Model-free bounds on bilateral counterparty valuation

Haase, Joerdis; Ilg, Melanie and Werner, Ralf
Finbridge AG, Deutsche Pfandbriefbank AG, Hochschule München

02. September 2010

Model-Free Bounds on Bilateral Counterparty Valuation Adjustment

Jördis Haase[*] Melanie Ilg[†] Ralf Werner[‡]

September 2, 2010

First version: December 7, 2009

Abstract

In the last years, counterparty default risk has experienced an increased interest both by academics as well as practitioners. This was especially motivated by the market turbulences and the financial crises over the past years which have highlighted the importance of counterparty default risk for uncollateralized derivatives. The following paper focuses on the pricing of derivatives subject to such counterparty risk. After a succinct introduction to the topic, a brief review of state-of-the-art methods for the calculation of bilateral counterparty value adjustments is presented. Due to some weaknesses of these models, a novel method for the determination of model-free tight lower and upper bounds on these adjustments is presented. It will be shown in detail how these bounds can be easily and efficiently calculated by the solution of a corresponding linear optimization problem. It will be illustrated how usual discretization methods like Monte Carlo methods allow to reduce the calculation of bounds to an ordinary finite dimensional transportation problem, whereas a continuous time approach will lead to a general mass transportation problem. The paper is closed with several applications of these model-free bounds, like stress-testing and estimation of model reserves.

1 Introduction

Recent events such as Lehman's default have drawn the attention to counterparty default risk. At the very latest after this default, it has become obvious to all market participants that the credit quality of both counterparties – usually a client and an investment bank – need to be considered in the pricing of uncollateralized OTC derivatives. Before, either both the client and the counterparty have been assumed to be risk-free, therefore completely neglecting default risk. More sophisticated models already considered unilateral counterparty risk, i.e. it has been assumed that the client is subject to default risk.

For the latter case, over the past years, several authors have been investigating most types of derivatives together with a variety of suitable models. For instance, the earliest[1] article referring to unilateral counterparty value adjustments of interest rate swaps is probably due to Cooper and Mello [12]. The bilateral case seemed to be first covered by Sorensen and Bollier [27] in a very simple setting and Duffie and Huang [14] who already considered wrong-way risk in a simplified fashion. More recently, state-of-the-art models have for example been applied by Brigo and Masetti [7] or Brigo and Pallavicini [8]. Additionally, other types of derivatives have been considered as well, cf. Brigo and Bakkar [4] who focused on counterparty risk for commodity

[*]Finbridge GmbH & Co. KG, Usinger Strasse 118a, 61440 Oberursel, joerdis.haase@finbridge.de

[†]Deutsche Pfandbriefbank AG, Freisinger Strasse 5, 85716 Unterschleissheim, melanie.ilg@pfandbriefbank.com

[‡]Hochschule München, Fakultät für Informatik und Mathematik, Lothstr. 64, 80335 München, werner@cs.hm.edu

[1]In Cooper and Mello [12] a few unpublished sources dating back to 1987 are cited.

derivatives. Further, counterparty risk for credit default swaps[2] was tackled by several authors, for example Brigo and Chourdakis [6], Leung and Kwok [21], Hull and White [19], Walker [30], Crépey et al. [13], Blanchet-Scalliet and Patras [3] or Lipton and Sepp [22] just to name a few. While the works of Brigo et al. ([4], [6] and [8]) concentrate on dynamic stochastic models and Leung and Kwok [21] model the default intensities as deterministic constants, Lipton and Sepp [22] introduce a structural model with jumps and Crépey et al. [13] use a Markov chain copula model in order to derive counterparty value adjustments. Finally, Hull and White [19] use barrier correlated models, whereas in Walter [30] contagion is considered by using transition rates.

In contrast to unilateral adjustments, bilateral counterparty valuation adjustments have been considered in more detail only recently by Brigo and Capponi [5] who use a stochastic intensity model and a trivariate copula function for pricing credit default swaps subject to default risk of both counterparties. For interest rate products Brigo et al. [9] generalized the works concerning unilateral adjustments (cf. Sorensen and Bollier [27] or Brigo and Masetti [7]) to the bilateral case where they also use dynamic stochastic models with correlation between the default times and the underlying risk factors, see for example Section 3.4.

In the following exposition, a conceptually completely different approach is proposed. In contrast to most existing work on bilateral counterparty value adjustments, the proposed model-free approach does not rely on any specific model for the joint evolution of the underlying risk factors. As only exception we note the paper by Cherubini [11] which provided the basis for this model-free approach: Following and generalizing Cherubini [11], the counterparty valuation adjustment will be decomposed into three main components: the first component is represented by the loss process which is usually assumed to be a constant (unless random recoveries are modelled); the second component consists of the default indicators of the two counterparties and the third component is comprised of the exposure-at-default of the OTC derivative, i.e. the risk-free present value of the outstanding amount in case of default. In such a manner, the proposed approach is able to cover all kind of derivatives[3] (interest rate, commodity, or credit default swap) in a unified way.

It will be argued that especially the marginal models for the latter two components are already (almost) fully specified by available market prices of associated financial instruments and that only the interdependence of the three components needs to be specified for a particular calculation of counterparty adjustments. It will be illustrated that the proposed model-free approach contains the cases of independent components as well as any coupling derived from dynamic stochastic models as special cases.

In addition to the existing literature on counterparty adjustments, this exposition makes the following main contributions to the topic of bilateral counterparty value adjustments: first, the three main building blocks of such an adjustment are clearly identified and separated, and it is shown how any coupling of these blocks leads to a feasible adjustment, and second, that tight bounds on the adjustments can be efficiently obtained by the solution of linear optimization problems. Third, several applications of these tight bounds are presented. Especially, in contrast to the approaches of Turnbull [29] or Cherubini [11], both the upper and lower bound derived here are indeed *tight* bounds, i.e. there exists a model which is consistent with all market prices and where these bounds are attained. Further, these bounds are not restricted to any specific financial instrument, but are valid for a large variety of instruments. Additionally, unlike Cherubini [11], where only a very specific setup[4] was considered, the following setup is completely general and covers all kind of derivatives, together with arbitrary multi-factor models. Further, in contrast to Cherubini, the order of defaults is adequately considered to obtain the correct counterparty value adjustment. Finally, by generalizing Cherubini's approach, provable tight upper and lower

[2] Counterparty credit risk for CDS represents the most complicated case as joint defaults of both counterparties and the reference entity have to be captured. Further, at least stochastic credit spreads need to be considered to cover the fluctuation in the market value of CDS contracts.

[3] This approach covers all kind of financial derivatives where the payoff, and thus the present value of the derivative, is not explicitly depending on the credit quality of any of the two counterparties.

[4] Cherubini [11] only considered the very specific case of an interest rate swap. Further, only one particular two-dimensional copula was used to couple each individual forward swap par rate with the default time. Obviously, a more general approach would couple all forward swap par rates with the default time. From there it is easy to see that the most general approach links all potential risk factors with both default times – which is the basic underlying idea of the presented model-free approach.

bounds on counterparty value adjustments are obtained.

The rest of the paper is organized as follows. In the first part of Section 2 a succinct introduction to the topic of bilateral counterparty risk is given, before the decomposition of the adjustment into its building blocks is carried out. For practical implementations, discretization methods are briefly mentioned. In Section 3 the two main approaches for the calculation of counterparty valuation adjustments are briefly introduced and discussed, with a special focus on the default modelling. Section 4 deals with the model-free approach: a general framework similar to mass transportation problems is presented, before the problem of the calculation of the adjustments is embedded into this framework. It is explained in detail, how the corresponding transportation problem is obtained for discrete models. An in-depth example is given in Section 5 which especially highlights the large gaps between upper and lower bound on the adjustments for the specific example of an ordinary interest rate swap, before Section 6 concludes the paper.

2 Counterparty default risk

As usual, to model financial transactions with default risk, let $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbf{Q})$ be a probability space where $\mathcal{G}_t$ models the flow of information and $\mathbf{Q}$ denotes the risk-neutral measure, see e.g. Bielecki and Rutkowski [2] for more details. Further, let the space be endowed with a right-continuous and complete sub-filtration $\mathcal{F}_t$ modelling the flow of information except default, such that $\mathcal{F}_t \subseteq \mathcal{G}_t := \mathcal{F}_t \vee \mathcal{H}_t$ with $\mathcal{H}_t$ being the right-continuous filtration generated by the default events.

Subsequently, we consider a transaction with maturity T between a client A and a counterparty B where both are subject to default. The respective random default times are denoted by τ_A and τ_B. In order to take into account counterparty default risk we distinguish three cases,

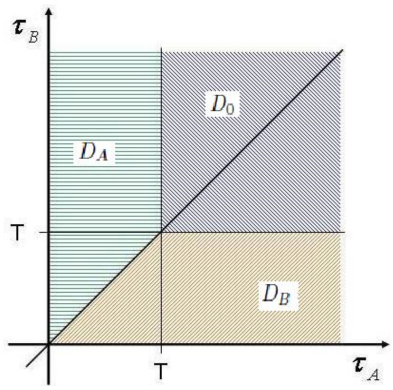

Figure 1: Decomposition of potential default events.

illustrated in Figure 1:

- neither A nor B defaults before T: $D_0 := \{\tau_A > T\} \cap \{\tau_B > T\} = \{T < \min(\tau_A, \tau_B)\}$,
- A defaults before B and before T: $D_A := \{\tau_A \leq T\} \cap \{\tau_A \leq \tau_B\} = \{\tau_A \leq \min(T, \tau_B)\}$,
- B defaults before A and before T: $D_B := \{\tau_B \leq T\} \cap \{\tau_B \leq \tau_A\} = \{\tau_B \leq \min(T, \tau_A)\}$.

In case that $\mathbf{Q}[\tau_A = T] = \mathbf{Q}[\tau_B = T] = \mathbf{Q}[\tau_A = \tau_B] = 0$, these sets yield a decomposition of one, i.e. it holds

$$\mathbf{1}_{D_0} + \mathbf{1}_{D_A} + \mathbf{1}_{D_B} = 1 \quad \mathbf{Q}\text{-almost-surely.}$$

Please note that Brigo et al. (in [5] and [9]) use different sets to order the default times, which are in essence reducible to the above three sets.

2.1 Valuation without counterparty default risk

Neglecting the possibility of default for the moment, i.e. assuming $\mathbf{Q}[D_0] = 1$, the value of the transaction for client A is the difference of the discounted cash flows from the client A to the counterparty B and vice versa:

$$\begin{aligned} V_A(t,T) &= \mathbb{E}\left[\sum_{i=1}^{m_B} P(t,T_i)\cdot C(B,A,T_i)\middle|\mathcal{F}_t\right] - \mathbb{E}\left[\sum_{j=1}^{m_A} P(t,T_j)\cdot C(A,B,T_j)\middle|\mathcal{F}_t\right] \\ &= -V_B(t,T) \end{aligned}$$

In the above equation $V_A(t,T)$ denotes the value of the transaction as seen by client A at time t. The investment consists of cash flows $C(B,A,T_i)$ paid by the counterparty B at times T_i, $i = 1\dots m_B$, and cash flows $C(A,B,T_j)$ paid by the client A at times T_j, $j = 1\dots m_A$, e.g. an interest rate swap. The risk-free discount factor at time t for time s is denoted as $P(t,s)$.

