK 2.1 To alleviate the notation, in this conceptual section of the paper, we only consider an FVA as the global cost of raising collateral for the bank, as opposed to a distinction, in the industry and in later sections in the paper, between

[†] See also Remark 2.1 regarding the meaning of the FVA in (2).

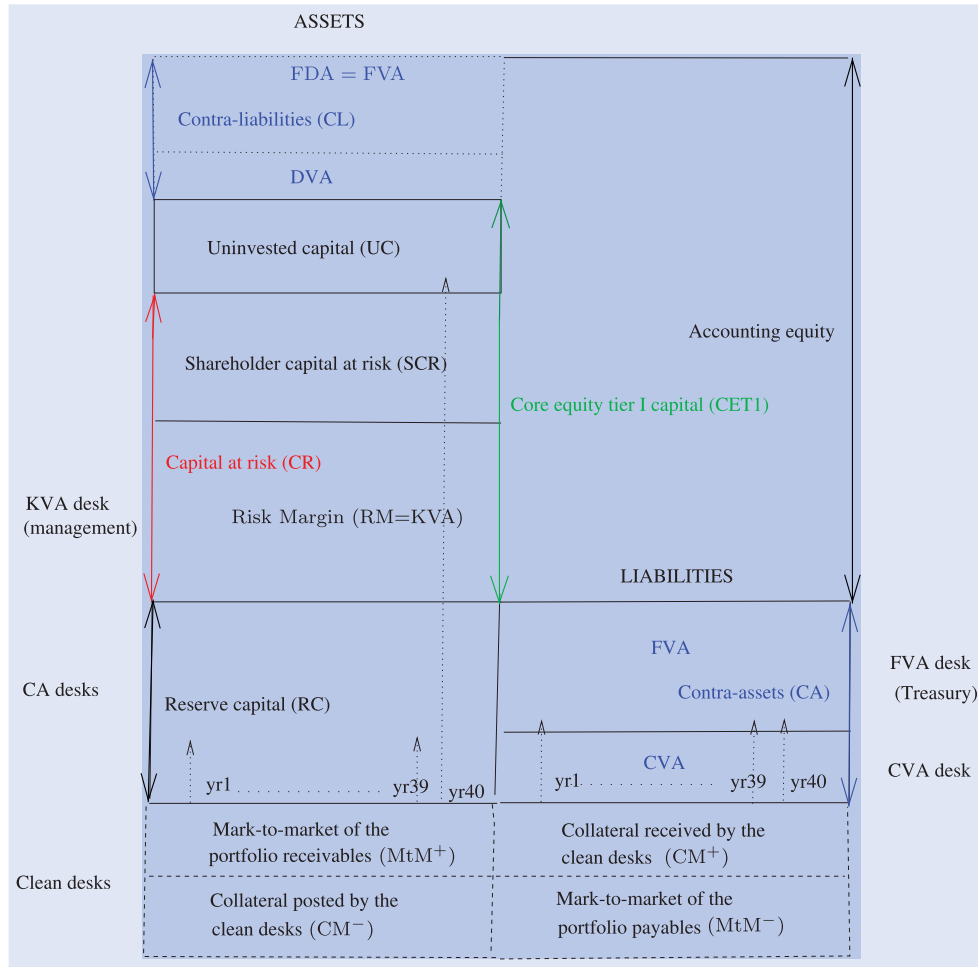


Figure 1. Balance sheet of a dealer bank. Contra-liability valuation (CL) at the top is shown in dotted boxes because it is only value to the bondholders (see section 3.5). Mark-to-market valuation (MtM) of the derivative portfolio of the bank by the clean desks, as well as the corresponding collateral (clean margin CM), are shown in dashed boxes at the bottom. Their role will essentially vanish in our setup, where we assume a perfect clean hedge by the bank. The arrows in the left column represent trading losses of the CA desks in ‘normal years 1 to 39’ and in an ‘exceptional year 40’ with full depletion (i.e. refill via UC, under Assumption 2.1(ii)) of RC, RM, and SCR. The numberings yr1 to yr40 are fictitious yearly scenarios in line with a 97.5% expected shortfall of the one-year-ahead trading losses of the bank that we use for defining its economic capital. The arrows in the right column symbolize the average depreciation in time of contra-assets between deals. The collateral between the bank and its counterparties is not shown to alleviate the picture.

an FVA, in the strict sense of the cost of raising variation margin, and an MVA for the cost of raising initial margin.

The CA desks guarantee the trading of the clean desks against counterparty defaults, through a *clean margin account*, which can be seen as (re-hypothecable) collateral exchanged between the CA desks and the clean desks. The corresponding clean margin amount (CM) also plays the role of the funding debt of the clean desks put at their disposal at a risk-free cost by the Treasury of the bank. This is at least the case when $CM > 0$ (clean desks clean margin receivers). In the case when $CM < 0$ (clean desks clean margin posters), $(-CM)$ corresponds to excess cash generated by the trading of the clean desks, usable by the Treasury for its other funding purposes. See the bottom, dashed boxes in figure 1.

In addition, the CA desks value the contra-assets (future counterparty default losses and funding expenditures), charge them to the (corporate) clients at deal inception, deposit the corresponding payments in a *reserve capital account*, and then are exposed to the corresponding payoffs. As time proceeds,

contra-assets realize and are covered by the CA desks with the reserve capital account.

On top of reserve capital, the so-called risk margin is sourced by the management of the bank from the clients at deal inception, deposited into a *risk margin account*, and then gradually released as KVA payments into the shareholder dividend stream.

Another account contains the *shareholder capital at risk* earmarked by the bank to deal with exceptional trading losses (beyond the expected losses that are already accounted for by reserve capital).

Last, there is one more bank account with shareholder *uninvested capital*.

All cash accounts are remunerated at the risk-free rate.

DEFINITION 2.1 We write CM , RC , RM , SCR , and UC for the respective (risk-free discounted) amounts on the clean margin, reserve capital, risk margin, shareholder capital at risk, and uninvested capital accounts of the bank. We also define

$$SHC = SCR + UC, \quad CET1 = RM + SCR + UC. \quad (3)$$

From a financial interpretation point of view, before bank default, SHC corresponds to shareholder capital (or equity); CET1 is the *core equity tier 1 capital* of the bank, representing the financial strength of the bank assessed from a regulatory, structural solvency point of view, i.e. the sum between shareholder capital and the risk margin (which is also loss-absorbing), but excluding the value CL of the so-called contra-liabilities (see figure 1). Indeed, the latter only benefits the bondholders (cf. section 3.5), hence it only enters accounting equity. Before the default of the bank, *shareholder wealth* and *bondholder wealth* are, respectively, given by $\text{SHC} + \text{RM}^{\text{sh}}$ and $\text{CL} + \text{RM}^{\text{bh}}$, for shareholder and bondholder components of RM to be detailed in Remark 3.3; shareholder and bondholder wealths sum up to the accounting equity $\text{RM} + \text{SCR} + \text{UC} + \text{CL}$, i.e. the wealth of the firm as a whole (see figure 1).

REMARK 2.2 The purpose of our capital structure model of the bank is *not* to model the default of the bank, like in a Merton (1974) model, as the point of negative equity (i.e. $\text{CET1} < 0$). In the case of a bank, such a default model would be unrealistic. For instance, at the time of its collapse in April 2008, Bear Stearns had billions of capital. In fact, the legal definition of default is an unpaid coupon or cash flow, which is a liquidity (as opposed to solvency) issue. Eventually we will model the default of the bank as a totally unpredictable event at some exogenous time τ calibrated to the credit default swap (CDS) curve referencing the bank. Indeed we view the latter as the most reliable and informative credit data regarding anticipations of markets participants about future recapitalization, government intervention, bail-in, and other bank failure resolution policies.

The aim of our capital structure model, instead, is to put in a balance sheet perspective the contra-assets and contra-liabilities of a dealer bank, items which are not present in the Merton model and play a key role in our XVA analysis.

In line with the Volcker rule banning proprietary trading for a bank, we assume a perfect market hedge of the derivative portfolio of the bank by the clean desks, in a sense to be specified below in the respective static and continuous-time setups. By contrast, as jump-to-default exposures (own jump-to-default exposure, in particular) cannot be hedged by the bank (cf. section 1.1), we conservatively assume no XVA hedge. We work on a measurable space (Ω, \mathfrak{A}) endowed with a probability measure \mathbb{Q}^* , with \mathbb{Q}^* expectation denoted by \mathbb{E}^* , which is used for the linear valuation task, using the risk-free asset as our numéraire everywhere.

REMARK 2.3 Regarding the nature of our reference probability measure \mathbb{Q}^* , ‘physical or risk-neutral’, one should view it as a blend between the two. For instance, even if we do not use this explicitly in the paper, one could conceptually think of \mathbb{Q}^* as the probability measure introduced by Dybvig (1992) to deal with incomplete markets that are a mix of financial traded risk factors and unhedgeable ones (jumps to default, in our setup), recently revisited in a finance and insurance context by Artzner *et al.* (2020). Namely, one could think of \mathbb{Q}^* as the unique probability measure on \mathfrak{A}^\dagger that coincides (i) with a

given risk-neutral pricing measure on the financial σ algebra $\subseteq \mathfrak{A}$, and (ii) with the physical probability measure *conditional* on the financial σ algebra (the risk-neutral and physical measures being assumed equivalent on the financial σ algebra). The risk-neutral pricing measure (hence, in view of (i), \mathbb{Q}^* itself) is calibrated to prices of fully collateralized transaction for which counterparty risk is immaterial. The physical probability measure expresses user views on the unhedgeable risk factors. The uncertainty about \mathbb{Q}^* can be dealt with by a Bayesian variation on our baseline XVA approach, whereby paths of alternative, co-calibrated models are combined in a global simulation (cf. Hoeting *et al.* 1999).

2.1. Run-off portfolio

Until section 4.2, we consider the case of a portfolio held on a run-off basis, i.e. set up at time 0 and such that no new unplanned trades enter the portfolio in the future.

The trading cash flows of the bank (cumulative cash flow streams starting from 0 at time 0) then consist of

- the contractually promised cash flows \mathcal{P} from counterparties,
- the counterparty credit cash flows \mathcal{C} to counterparties (i.e. because of counterparty risk, the effective cash flows from counterparties are $\mathcal{P} - \mathcal{C}$),
- the risky funding cash flows \mathcal{F} to the external funder, and
- the hedging cash flows \mathcal{H} of the clean desks to financial hedging markets

(note that all cash flow differentials can be positive or negative). See section 3.1 and (A1)–(A2) for concrete specifications in respective one-period and continuous-time setups.

ASSUMPTION 2.1 (i) (*Self-financing condition*) $\text{RC} + \text{RM} + \text{SCR} + \text{UC} - \text{CM}$ evolves like the received trading cash flows $\mathcal{P} - \mathcal{C} - \mathcal{F} - \mathcal{H}$.

(ii) (*Mark-to-model*) The amounts on all the accounts but UC are marked-to-model (hence the last, residual amount, UC, plays the role of an adjustment variable). Specifically, we assume that the following shareholder balance conditions hold at all times:

$$\text{CM} = \text{MtM}, \quad \text{RC} = \text{CA}, \quad \text{RM} = \text{KVA}, \quad (4)$$

for theoretical target levels MtM, CA, and KVA to be specified in later sections of the paper (which will also determine the theoretical target level for SCR).

(iii) (*Agents*) The initial amounts MtM_0 , CA_0 , and KVA_0 are provided by the clients at portfolio inception time 0. Resets between time 0 and the bank default time τ (excluded) are on bank shareholders. At the (positive) bank default time τ , the property of the residual amount on the reserve capital and risk margin accounts is transferred from the shareholders to the bondholders of the bank.

REMARK 2.4 In an asymmetric setup with a price maker and a price taker, the price maker passes his costs to the price taker. Accordingly, in our setup, the (corporate) clients provide all

[†] See Artzner *et al.* (2020, Proposition 2.1) for a proof.

the amounts to the clean margin, reserve capital, and risk margin accounts of the bank required for resetting the accounts to their theoretical target levels (4) corresponding to the updated portfolio.

Under a cost-of-capital XVA approach, we define valuation so as to make shareholder trading losses (that include marked-to-model liability fluctuations) centred, then we add a KVA risk premium in order to ensure to bank shareholders some positive hurdle rate h on their capital at risk.

In what follows, such an approach is developed, first, in a static setup, which can be solved explicitly, and then, in a dynamic and trade incremental setup, as suitable for dealing with a real derivative banking portfolio.

3. XVA analysis in a static setup

In this section, we apply the cost-of-capital XVA approach to a portfolio made of a single deal, \mathcal{P} (random variable promised to the bank), between a bank and a client, without prior endowment, in an elementary one-period (one year) setup. All the trading cash flows \mathcal{P} , \mathcal{C} , \mathcal{F} , and \mathcal{H} are then random variables (as opposed to processes in a multi-period setup later in the paper). We first assume no collateral exchanged between the bank and its client (but collateral exchanged as always between the CA and the clean desks as well as collateral on the market hedge of the bank, the way explained after the respective Remarks 2.1 and 2.2). Risky funding assets are assumed fairly priced by the market, in the sense that $\mathbb{E}^*\mathcal{F} = 0$.

The bank and client are both default prone with zero recovery to each other. The bank also has zero recovery to its external funder. We denote by J and J_1 the survival indicators (random variables) of the bank and client at time 1, with default probability of the bank $\mathbb{Q}^*(J = 0) = \gamma$.

Since prices and XVAs only matter at time 0 in a one-period setup, we identify all the XVA processes, as well as the mark-to-market (valuation by the clean desks) MtM of the deal, with their values at time 0.

For any random variable \mathcal{Y} , we define

$$\mathcal{Y}^\circ = J\mathcal{Y} \quad \text{and} \quad \mathcal{Y}^\bullet = -(1 - J)\mathcal{Y}, \quad \text{hence } \mathcal{Y} = \mathcal{Y}^\circ - \mathcal{Y}^\bullet. \quad (5)$$

Let \mathbb{E} denote the expectation with respect to the bank survival measure, say \mathbb{Q} , associated with \mathbb{Q}^* , i.e. for any random variable \mathcal{Y} ,

$$\mathbb{E}\mathcal{Y} = (1 - \gamma)^{-1}\mathbb{E}^*(\mathcal{Y}^\circ) \quad (6)$$

(which is also equal to $\mathbb{E}\mathcal{Y}^\circ$). The notion of bank survival measure was introduced in greater generality by Schönbucher (2004). In the present static setup, (6) is nothing but the \mathbb{Q}^* expectation of \mathcal{Y} conditional on the survival of the bank (note that, whenever \mathcal{Y} is independent from J , the right-hand-side in (6) coincides with $\mathbb{E}^*\mathcal{Y}$).

LEMMA 3.1 *For any random variable \mathcal{Y} and constant Y , we have*

$$Y = \mathbb{E}^*(\mathcal{Y}^\circ + (1 - J)Y) \iff Y = \mathbb{E}\mathcal{Y}. \quad (7)$$

Proof Indeed,

$$\begin{aligned} Y = \mathbb{E}^*(J\mathcal{Y} + (1 - J)Y) &\iff \mathbb{E}^*(J(\mathcal{Y} - Y)) = 0 \\ &\iff \mathbb{E}(\mathcal{Y} - Y) = 0 \iff Y = \mathbb{E}\mathcal{Y}, \end{aligned} \quad (8)$$

where the equivalence in the middle is justified by (6). ■

REMARK 3.1 For simplicity in a first stage, we will ignore the possibility of using capital at risk for funding purposes, only considering in this respect reserve capital $\text{RC} = \text{CA}$ (cf. (4)). The additional free funding source provided by capital at risk will be introduced later, as well as collateral between bank and client, in section 3.4.

