Rethinking Margin Period of Risk *†

Leif Andersen
Bank of America Merrill Lynch

Michael Pykhtin
Federal Reserve Board

Alexander Sokol
CompatibL

January 21, 2016

Abstract

We describe a new framework for collateralized exposure modelling under an ISDA Master Agreement with a Credit Support Annex. The proposed model captures legal and operational aspects of default in considerably greater detail than models currently used by most practitioners, while remaining fully tractable and computationally feasible. Specifically, it considers the remedies and suspension rights available within these legal agreements; the firm's policies in availing itself of these rights; and the typical time it takes to exercise them in practice. The inclusion of these effects is shown to produce significantly higher credit exposure for representative portfolios compared to the currently used models. The increase is especially pronounced when dynamic initial margin is also present.

1 Introduction

In modelling the exposure of collateralized positions, it is well recognized that credit default cannot be treated as a one-time event. Rather, the entire sequence of events leading up to and following the default must be considered, from the last successful margin call in advance of the eventual default to the time when the amount of loss becomes known (in industry parlance, "crystallized"). These events unfold over a period of time called the *margin period of risk* (MPR).

To properly identify exposures that arise during the MPR, a detailed understanding of contractual obligations is essential. In this paper, we focus on collateralized exposure under bilateral trading relationships governed by the *ISDA Master Agreement* (IMA) and its *Credit Support Annex* (CSA). The IMA is, by far, the most common legal contract for bilateral over-the-counter (OTC) derivatives trading, although other agreements are sometimes used (such as the national forms of agreement used in some jurisdictions for domestic trading). We expect our analysis to apply

*Source code to the numerical examples in this paper is available at http://modval.org/models/mpr/ and http://modval.org/papers/aps2016/.

†The authors wish to acknowledge helpful comments from colleagues and from participants at a number of industry conferences, including 2014 Risk Minds, 2015 WBS Fixed Income USA, 2015 Global Derivatives, 2015 Risk Minds Americas, and 2015 WBS Fixed Income. All remaining errors are our own.

to a broad class of contracts, although model assumptions should be re-examined to confirm that key legal provisions remain substantially the same as the IMA.

We note that the modeling of default exposure and close-out risk arising from a non-zero MPR length has received a fair amount of attention in the past (see, e.g., [3], [4], [8], and [9]), although most[1] past analysis has been conducted under strong simplifying assumptions about trade and margin flows during the MPR. Here, we use a more detailed framework for legal and operational behavior to refine the classical models for collateralized exposure modeling.

The rest of this paper is organized as follows. First, we outline the fundamentals of variation margin posting in Section 2 and present the classical collateralized exposure model in Section 3. In Section 4, we discuss the full timeline of events that are likely to transpire during a credit default, from both operational and legal perspectives. This sets the stage for Section 5 where we propose a condensed representation of the timeline suitable for analytical and numerical work. The resulting setup leads to a significantly more nuanced and flexible definition of collateralized trading exposure. As fixing the actual values of model parameters ("calibrating" the MPR model) requires taking a stance on corporate behavior and operational procedures, Section 6 discusses how such parameterizations may be done in practice, for various levels of overall model prudence and counterparty types.

In the second part of the paper, we flesh out the model in more detail, especially as pertains to numerical implementation and quantitative comparisons with the classical model. As a starting point, Section 7 first formulates our exposure model in mathematical terms, and highlights the key differences to classical models by means of numerical results computed by brute-force Monte Carlo simulations. Computational techniques permitting efficient model implementation are introduced in Section 8, along with several test results. Section 9 briefly discusses applications for portfolios with risk-based initial margin, and Section 10 concludes the paper.

2 The Fundamentals of Variation Margin

2.1 Basic Definitions

In bilateral OTC derivatives trading, it is common for parties to require posting of collateral to mitigate excessive credit exposures. In the ISDA legal framework, the collateral mechanism is specified in the CSA as a combination of two types of margin: *initial margin* or *variation margin*. Although we briefly discuss initial margin in Section 9, we shall primarily focus on variation margin (VM), a form of collateral that is regularly adjusted based on the changing value of the bilateral portfolio. The VM is calculated and settled through time according to a set of CSA rules that we review in Section 4.

For concreteness, throughout the paper, we consider the exposure of a bank B to a counterparty C with whom B engages in bilateral OTC trading under the IMA/CSA legal framework. We will refer to C as the "defaulting party" and to B as "the bank" or the "non-defaulting party". All present value and exposure amounts throughout the paper will be calculated from the viewpoint of B. Let the net default-free market value (to B) of the securities portfolio at time t be $V(t)$, and let $A_B(t)$ and $A_C(t)$ be the collateral support amounts stipulated by the CSA to be posted by B and C, respectively. In the absence of initial margin, it is virtually always the case that only one

[1]One exception is the conference presentation [2] which contains elements of a more detailed framework, including recognition of the role played by cash flows close to the default event.

of A_B and A_C are positive, i.e., only one party will be required to post margin at a given point in time.