2.2 Valuation under bilateral counterparty risk

Let us point out that the above relationship is only true under the assumption $\mathbf{Q}[D_0] = 1$, which for example holds if both client and counterparty are supposed to be (default) risk-free. By allowing for default risk, we have to incorporate the remaining two cases into the valuation of the transaction, where A defaults before the investment's maturity, i.e. D_A, and, respectively, the case of the counterparty B's default, i.e. D_B. These cases then represent what is called *counterparty default risk* or *counterparty risk* in short. If only one party A or B is assumed to be default free then we talk about *unilateral counterparty risk*, if both parties are assumed to be subject to default risk, we are concerned with *bilateral counterparty risk*. For more details on counterparty risk in general and especially on the pros and cons of unilateral vs. bilateral counterparty risk let us refer to the extensive books by Pykhtin [23] and Gregory [17], as well as the two articles by Gregory [16] and [18].

Let us consider the case that counterparty B defaults and that she is not able to fulfill her obligations. Therefore, if the value of the transaction at default time τ_B is in favour of the client A, i.e. $V_A(\tau_B,T) > 0$, client A will suffer a loss, as only a fraction of the outstanding value is received, namely the recovery rate of the outstanding value. Otherwise, if the value at time τ_B is negative for client A, i.e. $V_A(\tau_B,T) < 0$, A has to pay the full outstanding value to the defaulted counterparty B (in accordance to the *full two-way payment* rule under ISDA Master contracts, see e.g. Bielecki and Rutkowski [2], Section 14.4.4). It is clear that only the first case requires an adjustment to the valuation of the transaction, while the second case does not impact valuation, as this can be considered as a buy-back at market price. Of course, this adjustment is only necessary in case B defaults before A and before the transaction has matured, i.e. only if $\mathbf{1}_{D_B} = 1$. Summarizing all these considerations, we obtain the following loss to the client A at time τ_B,

$$\mathbf{1}_{D_B} \cdot L^B_{\tau_B} \cdot \max(0, V_A(\tau_B,T)),$$

which leads to the first term in Equation (2.1) below. If only default of one party (in the above case party B) is considered, i.e. if only unilateral counterparty risk is taken into account, the above loss is the only loss which can occur. However, if also the client A is subject to default risk, i.e. if bilateral counterparty risk should be taken into account, then switching the roles of A and B in the above argumentation yields the corresponding loss to party B at time τ_A,

$$\mathbf{1}_{D_A} \cdot L^A_{\tau_A} \cdot \max(0, V_B(\tau_A,T)),$$

in case of default of the client A. Similarily to above, this loss is covered by the second term in Equation (2.1) below.

In the above paragraph, we have assumed that at time of default τ_B only a fraction of the outstanding market value $\max(0, V_A(\tau_B,T))$ could be recovered, i.e. the loss to party A is given as

$$L^B_{\tau_B} \cdot \max(0, V_A(\tau_B,T)),$$

where $L^B_t \in [0,1]$ represents the continuous random *loss process* at time t, which is usually assumed to be a constant $L^B_t = l^B$, but could also depend on time t as well as on the default time τ_B. The corresponding recovery process is then given by $R_t^B = 1 - L_t^B$. Although alternative models[5] for recovery payments (like recovery of face value) might be even more popular, see for example Duffie and Singleton [15] for a thorough review, or Bielecki and Rutkowsi [2], Sections 8.2.5 and 8.3.2, for derivative transactions subject to default risk, recovery of present value of a default free but otherwise identical transaction is the most reasonable assumption (which is also fixed in the ISDA master agreement for derivatives). Let us summarize the above considerations in the following well-known theorem, which in essence goes back to Sorensen and Bollier [27]:

Theorem 2.1. *Conditional on the event $\{t < \min(\tau_A, \tau_B)\}$, i.e. no default has occurred until time t, the value of the transaction under consideration of counterparty risk at time t is given by*

$$\begin{aligned} V_A^D(t,T) &= V_A(t,T) - CVA_A(t,T) \\ &= -\big(V_B(t,T) - CVA_B(t,T)\big) \\ &= -V_B^D(t,T) \end{aligned}$$

where the bilateral counterparty value adjustment $CVA_A(t,T)$ *is defined as*

$$\begin{aligned} CVA_A(t,T) &:= \mathbf{E}\left[\mathbf{1}_{D_B} \cdot P(t,\tau_B) \cdot L_{\tau_B}^B \cdot \max(0, V_A(\tau_B,T)) \,|\, \mathcal{G}_t\right] \\ &- \mathbf{E}\left[\mathbf{1}_{D_A} \cdot P(t,\tau_A) \cdot L_{\tau_A}^A \cdot \max(0, V_B(\tau_A,T)) \,|\, \mathcal{G}_t\right] \\ &= -CVA_B(t,T). \end{aligned} \tag{2.1}$$

Proof. A proof of Theorem 2.1 can for example be found in Bielecki and Rutkowski [2], Formula (14.25) or Brigo and Capponi [5], Proposition 2.1 and Appendix A, resp. □

We note from the preceding theorem that the bilateral counterparty valuation adjustment is indeed symmetric, i.e. it holds

$$CVA_A(t,T) = -CVA_B(t,T)$$

as opposed to the unilateral counterparty value adjustment. We further note that the bilateral counterparty valuation adjustment can actually become negative causing the overall value of the defaultable transaction being greater than the corresponding risk-free value. As already mentioned, a detailed discussion of the implications of considering bilateral counterparty risk instead of unilateral one can for example be found in Gregory [16] and [18].

2.3 Decomposition of bilateral CVA

Based on Theorem 2.1, the general approach for the calculation of the counterparty risk adjusted value $V_A^D(t,T)$ is to determine first the risk-free value $V_A(t,T)$ of the transaction. This can be done by any common valuation method for this kind of transaction. In a second step the counterparty value adjustment $CVA_A(t,T)$ needs to be determined. So far, two main approaches have emerged in the academic literature, which will be reviewed in more detail in the following section. Before we do so, let us consider Equation (2.1) in more detail, to obtain a more compact and more handy version of formula (2.1).

To be able to separate the default dynamics from the market value dynamics, let us introduce the auxiliary time s, $s \in [t,T]$ to rewrite Equation (2.1) as:

$$\begin{aligned} CVA_A(t,T) &= \mathbf{E}\left[L_s^B \cdot \mathbf{1}_{D_B} \cdot \mathbf{1}_s(\tau_B) \cdot P(t,s) \cdot \max(0, V_A(s,T)) \,|\, \mathcal{G}_t\right] \\ &- \mathbf{E}\left[L_s^A \cdot \mathbf{1}_{D_A} \cdot \mathbf{1}_s(\tau_A) \cdot P(t,s) \cdot \max(0, V_B(s,T)) \,|\, \mathcal{G}_t\right]. \end{aligned} \tag{2.2}$$

With a further bit of notation, if we introduce the following processes in time s, $s \in [t,T]$,

$$\begin{aligned} \delta_s^i &:= \mathbf{1}_{D_i} \cdot \mathbf{1}_s(\tau_i), \\ \widetilde{V}_i^+(t,s,T) &:= P(t,s) \cdot \max(0, V_i(s,T)), \end{aligned}$$

[5]Let us point out that the choice of the recovery assumptions needs to be carefully considered in the calibration of default models to market data to avoid technical inconsistencies in the subsequent calculation. For instance, default models are often calibrated to CDS prices (see for example Brigo and Capponi [5]), where a recovery of face value is assumed – which may question the assumption that the very same recovery rate can be applied in the CVA calculation. In more detail, Cooper and Mello [12] also considered different settlement rules and their impact on CVA.

where $i \in \{A, B\}$, then the bilateral counterparty value adjustment can be rewritten in a more compact manner (skipping the integral over s for notational convenience) as

$$CVA_A(t,T) \quad = \quad \mathbf{E}\left[L_s^B \cdot \delta_s^B \cdot \widetilde{V}_A^+(t,s,T) \,|\, \mathcal{G}_t \right] - \mathbf{E}\left[L_s^A \cdot \delta_s^A \cdot \widetilde{V}_B^+(t,s,T) \,|\, \mathcal{G}_t \right]. \tag{2.3}$$

From Equation (2.3) we immediately see that the bilateral counterparty value adjustment is composed of six processes:

- two *default indicator processes* δ_s^A and δ_s^B,
- two *loss processes* L_s^A and L_s^B, and
- two *discounted exposure processes* $\widetilde{V}_A^+(t,s,T)$ and $\widetilde{V}_B^+(t,s,T)$.

In this way, we are able to separate the default dynamics δ from the loss process L and the exposure process $\widetilde{V}$. We will see in the following that this separation is the key to the model-free calculation of the bilateral counterparty value adjustment. From this decomposition, it also becomes obvious that the bilateral counterparty value adjustment is completely determined by the joint distribution of these six processes. Therefore, any model for the joint evolution of these six processes fully determines the bilateral counterparty value adjustment.

2.4 Time and state space discretization

As we will see in Section 4, for computational purposes it is usually easier to work with discrete processes in discrete time. If the model is not as such, i.e. if we are given a continuous time model with continuous state space, both the time domain as well as the state space need to be properly discretized to obtain a discrete process in discrete time as a reasonably good approximation of the original process.

For the purpose of time discretization, we use the idea of *default bucketing*, which can for example be found in Brigo and Chourdakis [6], and which is applied in several approaches for CVA computation. In the default bucketing approach, the remaining time interval $(t, T]$ is decomposed into K disjoint time intervals $\Delta_k := (t_{k-1}, t_k]$, $k = 1, \ldots, K$, with $t_0 = t$ and $t_K = T$. This leads to the discrete time version of the above defined processes,

$$\begin{aligned} \delta_k^i &:= \delta_{t_k}^i := \mathbf{1}_{D_i} \cdot \mathbf{1}_{\Delta_k}(\tau_i), \\ L_k^i &:= L_{t_k}^i, \\ X_k^i &:= \tilde{V}_i^+(t, t_k, T), \end{aligned}$$

for $k = 1, \ldots, K$ and $i \in \{A, B\}$. It holds that the discrete time version is a sufficiently good approximation of the continuous time setting for large enough K, i.e. it holds

$$CVA_A(t,T) \quad = \quad \lim_{K\to\infty} \sum_{k=1}^{K} \left(\mathbf{E}\left[L_k^B \cdot \delta_k^B \cdot X_k^A \,|\, \mathcal{G}_t \right] - \mathbf{E}\left[L_k^A \cdot \delta_k^A \cdot X_k^B \,|\, \mathcal{G}_t \right] \right) \tag{2.4}$$

if the time intervals Δ_k become smaller and smaller, i.e. if $\operatorname{diam}(\Delta_k) \to 0$ for $K \to \infty$. Note that for this result it is necessary that both the loss processes L_s^A and L_s^B as well as the discounted exposure processes $\widetilde{V}_A^+(t,s,T)$ and $\widetilde{V}_B^+(t,s,T)$ need to have continuous paths, i.e. need to be continuous in time s.