3.1. Cash flows

LEMMA 3.2 *Given the (to be specified) MtM and CA amounts (cf. Assumption 2.1(ii)), the credit and funding cash flows \mathcal{C} and \mathcal{F} of the bank and its trading loss (and profit) L are such that*

$$\begin{aligned} \mathcal{C}^\circ &= J(1 - J_1)\mathcal{P}^+, \quad \mathcal{F}^\circ = J\gamma(\text{MtM} - \text{CA})^+ \\ \mathcal{C}^\bullet &= (1 - J)(\mathcal{P}^- - (1 - J_1)\mathcal{P}^+), \\ \mathcal{F}^\bullet &= (1 - J)((\text{MtM} - \text{CA})^+ - \gamma(\text{MtM} - \text{CA})^+) \\ L^\circ &= \mathcal{C}^\circ + \mathcal{F}^\circ - J\text{CA}, \quad L^\bullet = \mathcal{C}^\bullet + \mathcal{F}^\bullet + (1 - J)\text{CA}, \\ L &= \mathcal{C} + \mathcal{F} - \text{CA}. \end{aligned} \quad (9)$$

Proof For the deal to occur, the bank needs to borrow $(\text{MtM} - \text{CA})^+$ unsecured or invest $(\text{MtM} - \text{CA})^-$ risk-free (cf. Remark 3.1). Having assumed zero recovery to the external funder, unsecured borrowing is fairly priced as $\gamma \times$ the amount borrowed by the bank (in line with our assumption that $\mathbb{E}^*\mathcal{F} = 0$), i.e. the bank must pay for its risky funding the amount

$$\gamma(\text{MtM} - \text{CA})^+.$$

Moreover, at time 1, under zero recovery upon defaults:

- If the bank is not in default (i.e. $J = 1$), then the bank closes its position with the client while receiving \mathcal{P} from its client if the latter is not in default (i.e. $J_1 = 1$), whereas the bank pays \mathcal{P}^- to its client if the latter is in default (i.e. $J_1 = 0$). In addition, the bank reimburses its funding debt $(\text{MtM} - \text{CA})^+$ or receives back the amount $(\text{MtM} - \text{CA})^-$ it had lent at time 0;
- If the bank is in default (i.e. $J = 0$), then the bank receives back $J_1\mathcal{P}^+$ on the derivative as well as the amount $(\text{MtM} - \text{CA})^-$ it had lent at time 0.

Also accounting for the hedging loss \mathcal{H} , the trading loss of the bank over the year is

$$\begin{aligned} L &= \gamma(\text{MtM} - \text{CA})^+ - J(J_1\mathcal{P} - (1 - J_1)\mathcal{P}^- \\ &\quad - (\text{MtM} - \text{CA})^+ + (\text{MtM} - \text{CA})^-) \\ &\quad - (1 - J)(J_1\mathcal{P}^+ + (\text{MtM} - \text{CA})^-) + \mathcal{H}. \end{aligned} \quad (10)$$

In the static setup, the perfect clean hedge condition (see after Remark 2.2) writes $\mathcal{H} = \mathcal{P} - \text{MtM}$. Inserting this into the

above yields

$$L = (1 - J_1)\mathcal{P}^+ + \gamma(\text{MtM} - \text{CA})^+ - \text{CA} - (1 - J)(\mathcal{P}^- + (\text{MtM} - \text{CA})^+), \quad (11)$$

as easily checked for each of the four possible values of the pair (J, J_1) . That is,

$$\begin{aligned} L^\circ &= \underbrace{J(1 - J_1)\mathcal{P}^+}_{\mathcal{C}^\circ} + \underbrace{J\gamma(\text{MtM} - \text{CA})^+}_{\mathcal{F}^\circ} - J\text{CA} \\ L^\bullet &= \underbrace{(1 - J)(\mathcal{P}^- - (1 - J_1)\mathcal{P}^+)}_{\mathcal{C}^\bullet} \\ &\quad + \underbrace{(1 - J)((\text{MtM} - \text{CA})^+ - \gamma(\text{MtM} - \text{CA})^+)}_{\mathcal{F}^\bullet} \\ &\quad + (1 - J)\text{CA}, \end{aligned} \quad (12)$$

where the identification of the different terms as part of \mathcal{C} or \mathcal{F} follows from their financial interpretation. ■

REMARK 3.2 The derivation (10) implicitly allows for negative equity (that arises whenever $L^\circ > \text{CET1}$, cf. (3)), which is interpreted as recapitalization. In a variant of the model excluding both recapitalization and negative equity, the default of the bank would be modelled in a structural fashion as the event $\{L = \text{CET1}\}$, where

$$L = ((1 - J_1)\mathcal{P}^+ + \gamma(\text{MtM} - \text{CA})^+ - \text{CA}) \wedge \text{CET1}, \quad (13)$$

and we would obtain, instead of (11), the following trading loss for the bank:

$$\mathbb{1}_{\{\text{CET1} > L\}}L + \mathbb{1}_{\{\text{CET1} = L\}}(\text{CET1} - \mathcal{P}^- - (\text{MtM} - \text{CA})^+). \quad (14)$$

In this paper, we consider a model with recapitalization for the reasons explained in Remark 2.2.

Structural XVA approaches in a static setup have been proposed in Andersen *et al.* (2019) (without KVA) and Kjaer (2019) (including the KVA). Their marginal, limiting results as a new deal size goes to zero are comparable to some of the results that we have here. But then, instead of developing a continuous time version of their corporate finance model and taking the small trade limit, these papers start the development of the continuous time model from the single period small trade limit model. By contrast, in our framework, we have end to end development in the continuous time model of section 4 and in the present single period model.

3.2. Contra-assets and contra-liabilities

To make shareholder trading losses centred (cf. the next-to-last paragraph of section 2), clean and CA desks value by \mathbb{Q}^* expectation their shareholder sensitive cash flows. These include, in case of default of the bank, the transfer of property from the CA desks to the clean desks of the collateral amount MTM on the clean margin account, as well as (cf. Assumptions 2.1(ii and iii)) the transfer from shareholders to bondholders of the residual value $\text{RC} = \text{CA}$ on the reserve capital account. Accordingly:

DEFINITION 3.1 We let

$$\text{MtM} = \mathbb{E}^*(\mathcal{P}^\circ + (1 - J)\text{MtM}) \quad (15)$$

and

$$\text{CA} = \text{CVA} + \text{FVA}, \quad (16)$$

where

$$\begin{aligned} \text{CVA} &= \mathbb{E}^*(\mathcal{C}^\circ + (1 - J)\text{CVA}) \\ \text{FVA} &= \mathbb{E}^*(\mathcal{F}^\circ + (1 - J)\text{FVA}), \end{aligned} \quad (17)$$

hence $\text{CA} = \mathbb{E}^*(\mathcal{C}^\circ + \mathcal{F}^\circ + (1 - J)\text{CA})$. We also define the contra-liabilities value

$$\text{CL} = \text{DVA} + \text{FDA}, \quad (18)$$

where

$$\text{DVA} = \mathbb{E}^*(\mathcal{C}^\bullet + (1 - J)\text{CVA}) \quad (19)$$

$$\text{FDA} = \mathbb{E}^*(\mathcal{F}^\bullet + (1 - J)\text{FVA}). \quad (20)$$

Finally we define the firm valuation of counterparty risk,

$$\text{FV} = \mathbb{E}^*(\mathcal{C} + \mathcal{F}). \quad (21)$$

The definitions of MtM, CVA, and FVA are in fact fix-point equations. However, the following result shows that these equations are well-posed and yields explicit formulas for all the quantities at hand.

PROPOSITION 3.1 We have

$$\begin{aligned} \text{MtM} &= \mathbb{E}\mathcal{P}^\circ \\ \text{CVA} &= \mathbb{E}((1 - J_1)\mathcal{P}^+) \\ \text{FVA} &= \gamma(\text{MtM} - \text{CA})^+ = \frac{\gamma}{1 + \gamma}(\text{MtM} - \text{CVA})^+ \end{aligned} \quad (22)$$

and

$$\begin{aligned} \mathbb{E}^*L^\circ &= \mathbb{E}L = 0 \\ \text{FDA} &= \text{FVA} \\ \text{FV} &= \mathbb{E}^*\mathcal{C} = \text{CVA} - \text{DVA} = \text{CA} - \text{CL}. \end{aligned} \quad (23)$$

Proof The first identities in each line of (22) follow from Definition 3.1 by Lemma 3.1 and definition of the involved cash flows in Lemma 3.2. Given (16), the formula $\text{FVA} = \gamma(\text{MtM} - \text{CA})^+$ in (22) is in fact a semi-linear equation

$$\text{FVA} = \gamma(\text{MtM} - \text{CVA} - \text{FVA})^+. \quad (24)$$

But, as γ (a probability) is nonnegative, this equation has the unique solution given by the right-hand side in the third line of (22).

Regarding (23), we have

$$\begin{aligned} \mathbb{E}^* L^\circ &= (1 - \gamma) \\ &\times \mathbb{E}((1 - J_1)\mathcal{P}^+ + \gamma(\text{MtM} - \text{CA})^+ - \text{CA}) = 0, \end{aligned}$$

by application of (6), the first line in (12), (22), and (16). Hence, using (6) again,

$$\mathbb{E}L = (1 - \gamma)^{-1}\mathbb{E}^*L^\circ = 0.$$

This is the first line in (23), which implies the following ones by definition of the involved quantities and from the assumption that $\mathbb{E}^*\mathcal{F} = 0$. ■

Note that $\text{MtM} = \mathbb{E}\mathcal{P}^\circ$ also coincides with $\mathbb{E}\mathcal{P}$ (cf. (22) and the parenthesis following (6)). In practice \mathcal{P}° has less terms than \mathcal{P} (that also includes cash flows from bank default onward), which is why we favor the formulation $\mathbb{E}\mathcal{P}^\circ$ in (22). The alternative formulation $\mathbb{E}\mathcal{P}$ may seem more in line with the intuition of MtM as value deprived from any credit/funding considerations. However, as the measure underlying \mathbb{E} is the survival one (see before Lemma 3.1), this intuition is in fact simplistic and only strictly correct in the case without wrong way risk between credit and market (cf. the parenthesis preceding Lemma 3.1).

3.3. Capital valuation adjustment

Economic capital (EC) is the level of capital at risk that a regulator would like to see on an economic, structural basis. Risk calculations are typically performed by banks ‘on a going concern’, i.e. assuming that the bank itself does not default. Accordingly:

DEFINITION 3.2 *The economic capital (EC) of the bank is given by the 97.5% expected shortfall[†] of the bank trading loss L under \mathbb{Q} , which we denote by[‡] $\mathbb{E}\mathcal{S}(L^\circ)$.*

The risk margin (sized by the to-be-defined KVA in our setup) is also loss-absorbing, i.e. part of capital at risk, and the KVA is originally sourced from the client (see Assumption 2.1(iii)). Hence, *shareholder* capital at risk only consists of the *difference* between the (total) capital at risk and the KVA. Accordingly (and also accounting, regarding (26), for the last part in Assumption 2.1(iii):

DEFINITION 3.3 *The capital at the risk (CR) of the bank is given by $\max(\text{EC}, \text{KVA})$ and the ensuing shareholder capital at risk (SCR) by*

$$\text{SCR} = \max(\text{EC}, \text{KVA}) - \text{KVA} = (\text{EC} - \text{KVA})^+, \quad (25)$$

where, given some hurdle rate (target return-on-equity) h ,

$$\text{KVA} = \mathbb{E}^*(h\text{SCR}^\circ + (1 - J)\text{KVA}). \quad (26)$$

[†] See e.g. Föllmer and Schied (2016, Section 4.4).

[‡] Note that, by definition of \mathbb{Q} , this quantity does not depend on L^\bullet .

REMARK 3.3 In view of (26) and of the last balance condition in (4), we have

$$\text{RM}^{sh} = \mathbb{E}^*(h\text{SCR}^\circ), \quad \text{RM}^{bh} = \mathbb{E}^*((1 - J)\text{KVA}). \quad (27)$$

We refer the reader to the last bullet point in Albanese and Crépey (2020, Definition A.1) for the analogous split of RM between shareholder and bondholder wealth in a dynamic, continuous-time setup.

PROPOSITION 3.2 *We have*

$$\text{KVA} = h\text{SCR} = \frac{h}{1+h}\text{EC} = \frac{h}{1+h}\mathbb{E}\mathcal{S}(L^\circ). \quad (28)$$

Proof The first identity follows from Lemma 3.1. The resulting KVA semi-linear equation (in view of (25)) is solved similarly to the FVA equation (24). ■

The KVA formula (28) (as well as its continuous-time analogue (A8)) can be used either in the direct mode, for computing the KVA corresponding to a given h , or in the reverse-engineering mode, for defining the ‘implied hurdle rate’ associated with the actual level on the risk margin account of the bank. Cost of capital proxies have always been used to estimate return-on-equity. The KVA is a refinement, fine-tuned for derivative portfolios, but the base return-on-equity concept itself is far older than even the CVA. In particular, the KVA is very useful in the context of collateral and capital optimization.

3.3.1. KVA risk premium and indifference pricing interpretation. The CA component of the FTP corresponds to the expected costs for the shareholders of concluding the deal. This CA component makes the shareholder trading loss L° centred (cf. the first line in (23)). On top of expected shareholder costs, the bank charges to the clients a risk margin (RM). Assume the bank shareholders endowed with a utility function U on \mathbb{R} such that $U(0) = 0$. In a shareholder indifference pricing framework, the risk margin arises as per the following equation:

$$\mathbb{E}^*U(J(\text{RM} - L)) = \mathbb{E}^*U(0) = 0 \quad (29)$$

(the expected utility of the bank shareholders without the deal), where

$$\begin{aligned} \mathbb{E}^*U(J(\text{RM} - L)) &= \mathbb{E}^*(JU(\text{RM} - L)) \\ &= (1 - \gamma)\mathbb{E}U(\text{RM} - L), \end{aligned}$$

by (6). Hence

$$\mathbb{E}U(\text{RM} - L) = 0. \quad (30)$$

The corresponding RM is interpreted as the minimal admissible risk margin for the deal to occur, seen from bank shareholders’ perspective.

Taking for concreteness $U(-\ell) = \frac{1 - e^{\rho\ell}}{\rho}$, for some risk aversion parameter ρ , (30) yields $\text{RM} = \rho^{-1} \ln \mathbb{E} e^{\rho L} = \rho^{-1} \ln \mathbb{E} e^{\rho L^\circ}$, by the observation following (6). In the limiting case where the shareholder risk aversion parameter $\rho \rightarrow 0$

and $\mathbb{E}U(-L) \rightarrow -\mathbb{E}(L) = 0$ (by the first line in (23)), then $\text{RM} \rightarrow 0$.

In view of (4) and (28), the corresponding implied KVA and hurdle rate h are such that

$$\text{KVA} = \rho^{-1} \ln \mathbb{E} e^{\rho L^\circ}, \quad \frac{h}{1+h} = \frac{\rho^{-1} \ln \mathbb{E} e^{\rho L^\circ}}{\mathbb{E}\mathbb{S}(L^\circ)}. \quad (31)$$

Hence, ‘for h and ρ small’,

$$h \approx \frac{\text{Var}(L^\circ)}{2\mathbb{E}\mathbb{S}(L^\circ)} \rho \quad (32)$$

(as $\mathbb{E}(L^\circ) = 0$), where Var is the \mathbb{Q} variance operator. The hurdle rate h in our KVA setup plays the role of a risk aversion parameter, like ρ in the exponential utility framework.

An indifference price has a competitive interpretation. Assume that the bank is competing for the client with other banks. Then, in the limit of a continuum of competing banks with a continuum of indifference prices, whenever a bank makes a deal, this can only be at its indifference price. Our stylized indifference pricing model of a KVA defined by a constant hurdle rate h exogenizes (by comparison with the endogenous hurdle rate h in (31)) the impact on pricing of the competition between banks. It does so in a way that generalizes smoothly to a dynamic setup (see section 4), as required to deal with a real derivative banking portfolio. It then provides a refined notion of return-on-equity for derivative portfolios, where a full-fledged optimization approach would be impractical.

3.4. Collateral with clients and fungibility of capital at risk as a funding source

In case of variation margin (VM) that would be exchanged between the bank and its client, and of initial margin that would be received (RIM) and posted (PIM) by the bank, at the level of, say, some \mathbb{Q} value-at-risk of $\pm(\mathcal{P} - \text{VM})$, then

- \mathcal{P} needs be replaced by $(\mathcal{P} - \text{VM} - \text{RIM})$ everywhere in the above, whence an accordingly modified (in principle: diminished) CVA,
- an additional initial margin-related cash flow in \mathcal{F}° given as $J\gamma\text{PIM}$, triggering an additional adjustment MVA in CA, where

$$\text{MVA} = \mathbb{E}^*(J\gamma\text{PIM} + (1-J)\text{MVA}) = \gamma\text{PIM}; \quad (33)$$

- additional initial margin-related cash flows in \mathcal{F}^\bullet given as $(1-J)(\text{PIM} - \gamma\text{PIM})$ and $(1-J)\text{MVA}$, triggering an additional adjustment MDA = MVA in CL;
- the second FVA formula in (22) modified into $\text{FVA} = \frac{\gamma}{1+\gamma}(\text{MtM} - \text{VM} - \text{CVA} - \text{MVA})^+$.

