

Tight Semi-model-free Bounds on (Bilateral) CVA

Jördis Helmers, Jan-J. Rückmann and Ralf Werner

Abstract In the last decade, 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 decade which have highlighted the importance of counterparty default risk for uncollateralized derivatives. After a succinct introduction to the topic, it is demonstrated that standard models can be combined to derive semi-model-free tight lower and upper bounds on bilateral CVA (BCVA). It will be shown in detail how these bounds can be easily and efficiently calculated by the solution of two corresponding linear optimization problems.

Keywords Counterparty credit risk · CVA · Tight bounds · Mass transportation problem

1 Introduction

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 qualities of both counterparties—usually a client and an investment bank—need to be considered in the pricing of uncollateralized OTC derivatives.

J. Helmers
Finbridge GmbH & Co. KG, Louisenstr. 100, 61348 Bad Homburg, Germany
e-mail: joerdis.helmers@finbridge.de

J.-J. Rückmann
Department of Informatics, University of Bergen,
P.O. Box 7803, 5020 Bergen, Norway
e-mail: Jan-Joachim.Ruckmann@ii.uib.no

R. Werner (✉)
Institut für Mathematik, Universität Augsburg, Universitätsstr. 14,
86159 Augsburg, Germany
e-mail: ralf.werner@math.uni-augsburg.de

Over the past years, several authors have been investigating the pricing of derivatives based on a variety of models which take into account these default risks. Most of these results are covered by a variety of excellent books, for example Pykhtin [16], Gregory [12], or Brigo et al. [7] just to name a few. For a profound discussion on the pros and cons of unilateral versus bilateral counterparty risk let us refer to the two articles by Gregory [11, 13].

In the following exposition, we are concerned with the quantification of the smallest and largest BCVA which can be obtained by any given model with predetermined marginal laws. This takes considerations of Turnbull [21] much further, who first derived weak bounds on CVA for certain types of products. Our approach extends first ideas from Hull and White [15], where the hazard rate determining defaults is coupled to the exposure or other risk factors in either deterministic or stochastic way. Still, Hull and White rely on an explicit choice of the default model and on an explicit coupling. More related is the work by Rosen and Saunders et al. [8, 17], on which we prefer to comment later in Remark 8. As the most related work we note the paper by Cherubini [9] which provided the basis for this semi-model-free approach. There, 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 couples each forward swap par rate with each other and the default time—which is in gist similar to Hull and White [15]. From there the final step to our approach is to observe that the most general approach directly links the whole stochastic evolution of the exposure with both random default times. We will illustrate in the following that these couplings can be readily derived by linear programming. For this purpose the BCVA will be decomposed into three main components: the first component is represented by the loss process, 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 amount¹ at time of default. This approach takes further early considerations of Haase and Werner [14], where comparable results were obtained from the point of view of generalized stopping problems.

In a very recent working paper by Scherer and Schulz [18], the above idea was analyzed in more detail. It was shown that the computational complexity of the problem is the same, no matter if only marginal distributions of defaults or the joint distribution of defaults are known.

After submission of this paper we became aware of related results by Glasserman and Yang, see [10]. Although the main idea of their exposition is similar in gist, Glasserman and Yang focus on the unilateral CVA instead of bilateral CVA. Besides an analysis of the convergence of finite samples to the continuous setup, their exposition is mainly focused on the penalization of deviation from some *base distribution*. In contrast, our focus is on bilateral CVA, with special attention to numerical solution and to the case that payoffs also depend on the credit quality.

¹In accordance with the *full two-way payment* rule under ISDA master contracts, see e.g. Bielecki and Rutkowski [2] (Sect. 14.4.4), we assume that the close-out value is determined by the then prevailing risk-free present value.

In summary, this exposition makes the following main contributions:

- 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. Unlike Cherubini, who only considered the very specific case of an interest rate swap, all kinds of derivatives (interest rate, FX, commodity, and even credit derivatives) are covered in a unified way—even if the payoff, and thus the present value of the derivative, is explicitly depending on the credit quality of any of the two counterparties.
- Second, by generalizing Cherubini’s approach, upper and lower bounds on unilateral and bilateral counterparty value adjustments are derived. It will be demonstrated that these bounds can be efficiently obtained by the solution of linear optimization problems, more specifically, by the solution of balanced transportation problems. In contrast to the approaches of Turnbull [21] or Cherubini [9], both the upper and lower bound derived here are *tight bounds*, i.e. there exists some stochastic model which is consistent with all given market prices in which these bounds are attained.

The rest of the paper is organized as follows. In Sect. 2 a succinct introduction to bilateral counterparty risk is given, before the decomposition of the BCVA into its building blocks is carried out in Sect. 3. In Sect. 4 the two main approaches for the calculation of counterparty valuation adjustments are briefly reviewed. Finally, the tight bounds on CVA are derived in Sect. 5, before the paper concludes.