Assuming that collateral is netted (rather than posted by both parties in full and held in segregated accounts or by a third party), the total collateral amount in B's possession may be calculated as of time t as $c(t) = A_C(t) - A_B(t)$. Assuming also that collateral may be treated as *pari passu* with the derivatives portfolio itself for the purposes of the bankruptcy claim[2], it is common to denote the positive part of the difference $V(t) - c(t)$ as the *exposure* $E(t)$:

$$E(t) = (V(t) - c(t))^+, \quad c(t) = A_C(t) - A_B(t), \tag{1}$$

where we use the notation $x^+ = max(x, 0)$. There are several time lags and practical complications that render (1) an imprecise measure for default exposure, and we shall refine it substantially later on. In particular, we emphasize that collateral computed from market and portfolio observations at time t is generally not transferred to B until several days after t.

The type of VM encountered in the CSA is typically designed to broadly track the value of the portfolio between the parties, thereby ensuring that $E(t)$ in (1) does not grow excessive. However, to avoid unnecessary operational expenses, it is common to introduce language in the CSA to relax margin transfer requirements if the amounts are sufficiently small. To that end, the typical CSA language for collateral calculations will stipulate:

- collateral posting thresholds by each party, h_B and h_C, representing the maximum allowed amount of exposure before B or C, respectively, is required to post any collateral;
- a minimum transfer amount (MTA), establishing the minimum valid amount of a margin call; and
- rounding, which rounds collateral movements to some reasonable unit (say, \$1,000).

Formally, the effects of thresholds on stipulated collateral may be written as

$$A_B(t) = (-V(t) - h_B))^+, \quad A_C(t) = (V(t) - h_C))^+ \tag{2}$$

with the net stipulated credit support amount assigned to B being $c(t) = A_C(t) - A_B(t)$ as before. The actual availability of this amount is then subject to the (path-dependent) effects on collateral by MTA and rounding, of which the former has significant effect only for zero or very small thresholds, and the latter is usually negligible. Both of these effects have been omitted in (2).

Most CSAs are bilateral in nature, but unilateral CSAs exist in which only one of the two parties is required to post collateral. A CSA may also be formally bilateral, but highly asymmetric, requiring both parties to post collateral but with vastly different thresholds (e.g. h_B = \$20M vs h_C = \$2M). Typically, even for an asymmetric CSA the MTA and rounding are the same for both parties.

2.2 Margin Calls and Cash Flows

From an exposure perspective, the frequency with which the amount of collateral is adjusted (the *re-margining frequency*) is a critical component of the CSA. Following the financial crisis,

[2] Normally, both the collateral and the portfolio would be treated together as a senior unsecured claim of B against the bankruptcy estate of C.

most new IMA/CSAs, especially between major financial institutions, use daily re-margining frequency in order to reduce the amount by which exposure can change relative to collateral between margin calls. However, many smaller financial institutions or buy-side clients may not be able to cope with the operational burden of frequent margin calls and will often negotiate longer re-margining frequencies, e.g., weekly, monthly, or even longer.

The amount of collateral held by the parties is adjusted to their stipulated values A_B and A_C via the mechanism of a margin call. Many models for exposure treat the margin call as an instantaneous event, taking place on the re-margining date and completed instantaneously. In reality, the margin call is a chain of events which takes several days to complete. With daily re-margining, several such chains are running concurrently in an "interlaced' manner: even as one margin call is yet to be settled, another may already be initiated. The time lag in this settlement process, along with the inherent lag of the re-margining schedule, means that the changes in VM are always running behind the changes in portfolio value. This, in turn, implies that idealized exposure expressions such as (1) are inaccurate. The detailed events involved in the initiation and eventual settlement of a margin call will be discussed in Section 4.

With both default settlement and margin transfers being non-instantaneous events, it becomes relevant to track what payment flows take place (or not) during the periods close to a default. Two types of payments are needed here. The first type, for which we will use the term *trade flows*, covers the contractual cash flows, physical settlements, and other forms of asset transfers related to the trades themselves[3]. We are using the term "trade flows" rather than "cash flows" to emphasize that term sheets may involve flows other than cash – such as transfers of non-cash assets (e.g. commodities) or physical settlements resulting in creation of new trades from the old ones (e.g, the exercise of a physically settled swaption into a swap). A missed trade flow is a serious event under the IMA, and a "failure to pay" can rapidly lead to default and trade termination unless cured promptly. Any missed trade flow is, of course, part of the non-defaulting party's claim.

The second type of flows are those that arise from exchange of collateral between the parties ("margin flows"). The legal treatment of margin flows is governed by the IMA/CSA, rather than by trade documentation between the parties. For our purposes, the most important aspect of the IMA/CSA is the relatively mild treatment it affords to a party who misses a margin flow. Indeed, partially missing a margin payment is a common occurrence, as disputes about margin amounts happen regularly (and can sometimes persist for years).