Let us note that after time discretization, the default indicator process can only take a finite number of values. More exactly, it holds that the joint (i.e. two-dimensional) default indicator process $\delta = (\delta_k)_{k=1,\ldots,K} \in \mathbb{R}^{2\times K}$, defined by

$$\delta_k \quad := \quad \begin{pmatrix} \delta_k^A \\ \delta_k^B \end{pmatrix}, \quad k = 1, \ldots, K,$$

takes only values in the finite set

$$\mathcal{Y} := \{\gamma \in \mathbb{R}^{2\times K} \,|\, \gamma_{i,k} \in \{0,1\}, \sum_{i,k} \gamma_{i,k} \le 1\}$$

which has exactly $J := 2K + 1$ elements, which will be denoted by $\mathcal{Y} = \{y_j,\ j = 1, \dots, J\}$ in the following. Therefore, the discrete time default indicator process is already a process with a finite state space.

For the purpose of state space discretization of the remaining processes, let us introduce the joint loss process and the joint exposure process in analogy to the above,

$$\begin{aligned} L_k &:= \begin{pmatrix} L_k^A \\ L_k^B \end{pmatrix}, \quad k = 1, \dots, K, \\ X_k &:= \begin{pmatrix} X_k^B \\ -X_k^A \end{pmatrix}, \quad k = 1, \dots, K, \end{aligned}$$

however, for the process X we have switched the order of appearance of A and B as this order is also reversed in the CVA formula and we have provided the second component with a minus sign in accordance with Equation (2.4), see also Equation 4.3 for a further motivation. In general, there exist (at least) two different approaches how a suitable discrete state space version of the continuous processes L and X could be obtained:

- In the first approach – completely similar to the default bucketing approach – the state spaces $[0, 1]^{2\times K}$ for the joint loss process and the space $\mathbb{R}^{2\times K}$ for the joint exposure process X is divided into M, resp. N, disjoint components. Then, X is replaced by some representative value on this component (usually an average value or some mid or corner point) on each of the components, and the probabilities of the discretized process are set in accordance with the original probabilities of each component (cf. the default bucketing approach).

- In the second approach, a Monte Carlo simulation, i.e. M, resp. N, different scenarios (i.e. realizations) of the processes L and X are used instead of the original process. Each realization is assumed to have probability $\frac{1}{M}$, resp. $\frac{1}{N}$.

For both approaches it is known that they converge to the original process (almost surely and in distribution). Although the convergence properties of the first approach are more appealing from a theoretical point of view, all we are interested in is the calculation of the bilateral CVA, and for this reason we have opted for the second approach, which is more easily implemented in practical settings.

For the above reasons, if not explicitly stated otherwise, we subsequently assume that we are given a model in discrete time and discrete state space. In addition, in order to keep the following exposition as simple as possible, let us fix the loss processes to constants, i.e. let us set both losses to $L_k^A = l^A$ and $L_k^B = l^B$ and let us further assume $l^A = l^B = 1$ for notational convenience, as then the loss process could be dropped from Equation (2.4):

$$CVA_A(t, T) \quad = \quad \lim_{K\to\infty} \sum_{k=1}^{K} \left(\mathbf{E}\left[\, \delta_k^B \cdot X_k^A \,|\, \mathcal{G}_t \right] - \mathbf{E}\left[\, \delta_k^A \cdot X_k^B \,|\, \mathcal{G}_t \right] \right). \tag{2.5}$$

3 Models for Counterparty Risk

As we have seen in the last section, the CVA calculation is based on the knowledge of the joint distribution of the processes δ, L and X. In the last years two main approaches have emerged in the literature how to model the individual, resp. joint distribution of these quantities:

- The most popular approach is based on the rather strong assumption of independence between all components appearing in Equation (2.3), see Section 3.1. Based on this independence assumption, only individual models for δ, L and X (see Section 3.2 and 3.3) need to be specified for the CVA calculation. This kind of independence assumption is quite standard in the market, see for example Canabarro and Duffie[10] or the Bloomberg CVA function (for more details on the Bloomberg model let us refer to Stein and Lee [28]).

- Alternatively and more recently, a more general approach is based on a joint model (also called *hybrid model*) for the building blocks δ, L and X of the CVA calculation, see Section 3.4. For example, Redon [26] was able to obtain an analytical expression for unilateral counterparty value adjustments in a very simple setting. The first hybrid model was probably considered by Duffie and Huang [14] where the credit quality of one counterparty was assumed to be a simple function of the floating rate of the swap.

3.1 Independence of CVA components

Let us focus on the first term in Equation (2.4), and let us assume that the exposure process X is independent of the other two, merely default related, processes. Then the expectation inside the summation can be split into two parts:

$$\sum_{k=1}^{K} \mathbf{E}\left[L_k^B \cdot \delta_k^B \cdot X_k^A \,|\, \mathcal{G}_t \right] = \sum_{k=1}^{K} \mathbf{E}\left[L_k^B \cdot \delta_k^B \,|\, \mathcal{G}_t \right] \cdot \mathbf{E}\left[X_k^A \,|\, \mathcal{G}_t \right]. \tag{3.1}$$

Interestingly, the expected value

$$\mathbf{E}\left[X_k^A \,|\, \mathcal{G}_t \right] = \mathbf{E}\left[\widetilde{V}_A^{+}(t, t_k, T) \middle| \mathcal{G}_t \right] = \mathbf{E}\left[P(t, t_k) \cdot \max(V_A(t, t_k, T), 0) \middle| \mathcal{F}_t \right] \tag{3.2}$$

matches exactly the price of a call option on the basis transaction at time t with strike 0 and exercise time t_k. This means that on top of the basis transaction, options have to be modeled additionally to calculate the CVA for a transaction. For example, if the basis transaction is an interest rate swap, a swaption model is needed in order to calculate the corresponding CVA. More details on the necessary modelling efforts in such a setup are given subsequently in Section 3.3

Further, if also the loss process and the default indicator process are independent as well, it holds that

$$\begin{aligned} \mathbf{E}\left[L_k^B \cdot \delta_k^B \,|\, \mathcal{G}_t \right] &= \mathbf{E}\left[L_k^B \,|\, \mathcal{G}_t \right] \cdot \mathbf{E}\left[\delta_k^B \,|\, \mathcal{G}_t \right] \\ &= \mathbf{E}\left[L_k^B \,|\, \mathcal{G}_t \right] \cdot \mathbf{Q}\left[\tau_B \in \Delta_k, \tau_B \le \tau_A \middle| \mathcal{G}_t \right] \end{aligned}$$

as we have

$$\mathbf{E}\left[\delta_k^B \,|\, \mathcal{G}_t \right] = \mathbf{Q}\left[\delta_k^B = 1 \middle| \mathcal{G}_t \right] = \mathbf{Q}\left[\tau_B \in \Delta_k, \tau_B \le \tau_A \middle| \mathcal{G}_t \right].$$

Under such an independence assumption, the loss process $(L_k^B)_k$ can hence be replaced by its expected value, i.e. we can work with $l_k^B := \mathbf{E}\left[L_k^B \middle| \mathcal{G}_t \right]$ instead of the stochastic loss process. In case independence cannot be reasonably assumed, the discrete nature of δ can still be exploited to obtain a similar result. In the general setting, it still holds that

$$\begin{aligned} \mathbf{E}\left[L_k^B \cdot \delta_k^B \,|\, \mathcal{G}_t \right] &= \mathbf{E}\left[L_k^B \,|\, \mathcal{G}_t,\, \delta_k^B = 1 \right] \cdot \mathbf{Q}\left[\delta_k^B = 1 \middle| \mathcal{G}_t \right] \\ &= \mathbf{E}\left[L_k^B \,|\, \mathcal{G}_t,\, \delta_k^B = 1 \right] \cdot \mathbf{Q}\left[\tau_B \in \Delta_k, \tau_B \le \tau_A \middle| \mathcal{G}_t \right] \end{aligned}$$

and therefore, conditional expected values $l_k^B := \mathbf{E}\left[L_k^B \middle| \mathcal{G}_t,\, \delta_k^B = 1 \right]$ could be employed instead of ordinary expected values. Summarizing everything, we can rewrite the CVA Equation (2.4) as

$$CVA_A(t,T) \approx \sum_{k=1}^{K} \left(l_k^B \cdot \mathbf{E}\left[\delta_k^B \,|\, \mathcal{G}_t \right] \cdot \mathbf{E}\left[X_k^A \,|\, \mathcal{G}_t \right] - l_k^A \cdot \mathbf{E}\left[\delta_k^A \,|\, \mathcal{G}_t \right] \cdot \mathbf{E}\left[X_k^B \,|\, \mathcal{G}_t \right] \right) \tag{3.3}$$

and thus the CVA can be calculated without any further problems as the corresponding default probabilities $\mathbf{E}\left[\delta_k^B \,|\, \mathcal{G}_t \right] = \mathbf{Q}\left[\tau_B \in \Delta_k, \tau_B \le \tau_A \middle| \mathcal{G}_t \right]$, see following Section 3.2, can be easily computed.

3.2 Modelling of dependent default times

In order to calculate the probability $\mathbf{Q}\left[\tau_B \in \Delta_k, \tau_B \le \tau_A \middle| \mathcal{G}_t \right]$, the default times τ_A and τ_B and their dependence structure have to be modeled. One of the most popular models for default times in general are intensity models, as for example described in Bielecki and Rutkowsi [2], Part III. Let us point out that since we only consider the situation that no default has occurred yet, all available information is already given by the filtration $\mathcal{F}_t$.

In the framework of intensity models, also called reduced form models, the time of default is interpreted as the first jump of a Poisson process, whose parameter is associated with the default intensity. This means that the survival probability at time t denoted by $G_t := \mathbf{Q}[\tau > t|\mathcal{F}_t]$, can be expressed by means of a so-called intensity process $\Lambda(t)$:

$$G_t = \exp(-\Lambda(t)).$$

A very popular setup is the framework of Cox processes, which is a setup where the stochastic default intensity processes have strictly positive stochastic intensities λ_i, $i \in \{A, B\}$ such that

$$\Lambda_i(t) = \int_0^t \lambda_i(s) ds$$

is invertible. In such a setting the default times τ_i are usually defined via standard exponential random variables ξ_i as

$$\tau_i = \Lambda_i^{-1}(\xi_i).$$

Within an intensity model dependence of default times τ_A and τ_B can now be achieved in two different ways: Either by using dependent intensity processes λ_A and λ_B or by using a copula approach. A detailed example for introducing dependence via intensity processes[6] will be given within the hybrid model framework in Section 3.4.

It is well-known, see for example Brigo and Capponi [5], that the linkage via copula functions results in a much stronger dependence compared to a linking via correlated intensity processes. For the linking by means of a copula function, one chooses a copula for modelling correlated uniform distributed random variables associated to ξ_i. For this purpose consider the specification

$$U_i := 1 - \exp(-\xi_i), \quad \text{and} \quad \mathbf{Q}[U_A \leq u_A, U_B \leq u_B] = C(u_A, u_B),$$

for some copula function C. Afterwards, dependent default times are obtained by

$$\tau_i = \Lambda_i^{-1}(-\ln(1 - U_i)).$$

The most commonly used copula function is the Gaussian copula with correlation parameter ρ, which will also be the copula of choice in our specific example later on. For our purposes, we do not consider stochastic intensity processes, but work with deterministic Λ_i (i.e. deterministic default intensities λ_i) which are obtained by the calibration of CDS model prices to market prices. Please note that in this calibration process, only the marginal distributions of τ_i can be calibrated to CDS prices, but no information about the dependence structure of the default times is hidden in the CDS market (see for example Brigo and Masetti [7]). Further note that a model with deterministic default intensities plus a suitable copula is sufficient for the arbitrary specification of the joint distribution of default times and that stochastic intensities do not add any value in this context. This is because we only focus on the default time distribution, but are not interested in correlations to other risk factors, like interest rates, stock prices, etc.