Accounting further for the additional free funding source provided by capital at risk (cf. Remark 3.1), then, in view of the specification given in the first sentence of Definition 3.3 for the latter, one needs replace $(\text{MtM} - \text{CA})^\pm$ by $(\text{MtM} - \text{VM} - \text{CA} - \max(\text{EC}, \text{KVA}))^\pm$ everywhere before. This results in the same CVA and MVA as in

the bullet points above, but in the following system for the random variable L° and the FVA and the KVA numbers (cf. the corresponding lines in (12), (22), (28), and recall (16)):

$$\begin{aligned} L^\circ &= J(1 - J_1)\mathcal{P}^+ + J\gamma(\text{MtM} - \text{VM} - \text{CA} \\ &\quad - \max(\text{EC}, \text{KVA}))^+ + J\gamma\text{PIM} - J\text{CA} \\ \text{FVA} &= \gamma(\text{MtM} - \text{VM} - \text{CA} - \max(\text{EC}, \text{KVA}))^+ \\ \text{KVA} &= \frac{h}{1+h}\mathbb{E}\mathbb{S}(L^\circ). \end{aligned} \quad (34)$$

This system entails a coupled dependence between, on the one hand, the FVA and KVA numbers and, on the other hand, the shareholder loss process L° . However, once CVA, PIM, RIM, and MVA computed as in the above, the system (34) can be addressed numerically by Picard iteration, starting from, say, $L^{(0)} = \text{KVA}^{(0)} = 0$ and $\text{FVA}^{(0)} = \frac{\gamma}{1+\gamma}(\text{MtM} - \text{VM} - \text{CVA} - \text{MVA})^+$ (cf. the last line in (22)), and then iterating in (34) until numerical convergence.

REMARK 3.4 The rationale for funding FVA but not MVA from $\text{CA} + \max(\text{EC}, \text{KVA})$ is set out before equation (15) in Albanese *et al.* (2017).

3.5. Funds transfer price

The funds transfer price (all-inclusive XVA rebate to MtM) aligning the deal to shareholder interest (in the sense of a given hurdle rate h , cf. the next-to-last paragraph of section 2) is

$$\begin{aligned} \text{FTP} &= \frac{\text{CVA} + \text{FVA}}{\text{Expected shareholder costs CA}} \\ &\quad + \frac{\text{KVA}}{\text{Shareholder risk premium}} \\ &= \frac{\text{CVA} - \text{DVA}}{\text{Firm valuation FV}} + \frac{\text{DVA} + \text{FDA}}{\text{Wealth transfer CL}} \\ &\quad + \frac{\text{KVA}}{\text{Shareholder Risk premium}}, \end{aligned} \quad (35)$$

where all terms are explicitly given in Propositions 3.1 and 3.2 (or the corresponding variants of section 3.4 in the refined setup considered there).

3.5.1. Wealth transfer analysis. The above results implicitly assumed that the bank cannot hedge jump-to-default cash flows (cf. section 1.1). To understand this, let us temporarily suppose, for the sake of the argument, that the bank would be able to hedge its own jump-to-default through a further deal, whereby the bank would deliver a payment L^\bullet at time 1 in exchange of a fee fairly valued as

$$\text{CL} = \mathbb{E}^*L^\bullet = \text{DVA} + \text{FDA}, \quad (36)$$

deposited in the reserve capital account of the bank at time 0.

We include this hedge and assume that the client would now contribute at the level of $\text{FV} = \text{CA} - \text{CL}$ (cf. (23)), instead of

CA before, to the reserve capital account of the bank at time 0. Then the amount that needs be borrowed by the bank for implementing its strategy is still $\gamma(\text{MtM} - \text{CA})^+$ as before (back to the baseline funding setup of Remark 3.1). But the trading loss of the bank becomes, instead of L before,

$$\begin{aligned} \mathcal{C} + \mathcal{F} - \text{FV} + (L^\bullet - \text{CL}) &= \mathcal{C} + \mathcal{F} - \text{CA} + L^\bullet \\ &= L + L^\bullet = L^\circ, \end{aligned} \quad (37)$$

where the last line in (23) and the last identity in (9) were used in the first and second equality. By comparison with the situation from previous sections without own-default hedge by the bank:

- the shareholders are still indifferent to the deal in expected counterparty default and funding expenses terms,
- the recovery of the bondholders becomes zero,
- the client is better off by the amount $\text{CA} - \text{FV} = \text{CL}$.

The CL originating cash flow L^\bullet has been hedged and monetized by the shareholders, who have passed the corresponding benefit to the client.

Under a cost-of-capital pricing approach, the bank would still charge to its client a KVA add-on $\frac{h}{1+h} \mathbb{E}S(L^\circ)$, as risk compensation for the nonvanishing shareholder trading loss L° still triggered by the deal. If, however, the bank could also hedge the (zero-valued, by the first line in (23)) loss L° , hence the totality of $L = L^\circ - L^\bullet$ (instead of L^\bullet only in the above), then the trading loss and the KVA would vanish. As a result, the all-inclusive XVA add-on (rebate from MtM valuation) would boil down to

$$\text{FV} = \text{CVA} - \text{DVA}$$

(cf. (1)), the value of counterparty risk and funding to the bank as a whole.

3.5.2. Connection with the Modigliani–Miller theory. The Modigliani–Miller invariance result, with Modigliani and Miller (1958) as a seminal reference, consists in various facets of a broad statement that the funding and capital structure policies of a firm are irrelevant to the profitability of its investment decisions. Modigliani–Miller (MM) irrelevance, as we put it for brevity hereafter, was initially seen as a pure arbitrage result. However, it was later understood that there may be market incompleteness issues with it. So quoting Duffie and Sharer (1986, p. 9), ‘generically, shareholders find the span of incomplete markets a binding constraint [...] shareholders are not indifferent to the financial policy of the firm if it can change the span of markets (which is typically the case in incomplete markets)’; or Gottardi (1995, p. 197): ‘When there are derivative securities and markets are incomplete the financial decisions of the firm have generally real effects’.

A situation where shareholders may ‘find the span of incomplete markets a binding constraint’ is when market completion is legally forbidden. This corresponds to the XVA

case, which is also at the crossing between market incompleteness and the presence of derivatives pointed out above as the MM non irrelevance case in Gottardi (1995). Specifically, the contra-assets and contra-liabilities that emerge endogenously from the impact of counterparty risk on the derivative portfolio of a bank cannot be ‘undone’ by shareholders, because jump-to-default risk cannot be replicated by a bank.

As a consequence, MM irrelevance is expected to break down in the XVA setup. In fact, as visible on the trade incremental FTP (counterparty risk pricing) formula (35) (cf. also (43) and Proposition 4.2 in a dynamic and trade incremental setup below), cost of funding and cost of capital are material to banks and need be reflected in entry prices for ensuring shareholder indifference to the trades, i.e. preserving their hurdle rate throughout trades.

4. XVA analysis in a dynamic setup

We now consider a dynamic, continuous-time setup, with model filtration \mathbb{G} and a (positive) bank default time τ endowed with an intensity γ . The bank survival probability measure associated with the measure \mathbb{Q}^* is then the probability measure \mathbb{Q} with $(\mathbb{G}, \mathbb{Q}^*)$ density process $J e^{\int_0^\cdot \gamma_s ds}$ (assumed integrable), where $J = \mathbb{1}_{[0, \tau]}$ is the bank survival indicator process (cf. Collin-Dufresne *et al.* 2004, Schönbucher 2004). In particular, writing $Y^\circ = JY + (1 - J)Y_{\tau-}$, for any left-limited process Y , we have by application of the results of Crépey and Song (2017) (cf. the condition (A) there):

LEMMA 4.1 *For every \mathbb{Q} (resp. sub-, resp. resp. super-) martingale Y , the process Y° is a \mathbb{Q}^* (resp. sub-, resp. resp. super-) martingale.*

REMARK 4.1 In the dynamic setup, the survival measure formulation is a light presentation, sufficient for the purpose of the present paper (skipping the related integrability issues), of an underlying reduction of filtration setup, which is detailed in the above-mentioned reference (regarding Lemma 4.1, cf. also Collin-Dufresne *et al.* 2004, Lemma 1).

4.1. Case of a run-off portfolio

First, we consider the case of a portfolio held on a run-off basis (cf. section 2.1). We denote by T the final maturity of the portfolio and we assume that all prices and XVAs vanish at time T if $T < \tau$. Then the results of Albanese and Crépey (2020) show that all the qualitative insights provided by the one-period XVA analysis of section 3 are still valid. The trading loss of the bank is now given by the process

$$L = \mathcal{C} + \mathcal{F} + \text{CA} - \text{CA}_0 \quad (38)$$

and the bank *shareholder* trading loss by the \mathbb{Q} (hence \mathbb{Q}^* , by Lemma 4.1) martingale

$$L^\circ = \mathcal{C}^\circ + \mathcal{F}^\circ + \text{CA} - \text{CA}_0. \quad (39)$$

In (38)–(39), we have $\text{CA} = \text{CVA} + \text{FVA}$ as in (16); the processes \mathcal{C} , \mathcal{F} , CVA , and FVA are continuous-time processes

analogs, detailed in the case of bilateral trade portfolios in section A.1–A.2, of the eponymous quantities in section 3 (which were constants or random variables there).

PROPOSITION 4.1 *The core equity tier 1 capital of the bank is given by*

$$\text{CET1} = \text{CET1}_0 - L. \quad (40)$$

Shareholder equity is given by

$$\text{SHC} = \text{SHC}_0 - (L + \text{KVA} - \text{KVA}_0). \quad (41)$$

Proof In the continuous-time setup, Assumption 2.1(i) is written as

$$\text{RC} + \text{RM} + \text{SCR} + \text{UC} - \text{CM} - (\text{RC} + \text{RM} + \text{SCR} + \text{UC} - \text{CM})_0 = \mathcal{P} - (\mathcal{C} + \mathcal{F} + \mathcal{H}).$$

Given the definition of CET1 in (3), the perfect clean hedge condition (see after Remark 2.2) written in the dynamic setup as $\mathcal{P} + \text{MtM} - \text{MtM}_0 - \mathcal{H} = 0$, and the balance conditions (4), this is equivalent to

$$\text{CA} + \text{CET1} - (\text{CA} + \text{CET1})_0 = -(\mathcal{C} + \mathcal{F}).$$

In view of (38), we obtain (40).

As $\text{SHC} = \text{CET1} - \text{RM}$ (cf. (3)), we have by (40):

$$\begin{aligned} \text{SHC} &= \text{CET1}_0 - L - \text{RM} = \text{CET1}_0 - \text{RM}_0 \\ &\quad - (L + \text{RM} - \text{RM}_0), \end{aligned}$$

which, by the third balance condition in (4), yields (41). ■

Moreover, by Lemma 4.1, the continuous-time process KVA that stems from (A6)–(A7) is a \mathbb{Q}^* supermartingale with terminal condition $\text{KVA}_T = 0$ on $\{T < \tau\}$ and drift coefficient $h\text{SCR}$, where SCR is given as in (25), but for EC there dynamically defined as the time- t conditional, 97.5% expected shortfall of $(L_{t+1}^\circ - L_t^\circ)$ under \mathbb{Q} , killed at τ .

REMARK 4.2 It is only before τ that the right-hand-sides in the definitions (3) really deserve the respective interpretations of shareholder equity of the bank and core equity tier 1 capital. Hence, it is only the parts of (40) and (41) stopped before τ , i.e.

$$\begin{aligned} \text{CET1}^\circ &= \text{CET1}_0 - L^\circ, \quad \text{SHC}^\circ = \text{SHC}_0 \\ &\quad - (L^\circ + \text{KVA} - \text{KVA}_0), \end{aligned} \quad (42)$$

which are interesting financially.

4.2. Trade incremental cost-of-capital XVA strategy

In Albanese and Crépey (2020) and in section 4.1 above, the derivative portfolio of the bank is assumed held on a run-off basis. By contrast, real-life derivative portfolios are incremental.

Assume a new deal shows up at time $\theta \in (0, \tau)$. We denote by $\Delta \cdot$, for any portfolio related process, the difference between the time θ values of this process for the run-off versions of the portfolio with and without the new deal.

DEFINITION 4.1 *We apply the following trade incremental pricing and accounting policy:*

- The clean desks pay ΔMtM to the client and the CA desks add an amount ΔMtM on[†] the clean margin account;
- The CA desks charge to the client an amount ΔCA and add it on[‡] the reserve capital account;
- The management of the bank charges the amount ΔKVA to the client and adds it on[§] the risk margin account.

The funds transfer price of a deal is the all-inclusive XVA add-on charged by the bank to the client in the form of a rebate with respect to the mark-to-market ΔMtM of the deal. Under the above scheme, the overall price charged to the client for the deal is $\Delta\text{MtM} - \Delta\text{CA} - \Delta\text{KVA}$, i.e.

$$\begin{aligned} \text{FTP} &= \Delta\text{CA} + \Delta\text{KVA} = \Delta\text{CVA} + \Delta\text{FVA} + \Delta\text{KVA} \\ &= \Delta\text{FV} + \Delta\text{CL} + \Delta\text{KVA}, \end{aligned} \quad (43)$$

by (16) and the last line in (23) (which still hold in continuous time, see Albanese and Crépey (2020, equations (1) and (66))) applied to the portfolios with and without the new deal.

REMARK 4.3 As opposed to the ΔXVA terms, which entail portfolio-wide computations, ΔMtM reduces to the so-called clean valuation of the new deal, by trade-additivity of MtM (as follows from Albanese and Crépey (2020, equations (25) and (37))).

Obviously, the legacy portfolio of the bank has a key impact on the FTP. It may very well happen that the new deal is risk-reducing with respect to the portfolio, in which case $\text{FTP} < 0$, i.e. the overall, XVA-inclusive price charged by the bank to the client would be $\Delta\text{MtM} - \text{FTP} > \Delta\text{MtM}$ (subject of course to the commercial attitude adopted by the bank under such circumstance).

In order to exclude for simplicity jumps of our L and KVA processes at θ (the ones related to the initial portfolio, but also those, starting at time θ , corresponding to the augmented portfolio), we assume a quasi-left continuous model filtration \mathbb{G} and a \mathbb{G} predictable stopping time θ . The first assumption excludes that martingales can jump at predictable times. It is satisfied in all practical models and, in particular, in all models with Lévy or Markov chain driven jumps. The second assumption is reasonable regarding the time at which a financial contract is concluded. Note that it was actually already assumed regarding the (fixed) time 0 at which the portfolio of the bank is supposed to have been set up in the first place.

LEMMA 4.2 *Assuming the new trade at time θ handled by the trade incremental policy of Definition 4.1 after that the balance conditions (4) have been held before θ , then shareholder equity SHC° (see Remark 4.2) is a \mathbb{Q}^* submartingale on $[0, \theta] \cap \mathbb{R}_+$, with drift coefficient $h\text{SCR}$ killed at τ .*

Proof In the case of a trade incremental portfolio, a priori, the second identity in (42) is only guaranteed to hold before θ .

[†] i.e. remove $(-\Delta\text{MtM})$ from, if $\Delta\text{MtM} < 0$.

[‡] i.e. remove $(-\Delta\text{CA})$ from, if $\Delta\text{CA} < 0$.

[§] i.e. removes $(-\Delta\text{KVA})$ from, if $\Delta\text{KVA} < 0$.