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 for a given risk-free numéraire process $N_t > 0$, 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 modeling 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:

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

For simplicity of presentation, we assume in the following that $\mathbf{Q}[\tau_A = T] = \mathbf{Q}[\tau_B = T] = \mathbf{Q}[\tau_A = \tau_B] = 0$. Under this assumption these sets² yield a decomposition of one, i.e. it holds

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

In the following, let us consider a transaction consisting 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 $T_j, j = 1, \dots, m_A$. Taking into account default risk of both counterparties, the quantification of the bilateral CVA is summarized in the following well-known theorem, which in essence goes back to Sorensen and Bollier [19].

Theorem 1 *Conditional on the event $\{t < \min(\tau_A, \tau_B)\}$, i.e. no default has occurred until time t , the value $V_A^D(t, T)$ of the transaction under consideration of bilateral counterparty risk at time t is given by*

$$V_A^D(t, T) = V_A(t, T) - CVA_A(t, T) = -(V_B(t, T) - CVA_B(t, T)) = -V_B^D(t, T)$$

where the risk-free present value of the transaction is given as

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

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

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

Here L_t^i denotes the random loss (between 0 and 1) of counterparty i at time t .

Proof A proof of Theorem 1 can be found in Bielecki and Rutkowski [2], Formula (14.25) or Brigo and Capponi [4], Proposition 2.1 and Appendix A, respectively.

Based on Theorem 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 briefly reviewed in Sect. 4.

²We note that Brigo et al. (in [4, 6]) use different sets to order the default times, which are in essence reducible to the above three events.

3 The Main Building Blocks of CVA

Subsequently, let us assume that the default times τ_i with $i \in \{A, B\}$ can only take a finite number of values $\{\bar{t}_1, \dots, \bar{t}_K\}$ in the interval $]0, T[$. For continuous time models this assumption can be justified by the *default bucketing* approach, which can, for example, be found in Brigo and Chourdakis [5], if K is chosen sufficiently large. To be able to separate the default dynamics from the market value dynamics, let us introduce the auxiliary time s , $s \in [t, T]$ and the discounted market value

$$\tilde{V}_i^+(t, s, T) := \frac{N_t}{N_s} \cdot \max(0, V_i(s, T)).$$

Then we can rewrite Eq. (1) as:

$$\begin{aligned} CVA_A(t, T) = & \mathbb{E} \left[\sum_{k=1}^K L_{\bar{t}_k}^B \cdot \mathbf{1}_{D_B} \cdot \mathbf{1}_{\bar{t}_k}(\tau_B) \cdot \tilde{V}_A^+(t, \bar{t}_k, T) \mid \mathcal{G}_t \right] \\ & - \mathbb{E} \left[\sum_{k=1}^K L_{\bar{t}_k}^A \cdot \mathbf{1}_{D_A} \cdot \mathbf{1}_{\bar{t}_k}(\tau_A) \cdot \tilde{V}_B^+(t, \bar{t}_k, T) \mid \mathcal{G}_t \right]. \end{aligned} \quad (2)$$

Here, $\mathbf{1}_M$ is the indicator function of the set M ; if $M = \{m\}$ we simply write $\mathbf{1}_m$ instead. Now, collecting all terms relating to the default in the default indicator process δ ,

$$\delta_k^i := \mathbf{1}_{D_i} \cdot \mathbf{1}_{\bar{t}_k}(\tau_i),$$

we can rewrite the BCVA in a more compact manner as

$$\begin{aligned} CVA_A(t, T) = & \mathbb{E} \left[\sum_{k=1}^K L_{\bar{t}_k}^B \cdot \delta_k^B \cdot \tilde{V}_A^+(t, \bar{t}_k, T) \mid \mathcal{G}_t \right] \\ & - \mathbb{E} \left[\sum_{k=1}^K L_{\bar{t}_k}^A \cdot \delta_k^A \cdot \tilde{V}_B^+(t, \bar{t}_k, T) \mid \mathcal{G}_t \right]. \end{aligned} \quad (3)$$

From Eq. (3) we immediately see that the BCVA at time t is composed of six discrete time³ 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* $\tilde{V}_A^+(t, s, T)$ and $\tilde{V}_B^+(t, s, T)$.

In this way, we are able to separate the default dynamics δ from the loss process L and the exposure process \tilde{V} . From this decomposition, it becomes obvious that the BCVA is completely determined by the joint distribution of these six processes.

³In the following, we replace the time index \bar{t}_k with k for notational convenience.

Remark 1 We note that in general it is even sufficient to model four processes (loss dynamics and market value dynamics) plus a two-dimensional random variable (τ_A, τ_B) . However, in the case of finitely many default times, it is more convenient to work with the default indicator process instead.

Remark 2 For simplicity of the subsequent exposition, we assume that the loss process is actually constant and equals 1: $L_t^i = l^i = 1$. The theory of the remainder of this exposition is not affected by this simplifying assumption, with one notable exception: the resulting two-dimensional transportation problems will become a multi-dimensional transportation problem which renders its numerical solution more complex, but still feasible.