In such a setup with deterministic default intensities and arbitrary copula, the distribution of the default indicator process δ is completely specified. For the corresponding probabilities of each of the $J = 2K + 1$ different realizations of the default indicator, denoted by

$$\begin{aligned}
q_k^\delta &:= \mathbf{Q}[\delta_k^A = 1 \,|\, \mathcal{F}_t], \quad k = 1, \ldots, K \\
q_{K+k}^\delta &:= \mathbf{Q}[\delta_k^B = 1 \,|\, \mathcal{F}_t], \quad k = 1, \ldots, K \\
q_J^\delta &:= 1 - \sum_{j=1}^{2K} q_j^\delta,
\end{aligned}$$

it holds that

$$\begin{aligned}
\mathbf{Q}[D_A] &= \sum_{k=1}^{K} \mathbf{Q}[\delta_k^A = 1 \,|\, \mathcal{F}_t] = \sum_{k=1}^{K} q_k^\delta, \\
\mathbf{Q}[D_B] &= \sum_{k=1}^{K} \mathbf{Q}[\delta_k^B = 1 \,|\, \mathcal{F}_t] = \sum_{k=K+1}^{2K} q_k^\delta, \text{ and} \\
\mathbf{Q}[D_0] &= q_J^\delta.
\end{aligned}$$

[6] The approach of dependent intensities has been originally introduced as conditionally independent default (CID) models, see for example Duffie and Singleton [15] or Lando [20].

These probabilities can be either calculated by a straightforward Monte Carlo simulation (derived from randomly sampled default times τ_A and τ_B) or by numerical integration: if we denote the density of the joint distribution of (τ_A, τ_B) by q, i.e.

$$\mathbf{Q}[\tau_A \le t_A, \tau_B \le t_B] = \int_0^{t_A} \int_0^{t_B} q(s_A, s_B) ds_B ds_A,$$

then it holds that

$$\begin{aligned} \mathbf{Q}[\delta_k^A = 1] &= \mathbf{Q}[\tau_A \le \tau_B,\, t_{k-1} < \tau_A \le t_k] \\ &= \int_{t_{k-1}}^{t_k} \int_{s_A}^{\infty} q(s_A, s_B) ds_B ds_A. \end{aligned}$$

Taking into account that $\mathbf{Q}[\tau_A \le t_A, \tau_B \le t_B]$ can be expressed by means of the copula as

$$\begin{aligned} &\mathbf{Q}[\tau_A \le t_A, \tau_B \le t_B] = \\ &\qquad = \mathbf{Q}[U_A \le 1 - e^{-\Lambda_A(t_A)},\, U_B \le 1 - e^{-\Lambda_B(t_B)}] = C(1 - e^{-\Lambda_A(t_A)},\, 1 - e^{-\Lambda_B(t_B)}) \end{aligned}$$

the above integral over the density q can be reformulated with the help of the copula density c as

$$\int_{t_{k-1}}^{t_k} \int_{s_A}^{\infty} q(s_A, s_B) ds_B ds_A = \int_{t_{k-1}}^{t_k} \int_{s_A}^{\infty} c(1 - e^{-\Lambda_A(s_A)},\, 1 - e^{-\Lambda_B(s_B)}) ds_B ds_A,$$

which can be evaluated by standard techniques from numerical integration.

The following Table 1 and Figure 2 illustrate the probabilities resulting from a Gaussian copula model with correlation parameter ρ varying in $(-1, 1)$. Two different intensity setups are shown in the figure: on the left identical default intensities $\lambda_A = \lambda_B = 150bps$, whereas the right graph illustrates the case of counterparties of different qualities, i.e. counterparty B having a higher default intensity and such a higher probability of defaulting before A, namely $\lambda_A = 150bps$ and $\lambda_B = 300bps$. In such a simple setting with constant default intensities the

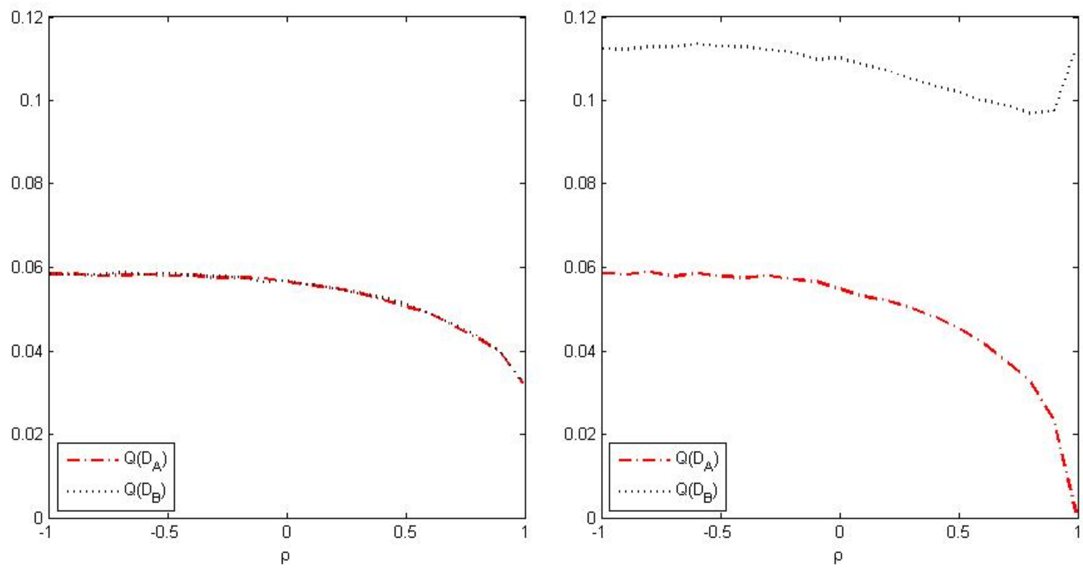

Figure 2: Probabilities of D_A and D_B for varying ρ.

default probabilities $\mathbf{Q}[\tau_i \in \Delta_k | \mathcal{F}_t]$ can be determined analytically:

$$\begin{aligned} \mathbf{Q}[\tau_i \in \Delta_k | \mathcal{F}_t] &= \mathbf{Q}[\tau_i > t_{k-1} | \mathcal{F}_t] - \mathbf{Q}[\tau_i > t_k | \mathcal{F}_t] = G_{t_{k-1}} - G_{t_k} \\ &= \exp(-\Lambda(t_{k-1})) - \exp(-\Lambda(t_k)) \\ &= \exp(-\lambda \cdot t_{k-1}) - \exp(-\lambda \cdot t_k). \end{aligned}$$

In Table 1 the default indicator probabilities are listed in detail which result from a simulation over a time horizon of $T = 4$ years divided into $K = 8$ disjoint time intervals Δ_k each covering half a year in case of different counterparties and a copula correlation of $\rho = 0.9$. For comparison the analytical probabilities $\mathbf{Q}[\delta_k^i = 1|\mathcal{F}_t]$ are shown as well. It can be clearly observed that due to the influence of the copula parameter ρ, defaults of party B are more than four times more likely than defaults for party A, although default probabilities are only twice as high.

k	1	2	3	4	5	6	7	8	Σ
$\mathbf{Q}[\delta_k^A = 1 \,\vert\, \mathcal{F}_t]$ in %	0.39	0.35	0.31	0.29	0.27	0.25	0.24	0.23	2.33
$\mathbf{Q}[\tau_A \in \Delta_k \,\vert\, \mathcal{F}_t]$ in %	0.75	0.74	0.74	0.73	0.73	0.72	0.71	0.71	5.83
$\mathbf{Q}[\delta_k^B = 1 \,\vert\, \mathcal{F}_t]$ in %	1.31	1.27	1.26	1.22	1.22	1.20	1.18	1.16	9.81
$\mathbf{Q}[\tau_B \in \Delta_k \,\vert\, \mathcal{F}_t]$ in %	1.49	1.47	1.44	1.42	1.40	1.38	1.36	1.34	11.30

Table 1: $\mathbf{Q}[\delta_k^i = 1 \,|\, \mathcal{F}_t]$ for $\rho = 0.9$ and $\lambda_A = 150bps$, $\lambda_B = 300bps$.

3.3 Modelling options on the basis transaction

Since it could be observed in Equation (3.2) that options on the basis transaction need to be priced, a suitable model for this option pricing task needs to be available. Depending on the type of derivative, any model which can be reasonably well calibrated to the market data is sufficient. For instance, for interest rate derivatives, any model ranging from a simple Vasicek or CIR model to sophisticated Libor market models or two-factor Hull-White models could be applied. In case of a credit default swap, any model which allows to price CDS options, i.e. any model with stochastic credit spread would be feasible. However, for CVA calculations, usually a trade-off between accuracy of the model and efficiency of calculations needs to be made. For this reason, usually more simple models are applied for CVA calculations than for other pricing applications. It needs to be noted that since the financial market usually provides sufficiently many prices of liquid derivatives, any reasonable model can be calibrated to these market prices, and therefore, we can assume in the following that the market implied distribution of the discounted exposure process is fully known and available.

3.4 Hybrid models – an example

Another way to calculate the CVA is to use a so-called *hybrid approach* which models all the involved underlying risk factors which influence the value of the transaction at the same time. Instances of such models, which are based on state-of-the-art models can for example be found in Brigo and Capponi [5] for the case of a credit default swap, or Brigo, Pallavicini and Papatheodorou [9] for interest rate derivatives. In Brigo, Pallavicini and Papatheodorou [9], an integrated framework is introduced, where a two-factor Gaussian interest-rate model is set up for a variety of interest rate derivatives[7] in order to deal with the option inherent in the CVA. Further, to model the possible default of the client and its counterparty their stochastic default intensities are given as CIR processes with exponentially distributed positive jumps. The Brownian motions driving those risk factors are assumed to be correlated. Additionally, the defaults of the client and the counterparty are linked by a Gaussian copula. In this approach the short rate is given in detail by the two-factor model

$$\begin{aligned} r(t) &= x(t) + z(t) + \phi(t;\alpha) \quad r(0) = r_0 \\ dx(t) &= -ax(t)dt + \sigma dZ_1(t) \quad x(0) = 0 \\ dz(t) &= -bz(t)dt + \mu dZ_2(t) \quad z(0) = 0 \end{aligned}$$

with correlated Brownian motions Z_1 and Z_2 with $d\langle Z_1, Z_2\rangle_t = \rho_{12}dt$, and further positive constants r_0, a, b, σ, μ and $\rho_{12} \in [-1, 1]$. The deterministic function ϕ depends on all these

[7]Although this modelling approach is a rather general one, it has to be noted that it links the dependence on tenors of swaption volatilities to the form of the initial yield curve. Therefore, the limits of such an approach became apparent as the yield curve steepened in conjunction with a movement of the volatility surface in the aftermath of the beginning financial crisis in 2008, when these effects could not be reproduced by such a model.

parameters which are expressed by α, in that $\alpha = (r_0, a, b, \sigma, \mu, \rho_{12})$ and it can be always chosen in such a way that the model is automatically calibrated to the initial zero rate curve observed in the market.