However, in view of the observation made in Remark 2.4 and because, under our (harmless) technical assumptions, there can be no dividends arising from the portfolio expanded with the new deal (i.e. jumps in the related processes L and KVA, defined on $[\theta, +\infty)$) at time θ itself, the process SHC does not jump at θ . The process L and KVA related to the legacy portfolio cannot jump at θ either. As a result, the second identity in (42) still holds at θ . It is therefore valid on $[0, \theta] \cap \mathbb{R}_+$. The result then follows from the respective martingale and supermartingale properties of the (original) processes L° and KVA recalled before and after Proposition 4.1. ■

The above XVA strategy can be iterated between and throughout every new trade. We call this approach the *trade incremental cost-of-capital XVA strategy*. By an iterated application of Lemma 4.2 at every new trade, we obtain the following:

PROPOSITION 4.2 *Under a dynamic and trade incremental cost-of-capital XVA strategy, shareholder equity SHC° is a \mathbb{Q}^* submartingale on \mathbb{R}_+ , with drift coefficient $h\text{SCR}$ killed at τ .*

Thus, a trade incremental cost-of-capital XVA strategy results in a sustainable strategy for profits retention, both between and throughout deals, which was already the key principle behind Solvency II (see section 1.1). Note that, without the KVA (i.e. for $h = 0$), the (risk-free discounted) shareholder equity process SHC° would only be a \mathbb{Q}^* martingale, which could only be acceptable to shareholders without risk aversion (cf. section 3.3.1).

4.3. Computational challenges

Figure 2 yields a picturesque representation, in the form of a corresponding XVA dependence tree, of the continuous-time XVA equations.

For concreteness, we restrict ourselves to the case of bilateral trading in what follows, referring the reader to Albanese et al. (2020, Section 6.2) for the more general and realistic situation of a bank also involved in centrally cleared trading. As visible from the corresponding equations in section A, the CVA of the bank can then be computed as the sum of its CVAs restricted to each netting set (or counterparty i of the bank, with default time denoted by τ_i in figure 2). The initial margins and the MVA are also most accurately calculated at each netting set level. By contrast, the FVA is defined in terms of a semilinear equation that can only be solved at the level of the overall derivative portfolio of the bank. The KVA can only be computed at the level of the overall portfolio and relies on conditional risk measures of future fluctuations of the shareholder trading loss process L° , which itself involves future fluctuations of the other XVA processes (as these are part of the bank liabilities).

Moreover, the fungibility of capital at risk with variation margin (cf. Remark 3.4) induces a coupling between, on the one hand, the ‘backward’ FVA and KVA processes and, on the other hand, the ‘forward’ shareholder loss process L° . As in the static case of section 3.4 (cf. the last paragraph there), the ensuing forward backward system can be decoupled by Picard iteration.

These are heavy computations encompassing all the derivative contracts of the bank. Yet these computations require accuracy so that trade incremental XVA computations, which are required as XVA add-ons to derivative entry prices (cf. section 4.2), are not in the numerical noise of the machinery.

As developed in Abbas-Turki et al. (2018, Section 3.2), computational strategies for (each Picard iteration of) the XVA equations involve a mix of nested Monte Carlo (NMC) and of simulation/regression schemes, optimally implemented on GPUs. In view of figure 2, a pure NMC approach would involve five nested layers of simulation (with respective numbers of paths $M_{xva} \sim \sqrt{M_{mm}}$, see Abbas-Turki et al. (2018, Section 3.3)). Moreover, nested Monte Carlo implies intensive repricing of the mark-to-market cube, i.e. pathwise MtM valuation for each netting set, or/and high dimensional interpolation. In this work, we use no nested Monte Carlo or conditional repricing of future MtM cubes: beyond the base MtM layer in the XVA dependence tree, each successive layer (from right to left in figure 2, at each Picard iteration) will be ‘learned’ instead.

4.4. Deep (quantile) regression XVA framework

We denote by \mathbb{E}_t , VaR_t , and $\mathbb{E}\mathbb{S}_t$ (and simply, in case $t = 0$, \mathbb{E} , VaR , and $\mathbb{E}\mathbb{S}$) the time- t conditional expectation, value-at-risk, and expected shortfall with respect to the bank survival measure \mathbb{Q} .

We compute the mark-to-market cube using CUDA routines. The pathwise XVAs are obtained by deep learning regression, i.e. extension of Longstaff and Schwartz (2001) kind of schemes to deep neural network regression bases as also considered in Huré et al. (2020) or Beck et al. (2019), based on the classical quadratic (also known as mean square error, MSE) loss function. The conditional value-at-risks and expected shortfalls involved in the embedded pathwise EC and IM computations are obtained by deep quantile regression, as follows.

Given features X and labels Y (random variables), we want to compute the conditional value-at-risk and expected shortfall functions $q(\cdot)$ and $s(\cdot)$ such that $\text{VaR}(Y|X) = q(X)$ and $\mathbb{E}\mathbb{S}(Y|X) = s(X)$. Recall from Fissler et al. (2016) and Fissler and Ziegel (2016) that value-at-risk is *elicitable*, expected shortfall is not, but their pair is *jointly elicitable*. Specifically, we consider loss functions ρ of the form (where in our notation Y is a signed loss, whereas it is a signed gain in their paper)

$$\begin{aligned} \rho(q(\cdot), s(\cdot); X, Y) &= (1 - \alpha)^{-1} (f(Y) - f(q(X)))^+ + f(q(X)) \\ &\quad + g(s(X)) - \dot{g}(s(X)) (s(X) - q(X)) \\ &\quad - (1 - \alpha)^{-1} (Y - q(X))^+. \end{aligned} \quad (44)$$

One can show (cf. also Dimitriadis and Bayer 2019) that, for a suitable choice of the functions f, g including $f(z) = z$ and $g = -\ln(1 + e^{-z})$ (our choice in our numerics), the pair of the conditional value-at-risk and expected shortfall functions is the minimizer, over all measurable pair-functions $(q(\cdot), s(\cdot))$,

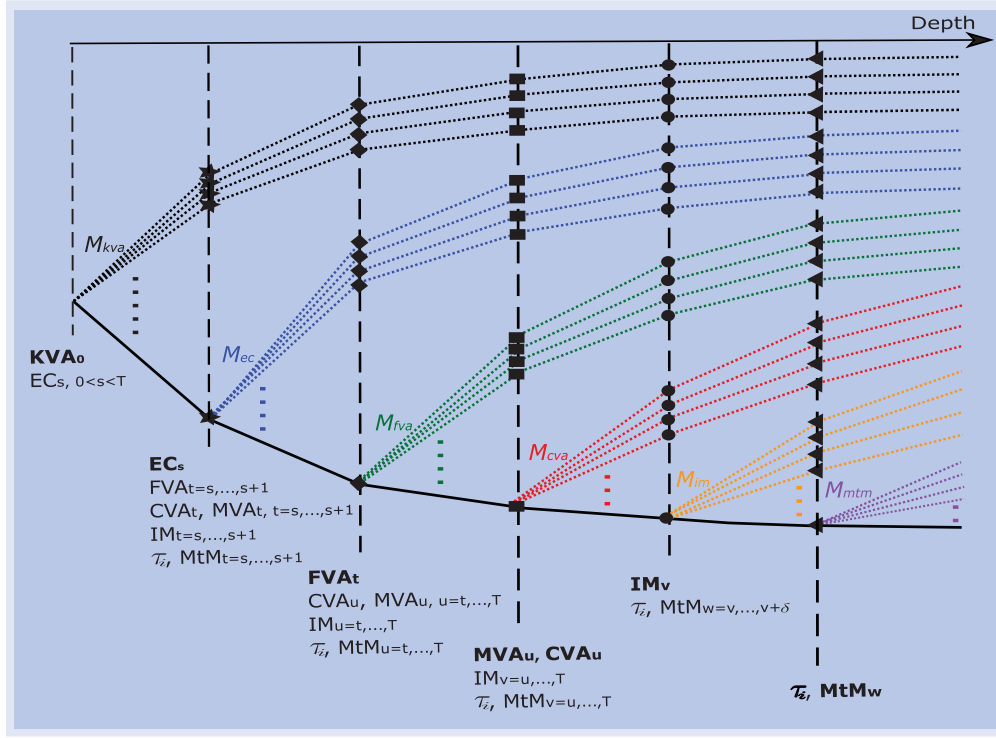


Figure 2. The XVA equations dependence tree (Source: Abbas-Turki *et al.* (2018)).

of the error

$$\mathbb{E}\rho(q(\cdot), s(\cdot); X, Y). \quad (45)$$

In practice, one minimizes numerically the error (45), based on m independent simulated values of (X, Y) , over a parametrized family of functions $(q, s)(x) \equiv (q, s)_\theta(x)$. Dimitriadis and Bayer (2019) restrict themselves to multilinear functions. In our case we use a feedforward neural network parameterization (see e.g. Goodfellow *et al.* 2016). The minimizing pair $(q, s)_\theta$ then represents the two scalar neural network approximations of the conditional value-at-risk and expected shortfall functions pair.

The left and right panels of figure 3 show the respective deep neural networks for pathwise value-at-risk/expected shortfall (with error (45)) and pathwise XVAs (with classical quadratic norm error). Deep learning methods often show particularly good generalization and scalability performances (cf. section 5.5). In the case of conditional value-at-risk and expected shortfall computations, deep learning quantile regression is also easier to implement than more naive methods, such as the resimulation and sort-based scheme of Barrera *et al.* (2019) for the value-at-risk and expected shortfall at each outer node of a nested Monte Carlo simulation.

The neural network topology and hyper-parameters used by default in our examples are detailed in table 2. We use hyperbolic tangent activation functions in all cases.

Algorithm 1 yields our fully (time and space) discrete scheme for simulating the Picard iteration (A10) until numerical convergence to the XVA processes. Note that, as opposed to more rudimentary, expected exposure based XVA computational approaches (see Section 1 in Abbas-Turki *et al.* 2018), this algorithm requires the simulation of the counterparty defaults.

5. Swap portfolio case study

We consider an interest rate swap portfolio case study with counterparties in different economies, first involving 10 one-factor Hull White interest-rates, 9 Black-Scholes exchange rates, and 11 Cox-Ingersoll-Ross default intensity processes. The default times of the counterparties and the bank itself are jointly modelled by a ‘common shock’ or dynamic Marshall-Olkin copula model as per Crépey *et al.* (2014, Chapt. 8–10) and Crépey and Song (2016) (see also Elouerkhaoui 2007, 2017).

This whole setup results in about 40 risk factors used as deep learning features (including the counterparty default indicators).

In this model, we consider a bank portfolio of 10K randomly generated swap trades, with

- trade currency and counterparty both uniform on $[1, 2, 3, \dots, 10]$,
- notional uniform on $[10K, 20K, \dots, 100K]$,
- collateralization (cf. section A.4): either ‘no CSA counterparty’ without initial margin (IM) nor variation margin (VM), or ‘CSA counterparty’ with $VM = MtM$ and posted initial margin (PIM) pledged at 99% gap risk value-at-risk, received initial margin (RIM) covering 75% gap risk and leaving excess as residual gap CVA,
- for economic capital, 97.5% expected shortfall of 1-year ahead trading loss of the bank shareholders.

By default we use Monte Carlo simulation with 50K paths of 16 coarse (pricing) and 32 fine (risk factors) time steps per year.

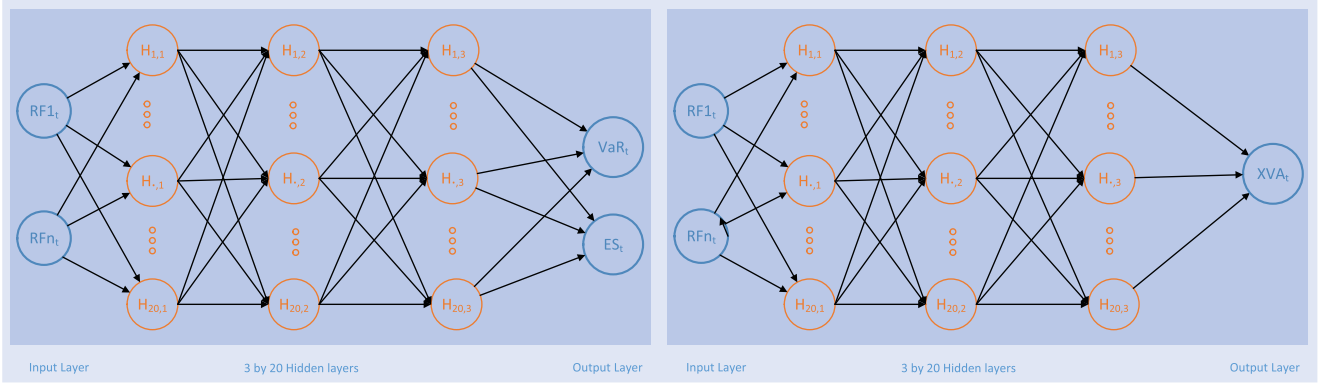


Figure 3. Neural networks with state variables (realizations of the risk factors at the considered pricing time) as features. (Left) Joint value-at-risk/expected shortfall neural network: output is joint estimate of pathwise conditional value-at-risk and expected shortfall, at a selected confidence level, of the label (inputs to initial margin or economic capital) given the features. (Right) XVAs neural network: output is estimate of pathwise conditional mean of the label (XVA generating cash flows) given the features.

Table 2. Neural network topology and learning parameters used by default in our numerics (portf. \equiv overall derivative portfolio of the bank).

	CVA	FVA	IM	MVA	Gap CVA [†]	EC	KVA
Hidden layers	3	5	3	3	3	3	3
Hidden layer size	20	6	20	20	20	20	20
Learning rate	0.025	0.025	0.05	0.1	0.1	0.025	0.1
Momentum	0.95	0.95	0.5	0.5	0.5	0.95	0.5
Iterations	100	50	150	100	100	100	100
Loss function	MSE	MSE	(44)	MSE	(44)	(44)	MSE
Application	Netting set	Portf.	Netting set	Netting set	Netting set	Portf.	Portf.

[†]See section 5.

5.1. Validation results

The validation of our deep learning methodology is done in the setup of a portfolio of swaps issued at par, with final maturity $T = 10$ years, without initial margin (IM) nor variation margin (VM).

We first focus on the CVA, as the latter is amenable to validation by a standard nested Monte Carlo ('NMC') methodology. Figures 4–6 show that the learned CVA is consistent with that obtained from a nested Monte Carlo simulation. Regarding figure 6 (and also later below), note the equivalence of

Algorithm 1 Deep XVAs algorithm.

- Simulate forward m realizations (Euler paths) of the market risk factor processes and of the counterparty survival indicator processes (i.e. default times) on a refined time grid;
- For each pricing time $t = t_i$ of a pricing time grid, with coarser time step denoted by h , and for each counterparty c :
 - Learn the corresponding VaR_t and ES_t terms visible in (A11) or (under the time-discretized outer integral in) (A13);
 - Learn the corresponding \mathbb{E}_t terms visible in (A12) through (A14);
 - Compute the ensuing pathwise CVA and MVA as per (A12)–(A14);
- For FVA⁽⁰⁾, consider the following time discretization of (A9) (in which λ is the risky funding spread process of the bank) with time step h :

$$\text{FVA}_t^{(0)} \approx \mathbb{E}_t[\text{FVA}_{t+h}^{(0)}] + h\lambda_t \left(\sum_c J_t^c (P_t^c - \text{VM}_t^c) - \text{CVA}_t - \text{MVA}_t - \text{FVA}_t^{(0)} \right)^+ \quad (1)$$

and, for each $t = t_i$, learn the corresponding \mathbb{E}_t in (46), then solve the semi-linear equation for FVA_t⁽⁰⁾;

- For each Picard iteration k (until numerical convergence), simulate forward $L^{(k)}$ as per the first line in (A10) (which only uses known or already learned quantities), and:
 - For economic capital $\text{EC}^{(k)}$, for each $t = t_i$, learn $\mathbb{E}_t((L^{(k)})_{t+1}^\circ - (L^{(k)})_t^\circ)$ (cf. Definition A.1);
 - KVA^(k) and FVA^(k) then require a backward recursion solved by deep learning approximation much like the one for FVA⁽⁰⁾ above.