Remark 3 As we have noted, the default indicator process can only take a finite number of values in the bucketing approach. More exactly, it holds that the joint (i.e. two-dimensional) default indicator process $\delta = (\delta_k)_{k=1, \dots, K} \in \mathbb{R}^{2 \times K}$, defined by

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

takes only values in the finite set

$$\mathcal{Y} := \left\{ \gamma \in \mathbb{R}^{2 \times K} \mid \gamma_{i,k} \in \{0, 1\}, \sum_{i,k} \gamma_{i,k} \leq 1 \right\}$$

which has exactly $2K + 1$ elements. Therefore, the discrete time default indicator process is also a process with a finite state space.

Let us further introduce the joint exposure process in analogy to the above,

$$X_k := \begin{pmatrix} \tilde{V}_A^+(t, \tilde{t}_k, T) \\ \tilde{V}_B^+(t, \tilde{t}_k, T) \end{pmatrix}, \quad k = 1, \dots, K.$$

Then it holds

$$CVA_A(t, T) = \sum_{k=1}^K (\mathbb{E} [\delta_k^B \cdot X_k^A \mid \mathcal{G}_t] - \mathbb{E} [\delta_k^A \cdot X_k^B \mid \mathcal{G}_t]). \quad (4)$$

To avoid technical considerations for brevity of presentation, we prefer to work with discrete processes (i.e. discrete state space) in discrete time. Thus, it may be necessary to discretize the state space of the remaining discounted exposure process. In general, there exist (at least) two different approaches how a suitable discrete state space version of the process X could be obtained:

- In the first approach—completely similar to the default bucketing approach—the state space $\mathbb{R}^{2 \times K}$ for the joint exposure process X is divided into N disjoint

components. Then X is replaced by some representative value on this component (usually an average value) 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).

- From a computational and practical point of view, a much more convenient approach relies on Monte Carlo simulation: N different scenarios (i.e. realizations) of the process X are used instead of the original process. Each realization is assumed to have probability $1/N$.

For both approaches it is known that they converge at least⁴ in distribution to the original process, which is sufficient for our purposes. For more details on the convergence, let us refer to the recent working paper by Glasserman and Yang [10].

4 Models for Counterparty Risk

In the last decade two main approaches have emerged in the literature how to model the individual, resp. joint distribution of the processes δ and X :

- The most popular approach is based on the rather strong assumption of independence between exposure and default. Based on this independence assumption, only individual models for δ and X need to be specified for the CVA calculation. This kind of independence assumption is quite standard in the market, see for example the Bloomberg CVA function (for more details on the Bloomberg model let us refer to Stein and Lee [20]).
- Alternatively, and more recently, a more general approach is based on a joint model (also called *hybrid model*) for the building blocks δ and X of the CVA calculation, see Sect. 4.3.

4.1 Independence of CVA Components

Let us assume that the exposure process X is independent of the default process δ . Then the expectation inside the summation can be split into two parts:

$$\sum_{k=1}^K \mathbb{E} [\delta_k^B \cdot X_k^A | \mathcal{G}_t] = \sum_{k=1}^K \mathbb{E} [\delta_k^B | \mathcal{G}_t] \cdot \mathbb{E} [X_k^A | \mathcal{G}_t]. \quad (5)$$

⁴The Monte Carlo approach converges in distribution due to the Theorem of Glivenko–Cantelli. For state space discretization, if for example conditional expectations are used on each bucket, then convergence is in fact almost surely and in L^1 due Lévy's 0–1 law.

It is well known that the expected value

$$\mathbb{E} [X_k^A | \mathcal{G}_t] = \mathbb{E} [\tilde{V}_A^+(t, \bar{t}_k, T) | \mathcal{G}_t] = \mathbb{E} \left[\frac{N_t}{N_{\bar{t}_k}} \cdot \max(V_A(t, \bar{t}_k, T), 0) \middle| \mathcal{F}_t \right] \quad (6)$$

matches exactly the price of a call option on the basis transaction at time t with strike 0 and exercise time \bar{t}_k . The CVA equation can hence be rewritten as

$$CVA_A(t, T) = \sum_{k=1}^K (\mathbb{E} [\delta_k^B | \mathcal{G}_t] \cdot \mathbb{E} [X_k^A | \mathcal{G}_t] - \mathbb{E} [\delta_k^A | \mathcal{G}_t] \cdot \mathbb{E} [X_k^B | \mathcal{G}_t]), \quad (7)$$

and thus the BCVA can be calculated without any further problems as the corresponding default probabilities⁵ $\mathbb{E} [\delta_k^B | \mathcal{G}_t] = \mathbf{Q} [\tau_B \in \Delta_k, \tau_B \leq \tau_A | \mathcal{G}_t]$ can be easily computed from any given credit risk model: in order to calculate the probability $\mathbf{Q} [\tau_B \in \Delta_k, \tau_B \leq \tau_A | \mathcal{G}_t]$, the default times τ_A and τ_B together with 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 Rutkowski [2], Part III.

Remark 4 It has to be noted that a model with deterministic default intensities plus a suitable copula is sufficient for the arbitrary specification of the joint distribution of default times. Stochastic intensities do not add any value in this context. This is true as long as the default risk-free discounted present value is independent of the credit quality of each counterparty. This means that the payoff itself is not allowed to be linked explicitly to the credit quality of any counterparty.