The stochastic default intensities of the client A and the counterparty B in the CIR approach are modelled as shifted square root diffusions processes with positive jumps

$$\begin{aligned}
\lambda_t^i &= y_t^i + \psi^i(t;\beta^i) \\
\psi(0;\beta) &= \lambda_0 - y_0 \\
dy_t^i &= \kappa^i(\mu^i - y_t^i)dt + dJ_t^i(\xi_1^i, \xi_2^i) + \nu^i \sqrt{y - t^i} dZ_3^i(t) \\
J_t^i(\xi_1^i, \xi_2^i) &:= \sum_{k=1}^{M_t^i(\xi_1^i)} X_k^i(\xi_2^i)
\end{aligned}$$

with $i \in \{A, B\}$. As before the positive deterministic constants $\kappa^i, \mu^i, \nu^i, y_0^i, \xi_1^i, \xi_2^i$ are summarized in a vector, i.e. $\beta^i = \left(\kappa^i, \mu^i, \nu^i, y_0^i, \xi_1^i, \xi_2^i\right)$. The Brownian motions Z_3^i are assumed to be correlated with the former ones, i.e. $d\left\langle Z_j, Z_3^i \right\rangle_t = \rho_{j,i} dt$, $j \in \{1, 2\}$, $i \in \{A, B\}$, but not correlated to each other ($d\left\langle Z_3^A, Z_3^B \right\rangle_t = 0$). Thus, the following correlation matrix for the model driving Brownian motions results:

$$\begin{pmatrix} 1 & \rho_{12} & \rho_{1A} & \rho_{1B} \\ \rho_{12} & 1 & \rho_{2A} & \rho_{2B} \\ \rho_{1A} & \rho_{2A} & 1 & 0 \\ \rho_{1B} & \rho_{2B} & 0 & 1 \end{pmatrix}$$

Brigo, Pallavicini and Papatheodorou [9] even assume y^i, for $i \in \{A, B\}$, to be independent (and thus also Z_3^i) to ease the calibration of the model and to focus on default correlation rather than spread correlation.

Further, the amount of wrong-way risk which can be modelled within such a framework depends strongly on the model choice, e.g. if solely correlations between default intensities (i.e. credit spreads) and interest rates are taken into account, only a rather weak relation will emerge between default and the exposure of interest rate derivatives, cf. Brigo, Pallavicini and Papatheodorou [9].

4 Model-free approach

In the last section, it became obvious that each of the two approaches presented so far has a major drawback:

- The independence assumption is a rather strong assumption and usually not fulfilled in practical settings. In such an approach *wrong-way risk*, i.e. a potential positive correlation between default (more exactly the default indicator process) and the exposure process is completely neglected.
- Although hybrid models overcome this drawback, it is not clear how hybrid models could be reasonably parametrized, as usually no appropriate market information is available, i.e. there are no markets for financial instruments depending on the correlation between default and exposure. In the specific example in Section 3.4 there are four parameters correlating the short rate process with the stochastic default intensities. With these four parameters, it is unfortunately not obvious how big or small the bilateral CVA could get for reasonable choices of the correlations, especially as these four parameters all have a nonlinear impact on the resulting CVA, see for example Section 5 in Brigo and Capponi [5].

For the above reasons, we propose a completely different approach to calculate the counterparty value adjustment. This new approach can be seen as a *model-free approach*, as it does not rely on any specific model for the dependence of the default times and the exposure process. Instead, our model-free approach directly links all components which appear in the general CVA formula (2.4): the joint loss process L, the joint default indicator process δ and the joint discounted exposure processes X. Although the loss process L was dropped for convenience, please note that – in contrast to the previous approaches in Section 3 – the following framework is completely flexible and would allow for a very easy consideration of general loss processes as well. As already mentioned in Section 3.3, we can reasonably assume that the distribution of the exposure process X is already completely determined by the available market information. In a similar manner, we have argued in Section 3.2 that also the distribution of the default indicator process δ can be assumed to be given by the market (together with some expert opinion on the default correlation). Nevertheless, let us point out that the following ideas and concepts could indeed be generalized to the case that only the marginal distributions of the default times are known. For these reasons, we can reasonably argue that the following approach is indeed model-free in the sense that no model needs to be specified which links the default indicator process with the discounted exposure processes – in clear contrast to the hybrid model – and we will see that also no independence assumption between these components needs to be imposed. Of course, as outlined in Sections 3.3 and 3.2, for each individual component, some pricing model consistent to the market data needs to be available.

In the following paragraphs we are going to introduce a general framework for discrete processes, before we embed the CVA calculations into this framework, which will lead to model-free lower and upper bounds on the bilateral CVA. We close this section with a few applications of these model-free bounds.

4.1 A general framework

In the following, let us assume that two discrete time processes $Y = (Y_k)_{k=1,\dots,K}$ and $Z = (Z_k)_{k=1,\dots,K}$ are given. Further assume that both processes live in a finite state space, i.e. process Y takes values in $\mathcal{Y} := \{y^{(j)} \in \mathbb{R}^{n_Y \times K},\ j = 1,\dots,J\}$ and process Z takes values $\mathcal{Z} := \{z^{(i)} \in \mathbb{R}^{n_Z \times K},\ i = 1,\dots,N\}$. Finally, we assume that the (marginal) distributions of the processes Y and Z are known, i.e. we are given

$$\begin{aligned} q_j^Y \quad &:= \quad \mathbf{Q}[Y = y^{(j)}] = \mathbf{Q}[Y_1 = y_1^{(j)}, \dots, Y_K = y_K^{(j)}], \quad j = 1,\dots,J, \\ q_i^Z \quad &:= \quad \mathbf{Q}[Z = z^{(i)}] = \mathbf{Q}[Z_1 = z_1^{(i)}, \dots, Z_K = z_K^{(i)}], \quad i = 1,\dots,N, \\ \text{with} \qquad & \sum_i q_i^Z = \sum_j q_j^Z = 1. \end{aligned}$$

In this framework, the distribution of the process $\binom{Y}{Z}$, i.e. the joint distribution of the processes Y and Z, is not known in advance. Since both processes have finite state space, the joint distribution of Y and Z can be fully characterized by a matrix $Q \in \mathbb{R}^{J\times N}$ where $Q_{i,j} := \mathbf{Q}[Z = z^{(i)}, Y = y^{(j)}]$. Therefore, any such Q has the following properties:

- the matrix Q is *proper*, which means that it represents a proper probability distribution, i.e. it holds
$$Q_{i,j} \geq 0, \quad \text{and} \quad \sum_{i,j} Q_{i,j} = 1,$$
- the matrix Q is *consistent* with the given marginal distributions, i.e. it holds
$$\begin{aligned} q_j^Y \quad &= \quad \sum_i Q_{i,j} \;=\; \sum_i \mathbf{Q}[Z = z^{(i)}, Y = y^{(j)}] \;=\; \mathbf{Q}[Y = y^{(j)}], \quad \forall j = 1,\dots J, \\ q_i^Z \quad &= \quad \sum_j Q_{i,j} \;=\; \sum_j \mathbf{Q}[Z = z^{(i)}, Y = y^{(j)}] \;=\; \mathbf{Q}[Z = z^{(i)}], \quad \forall i = 1,\dots N. \end{aligned}$$

Let $\mathcal{Q}$ denote the set of all feasible and consistent matrices Q which represent possible joint distributions for the processes Y and Z, i.e. let

$$\mathcal{Q} := \{Q \in \mathbb{R}_+^{J\times N} \mid \sum_i Q_{i,j} = q_j^Y\,,\ \sum_j Q_{i,j} = q_i^Z\ \forall i,\, \forall j\,\}.$$

Since $\mathcal{Q}$ is a convex and compact polyhedron in $\mathbb{R}^{J\times N}$, this means that the set of all proper and consistent joint probability distributions of Y and Z can be represented by a compact set due to the discrete nature of the processes Y and Z. Let us note that such a (much more technical) construction could have also been carried out for continuous time and continuous state spaces. Since these constructions resemble the setup of so-called (Monge-Kantorovich) *mass transportation problems* let us refer to Rachev and Rüschendorf [24, 25] for more details on the general concept of these mass transportation problems.

For any arbitrary function $w : \mathcal{Y} \times \mathcal{Z} \to \mathbb{R}$ we can then define a new random variable W based on the two processes Y and Z by $W := w(X, Z)$. Obviously W takes only finitely many values $W_{i,j} := w(y^{(j)}, z^{(i)})$ with probability $\mathbf{Q}[W = W_{i,j}] = \mathbf{Q}[Z = z^{(i)}, Y = y^{(j)}] = Q_{i,j}$, i.e. each mapping w can be equivalently represented by a corresponding *value matrix* $W \in \mathbb{R}^{J \times N}$. Now, for given mapping w, let us consider the following linear optimization problem,

$$\min_{Q \in \mathcal{Q}} \mathbf{E}_Q[w(Y, Z)] \;=\; \min_{Q \in \mathcal{Q}} \mathbf{E}_Q[W] \;:=\; \min_{Q \in \mathcal{Q}} \sum_{i,j} Q_{i,j} \cdot W_{i,j}, \tag{4.1}$$

where the expected value of W is minimized over all proper and consistent distributions Q. Problem (4.1) is indeed a linear optimization problem in the variable $Q \in \mathbb{R}^{J \times N}$ – for any arbitrary function w – as both the objective function and the constraints defining $\mathcal{Q}$ are obviously linear in Q. Further, due to compactness of $\mathcal{Q}$, the minimum exists. Obviously, for any $Q \in \mathcal{Q}$,

$$\min_{P \in \mathcal{Q}} \mathbf{E}_P[w(Y, Z)] \leq \mathbf{E}_Q[w(Y, Z)] \leq \max_{P \in \mathcal{Q}} \mathbf{E}_P[w(Y, Z)].$$

As a specific example for the mapping w let $n_Y = n_Z = n$ and let us consider the usual scalar product $\circ$ for matrices $y, z \in \mathbb{R}^{n \times K}$,

$$w(y, z) \;:=\; y \circ z := \sum_{l=1}^{n} \sum_{k=1}^{K} y_{l,k} \cdot z_{l,k}.$$

Then the value matrix $W \in \mathbb{R}^{J \times N}$ can be computed from the given data $y^{(j)}$ and $z^{(i)}$ by

$$W_{i,j} \;=\; y^{(j)} \circ z^{(i)} \;=\; \sum_{l=1}^{n} \sum_{k=1}^{K} y^{(j)}_{l,k} \cdot z^{(i)}_{l,k}$$

and it holds that

$$\begin{aligned} \mathbf{E}_Q[w(Y, Z)] \;=\; \mathbf{E}_Q[Y \circ Z] \;=\mathbf{E}_Q &\left[\sum_{l=1}^{n} \sum_{k=1}^{K} Y_{l,k} \cdot Z_{l,k} \right] \\ &= \sum_{i,j} Q_{i,j} \cdot \left(\sum_{l=1}^{n} \sum_{k=1}^{K} y^{(j)}_{l,k} \cdot z^{(i)}_{l,k} \right) \;=\; \sum_{i,j} Q_{i,j} W_{i,j}. \end{aligned} \tag{4.2}$$

A few interesting things about Problem (4.1) can be noted.

- First of all, the structure of the LP (4.1) coincides with the structure of so-called *balanced linear transportation problems.* Transportation problems constitute a very important subclass of linear programming problems, see for example Bazaraa, Jarvis and Sherali [1], Chapter 10, for more details. Further, there exist several very efficient algorithms for the numerical solution of such transportation problems, see also Bazaraa, Jarvis and Sherali [1], Chapter 10 to 12.