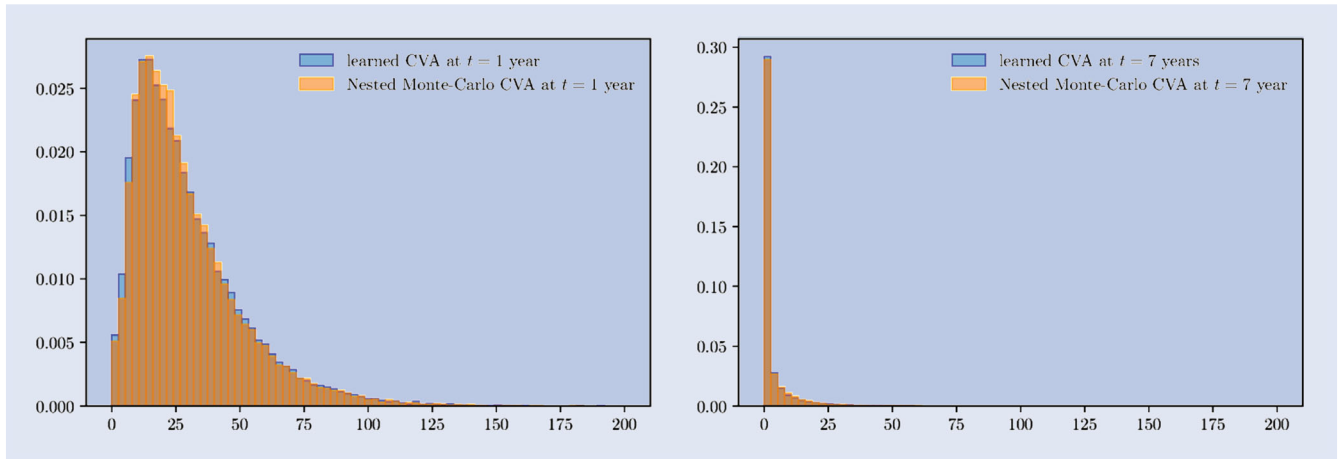


Figure 4. Random variables CVA_1^c and CVA_7^c (in the case of a no CSA netting set c , respectively observed after 1 and 7 years) obtained by learning (blue histogram) versus nested Monte Carlo (orange histogram). All histograms are based on out-of-sample paths.

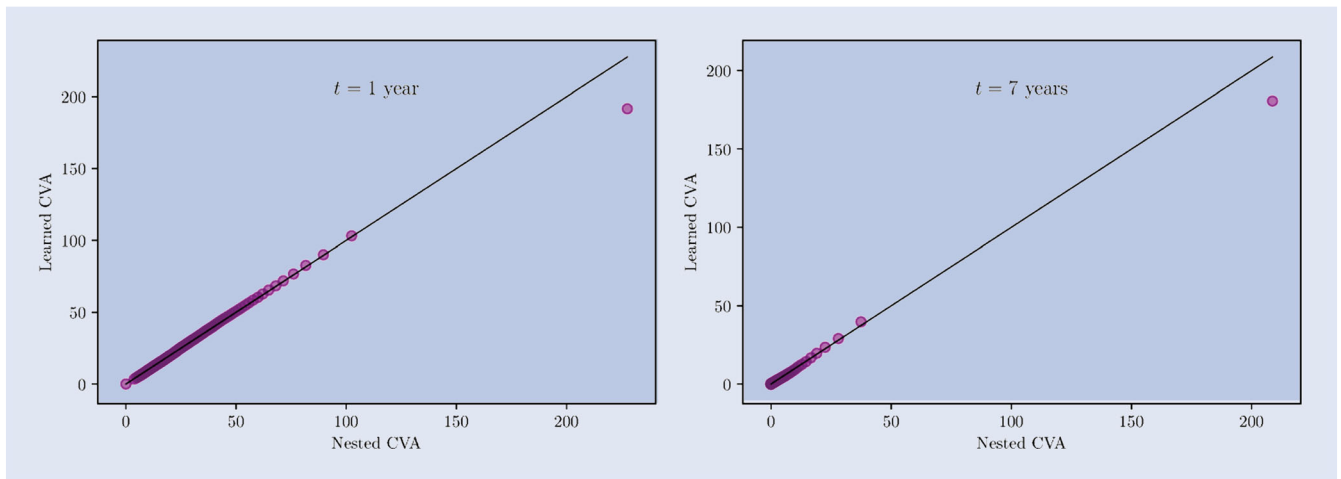


Figure 5. QQ-plot of learned versus nested Monte Carlo CVA for the random variables CVA_1^c (left) and CVA_7^c (right). Paths are out-of-sample.

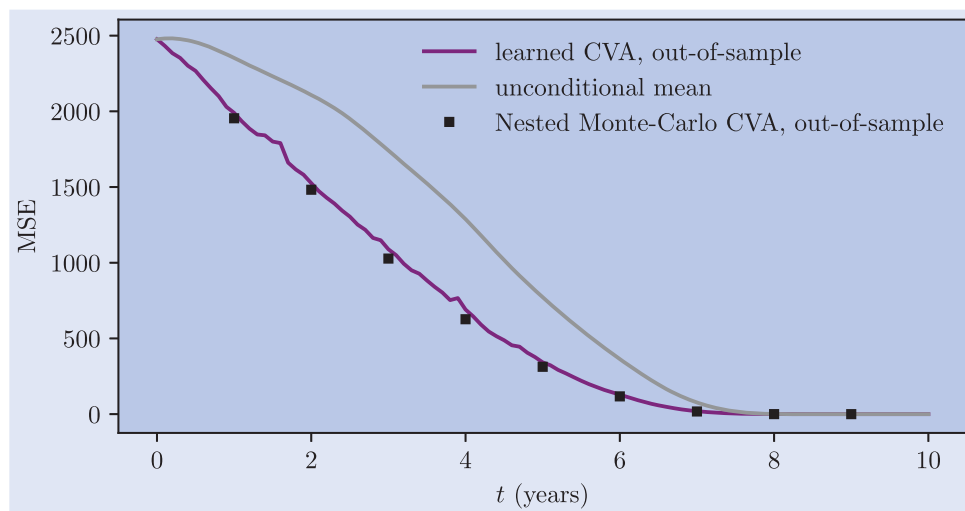


Figure 6. Empirical quadratic loss of each CVA estimator at all coarse time-steps. The lower, the closer to the true conditional expectation (cf. (47)). Since the nested Monte Carlo method is computationally expensive, it was carried out only once every 10 coarse time-steps.

optimising the mean quadratic error:

- between the neural network-learned estimator $h(X)$ and the labels Y ('MSE'), $\mathbb{E}[(h(X) - Y)^2]$, and
- between the neural network-learned estimator and the conditional expectation $\mathbb{E}[Y|X]$ (in our case estimated by NMC), $\mathbb{E}[(h(X) - \mathbb{E}[Y|X])^2]$.

The equivalence stems from the following identities, which hold for any random variables X, Y and hypothesis function h such that Y and $h(X)$ are square integrable:

$$\begin{aligned} & \mathbb{E}[(h(X) - Y)^2] \\ &= \mathbb{E}[(h(X) - \mathbb{E}[Y|X])^2] + \mathbb{E}[(\mathbb{E}[Y|X] - Y)^2] \\ & \quad + 2\mathbb{E}[(h(X) - \mathbb{E}[Y|X])(\mathbb{E}[Y|X] - Y)] \\ &= \mathbb{E}[(h(X) - \mathbb{E}[Y|X])^2] + \mathbb{E}[\text{Var}(Y|X)] \end{aligned} \quad (47)$$

(as the second line vanishes), where $\mathbb{E}[\text{Var}(Y|X)]$ does not depend on h .

The CVA error profile on figure 6 reveals slightly more difficulty in learning the earlier CVAs. This is because of a higher variance of the corresponding cash flows (integrated over longer time frames) in conjunction with a lower variance of the features (risk factors diffused over shorter time horizons).

Table 3 shows the computational cost and accuracy of the nested Monte Carlo method for different number of inner paths, using 32 768 outer paths. The convergence is already achieved for approximately 128 inner paths, in line with the NMC square root rule that is recalled in an XVA setup in Abbas-Turki *et al.* (2018, Section 3.3). Figure 7 and table 4 show that a good accuracy can be achieved through learning at a lower computational cost than through nested Monte Carlo, while also enjoying the advantages of the approach being parametric. Indeed, once the CVA is learned, one would pay only the cost of inference later on, which is generally negligible compared to training time. By contrast, a nested Monte Carlo approach would require to relaunch the nested simulations every time the CVA estimator is needed on new paths. Early stopping could be used to help reduce training time further while improving regularization.

More generally, in the presence of a multiple number of XVA layers (cf. figure 2), a purely nested Monte Carlo approach would require multiple layers of nested simulations, which would amount to a computational time that is exponential in the number of XVA layers, while the computational complexity for the learning approach is linear.

As with mainstream interpolation (as opposed to regression in our case) learning problems, a good architecture is key to better learning and hence better approximation of our XVA metrics. As expected, increasing the model capacity reduces the in-sample error as shown in the bottom panel of figure 8. Although fine-tuning in our case suggests a single layer yields the best out-of-sample performance for the CVA, a standard guess such as three layers can also be considered good enough as shown in the top panel. Of course such conclusions may depend on the complexity of the portfolio and the number of counterparties and risk factors.

Figure 9 shows the learned FVA⁽⁰⁾ profile as per (46). The orange FVA curve represents the mean FVA originating cash

Table 3. Accuracy and computation times for the estimation of a CVA at a given coarse time-step using the nested Monte Carlo procedure.

# of inner paths	MSE (vs labels)	Computational time (s)
2	0.523	37.562
4	0.427	37.815
8	0.393	37.819
16	0.370	38.988
32	0.360	40.707
64	0.353	57.875
128	0.348	157.536
256	0.349	301.406
512	0.348	584.475
1024	0.348	1213.756

Note: The MSE here is the mean quadratic error between the nested Monte Carlo estimator and the labels, and hence quantifies how well it is doing as a projection.

flows, which, in principle as on the picture, matches the blue mean FVA itself learned from these cash flows. The 5th and 95th percentiles FVA estimates are a bit less smooth in time then the mean profiles, as expected.

Figure 10 (left) is a sanity check that the profiles of the successive iterates $L^{(k)}$ of the shareholder trading loss process L° in Algorithm 1 converge rapidly with k . Figure 10 (right) shows the loss process $L^{(3)}$, displayed as its mean and mean ± 2 stdev profiles. Consistent with its martingale property, the loss process $L^{(3)}$ appears numerically centred around zero. The latter holds, at least, beyond $t \sim 5$ years. For earlier times, the regression errors, accumulated backward across pricing times since the final maturity of the portfolio, induce a non-negligible bias (the corresponding confidence intervals no longer contains 0). This is the reason why we use a coarser pricing time step than simulation time step in Algorithm 1.

5.2. Portfolio-wide XVA profiles

For the financial case study that follows, we consider

- swap rates uniformly distributed on $[0.005, 0.05]$ (hence swaps already in-the-money or out-of-the-money at time 0),
- number of six-monthly coupon resets uniform on $[5 \dots 60]$ (final maturity of the portfolio $T = 30$ years),
- portfolio direction: either 'asset heavy' bank mostly in the receivables in the future, or 'liability-heavy' bank mostly in the payables in the future (respectively corresponding, with our data, to a bank 75% likely to pay fixed in the swaps, or 75% likely to receive fixed).

The figures that follow only display profiles, i.e. term structures, that is, expectations as a function of time of the corresponding processes. But all these processes are computed pathwise, based on the deep learning regression and quantile regression methodology of section 4.4, allowing for all XVA inter-dependencies. Of course, XVA profiles (or pathwise XVAs if wished) are much more informative for traders than

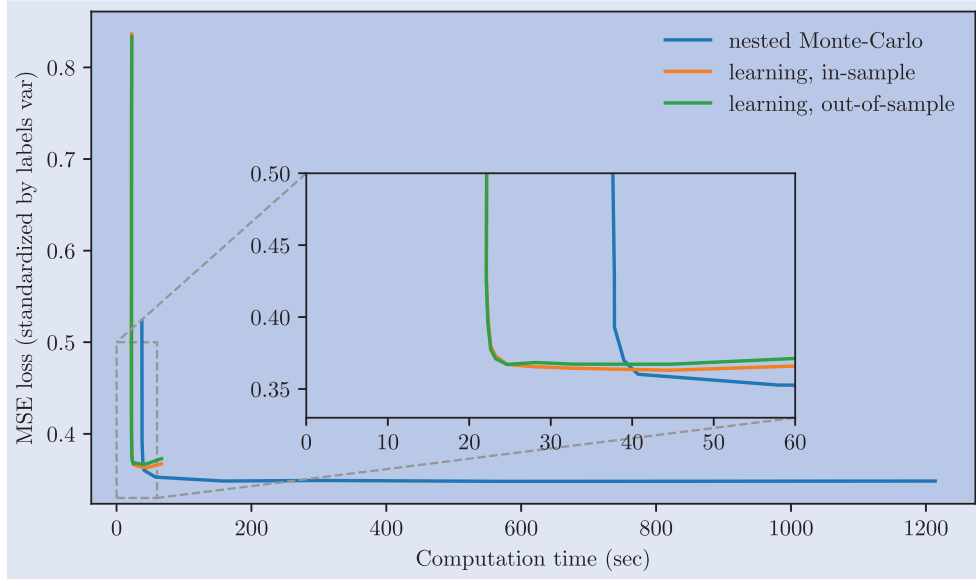


Figure 7. Speed versus accuracy in the case of a CVA at a given pricing time. We kept varying the number of inner paths for the nested Monte Carlo estimator and the number of epochs for the learning approach and recorded the computation time and the empirical quadratic loss.

Table 4. Accuracy and computation times (in sec) for the calculation of a CVA at a given coarse time-step using the learning approach.

# of epochs	MSE (vs NMC CVA)	MSE (vs labels)	Simulation time	Training time
1	0.977	0.979	21.992	0.880
2	0.729	0.729	21.992	0.434
4	0.423	0.425	21.992	0.524
8	0.399	0.401	21.992	0.719
16	0.371	0.369	21.992	1.088
32	0.369	0.365	21.992	1.800
64	0.370	0.363	21.992	3.243
128	0.371	0.363	21.992	6.227
256	0.370	0.361	21.992	10.883
512	0.370	0.362	21.992	20.096
1024	0.371	0.362	21.992	39.338

Notes: MSE against NMC CVA is the mean quadratic error between the learned CVA and a CVA obtained using a nested Monte Carlo with 512 inner paths, while MSE against labels designates the mean quadratic error between the learned CVA and the labels that were used during training and thus quantifies how well it is doing as a projection. Both errors are, respectively, normalized by the variances of the nested Monte Carlo estimator and of the labels. The paths used here are out-of-sample.

the spot XVA values (or time 0 confidence intervals) returned by most XVA systems.

Assuming 10 counterparties, figure 11 shows the GPU-generated profiles of

$$\text{MtM} = \sum_c P^c \mathbb{1}_{[0, \tau_c^s)} \quad (48)$$

in the case of the asset-heavy portfolio and of the liability-heavy portfolio.

Figure 12 shows the portfolio-wide XVA profiles of the asset-heavy (*top*) vs. liability-heavy (*bottom*) portfolio and of the no CSA (*left*) vs. CSA portfolio (*right*). Obviously, asset-heavy or no CSA means more CVA. The corresponding curves also emphasize the transfer from counterparty credit into liquidity funding risk prompted by extensive collateralisation. Yet FVA/MVA risk is ignored in current derivatives capital regulation.

Figure 13 shows that (*top left*) capital at risk as funding (cf. section 3.4) has a material impact on the already (reserve capital as funding) reduced FVA, (*top right*) treating KVA as a risk margin (cf. (26)) gives a huge discounting impact, (*bottom left*) deep learning detects material initial margin convexity in the asset-heavy CSA portfolio, and (*bottom right*) deep learning detects material economic capital convexity in the asset-heavy no CSA portfolio.

The above findings demonstrate the necessity of pathwise capital and margin calculations for accurate FVA, MVA, and KVA calculations.