Remark 5 Let us point out that the intensity model is just one specific example how default times could be modeled. The big advantage of our approach is that any arbitrary credit risk model can be used instead, as only the distribution of the default indicator δ finally matters. In case only marginal default models are available, we can also take into account the remaining unknown dependence between the default times, however, at the price of a higher dimensional transportation problem.

4.2 Modeling Options on the Basis Transaction

Since it could be observed in Eq. (6) 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,

⁵With $\Delta_k :=]\bar{t}_{k-1}, \bar{t}_k]$ if the default bucketing approach has been used, otherwise $\Delta_k := \{\bar{t}_k\}$.

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 simpler 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.

4.3 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. Instances of such models can for example be found in Brigo and Capponi [4] for the case of a credit default swap, or Brigo et al. [6] for interest rate derivatives. In Brigo et al. [6], an integrated framework is introduced, where a two-factor Gaussian interest-rate model is set up for a variety of interest rate derivatives⁶ 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 summary, the amount of wrong-way risk which can be modeled within such a framework strongly depends on the model choice. 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 et al. [6]. Figure 5 in Scherer and Schulz [18] provides an overview of potential CVA values for different models which illustrates that models can differ quite significantly.

5 Tight Bounds on CVA

From the previous section it becomes obvious that hybrid models yield different CVAs depending on the (model and parameter implied) degree of dependence between default and exposure. However, it remains unclear how large the impact of this dependence can be. In other words: *Is it possible to quantify, how small or large the CVA can get for any model, given that the marginal distributions for expo-*

⁶Although this modeling 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 of financial crisis in 2008, when these effects could not be reproduced by such a model.

sure and default are already given? In the following, we want to address this question based on our initially given decomposition of the CVA in building blocks.

As mentioned in Sect. 4.2, 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 that also the distribution of the default indicator process δ can be assumed to be given by the market. 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. Further, we can even consider the case that the dependence structure between different market risk factors is not known but remains uncertain. However, all these generalizations come at the price that the resulting two-dimensional transportation problem will become multi-dimensional.

For the above reasons, we argue that the following approach is indeed *semi-model-free* in the sense that no model needs to be specified which links the default indicator process with the discounted exposure processes.

5.1 Tight Bounds on CVA by Mass Transportation

Let us reconsider Eq. (4) and let us highlight the dependence of the BCVA on the measure \mathbf{P} .

$$\text{CVA}_A^{\mathbf{P}}(t, T) = \sum_{k=1}^K (\mathbb{E}_{\mathbf{P}} [\delta_k^B \cdot X_k^A | \mathcal{G}_t] - \mathbb{E}_{\mathbf{P}} [\delta_k^A \cdot X_k^B | \mathcal{G}_t]).$$

With some abuse of notation, the measure \mathbf{P} denotes the joint distribution of the default process δ and the exposure process X . Since both processes have finite support, \mathbf{P} can be represented as a $(2K + 1) \times N$ matrix with entries in $[0, 1]$. We note that the marginals of \mathbf{P} , i.e. the distributions of δ and X (denoted by the probability vectors $\mathbf{p}^{(X)} \in \mathbb{R}^N$ and $\mathbf{p}^{(\delta)} \in \mathbb{R}^{2K+1}$) are already predetermined from the market. Therefore, \mathbf{P} has to satisfy

$$\mathbf{1}^\top \mathbf{P} = \mathbf{p}^{(X)}, \quad \text{and} \quad \mathbf{P} \mathbf{1} = \mathbf{p}^{(\delta)}.$$

Remark 6 In case of independence between δ and X , \mathbf{P} is given by the product distribution of δ and X , whereas in hybrid models the joint distribution \mathbf{P} is determined by the specification and parametrization of the hybrid model. In the independent case, \mathbf{P} is hence given by the dyadic product

$$\mathbf{P} = \mathbf{p}^{(\delta)} \mathbf{p}^{(X)\top}.$$

Obviously, the smallest and largest CVA which can be obtained by any \mathbf{P} which is consistent with the given marginals, is given by

$$\begin{aligned} CVA_A^l(t, T) &:= \min_{\mathbf{P} \in \mathcal{P}} CVA_A^{\mathbf{P}}(t, T), \\ CVA_A^u(t, T) &:= \max_{\mathbf{P} \in \mathcal{P}} CVA_A^{\mathbf{P}}(t, T), \end{aligned}$$

where

$$\mathcal{P} := \{\mathbf{P} \in [0, 1]^{(2K+1) \times N} \mid \mathbf{1}^\top \mathbf{P} = \mathbf{p}^{(X)}, \mathbf{P}\mathbf{1} = \mathbf{p}^{(\delta)}\}.$$

It can be easily noted that the set \mathcal{P} is a convex polytope. Thus, the computation of $CVA_A^l(t, T)$ and $CVA_A^u(t, T)$ essentially requires the solution of a linear program, as the objective functions are linear in \mathbf{P} .

Remark 7 The structure of the above LPs 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 et al. [1], Chap. 10, for more details. There exist several very efficient algorithms for the numerical solution of such transportation problems, see also Bazaraa et al. [1], Chaps. 10, 11 and 12.