- Second, it is known that each LP attains its optimal solution also in a corner of the feasible polyhedron. As the polyhedron of feasible probability distributions is embedded in the $\mathbb{R}^{J \times N}$, $J \cdot N$ of the $J \cdot N + N + J$ linear constraints have to be active to define a corner. This means that at most $N + J$ entries of the optimal solution Q^* to Problem (4.1) are strictly larger than zero, the rest has to be equal to zero. Thus, for finer discretizations, i.e. increasing N and J, the matrix Q becomes (relatively) sparse.

- Finally, if the value matrix W satisfies certain conditions, the solution to Problem (4.1) could be determined without the numerical solution of the transportation problem. For example, if W satisfies the so-called *Monge condition*, see Rachev and Rüschendorf [24] (1.2.2),

$$W_{i,j} + W_{i+1,j+1} \leq W_{i+1,j} + W_{i,j+1}$$

then the well-known greedy north-west-corner-rule for the feasible initialization of the LP already gives the optimal solution, known to be the Hoeffding H distribution. However, in the given framework, the Monge condition is only very rarely satisfied, for example, if $K = n = 1$ and y and z are ordered increasingly.

4.2 Embedding the CVA calculation into the general framework

For the calculation of model-free bilateral counterparty value adjustments let us embed the previous setup from Equation (2.5) into the framework from the last section. For this purpose, let

$$\begin{aligned} Y &:= \delta \\ Z &:= X \end{aligned}$$

and let $w(Y, Z) = Y \circ Z$. Let us further recall that the the marginal probabilities are predetermined by the individual models for the default times and the exposure process. For example, the probabilities of the default process can be derived as in Table 1, and, assuming the scenario (i.e. Monte Carlo) approach is used for the discretization of X, see Section 2.4, each probability q_i^X equals $\frac{1}{N}$ by construction. With this notation, we can embedd Equation (2.5) into the general setup:

$$\begin{aligned} CVA_A(t,T) &\approx \mathbf{E}_Q\left[\,\delta_k^B \cdot X_k^A \,|\, \mathcal{G}_t\right] - \mathbf{E}_Q\left[\,\delta_k^A \cdot X_k^B \,|\, \mathcal{G}_t\right] \\ &= \mathbf{E}_Q\left[\,\delta \circ X \,|\, \mathcal{G}_t\right] \\ &= \mathbf{E}_Q\left[\,Y \circ Z \,|\, \mathcal{G}_t\right], \end{aligned} \tag{4.3}$$

where we have now highlighted the dependence of the CVA figure on the joint distribution Q. In case of the independence assumption, Q is given by the product distribution of δ and Z, whereas in hybrid models the joint distribution Q is determined by the specification and parametrization of the hybrid model. In contrast to these specific approaches, we are now in a situation where we can define lower and upper bounds on the CVA by varying the joint distribution over the set $\mathcal{Q}$ of all possible joint distributions which are consistent with the marginal distributions:

$$CVA_A^l(t,T) := \min_{Q\in\mathcal{Q}} \mathbf{E}_Q[\delta \circ X] \le CVA_A(t,T) \le \max_{Q\in\mathcal{Q}} \mathbf{E}_Q[\delta \circ X] =: CVA_A^u(t,T). \tag{4.4}$$

These bounds represent the lowest and the highest CVA which can be obtained by any (hybrid) model which is consistent with the market data. Further, there exists at least one model which reaches these bounds, i.e. the bounds are sharp. This is in contrast to Turnbull [29], where only weak upper and lower bounds were derived. Of course, bounds always represent a best-case and a worst-case estimate only, which may strongly under- and overestimate the true CVA. However, as we will see in Section 4.3, there are several applications for these estimates.

As a final remark on this embedding, let us mention that in this context the computation of the value matrix W has a nice interpretation: W contains the contributions to the bilateral CVA of all possible combinations of default scenarios applied to all different exposure scenarios, see also Figure 4 for a similar illustration.

4.3 Application to CVA calculation

The lower and upper bound for the bilateral counterparty value adjustment, $CVA_A^l(t,T)$ and $CVA_A^u(t,T)$, bracket $CVA_A^i(t,T)$, the bilateral counterparty value adjustment in case of independence, i.e. it holds

$$CVA_A^l(t,T) \le CVA_A^i(t,T) \le CVA_A^u(t,T).$$

For the specific example introduced in Section 3, this is illustrated in Figure 3, where the bounds are given together with the result of the independent case. As Figure 3 shows, we can introduce a piecewise linear and monotone function

$$CVA : \begin{cases} [-1;1] \to [CVA_A^l(t,T), CVA_A^u(t,T)] \\ \kappa \mapsto CVA(\kappa) \end{cases}$$

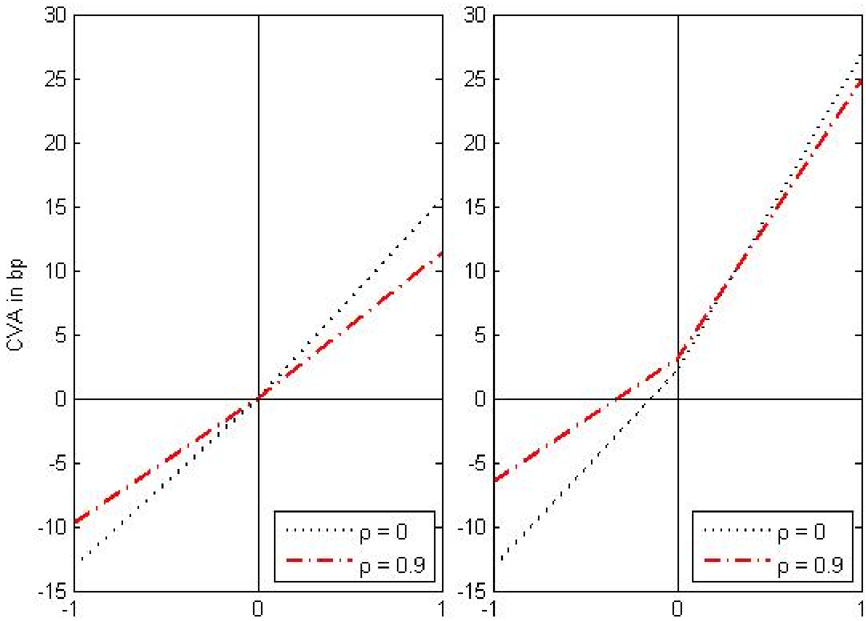

Figure 3: Minimal, maximal and independent bilateral CVA for symmetric and asymmetric counterparties.

with $CVA(-1) = CVA^l_A(t,T)$, $CVA(0) = CVA^i_A(t,T)$ and $CVA(1) = CVA^u_A(t,T)$. This allows to view the parameter κ as the *total default correlation*. Instead of basing CVA calculations on several nonlinear correlation parameters (as in the hybrid model), it is now possible to summarize the CVA behaviour in one easy-to-interpret variable. Such a setup can be used for several practical applications:

- Setting $\kappa = 1$, the upper bound for the CVA can be used in stress tests on counterparty default risk.
- Setting κ to a realistic, but still rather high level, could be used for profitability and economic capital calculations.
- Comparisons to the implied κ from hybrid models allows to judge how strong the collection of correlation parameter actually influences the bilateral CVA.
- Translating correlation parameters from hybrid models into one implied κ allows for the consideration of model risk in a simplified manner.

4.4 An alternative formulation as assignment problem

For the above setup we have assumed that the probabilities q^δ for all possible realizations of the default indicator process could be precomputed from a suitable default model. If for some default model this should not be the case, but only scenarios (with repeated outcomes for the default indicator) could be obtained by a simulation, an alternative LP formulation could be formulated. In such a scenario setting, it is advisable that for both Monte Carlo simulations, the same number N of scenarios is chosen. Then for both given marginal distributions we have

$q_j^\delta = q_i^X = \frac{1}{N}$. If we apply the same arguments as in Section 4.1 we obtain the problem

$$\begin{aligned} \min_{Q \in \mathbb{R}_+^{N \times N}} \quad & \frac{1}{N} \sum_{i,j} Q_{i,j} W_{i,j} \\ \text{subject to} \quad & \sum_j Q_{i,j} = 1, \quad i = 1, ..., N \\ & \sum_i Q_{i,j} = 1, \quad j = 1, ..., N, \end{aligned} \tag{4.5}$$

where we have multiplied the probability distribution Q by N. Problem (4.5) again represents a balanced transportation problem, but now with a even more specific structure. If we have a closer look at Problem (4.5), we see that the optimization actually runs over all $N \times N$ permutation matrices – since each default scenario is mapped onto exactly one exposure scenario, see Figure 4 for illustration. This means that Problem (4.5) eventually belongs to the class

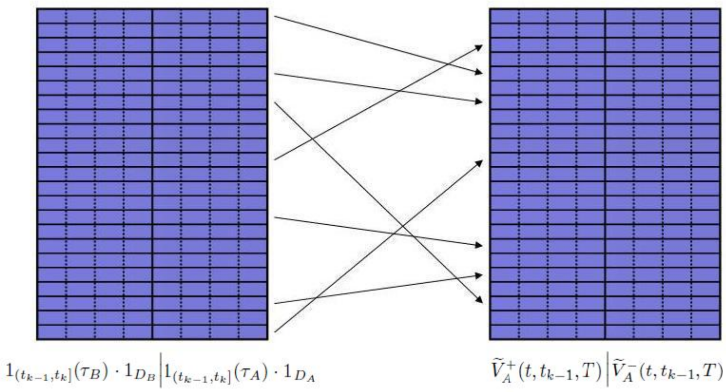

Figure 4: Matching the realisations of the default indicators to the corresponding present values of the transaction's outstanding amounts in case of default.

of assignment problems, for which very efficient algorithms are available, cf. Bazaraa, Jarvis and Sherali [1]. From the classification of Problem (4.5) as an assignment problem, we also obtain the interesting fact that the optimal solution is represented by a permutation matrix. Unfortunately, although this fact is exploited in the efficient numerical solution, this information cannot be transferred back to the original problem, as no structure is available in the Monte Carlo simulations. Please note that although assignment problems can be solved more efficiently than transportation problems, it is still advisable to solve the transportation problem due to its lower dimensionality, as usually $J = 2K + 1 < N$ (i.e. time discretization is usually much coarser than exposure discretization).

5 Example

5.1 Setup

To illustrate the model-free approach we will give a detailed example in this section. For this purpose let us consider a standard payer swap with a remaining lifetime of $T = 4$ years analyzed within a Cox-Ingersoll-Ross (CIR) model at time $t = 0$. In accordance to the example in Section 3.2 the time interval $(0, 4]$ is split up into $K = 8$ disjoint time intervals each covering half a year. Thus we have $t_k \in \{0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4\}$ for $\Delta_k = (t_{k-1}, t_k], k = 1 \ldots 8$. We have already observed that this means that the individual default indicator processes δ_s^i only take values in $\{0, 1\}^K$ with $\sum_{k=1}^K \delta_k^i \leq 1$. For simplicity, the loss process is again assumed to be 1.