5.3. Trade incremental XVA profiles

Next, we consider, on top of the previous portfolios, an incremental trade given as a par 30 year (receive fix or pay

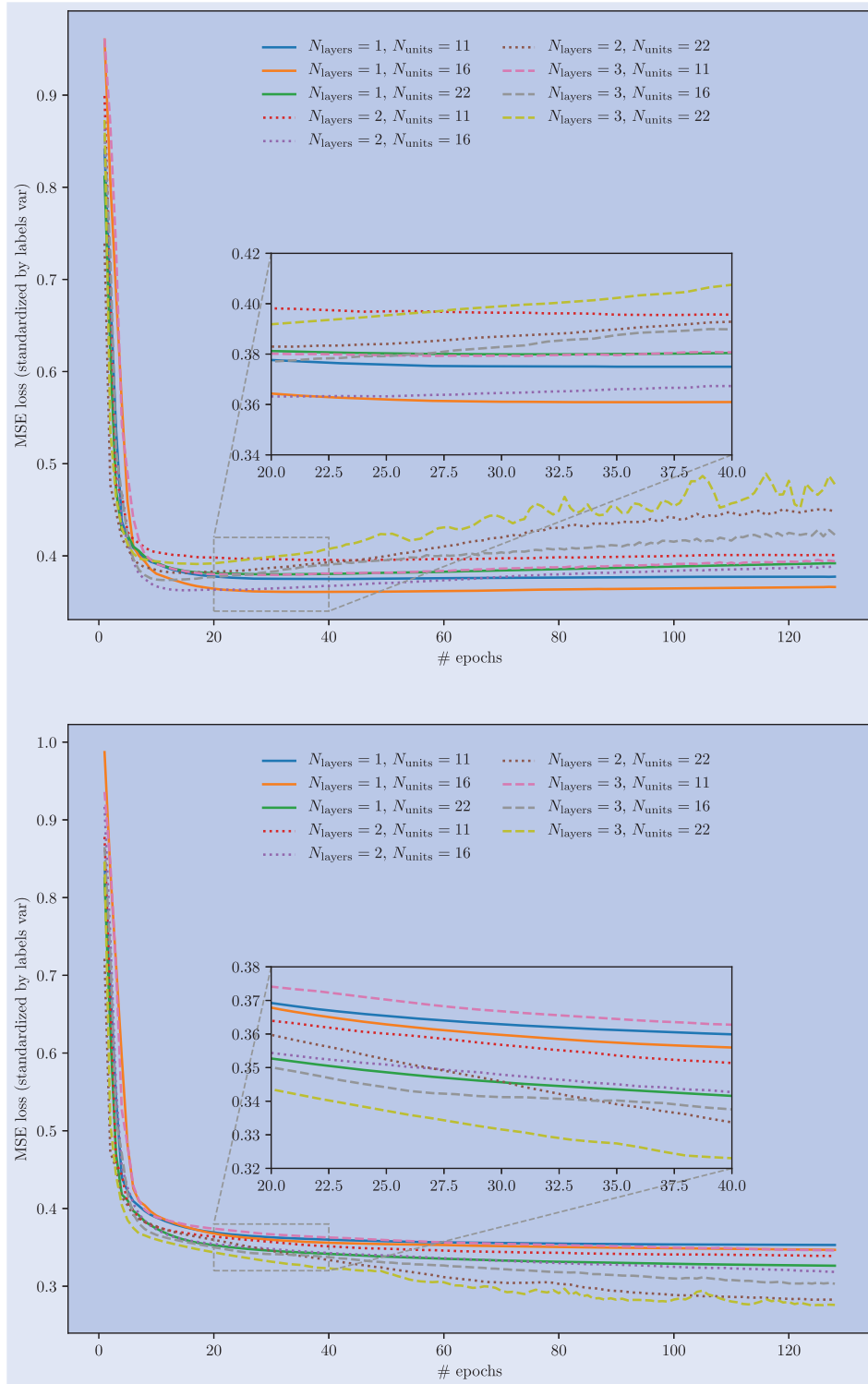


Figure 8. Empirical quadratic loss during CVA learning at time-step $t = 5$ years, standardized by the variance of the labels. (Bottom) Paths are in-sample. (Top) Paths are out-of-sample.

fix) swap with 100K notional. Figure 14 shows the trade incremental XVA profiles produced by our deep learning approach. Note that, for obtaining such smooth incremental profiles, it has been key to use common random numbers, as much as possible, between the original portfolio XVA computations and the ones regarding the portfolio expanded with the new trade.

5.4. Trade and hedge incremental XVA profiles

Our model assumes the market risk of trades to be fully hedged (see the paragraph following Remark 2.2 and the proofs of Lemma 3.2 and Proposition 4.1). In the previous subsection, the new swap was implicitly meant to be hedged, in terms of market risk, by the clean desks,

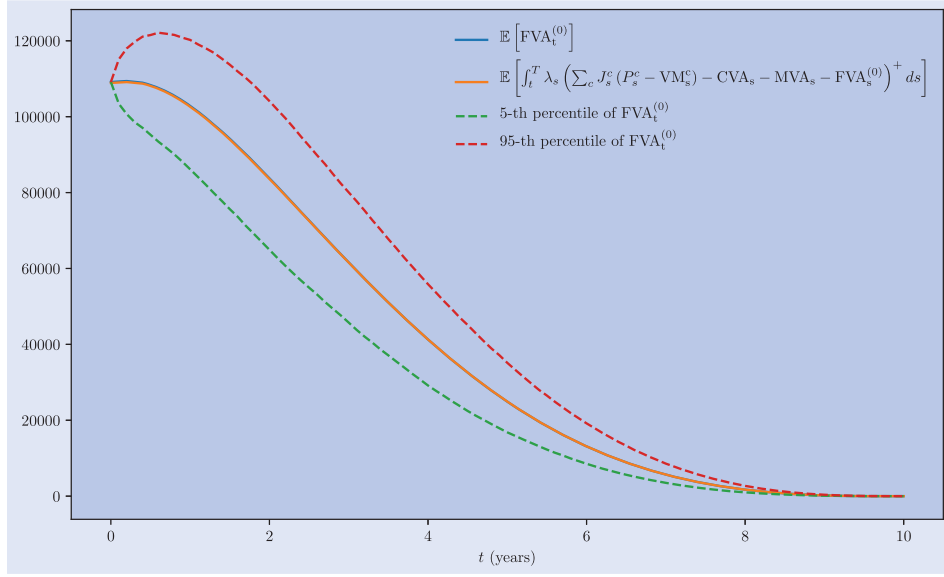
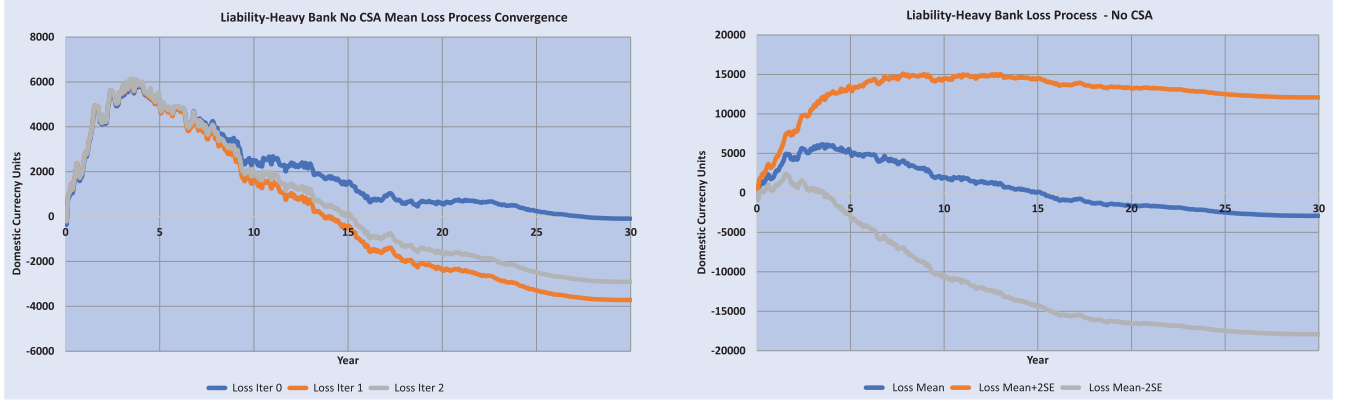
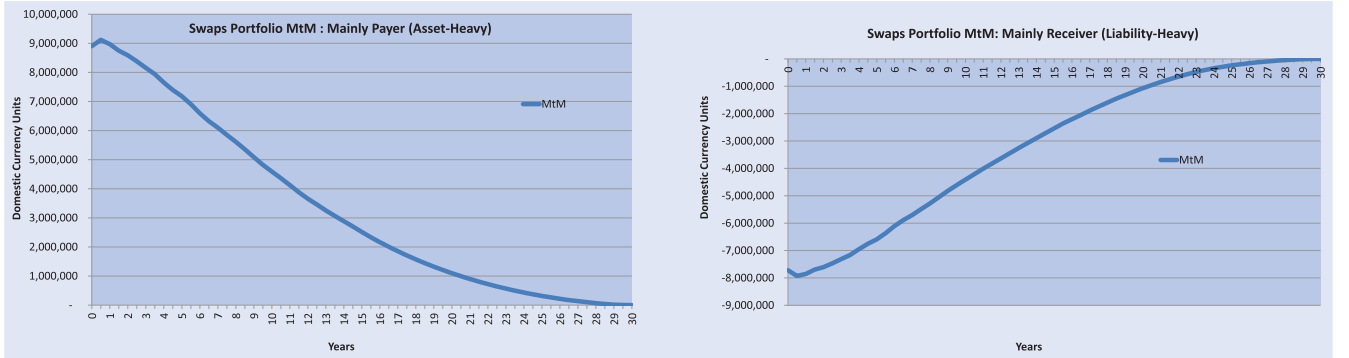
Figure 9. Learned $FVA^{(0)}$.Figure 10. (Left) Profiles of the processes $L^{(k)}$, for $k = 1, 2, 3$; (Right) Mean ± 2 stdev profiles of the process $L^{(3)}$.

Figure 11. MtM profiles. (Left) Asset-heavy portfolio. (Right) Liability-heavy portfolio.

through an accordingly modified hedging loss process \mathcal{H} (see section 2.1). Here we consider an alternative situation where the market risk of the new swap is back-to-back hedged via a financial, hedge counterparty. Specifically, we deal with

- 10 counterparties: 8 no CSA clients and 2 bilateral VM/IM CSA hedge counterparties,
- portfolios of 5K randomly generated swap trades as before, plus 5K corresponding hedge trades,

- an incremental trade given as a par 30 year swap with 100K notional, along with the corresponding hedge trade.

In particular, $MtM_0 = 0$ (cf. (48)), in both portfolios excluding or including the new swap. In case a client or hedge counterparty defaults, the corresponding market hedge is assumed to be rewired through the clean desks via an accordingly modified hedging loss process \mathcal{H} .

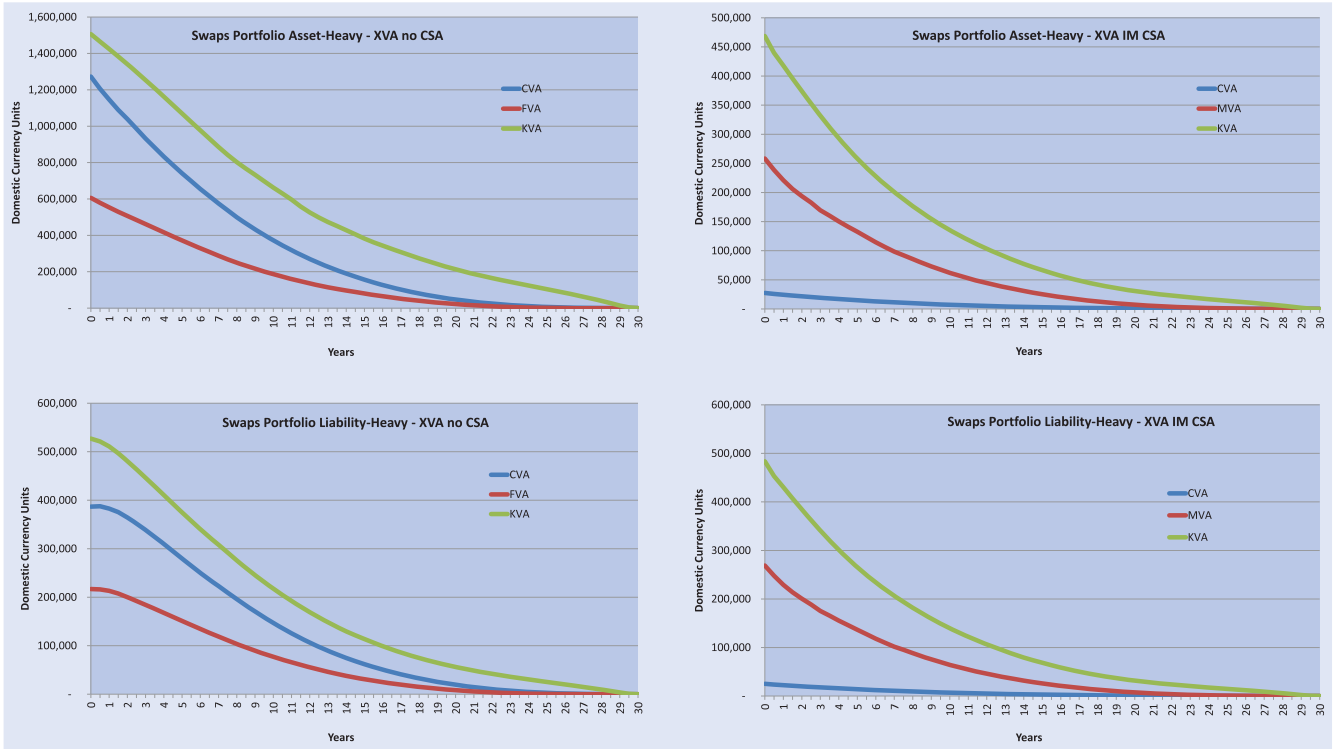


Figure 12. (Top left) Asset-heavy portfolio, no CSA. (Top right) Asset-heavy portfolio under CSA. (Bottom left) Liability-heavy portfolio, no CSA. (Bottom right) Liability-heavy portfolio under CSA.

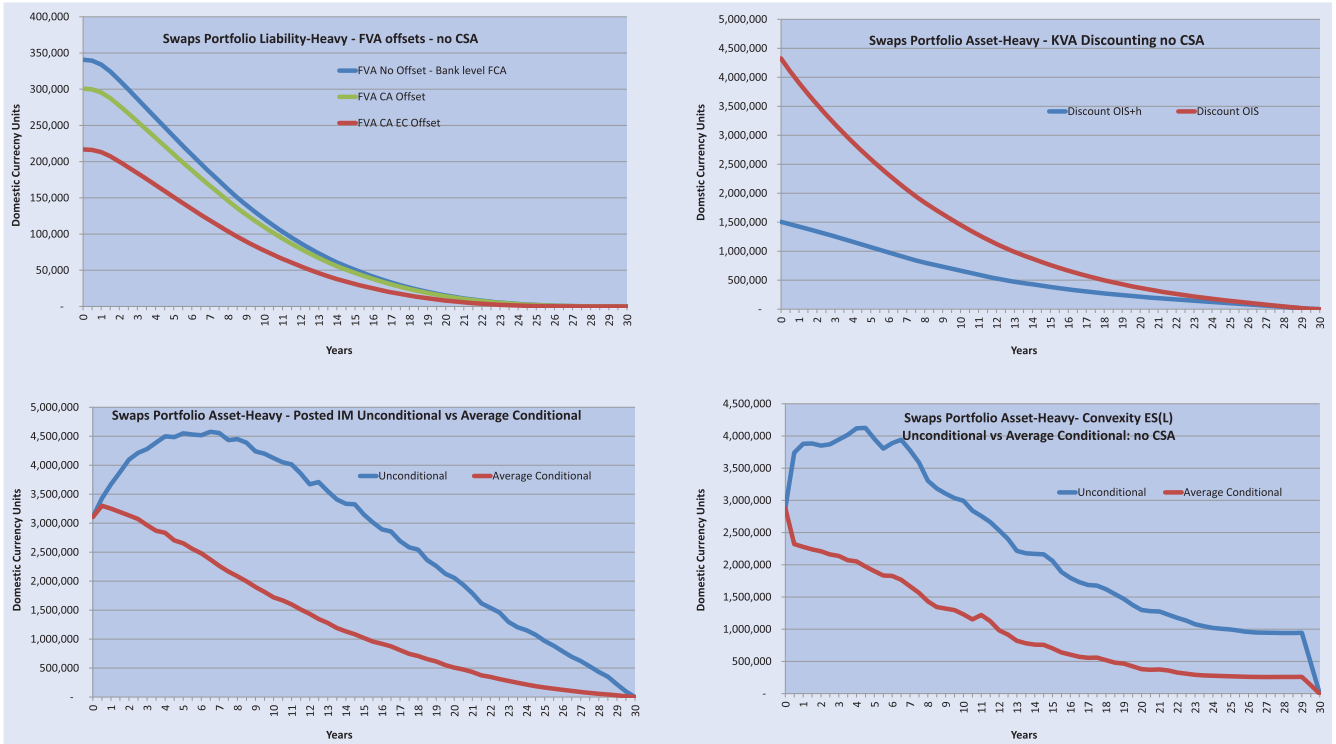


Figure 13. (Top left) FVA ignoring the off-setting impact of reserve capital and capital at risk, cf. section 3.4 (blue), FVA as per (A9) accounting for the off-setting impact of reserve capital but ignoring the one of capital at risk (green), refined FVA as per (A4) accounting for both impacts (red). (Top right) KVA ignoring the off-setting impact of the risk margin, i.e. with CR instead of $(CR - KVA)$ in (A8) (red), refined KVA as per (A6)–(A7) (blue). (Bottom left) In the case of the asset-heavy portfolio under CSA, unconditional PIM profile, i.e. with VaR_t replaced by VaR in (A11) (blue), vs. pathwise PIM profile, i.e. mean of the pathwise PIM process as per (A11) (red). (Bottom right) In the asset-heavy portfolio no CSA case, unconditional economic capital profile, i.e. EC profile ignoring the words ‘time- t conditional’ in Definition A.1 (blue), vs. pathwise economic capital profile, i.e. mean of the pathwise EC process as per Definition A.1 (red).