Let us summarize our results in the following theorem:

Theorem 2 *Under the given prerequisites, it holds:*

1. $CVA_A^l(t, T) \leq CVA_A(t, T) \leq CVA_A^u(t, T)$.
2. *These bounds are tight, i.e. they represent the lowest and the highest CVA which can be obtained by any (hybrid) model which is consistent with the market data and there exists at least one model which reaches these bounds.*

The tightness of our bounds is in contrast to Turnbull [21], where only weak 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.

Remark 8 We note that a related approach of coupling default and exposure via copulas was presented by Rosen and Saunders [17] and Crepedes et al. [8]. However, their approach differs from ours in some significant aspects. First, exposure scenarios are sorted by a single number (e.g. effective exposure) to be able to couple exposure scenarios with risk factors of defaults by copulas. Second, risk factors of some credit risk model are employed instead of working with the default indicator directly. Third, their approach is restricted to the real-world setting and does not consider restrictions on the marginal distributions in the coupling process, which is e.g. necessary if stochastic credit spreads should be considered.

5.2 An Alternative Formulation as Assignment Problem

For the above setup we have assumed that the probabilities 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 obtained. 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 $\mathbf{p}_j^{(\delta)} = \mathbf{p}_i^{(X)} = 1/N$. If we apply the same arguments as above we obtain again a transportation problem, however, with probabilities $1/N$ each. If we have a closer look at this problem, 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. This means that this problem eventually belongs to the class of assignment problems, for which very efficient algorithms are available, cf. Bazaraa et al. [1]. Nevertheless, 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 $2K + 1 \ll N$ (i.e. time discretization is usually much coarser than exposure discretization). However, if stochastic credit spreads have to be considered, they have to be part of the default simulation and thus assignment problems (with additional linear constraints to guarantee consistency of exposure paths and spreads) become unavoidable.

6 Example

6.1 Setup

To illustrate these semi-model-free CVA bounds let us give a brief example. 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 $t = 0$. The time interval $]0, 4[$ is split up into $K = 8$ disjoint time intervals each covering half a year. For simplicity, the loss process is again assumed to be 1.

6.1.1 Counterparty’s Default Modeling

To model the defaults we have chosen the well-known copula approach with constant intensities using the Gaussian copula. 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 counterparties’ default intensities are assumed to be deterministic. We will distinguish between symmetric counterparties with identical default intensities and asymmetric counterparties. Thus, four different settings

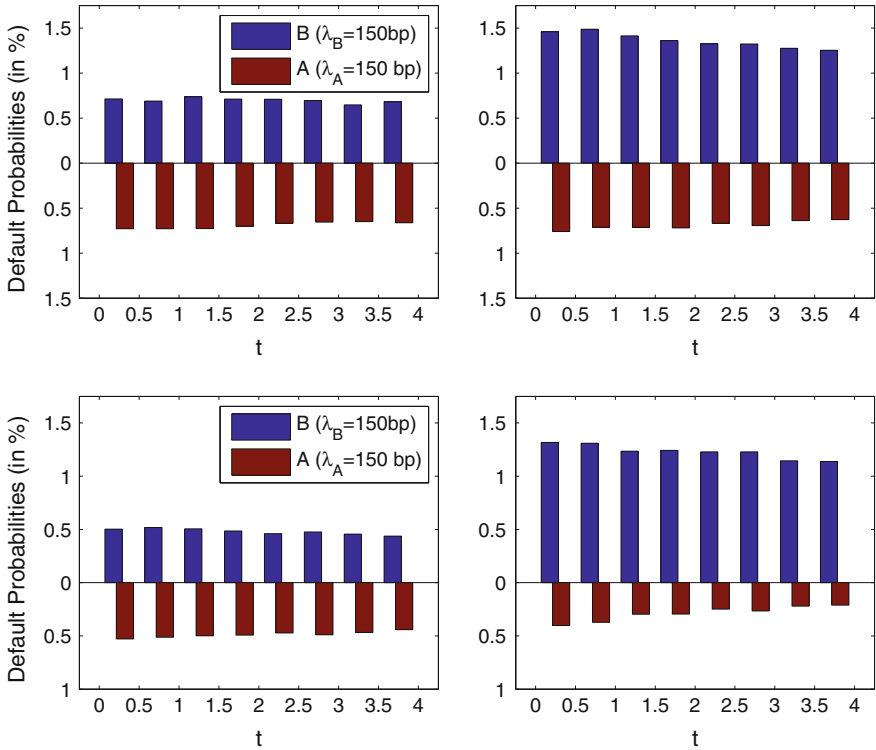


Fig. 1 Probabilities $\mathbb{E}_{\mathbf{Q}}[\delta_k^i]$ in % for Case 1 to Case 4

result: Fig. 1 shows the probabilities $\mathbf{Q}[\delta_k^i = 1] = \mathbb{E}_{\mathbf{Q}}[\delta_k^i]$ in each of the four cases under the risk-neutral measure \mathbf{Q} implied from the market. To be in line with the following figures, the probabilities for a default of counterparty B in Δ_k , i.e. $\mathbb{E}_{\mathbf{Q}}[\delta_k^B]$, correspond to the positive bars and defaults of counterparty A to the negative bars. The left plots 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 $\rho = 0.9$.