5.1.1 Counterparty's default modelling

To model the defaults we have chosen the copula approach described in Section 3.2 using the Gaussian copula. In Section 3.2 we also already illustrated the default probabilities for a varying copula correlation parameter ρ which models the dependence of the counterparties. For further analyses in this example we will focus on the case of uncorrelated counterparties ($\rho = 0$) and highly correlated counterparties ($\rho = 0.9$) . Furthermore, the counterparty's default intensities are assumed to be deterministic and we will distinguish between symmetric counterparties with identical default intensities and asymmetric counterparties. Thus, four different settings result: Figure 5 shows the probabilities $\mathbf{Q}[\delta_k^i = 1] = \mathbf{E}[\delta_k^i]$ in each of the four cases (compare also

Case 1: symmetric, uncorrelated	$\lambda_A = 150bps$	$\lambda_B = 150bps$	$\rho = 0$
Case 2: symmetric, correlated	$\lambda_A = 150bps$	$\lambda_B = 150bps$	$\rho = 0.9$
Case 3: asymmetric, uncorrelated	$\lambda_A = 150bps$	$\lambda_B = 300bps$	$\rho = 0$
Case 4: asymmetric, correlated	$\lambda_A = 150bps$	$\lambda_B = 300bps$	$\rho = 0.9$

to Table 1 in Section 3.2, where values for Case 4 with higher accuracy are stated). To be in line with following figures, the probabilities for a default of counterparty B in Δ_k, i.e. $\mathbf{E}[\delta_k^B]$, correspond to the positive bars and defaults of counterparty A to the negative bars. The left plots

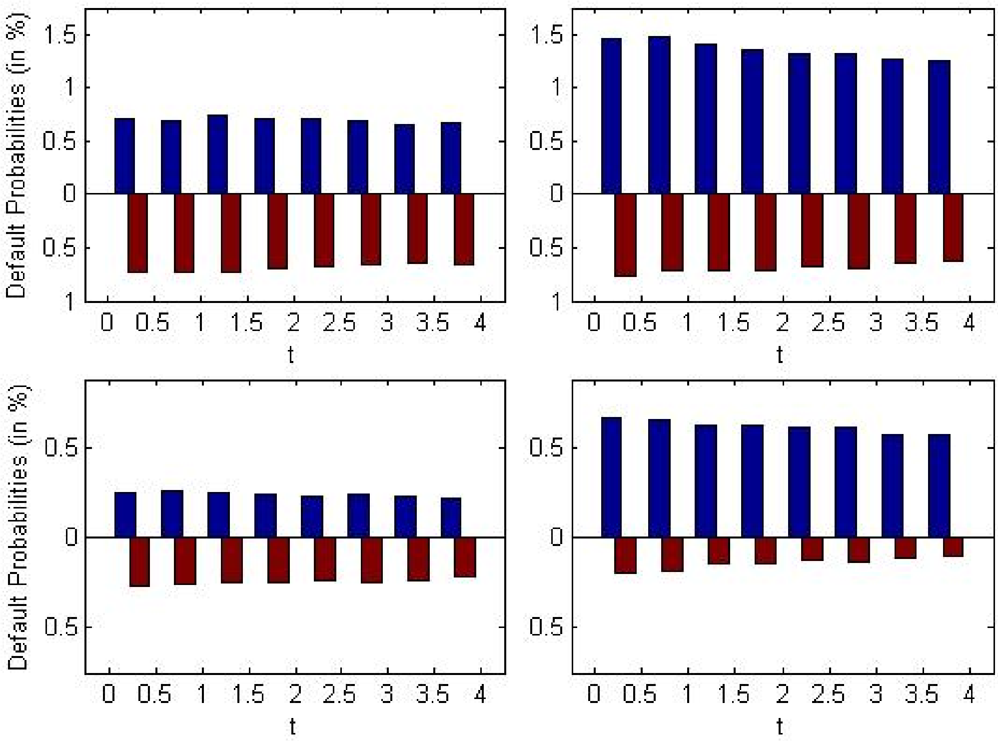

Figure 5: Probabilities $\mathbf{E}[\delta_k^i]$ in % for cases 1 to 4.

show identical counterparties (cases 1 and 2) and the right ones the cases, where counterparty B has a higher default intensity (cases 3 and 4). Furthermore, the upper plots correspond to uncorrelated defaults and for the ones below we have $\rho = 0.9$.

5.1.2 Counterparty exposure modelling

As already mentioned, a simple CIR model is applied for the valuation of the payer swap. As our focus is on the coupling of the default and the exposure model, we have opted for such a simple model for ease of presentation, although any more sophisticated model could have been applied as well. In the CIR model, the short rate r_t follows the stochastic differential equation

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dZ_t,$$

where the parameters have been calibrated to market data (yield curve plus selected swaption prices) yielding

$$\kappa = 0.0156 \quad \theta = 0.0311 \quad \sigma = 0.0313 \quad r_0 = 0.030.$$

Figure 6 shows exemplarily some of the generated short rate scenarios on the left and the corresponding resulting discounted swap values $P(0, s) \cdot V_A(s, T)$ on the right. Considering now the

k	1	2	3	4	5	6	7	8
$\mathbf{E}[X_k^A]$ in bp	49.2	59.2	60.1	55.2	45.9	33.4	17.9	0
$\mathbf{E}[X_k^B]$ in bp	48.9	58.5	59.1	54.2	45.1	32.6	17.5	0

Table 2: $\mathbf{E}\left[X_k^A\right]$ and $\mathbf{E}\left[X_k^B\right]$ in basis points.

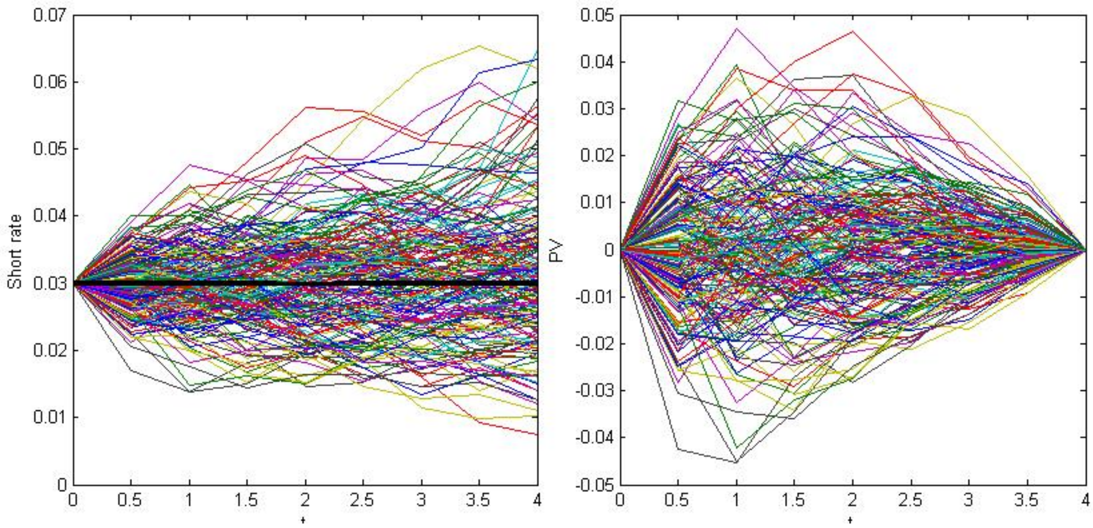

Figure 6: Simulated short rate scenarios and corresponding swap values $P(0, s) \cdot V_A(s, T)$.

exposure of each counterparty, i.e. $P(0, s) \cdot \max(0, V_i(s, T)) = \widetilde{V}_i^+(0, s, T)$ or within the discrete time framework of our example $P(0, t_k) \cdot \max(0, V_i(t_k, T)) = X_k^i$, we can easily compute $\mathbf{E}\left[X_k^i\right]$ as the average of all generated scenarios from a Monte Carlo simulation. Figure 7 illustrates the results of a simulation, which are also given in Table 2. Positive bars correspond to $\mathbf{E}\left[X_k^A\right]$, negative bars to $\mathbf{E}\left[X_k^B\right]$ and the small bars correspond to $\mathbf{E}[P(0, t_k) \cdot V_A(t_k, T)]$. Since payer and receiver swap are not completely symmetric instruments, there remains a residual expectation, as can be observed from Figure 7.

5.2 Results

In case of independence between default and exposure the bilateral CVA is easily obtained by multiplying the default probabilities (as shown in Figure 5) with the corresponding exposures (as shown in Figure 7) and summation. Besides the independent CVA^i the minimal and maximal CVA^l and CVA^u have been calculated as well by solving the transportation problem 4.4.

The results of these calculations are illustrated in Figure 8 and Table 3 for each time interval Δ_k. Analogously to Figure 5 we have for each of the four cases a separate subplot and the left plots belong again to cases 1 and 2. The positive bars now correspond to $\mathbf{E}\left[\delta_k^B \cdot X_k^A\right]$ and the negative ones to $\mathbf{E}\left[\delta_k^A \cdot X_k^B\right]$. In the case of the minimal CVA $\mathbf{E}\left[\delta_k^B \cdot X_k^A\right]$ vanishes, meaning that for counterparty A in case of a default of counterparty B the exposure is zero, as the present

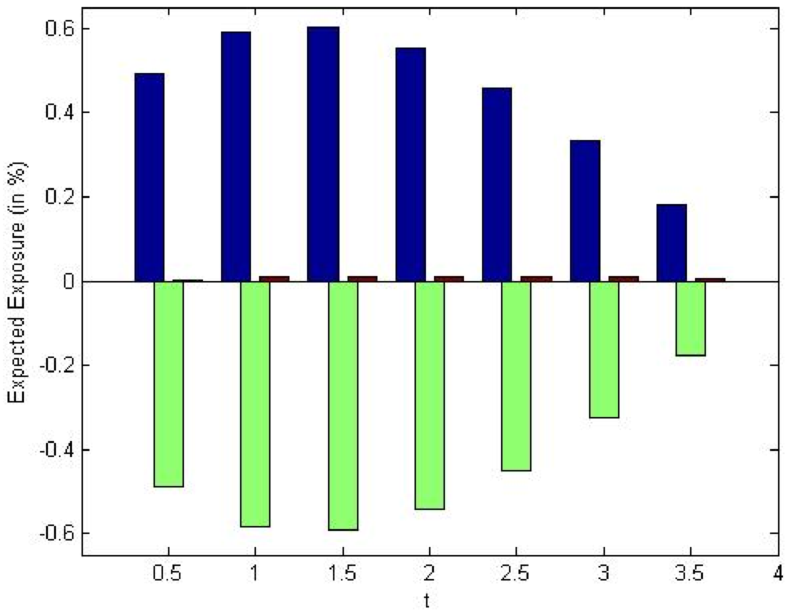

Figure 7: Expected exposures $\mathbf{E}\left[X_k^A\right]$, $\mathbf{E}\left[X_k^B\right]$ and $\mathbf{E}[P(0,t_k)\cdot V_A(t_k,T)]$.