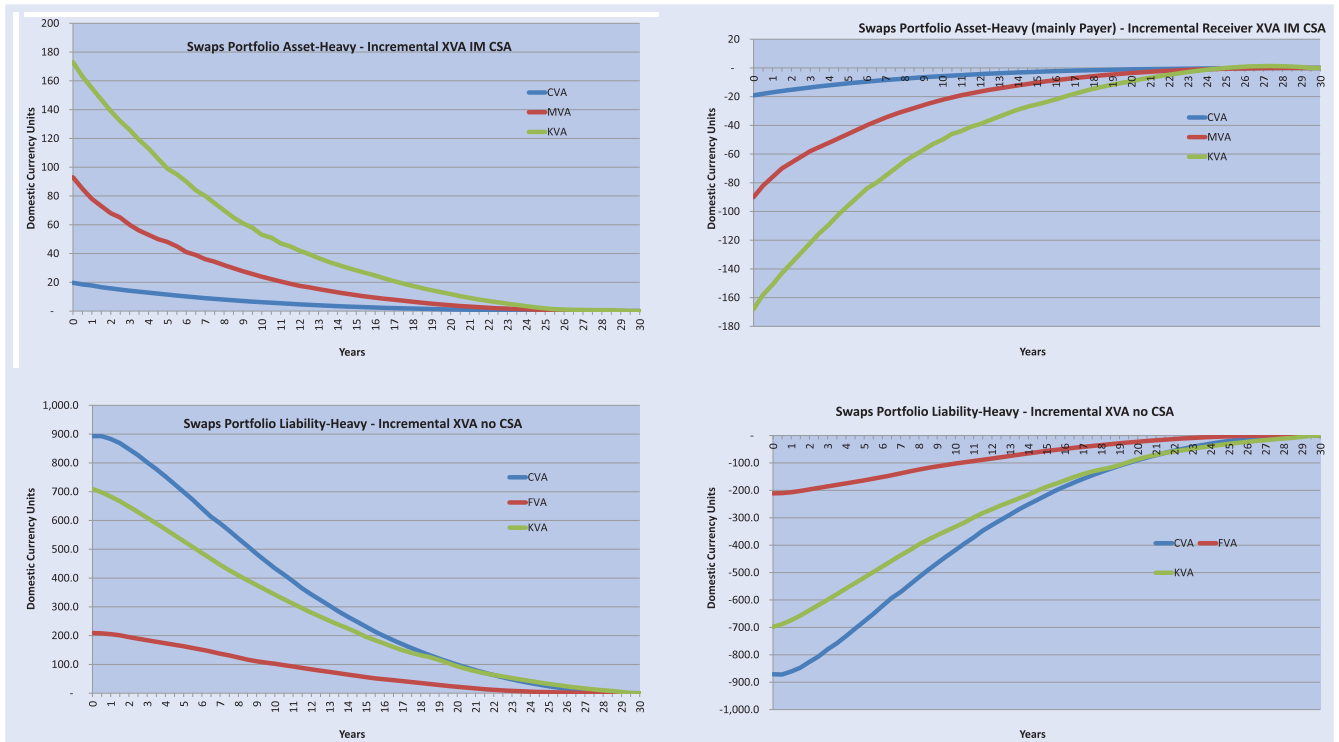


Figure 14. (Top left) Asset-heavy portfolio, no CSA. Incremental receive fix trade. (Top right) Liability-heavy portfolio, no CSA. Incremental pay fix trade. (Bottom left) Asset-heavy portfolio under CSA. Incremental Pay Fix Trade. (Bottom right) Liability-heavy portfolio under CSA. Incremental receive fix trade.

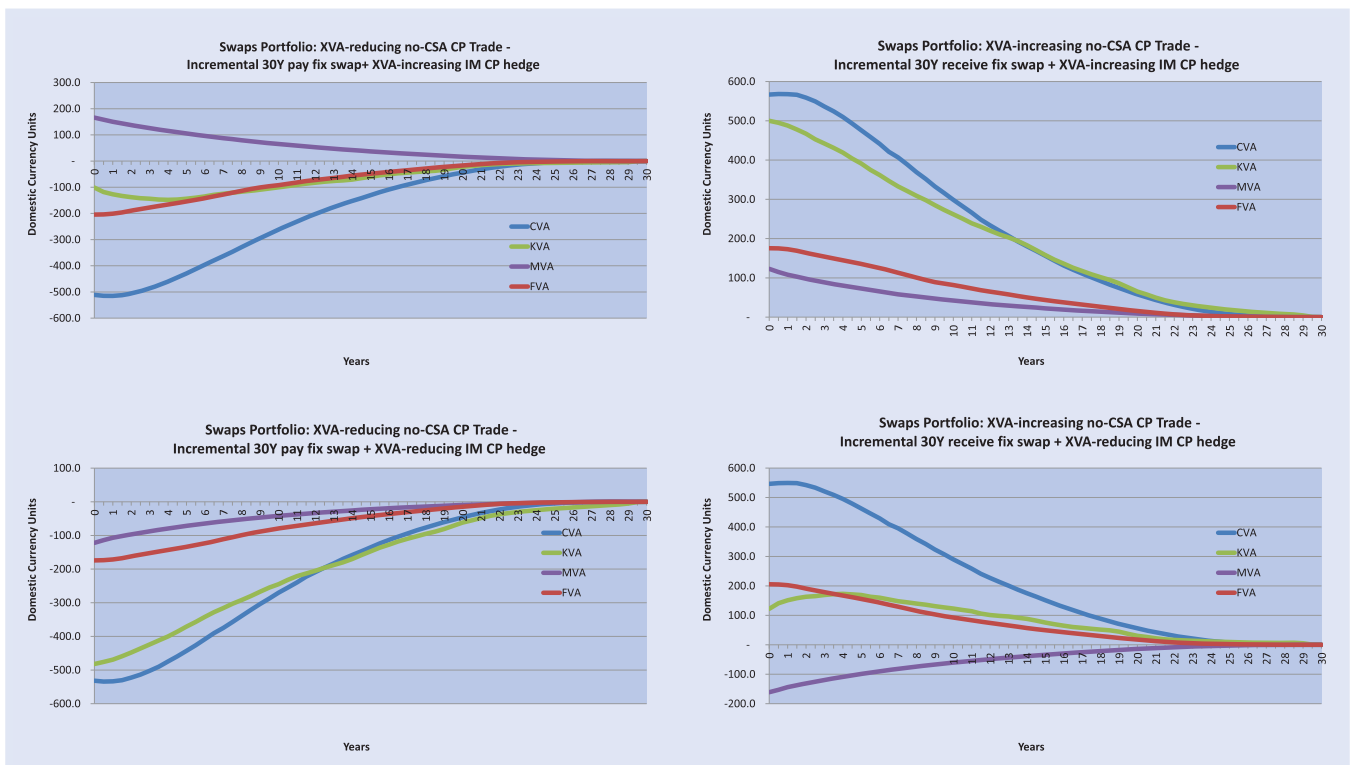


Figure 15. (Top left) XVA-reducing trade + XVA-increasing bilateral hedge (Top right) XVA-increasing trade + XVA-increasing bilateral hedge. (Bottom left) XVA-reducing trade + xva-reducing bilateral hedge (Bottom right) XVA-increasing trade + XVA-reducing bilateral hedge.

The 8 no CSA counterparties are primarily asset or liability heavy. One bilateral CSA hedge counterparty is asset-heavy and one liability-heavy. Figure 15 provides the trade

incremental XVA profiles of the bilateral hedge alternatives in combination with those for the initial counterparty trade. The main XVA impact of the hedge is then a corresponding

Table 5. XVA deep learning computation timings (s).

	10 CP 40 risk factors		20 CP 80 risk factors	
	No CSA	IM CSA	No CSA	IM CSA
Initial risk factor & trade pricing simulation Cuda	352	352	426	426
Counterparty and bank level learning calculations	4529	13 466	19 154	59 342
Total initial batch	4881	13 818	19 580	59 768
Re-simulate 1 counterparty trade pricing Cuda	40	40	51	51
Counterparty and bank level learning calculations	2785	2736	7694	6628
Total incremental trade	2825	2776	7745	6679

incremental MVA term, which can contribute to make the global FTP related to the trade + hedge package more or less positive or negative, depending on the data (cf. the four panels in figure 15), as can only be inferred by a refined XVA computation.

REMARK 5.1 In the above, we do not include the XVA costs/benefits of the bilateral hedge counterparty itself. Given Remark 2.4, in different circumstances it may be possible to attribute them to client trades of the original or hedge bank. Space is lacking for a fuller discussion of economics of XVA trading in different setups. In particular, many hedge trades now face central instead of bilateral counterparties. This occurs at additional XVA costs for the client of the initial swap that can be computed the way explained in Albanese *et al.* (2020).

5.5. Scalability

Our deep learning XVA implementation uses CNTK, the Microsoft Cognitive Toolkit. CNTK is written in core C++/CUDA (with wrappers for Python, C#, and Java). This is convenient for XVA applications, which are usually developed in C++: CNTK automatic differentiation in C++/CUDA enables C++ in-process training. This allows embedding the deep learning task within XVA processing.

Table 5 sets out computation times, including additional results obtained by doubling the numbers of counterparties and risk factors (to 20 counterparties and 80 risk factors).

All these results were based on 50K simulation paths, 32 time steps per year for risk factor simulation, and 16 time steps per year for all XVA calculations and deep learning. They were computed on a Lenovo P52 laptop with NVidia Quadro P3200 GPU @ 5.5 Teraflops peak FP32 performance, and 14 streaming multiprocessors.

The computations for 20 counterparties took more than twice as long as those for 10 counterparties. However, our deep learning calculations achieved around 80–90% Cuda occupancy for 10 counterparties and at times fell to half that level for 20 counterparties. Scaling to realistically high dimensions should be achievable, but acceptable trade incremental pricing performance in production would require server-grade GPU hardware, performance tuning for high GPU utilisation, and, possibly, caching computations.

Acknowledgments

The authors are grateful for useful discussions with Lokman Abbas-Turki, Agostino Capponi, Karl-Theodor Eisele, Chris Kenyon, Marek Rutkowski, and Michael Schmutz.

Disclosure statement

The views expressed herein by Rodney Hoskinson are his personal views and do not reflect the views of ANZ Banking Group Limited ('ANZ'). No liability shall be accepted by ANZ whatsoever for any direct or consequential loss from any use of this paper and the information, opinions and materials contained herein.

Funding

The PhD thesis of Bouazza Saadeddine is funded by a CIFRE grant from CA-CIB and French ANRT.

References

- Abbas-Turki, L., Diallo, B. and Crépey, S., XVA principles, nested Monte Carlo strategies, and GPU optimizations. *Int. J. Theor. Appl. Finance*, 2018, **21**, 1850030.
- Albanese, C. and Crépey, S., The cost-of-capital XVA approach in continuous time. Working Paper, 2020. Available online at: <https://math.maths.univ-evry.fr/crepey>.
- Albanese, C., Caenazzo, S. and Crépey, S., Credit, funding, margin, and capital valuation adjustments for bilateral portfolios. *Probab. Uncertain. Quant. Risk*, 2017, **2**(7), 26 pages.
- Albanese, C., Armenti, Y. and Crépey, S., XVA Metrics for CCP optimisation. *Stat. Risk Model.*, 2020, **37**(1-2), 25–53.
- Andersen, L., Duffie, D. and Song, Y., Funding value adjustments. *J. Finance*, 2019, **74**(1), 145–192.
- Artzner, P., Eisele, K.-T. and Schmidt, T., No arbitrage in insurance and the QP-rule. Working Paper, 2020. Available at: [arXiv:2005.11022](https://arxiv.org/abs/2005.11022).
- Barrera, D., Crépey, S., Diallo, B., Fort, G., Gobet, E. and Staszynski, U., Stochastic approximation schemes for economic capital and risk margin computations. *ESAIM: Proc. Surv.*, 2019, **65**, 182–218.
- Beck, C., Becker, S., Cheridito, P., Jentzen, A. and Neufeld, A., Deep splitting method for parabolic PDEs. 2019. [arXiv:1907.03452](https://arxiv.org/abs/1907.03452).
- Bichuch, M., Capponi, A. and Sturm, S., Arbitrage-free XVA. *Math. Finance*, 2018, **28**(2), 582–620.