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

Fig. 2 Expected exposures $\mathbb{E}_{\mathbf{Q}}[X_k^A]$, $\mathbb{E}_{\mathbf{Q}}[X_k^B]$ and $\mathbb{E}_{\mathbf{Q}}[\tilde{V}_A(t_k, T)]$

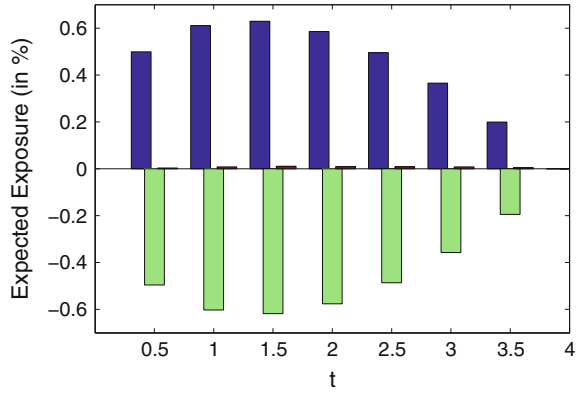


Table 1 $\mathbb{E}_{\mathbf{Q}}[X_k^A]$ and $\mathbb{E}_{\mathbf{Q}}[X_k^B]$ in basis points

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

6.1.2 Counterparty Exposure Modeling

As already mentioned, a simple CIR model is applied for the valuation of the payer swap. Since 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. In the CIR model, the short rate r_t follows the stochastic differential equation

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

where $(W_t)_{t \geq 0}$ denotes a standard Brownian motion. Instead of calibrating the parameters to market data (yield curve plus selected swaption prices) on one specific day, we have set the parameters in the following way

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

to obtain an interest rate market which is typical for the last years. Considering now the discounted exposure of each counterparty within the discrete time framework of our example, we can easily compute $\mathbb{E}_{\mathbf{Q}}[X_k^i]$ as the average of all generated scenarios from a Monte Carlo simulation. Figure 2 illustrates the results of a simulation, which are also given in Table 1. Positive bars correspond to $\mathbb{E}_{\mathbf{Q}}[X_k^A]$, negative bars to $\mathbb{E}_{\mathbf{Q}}[X_k^B]$, and the small bars correspond to $\mathbb{E}_{\mathbf{Q}}[\tilde{V}_A(t_k, T)]$. Since payer and receiver swap are not completely symmetric instruments, there remains a residual expectation, as can be observed from Fig. 2.

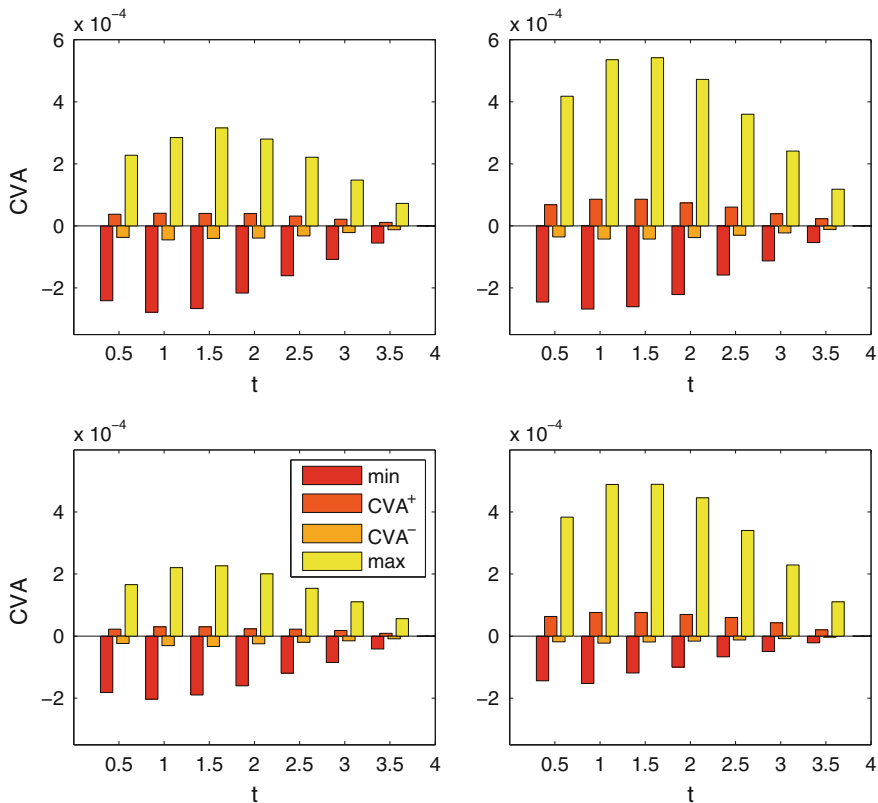


Fig. 3 Minimal and maximal CVA_A , $\mathbb{E}_{\mathbf{Q}_i} [\delta_k^B \cdot X_k^A | \mathcal{G}_t]$ and $-\mathbb{E}_{\mathbf{Q}_i} [\delta_k^A \cdot X_k^B | \mathcal{G}_t]$ in bps

6.2 Results

In case of independence between default and exposure, the bilateral CVA is easily obtained by multiplying the default probabilities (as shown in Fig. 1) with the corresponding exposures (as shown in Fig. 2) and summation. Besides the independent CVA^i , the minimal and maximal CVA^l and CVA^u have been calculated as well.