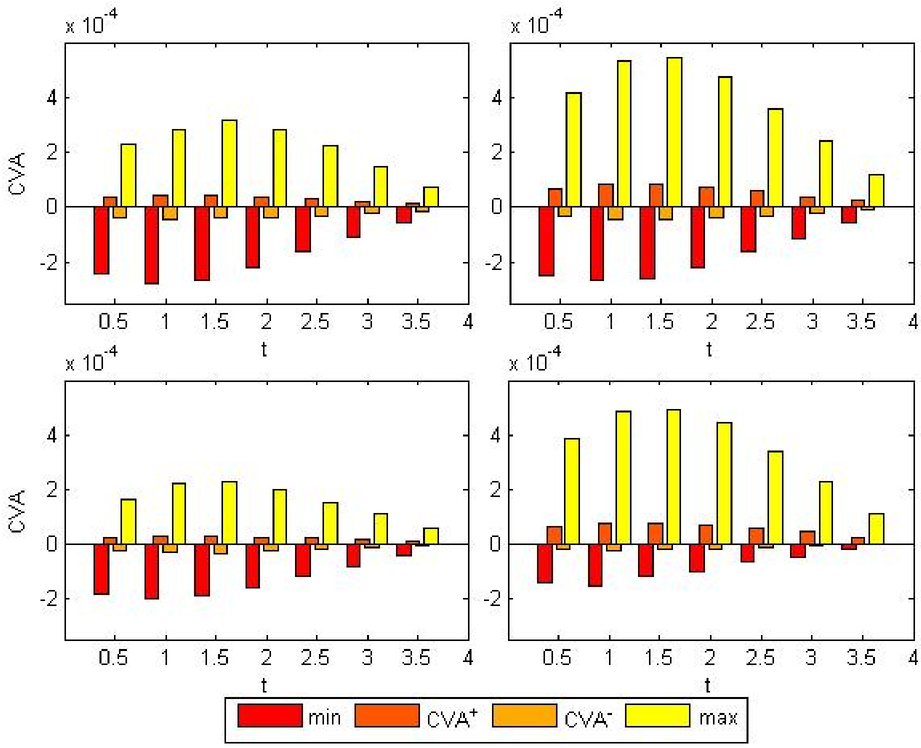

Figure 8: Minimal, maximal CVA_A, $\mathbf{E}\left[\delta_k^B\cdot X_k^A\,|\,\mathcal{G}_t\right]$ and $-\mathbf{E}\left[\delta_k^A\cdot X_k^B\,|\,\mathcal{G}_t\right]$

	k	1	2	3	4	5	6	7	8	$\sum$
Case 1	min	−2.41	−2.78	−2.67	−2.16	−1.61	−1.08	−0.55	0.00	−13.26
	$\mathbf{E}\left[\delta_k^B \cdot X_k^A \mid \mathcal{G}_t\right]$	0.37	0.41	0.41	0.40	0.32	0.22	0.12	0.00	2.24
	$\mathbf{E}\left[\delta_k^A \cdot X_k^B \mid \mathcal{G}_t\right]$	−0.37	−0.45	−0.40	−0.39	−0.32	−0.21	−0.12	0.00	−2.27
	max	2.28	2.85	3.16	2.80	2.21	1.48	0.73	0.00	15.51
Case 2	min	−1.82	−2.03	−1.90	−1.60	−1.19	−0.85	−0.42	0.00	−9.80
	$\mathbf{E}\left[\delta_k^B \cdot X_k^A \mid \mathcal{G}_t\right]$	0.23	0.30	0.30	0.24	0.22	0.18	0.09	0.00	1.56
	$\mathbf{E}\left[\delta_k^A \cdot X_k^B \mid \mathcal{G}_t\right]$	−0.24	−0.31	−0.34	−0.25	−0.20	−0.15	−0.09	0.00	−1.57
	max	1.66	2.21	2.26	2.01	1.54	1.10	0.56	0.00	11.34
Case 3	min	−2.45	−2.68	−2.60	−2.21	−1.58	−1.13	−0.54	0.00	−13.19
	$\mathbf{E}\left[\delta_k^B \cdot X_k^A \mid \mathcal{G}_t\right]$	0.68	0.86	0.86	0.74	0.61	0.39	0.23	0.00	4.38
	$\mathbf{E}\left[\delta_k^A \cdot X_k^B \mid \mathcal{G}_t\right]$	−0.35	−0.43	−0.42	−0.37	−0.30	−0.22	−0.11	0.00	−2.21
	max	4.18	5.36	5.42	4.72	3.60	2.41	1.18	0.00	26.88
Case 4	min	−1.44	−1.53	−1.19	−1	−0.67	−0.49	−0.21	0.00	−6.53
	$\mathbf{E}\left[\delta_k^B \cdot X_k^A \mid \mathcal{G}_t\right]$	0.63	0.76	0.76	0.69	0.60	0.43	0.20	0.00	4.08
	$\mathbf{E}\left[\delta_k^A \cdot X_k^B \mid \mathcal{G}_t\right]$	−0.18	−0.23	−0.19	−0.16	−0.12	−0.08	−0.03	0.00	−1.00
	max	3.83	4.88	4.89	4.46	3.40	2.29	1.11	0.00	24.87

Table 3: Minimal, maximal CVA, $\mathbf{E}\left[\delta_k^B \cdot X_k^A \mid \mathcal{G}_t\right]$ and $-\mathbf{E}\left[\delta_k^A \cdot X_k^B \mid \mathcal{G}_t\right]$ in bps

value of the swap at that time is negative from counterparty A's point of view. Contrarily, for the maximal CVA $\mathbf{E}\left[\delta_k^A \cdot X_k^B\right]$ is zero.

As main observation we note that there are large gaps between the lower and the independent CVA, as well as between the independent CVA and the upper bound. This means that wrong way exposure can have a significant impact on the bilateral CVA. Interestingly, this observation holds true for all four cases, of course, with different significance depending on the specific setup. Although it is clear that our analysis naturally shows more extreme gaps than any hybrid model, it has to be mentioned that these bounds are indeed tight. As observed in the last section, there exist some joint distributions – hence some corresponding model with appropriate parameterizations – which attain both bounds. Therefore, the impact of wrong-way risk may be much larger in the extreme case than expected. For example, Duffie and Huang [14] obtain factors around two for the increase in CVA by wrong-way risk, whereas the factors which could be observed here are more in line with the extreme examples of Cherubini [11]. We further note that the bilateral CVA does not even vanish for symmetric counterparties. The reason for this is that the exposure of a swap is not completely symmetric, as receiver is not equal to payer exposure, as observable from the exposures in Figure 7. Nevertheless, in case of independence of defaults and exposure, the CVA is usually rather small for interest rate swaps, even for asymmetric counterparties. However, it can be observed that asymmetry of the counterparties has a significant impact on the difference of the absolute values of the lower and the upper bound, as expected.

6 Conclusion and outlook

In this paper we have proposed a new model-free approach for the calculation of unilateral and bilateral counterparty valuation adjustment. This approach has the advantage that simulations of the uncertain default times on one hand and of the uncertain value of a transaction during her remaining life on the other hand can be completely separated, before they are linked via the easy solution of a standard linear program. Although this exposition is restricted to the case of two counterparties and does not include a random recovery, the model can easily be extended to more general cases. Further, as exposure is simulated separately from default, all risk mitigating components like CSAs or netting agreements can be easily included in a such a framework.

References

[1] M. S. Bazaraa, J. J. Jarvis, and H. D. Sherali. *Linear Programming and Network Flows.* Wiley & Sons, New Jersey, 4th edition, 2010.

[2] T. R. Bielecki and M. Rutkowski. *Credit Risk: Modeling, Valuation and Hedging.* Springer, Berlin, 2002.

[3] C. Blanchet-Scalliet and F. Patras. Counterparty risk valuation for CDS. working paper, 2008.

[4] D. Brigo and I. Bakkar. Accurate counterparty risk valuation for energy-commodities swaps. *Energy Risk*, March, 2009.

[5] D. Brigo and A. Capponi. Bilateral counterparty risk valuation with stochastic dynamical models and application to credit default swaps. working paper, 2008.

[6] D. Brigo and K. Chourdakis. Counterparty risk for credit default swaps: impact of spread volatility and default correlation. *International Journal of Theoretical and Applied Finance*, 2009. forthcoming.

[7] D. Brigo and M. Masetti. Risk-neutral pricing of counterparty risk. In M. Pykhtin, editor, *Counterparty Credit Risk Modeling.* Risk Books, London, 2005.

[8] D. Brigo and A. Pallavicini. Counterparty risk under correlation between default and interest rates. In J. Miller, D. Edelman, and J. Appleby, editors, *Numerical Methods for Finance.* Chapman Hall, 2007.

[9] D. Brigo, A. Pallavicini, and V. Papatheodorou. Bilateral counterparty risk valuation for interest-rate products: impact of volatilities and correlations. working paper, 2010.

[10] E. Canabarro and D. Duffie. Measuring and marking counterparty risk. In L. M. Tilman, editor, *Asset/Liability Management of Financial Institutions.* Euromoney Books, London, 2003.

[11] U. Cherubini. Pricing Swap Credit Risk with Copulas. EFMA, 2004 Basel Meeting Paper, 2004.

[12] I. A. Cooper and A. S. Mello. The Default Risk of Swaps. *Journal of Finance*, 46:597–620, 1991.

[13] S. Crépey, M. Jeanblanc, and B. Zargari. CDS with counterparty risk in a Markov chain copula model with joint defaults. working paper, 2009.

[14] D. Duffie and M. Huang. Swap rates and credit quality. *Journal of Finance*, 51:921–949, 1996.

[15] D. Duffie and K. J. Singleton. Modeling Term Structures of Defaultable Bonds. *The Review of Financial Studies Special*, 12(4):687–720, 1999.

[16] J. K. Gregory. Being two faced over counterparty risk. *Risk*, 22(2):86–90, 2009.

[17] J. K. Gregory. *Counterparty Credit Risk: The New Challenge for Global Financial Markets.* Wiley & Sons, 2009.

[18] J. K. Gregory. Gaining from your own default – counterparty credit risk and DVA. working paper, 2009.

[19] J. Hull and A. White. Valuing credit default swaps: modeling default correlations. working paper, 2000.

[20] D. Lando. On Cox processes and credit risky securities. *Review of Derivatives Research*, 2:99–120, 1998.

[21] S. Leung and Y. Kwok. Credit default swap valuation with counterparty risk. *The Kyoto Economic Review*, 74(1):25–45, 2005.

[22] A. Lipton and A. Sepp. Credit value adjustment for credit default swaps via the structural default model. *The Journal of Credit Risk*, 5(2):123–146, 2009.

[23] M. Pykhtin. *Counterparty Credit Risk Modelling: Risk Management Pricing and Regulation.* Risk Books, London, 2005.

[24] S. T. Rachev and L. Rüschendorf. *Mass Transportation Problems: Volume I: Theory.* Springer, New York, 1998.

[25] S. T. Rachev and L. Rüschendorf. *Mass Transportation Problems: Volume II: Applications.* Springer, New York, 1998.

[26] C. Redon. Wrong way risk modelling. *Risk Magazine*, April, 2006.

[27] E. Sorensen and T. Bollier. Pricing swap default risk. *Financial Analysts Journal*, 50(3):23–33, 1994.

[28] H. Stein and K. P. Lee. Counterparty valuation adjustments. In T Bielecki, D. Brigo, and F. Patras, editors, *The Handbook of Credit Derivatives.* Wiley & Sons, 2011.

[29] S. M. Turnbull. The Pricing Implications of Counterparty Risk for Non-Linear Credit Products. In M. Pykhtin, editor, *Counterparty Credit Risk Modeling.* Risk Books, London, 2005.

[30] M. Walker. Credit default swaps with counterparty risk: A calibrated markov model. *Journal of Credit Risk*, 2(2):31–49, 2006.