- Bielecki, T. and Rutkowski, M., *Credit Risk: Modeling, Valuation and Hedging*, 2002 (Springer Finance: Berlin).
- Bielecki, T.R. and Rutkowski, M., Valuation and hedging of contracts with funding costs and collateralization. *SIAM J. Financ. Math.*, 2015, **6**, 594–655.
- Brigo, D. and Capponi, A., Bilateral counterparty risk with application to CDSs. *Risk Magazine*, March 2010, 85–90. Preprint version available at: <https://arxiv.org/abs/0812.3705>.
- Brigo, D. and Pallavicini, A., Nonlinear consistent valuation of CCP cleared or CSA bilateral trades with initial margins under credit, funding and wrong-way risks. *J. Financ. Eng.*, 2014, **1**, 1–60.
- Burgard, C. and Kjaer, M., In the balance. *Risk Magazine*, October 2011, 72–75.
- Burgard, C. and Kjaer, M., Funding costs, funding strategies. *Risk Magazine*, December 2013, 82–87. Preprint version available at: <https://ssrn.com/abstract=2027195>.
- Burgard, C. and Kjaer, M., Derivatives funding, netting and accounting. *Risk Magazine*, March 2017, 100–104. Preprint version available at: <https://ssrn.com/abstract=2534011>.
- Castagna, A., Towards a theory of internal valuation and transfer pricing of products in a bank: Funding, credit risk and economic capital. 2014. Available at: <http://ssrn.com/abstract=2392772>.
- Collin-Dufresne, P., Goldstein, R. and Hugonnier, J., A general formula for valuing defaultable securities. *Econometrica*, 2004, **72**(5), 1377–1407.
- Committee of European Insurance and Occupational Pensions Supervisors, QIS5 technical specifications. 2010. Available online at: https://eiopa.europa.eu/Publications/QIS/QIS5-technical_specifications_20100706.pdf.
- Crépey, S., Bilateral counterparty risk under funding constraints. Part I: Pricing, followed by Part II: CVA. *Math. Finance*, 2015, **25**(1), 1–22 and 23–50. First published online on 12 December 2012.
- Crépey, S. and Song, S., Counterparty risk and funding: Immersion and beyond. *Finance Stoch.*, 2016, **20**(4), 901–930.
- Crépey, S. and Song, S., Invariance times. *Ann. Probab.*, 2017, **45**(6B), 4632–4674.
- Crépey, S., Bielecki, T.R. and Brigo, D., *Counterparty Risk and Funding: A Tale of Two Puzzles*, 2014 (Taylor & Francis: New York).
- Crépey, S., Sabbagh, W. and Song, S., When capital is a funding source: The anticipated backward stochastic differential equations of X-Value Adjustments. *SIAM J. Financ. Math.*, 2020, **11**(1), 99–130.
- Dimitriadis, T. and Bayer, S., A joint quantile and expected shortfall regression framework. *Electron. J. Stat.*, 2019, **13**(1), 1823–1871.
- Duffie, D. and Huang, M., Swap rates and credit quality. *J. Finance*, 1996, **51**, 921–950.
- Duffie, D. and Sharer, W., Equilibrium and the role of the firm in incomplete market. Working Paper No. 915, Stanford University, 1986. Available online at: <https://www.gsb.stanford.edu/faculty-research/working-papers/equilibrium-role-firm-incomplete-markets>.
- Dybvig, P., Hedging non-traded wealth: When is there separation of hedging and investment. In *Options: Recent Advances in Theory and Practice*, edited by S. Hodges, Vol. 2, pp. 13–24, 1992 (Manchester University Press: Manchester).
- Elouerkhaoui, Y., Pricing and hedging in a dynamic credit model. *Int. J. Theor. Appl. Finance*, 2007, **10**(4), 703–731.
- Elouerkhaoui, Y., *Credit Correlation: Theory and Practice*, 2017 (Palgrave Macmillan: London).
- Fissler, T. and Ziegel, J., Higher order elicibility and Osband’s principle. *Ann. Statist.*, 2016, **44**(4), 1680–1707.
- Fissler, T., Ziegel, J. and Gneiting, T., Expected shortfall is jointly elicitable with value at risk—Implications for backtesting. *Risk Magazine*, January, 2016.
- Föllmer, H. and Schied, A., *Stochastic Finance: An Introduction in Discrete Time* (4th ed.), 2016 (De Gruyter Graduate: Berlin).
- Goodfellow, I., Bengio, Y. and Courville, A., *Deep Learning*, 2016 (MIT Press: Cambridge).
- Gottardi, P., An analysis of the conditions for the validity of Modigliani-Miller theorem with incomplete markets. *Econ. Theory*, 1995, **5**, 191–207.
- Green, A., Kenyon, C. and Dennis, C., KVA: Capital valuation adjustment by replication. *Risk Magazine*, December 2014, 82–87. Preprint version “KVA: capital valuation adjustment”. Available at: ssrn.2400324.
- Hoeting, J.A., Madigan, D., Raftery, A.E. and Volinsky, C.T., Bayesian model averaging: A tutorial. *Stat. Sci.*, 1999, **14**(4), 382–417.
- Hull, J. and White, A., The FVA debate, followed by The FVA debate continued. *Risk Magazine*, July 2012, 83–85, and October 2012, 52.
- Huré, C., Pham, H. and Warin, C., Some machine learning schemes for high-dimensional nonlinear PDEs. *Math. Comput.*, 2020, **89**(324), 1547–1580.
- International Financial Reporting Standards, IFRS 4 insurance contracts exposure draft, 2013.
- Kjaer, M., In the balance redux. *Risk Magazine*, November, 2019.
- Longstaff, F.A. and Schwartz, E.S., Valuing American options by simulation: A simple least-squares approach. *Rev. Financ. Stud.*, 2001, **14**(1), 113–147.
- Merton, R., On the pricing of corporate debt: The risk structure of interest rates. *J. Finance*, 1974, **29**, 449–470.
- Modigliani, F. and Miller, M., The cost of capital, corporation finance and the theory of investment. *Econ. Rev.*, 1958, **48**, 261–297.
- Myers, S., Determinants of corporate borrowing. *J. Financ. Econ.*, 1977, **5**, 147–175.
- Piterbarg, V., Funding beyond discounting: Collateral agreements and derivatives pricing. *Risk Magazine*, August 2010, 57–63.
- Schönbucher, P., A measure of survival. *Risk Magazine*, 2004, **17**(8), 79–85.
- Swiss Federal Office of Private Insurance, Technical document on the Swiss solvency test. 2006. Available online at: https://www.finma.ch/FinmaArchiv/bpv/download/e/SST_techDok_061002_E_wo_Li_20070118.pdf.

A. Appendix. Continuous-time XVA equations

We recall from Crépey *et al.* (2020) the continuous-time XVA equations for bilateral trade portfolios when capital at risk is deemed fungible with variation margin, also adding here initial margin and MVA as in the refined static setup of section 3.4.

We write $\delta_\eta(dt) = d\mathbb{1}_{\{\eta \leq t\}}$ for the Dirac measure at a random time η .

A.1. Cash flows

We suppose that the derivative portfolio of the bank is partitioned into bilateral netting sets of contracts which are jointly collateralized and liquidated upon bank or counterparties (whether these are clients or market hedge counterparties) default. Given a netting set c of the bank portfolio, we denote by:

- \mathcal{P}^c and P^c , the corresponding contractually promised cash flows and clean value processes;
- τ_c , J^c , and R_c , the corresponding default times, survival indicators, and recovery rates, whereas τ , J , and R are the analogous data regarding the bank itself, with bank credit spread process $\lambda = (1 - R)\gamma$ taken as a proxy of its risky funding spread process;[†]
- $\tau_c^\delta = \tau_c + \delta$ and $\tau^\delta = \tau + \delta$, where δ is a positive margin period of risk, in the sense that the liquidation of the netting set c happens at time $\tau_c^\delta \wedge \tau^\delta$;

[†] See Albanese *et al.* (2020, Section 5) for the discussion of cheaper funding schemes for initial margin.

- VM^c , the variation margin (re-hypothecable collateral) exchanged between the bank and counterparty c , counted positively when received by the bank;
- PIM^c and RIM^c , the related initial margin (segregated collateral) posted and received by the bank;
- RC and CR , the reserve capital and capital at risk of the bank.

The contractually promised cash flows are supposed to be hedged out by the bank but one conservatively assumes no XVA hedge, so that the bank is left with the following trading cash flows \mathcal{C} and \mathcal{F} (cf. (38) and see Albanese and Crépey (2020, Lemmas 5.1 and 5.2) for detailed derivations of analogous equations in a slightly simplified setup):

- The (counterparty) credit cash flows

$$\begin{aligned} dC_t = & \sum_{c; \tau_c \leq \tau^\delta} (1 - R_c) \left((P^c + \mathcal{P}^c)_{\tau_c^\delta \wedge \tau^\delta} - (P^c + VM^c + RIM^c)_{(\tau_c \wedge \tau)^-} \right)^+ \delta_{\tau_c^\delta \wedge \tau^\delta}(dt) \\ & - (1 - R) \sum_{c; \tau \leq \tau_c^\delta} \left((P^c + \mathcal{P}^c)_{\tau^\delta \wedge \tau_c^\delta} - (P^c + VM^c + PIM^c)_{(\tau \wedge \tau_c)^-} \right)^- \delta_{\tau^\delta \wedge \tau_c^\delta}(dt); \end{aligned} \quad (A1)$$

- The (risky) funding cash flows

$$\begin{aligned} d\mathcal{F}_t = & J_t \lambda_t \left(\sum_c J^c (P^c - VM^c) - RC - CR \right)_t^+ dt \\ & - (1 - R) \left(\sum_c J^c (P^c - VM^c) - RC - CR \right)_{\tau^-}^+ \times \delta_\tau(dt) \\ & + J_t \lambda_t \sum_c J_t^c PIM_t^c dt - (1 - R) \times \sum_c J_{\tau^-}^c PIM_{\tau^-}^c \delta_\tau(dt), \end{aligned} \quad (A2)$$

where the RC and CR terms account for the fungibility of reserve capital and capital at risk with variation margin.

A.2. Valuation

Here (as in our numerics) we distinguish between a (strict) FVA, in the strict sense of the cost of raising variation margin, and an MVA for the cost of raising initial margin (see Remark 2.1). The (other than K)VA equations are then

$$RC = CA = CVA + FVA + MVA, \quad (A3)$$

the so-called ‘contra-assets valuation’ sourced from the clients and deposited in the reserve capital account of the bank, where, for $t < \tau$,

$$\begin{aligned} CVA_t = & \mathbb{E}_t \sum_{t < \tau_c^\delta} (1 - R_c) \left((P^c + \mathcal{P}^c)_{\tau_c^\delta} - (P^c + VM^c + RIM^c)_{\tau_c^-} \right)^+ \\ FVA_t = & \mathbb{E}_t \int_t^T \lambda_s \left(\sum_c J^c (P^c - VM^c) - CA - CR \right)_s^+ ds \\ MVA_t = & \mathbb{E}_t \int_t^T \lambda_s \sum_c J_s^c PIM_s^c ds. \end{aligned} \quad (A4)$$

The corresponding trading loss and profit process L of the bank is such that

$$\begin{aligned} L_0 = & 0 \quad \text{and, for } t < \tau, \\ dL_t = & \sum_c (1 - R_c) \\ & \times \left((P^c + \mathcal{P}^c)_{\tau_c^\delta} - (P^c + VM^c + RIM^c)_{\tau_c^-} \right)^+ \delta_{\tau_c^\delta}(dt) \\ & + \lambda_t \left(\sum_c J^c (P^c - VM^c) - CA - CR \right)_t^+ dt \\ & + \lambda_t \sum_c J_t^c PIM_t^c dt \\ & + dCA_t, \end{aligned} \quad (A5)$$

so that L is a \mathbb{Q} martingale, hence (by Lemma 4.1) L° is a \mathbb{Q}^* martingale.

By the same rationale as Definitions 3.2 and 3.3 in the static setup:

DEFINITION A.1 EC_t is the time- t conditional 97.5% expected shortfall of $(L_{t+1}^\circ - L_t^\circ)$ under \mathbb{Q} .

Given a positive target hurdle rate h :

DEFINITION A.2 We set

$$CR = \max(EC, KVA), \quad (A6)$$

for a KVA process such that, for $t < \tau$,

$$KVA_t = \mathbb{E}_t \left[\int_t^T h(CR_s - KVA_s) ds \right]. \quad (A7)$$

Hence, for $t < \tau$,

$$\begin{aligned} KVA_t = & \mathbb{E}_t \left[\int_t^T h e^{-h(s-t)} CR_s ds \right] \\ = & \mathbb{E}_t \left[\int_t^T h e^{-h(s-t)} \max(EC_s, KVA_s) ds \right]. \end{aligned} \quad (A8)$$

The next-to-last identity is the continuous-time analog of the risk margin formula under the Swiss solvency test cost of capital methodology: see Swiss Federal Office of Private Insurance (2006, Section 6, middle of page 86 and top of page 88).

A.3. The XVA equations are well-posed

In view of (A3), the second line in (A4) is in fact an FVA equation. Likewise, the second line in (A8) is a KVA equation. Moreover, as capital at risk is fungible with variation margin (cf. section 3.4), i.e. in consideration of the CR term in (A4)–(A5), where $CR = \max(EC, KVA)$, we actually deal with an (FVA, KVA) system, and even, as EC depends on L (cf. Definition A.1), with a forward backward system for the forward loss process L and the backward pair (FVA, KVA).

However, as in the refined static setup of section 3.4, the coupling between (FVA, KVA) and L can be disentangled by the following Picard iteration:

- Let CVA and MVA be as in (A4), $L^{(0)} = KVA^{(0)} = 0$, and, for $t < \tau$,

$$FVA_t^{(0)} = \mathbb{E}_t \int_t^T \lambda_s \left(\sum_c J^c (P^c - VM^c) - CA^{(0)} \right)_s^+ ds, \quad (A9)$$

where $CA^{(0)} = CVA + FVA^{(0)} + MVA$;

- For $k \geq 1$, writing explicitly $EC = EC(L)$ to emphasize the dependence of EC on L , let $L_0^{(k)} = 0$ and, for $t < \tau$,

$$\begin{aligned}
dL_t^{(k)} &= \sum_c (1 - R_c) \\
&\times \left((P^c + \mathcal{P}^c)_{\tau_c^\delta} - (\mathcal{P}^c + VM^c + RIM^c)_{\tau_c^-} \right)^+ \\
&\times \delta_{\tau_c^\delta}(dt) \\
&+ \lambda_t \left(\sum_c J^c (P^c - VM^c) - CA^{(k-1)} \right. \\
&\quad \left. - \max(EC(L^{(k-1)}), KVA^{(k-1)}) \right)_t^+ dt \\
&+ \lambda_t \sum_c J_t^c PIM_t^c dt + dCA_t^{(k-1)}, \\
KVA_t^{(k)} &= h \mathbb{E}_t \int_t^T e^{-h(s-t)} \\
&\quad \times \max(EC_s(L^{(k)}), KVA_s^{(k)}) ds, \\
CA_t^{(k)} &= CVA_t + FVA_t^{(k)} + MVA_t \text{ where } FVA_t^{(k)} \\
&= \mathbb{E}_t \int_t^T \lambda_s \left(\sum_c J^c (P^c - VM^c) - CA^{(k)} \right. \\
&\quad \left. - \max(EC(L^{(k)}), KVA^{(k)}) \right)_s^+ ds. \tag{A10}
\end{aligned}$$

Theorem 4.1 in Crépey *et al.* (2020) Assuming square integrable data, the XVA equations are well-posed within square integrable solution (including when one accounts for the fact that capital at risk can be used for funding variation margin). Moreover, the above Picard iteration converges to the unique square integrable solution of the XVA equations.

A.4. Collateralization schemes

We denote by $\Delta_t^c = \mathcal{P}_t^c - \mathcal{P}_{(t-\delta)-}^c$ the cumulative contractual cash flows with the counterparty c accumulated over a past period of

length δ . In our case study, we consider both ‘no CSA’ netting sets c , with $VM = RIM = PIM = 0$, and ‘(VM/IM) CSA’ netting sets c , with $VM_t^c = P_t^c$ and, for $t \leq \tau_c$,

$$\begin{aligned}
RIM_t^c &= \mathbb{V}a\mathbb{R}_t(P_{t^\delta}^c + \Delta_{t^\delta}^c - P_t^c), \\
PIM_t^c &= \mathbb{V}a\mathbb{R}_t(-(P_{t^\delta}^c + \Delta_{t^\delta}^c) + P_t^c), \tag{A11}
\end{aligned}$$

for some PIM and RIM quantile levels a_{pim} and a_{rim} (and $t^\delta = t + \delta$).

The following result can be derived by similar computations as the ones in Albanese *et al.* (2020, Section A).

PROPOSITION A.1 *In a common shock default model of the counterparties and the bank itself (see the beginning of section 5), with pre-default intensity processes γ^c of the counterparties and γ of the bank, then $CVA = CVA^{nocsa} + CVA^{csa}$, where, for $t < \tau$,*

$$\begin{aligned}
CVA_t^{nocsa} &= \sum_{c \text{ nocsa}} \mathbb{1}_{t < \tau_c} (1 - R_c) \mathbb{E}_t \\
&\int_t^T (P_{s^\delta}^c + \Delta_{s^\delta}^c)^+ \gamma_s^c e^{-\int_t^s \gamma_u^c du} ds \\
&+ \sum_{c \text{ nosca}} \mathbb{1}_{\tau_c < t < \tau_c^\delta} (1 - R_c) \mathbb{E}_t (P_{\tau_c^\delta}^c + \Delta_{\tau_c^\delta}^c)^+, \tag{A12}
\end{aligned}$$

$$\begin{aligned}
CVA_t^{csa} &= \sum_{c \text{ csa}} \mathbb{1}_{t < \tau_c} (1 - R_c) (1 - a_{rim}) \\
&\times \mathbb{E}_t \int_t^T (\mathbb{E}S_s - \mathbb{V}a\mathbb{R}_s) ((P_{s^\delta}^c + \Delta_{s^\delta}^c) - P_s^c) \gamma_s^c e^{-\int_t^s \gamma_u^c du} ds \\
&+ \sum_{c \text{ csa}} \mathbb{1}_{\tau_c < t < \tau_c^\delta} (1 - R_c) \mathbb{E}_t \\
&\times \left((P_{\tau_c^\delta}^c + \Delta_{\tau_c^\delta}^c) - (P_{\tau_c}^c + RIM_{\tau_c}^c) \right)^+, \tag{A13}
\end{aligned}$$

where $(\mathbb{E}S_s - \mathbb{V}a\mathbb{R}_s)$ in (A13) is computed at the a_{rim} confidence level. Assuming its posted initial margin borrowed unsecured by the bank, then $MVA = MVA^{csa}$, where, for $t < \tau$,

$$MVA_t^{csa} = \sum_{c \text{ csa}} J_t^c \mathbb{E}_t \int_t^T (1 - R) \gamma_s PIM_s^c e^{-\int_t^s \gamma_u^c du} ds. \tag{A14}$$