The results of these calculations are illustrated in Fig. 3 and Table 2 for each time interval Δ_k . Analogously to Fig. 1 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 $\mathbb{E}_{\mathbf{Q}_i} [\delta_k^B \cdot X_k^A]$ and the negative ones to $\mathbb{E}_{\mathbf{Q}_i} [\delta_k^A \cdot X_k^B]$. In the case of the minimal CVA, $\mathbb{E}_{\mathbf{Q}_i} [\delta_k^B \cdot X_k^A]$ vanishes, meaning that for counterparty A in case of a default of counterparty B the exposure is zero, as the present value of the swap at that time is negative from counterparty A's point of view. Contrarily, for the maximal CVA, $\mathbb{E}_{\mathbf{Q}_u} [\delta_k^A \cdot X_k^B]$ is zero. Here, \mathbf{Q}_u , \mathbf{Q}_l , and \mathbf{Q}_i denote the optimal measures for the maximal, the minimal, and the independent CVA, respectively. As expected there

Table 2 Minimal and maximal CVA, $\mathbb{E}_{\mathbf{Q}_i} [\delta_k^B \cdot X_k^A | \mathcal{G}_i]$ and $-\mathbb{E}_{\mathbf{Q}_i} [\delta_k^A \cdot X_k^B | \mathcal{G}_i]$ in bps

	k	1	2	3	4	5	6	7	8	Σ
Case 1	min	-2.41	-2.78	-2.67	-2.16	-1.61	-1.08	-0.55	0.00	-13.26
	$\mathbb{E}_{\mathbf{Q}_i} [\delta_k^B \cdot X_k^A \mathcal{G}_i]$	0.37	0.41	0.41	0.40	0.32	0.22	0.12	0.00	2.24
	$-\mathbb{E}_{\mathbf{Q}_i} [\delta_k^A \cdot X_k^B \mathcal{G}_i]$	-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
	$\mathbb{E}_{\mathbf{Q}_i} [\delta_k^B \cdot X_k^A \mathcal{G}_i]$	0.23	0.30	0.30	0.24	0.22	0.18	0.09	0.00	1.56
	$-\mathbb{E}_{\mathbf{Q}_i} [\delta_k^A \cdot X_k^B \mathcal{G}_i]$	-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
	$\mathbb{E}_{\mathbf{Q}_i} [\delta_k^B \cdot X_k^A \mathcal{G}_i]$	0.68	0.86	0.86	0.74	0.61	0.39	0.23	0.00	4.38
	$-\mathbb{E}_{\mathbf{Q}_i} [\delta_k^A \cdot X_k^B \mathcal{G}_i]$	-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
	$\mathbb{E}_{\mathbf{Q}_i} [\delta_k^B \cdot X_k^A \mathcal{G}_i]$	0.63	0.76	0.76	0.69	0.60	0.43	0.20	0.00	4.08
	$-\mathbb{E}_{\mathbf{Q}_i} [\delta_k^A \cdot X_k^B \mathcal{G}_i]$	-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 Computation times for the two-dimensional transportation problem

K	10	20	20	20
N	1024	1024	2048	4096
Time in seconds	0.2	0.5	1.5	6

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 risk (i.e. higher exposure comes with higher default rates) 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.

6.3 Computation Time, Choice of Algorithm, and Impact of Assumptions

Theoretically, the computation of the bounds boils down to the solution of a linear programming problem. From this it can be expected that state-of-the-art solvers like CPLEX or Gurobi will yield the optimal solution within reasonable computation time. Using CPLEX, we have obtained the following computation times on a standard workstation (Table 3).

It can be observed that the problem can be solved for reasonable discretization levels within decent time. Rather similar computation times have been obtained with an individual implementation of the standard network simplex based on Fibonacci heaps. However, for larger sizes, the performance of standard solvers begins to deteriorate. To dampen the explosion of computation time, we have resorted to a special purpose solver for min cost network flows (which are a general case of the transportation problem) for highly asymmetric problems, as in our case $2K + 1 \ll N$. Based on Brenner's min cost flow algorithm, see Brenner [3], we could still solve problems with $K = 40$ and $N = 8192$ beneath a minute.

If one has to resort to the assignment formulation (to consider credit spreads accordingly), computation times increase due to the fact that now assignment problems have to be solved. Here, a factor 100 compared to the above computation times cannot be avoided.

If the coupling of the two default times is left flexible, the problem becomes a transportation problem with three margins, i.e. of size $K + 1 \times K + 1 \times N$. For these types of problems, no special purpose solver is available and one has to resort to CPLEX. Scherer and Schultz [18] have exploited the structure of this three-dimensional transportation problem to reduce computational complexity. They were able to reduce the problem to a standard two-dimensional transportation problem, hence rendering the computation of bounds similarly easy, no matter if default times are already coupled or not.

7 Conclusion and Outlook

In this paper we have shown how tight bounds on unilateral and bilateral counterparty valuation adjustment can be derived by a linear programming approach. This approach has the advantage that simulations of the uncertain loss, of the default times and of the uncertain value of a transaction during her remaining life can be completely separated. Although we have restricted the exposition to the case of two counterparties and one derivative transaction, the model can easily be extended to more counterparties and a whole netting node of trades. Further, as exposure is simulated separately from default, all risk-mitigating components like CSAs, rating triggers, and netting agreements can be easily included in a such a framework.

Interesting open questions for future research include the analogous treatment in continuous time, which requires much more technically involved arguments. Further, this approach yields a new motivation to consider efficient algorithms for transportation or assignment problems with more than two marginals, which did not yet get much attention so far.

Acknowledgements The KPMG Center of Excellence in Risk Management is acknowledged for organizing the conference “Challenges in Derivatives Markets - Fixed Income Modeling, Valuation Adjustments, Risk Management, and Regulation”.

Open Access This chapter is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, duplication, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, a link is provided to the Creative Commons license and any changes made are indicated.

The images or other third party material in this chapter are included in the work’s Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in the work’s Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt or reproduce the material.

References

1. Bazaraa, M.S., Jarvis, J.J., Sherali, H.D.: *Linear Programming and Network Flows*, 4th edn. Wiley, New Jersey (2010)
2. Bielecki, T.R., Rutkowski, M.: *Credit Risk: Modeling, Valuation and Hedging*. Springer, Berlin (2002)
3. Brenner, U.: A faster polynomial algorithm for the unbalanced hitchcock transportation problem. *Oper. Res. Lett.* **36**(4), 408–413 (2008)
4. Brigo, D., Capponi, A.: Bilateral counterparty risk valuation with stochastic dynamical models and application to credit default swaps. Working paper (2009)
5. Brigo, D., Chourdakis, K.: Counterparty risk for credit default swaps: impact of spread volatility and default correlation. *Int. J. Theor. Appl. Financ.* **12**(7), 1007–1026 (2009)
6. Brigo, D., Pallavicini, A., Papatheodorou, V.: Arbitrage-free valuation of bilateral counterparty risk for interest-rate products: impact of volatilities and correlations. *Int. J. Theor. Appl. Financ.* **14**(6), 773–802 (2011)

7. Brigo, D., Morini, M., Pallavicini, A.: Counterparty Credit Risk, Collateral and Funding: With Pricing Cases for all Asset Classes. Wiley, New Jersey (2013)
8. Cespedes, J.C.G., Herrero, J.A., Rosen, D., Saunders, D.: Effective modeling of wrong way risk, counterparty credit risk capital and alpha in Basel II. *J. Risk Model Valid.* **4**(1), 71–98 (2010)
9. Cherubini, U.: Pricing swap credit risk with copulas. EFMA, 2004 Basel Meeting Paper (2004)
10. Glasserman, P., Yang, L.: Bounding wrong-way risk in CVA calculation. Columbia Business School Research Paper No. 15-59. Available at SSRN: <http://ssrn.com/abstract=2607649> (2015)
11. Gregory, J.K.: Being two faced over counterparty risk. *Risk* **22**(2), 86–90 (2009)
12. Gregory, J.K.: Counterparty Credit Risk: The New Challenge for Global Financial Markets. Wiley, New Jersey (2009)
13. Gregory, J.K.: Gaining from your own default—counterparty credit risk and DVA. Working paper (2009)
14. Haase, J., Werner, R.: A generalized stopping problem with application in counterparty risk modelling. In: Piunovsky, A.B. (ed.) *Modern Trends in Controlled Stochastic Processes: Theory and Applications*, pp. 196–215. Luniver, UK (2010)
15. Hull, J., White, A.: CVA and wrong-way risk. *Financ. Anal. J.* **68**(5) (2012)
16. Pykhtin, M.: Counterparty Credit Risk Modelling: Risk Management Pricing and Regulation. Risk Books, London (2005)
17. Rosen, D., Saunders, D.: Computing and stress testing counterparty credit risk capital. In: Canabarro, E. (ed.) *Counterparty Credit Risk*, pp. 245–292. Risk Books, London (2009)
18. Scherer, M., Schulz, T.: Extremal dependence for bilateral credit valuation adjustments. Working paper (2015)
19. Sorensen, E., Bollier, T.: Pricing swap default risk. *Financ. Anal. J.* **50**(3), 23–33 (1994)
20. Stein, H., Lee, K.P.: Counterparty valuation adjustments. In: Bielecki, T., Brigo, D., Patras, F. (eds.) *The Handbook of Credit Derivatives*. Wiley, New Jersey (2011)
21. Turnbull, S.M.: The pricing implications of counterparty risk for non-linear credit products. In: Pykhtin, M. (ed.) *Counterparty Credit Risk Modeling*. Risk Books, London (2005)