During a collateral dispute, the CSA protocol calls for continued payments of the undisputed component of the collateral, but there is of course the possibility that there will be no undisputed component at all, if one party's counter-proposals are sufficiently frivolous. Should suspicions about "gaming" arise, the CSA does contain a methodology to stop disputes through market quotations, but the resulting leakage of position information to competitors is often a strong deterrent to its use. As such, there is a potential for abuse by firms that are experiencing financial difficulties, and a good possibility that such abuse can go on for some time before B takes further steps to end it. This, in turn, may result in a fairly long period of time between the last fully settled margin call and the eventual termination of the portfolio due to a default.

[3]These terms are spelled out in trade documentation and *term sheets*, for each trade.

2.3 Revised Exposure Definition

In light of the discussion above, let us make a first try at improving equation (1). For this, consider a default of C at time τ, followed by an early termination of the trade portfolio at time $t \geq \tau$. At time t, let $K(t)$ denote the collateral that B can *actually* rely on for portfolio termination; this amount very likely will differ from the CSA-stipulated amount $c(t)$ (and from $c(\tau)$, for that matter), due to margin transfer time lags and some degree of non-performance by C. In addition, it is possible that some trade flows were missed; let us denote their value at time t, including accrued interest, as $\mathrm{UTF}(t)$. Then we may redefine exposure generated by a default at time $\tau \leq t$ as

$$E(t) = (V(t) + \mathrm{UTF}(t) - K(t))^+ . \tag{3}$$

Notice that (3) anchors exposure at the termination date, rather than at the default date τ; we return to this topic in Section 5.3. For later use we also define time-0 expectation of future time-t exposure as

$$\mathrm{EE}(t) = \mathbb{E}_0(E(t)),$$

where $\mathbb{E}$ is the expectations operator in a relevant probability measure.

Determining how $K(t)$ may differ from $c(t)$, and how large $\mathrm{UTF}(t)$ can realistically be, will require a more detailed understanding of the settlement and margin processes, a topic we return to in Section 4. First, however, we examine in Section 3 below how classical approaches go about modeling $K(t)$ and $\mathrm{UTF}(t)$ in (3).

3 Classical Model for Collateralized Exposure

3.1 Assumptions About Margin Flows

A naive, and now outdated, model for collateralized exposure follows the definition (1) literally, and assumes that the available collateral is exactly equal to its prescribed value at time t. That is, in the language of (3), we assume $K(t) = c(t)$. In addition, the parties are assumed to pay all of the trade flows as prescribed ($\mathrm{UTF}(t) = 0$) and it is assumed that the termination date t in (3) equals the default time τ, i.e., there is no lag between the default date and the termination date. In this model, the amount of loss crystallized at time t is a function of portfolio value at a single time point, $V(t)$, and does not depend on the earlier history of $V(\cdot)$. In the limit of "perfect CSA" where $c(t) = V(t)$, the collateralized exposure in such a model is exactly zero.

Assuming $K(t) = c(t)$ is an idealization that ignores the non-instantaneous nature of collateral settlement protocols and does not capture the fact that firms under stress may stop fully honoring margin calls, resulting in a divergence between the portfolio value and collateral value during some period δ prior to termination of the portfolio. In what we here denote the *classical model*[4], this particular lag effect is captured by modifying (1) to

$$E(t) = (V(t) - K(t))^+ , \quad K(t) = c(t - \delta). \tag{4}$$

So, for instance, for a CSA with thresholds h_B and h_C we get, from (2),

$$K(t) = (V(t-\delta) - h_C(t-\delta))^+ - (-V(t-\delta) - h_C(t-\delta))^+ . \tag{5}$$

[4] See, for example, [9].

Having a mechanism for capturing divergence between collateral and portfolio value is an important improvement over the older model described above, and the classical model has gained widespread acceptance for both CVA (Credit Valuation Adjustment) and regulatory capital calculations. Nevertheless, it hinges on a number of assumptions that are unrealistic. For instance, (4) assumes that both B and C will simultaneously stop paying margin at time $t-\delta$, freezing the margin level over the entire MPR. In reality, if the party due to post collateral at $t-\delta$ happens to be the non-defaulting party B, it will often continue doing so for some time even in the presence of news about the possible impending default of C. And should C miss a few margin payments (maybe under the guise of a dispute), B would often continue to post collateral while it evaluates its options. This creates an asymmetry between posting and receiving collateral that the classical model fails to recognize.

In (4), the lag parameter δ is clearly critical: the larger δ is, the more $V(t)$ may pull from the frozen margin value at time $t-\delta$ and the bigger expected exposure will become. In practice, determination of δ will often be done in a fairly simplistic manner, e.g., by using a fixed lag (10 business days are common) or, more realistically, by adding a universal time delay to the re-margining frequency of the CSA in question. This practice is echoed in regulatory guidelines, e.g., in the Basel 3 accord where the MPR is set equal to the re-margining frequency minus 1 business day plus an MPR floor that defaults to 10 business days[5]. With the high proportion of individually negotiated and amended features in real life IMA/CSAs, using a "one size fits all" assumption may, however, lead to significant inaccuracies.

3.2 Assumptions About Trade Flows

Because large trade flows after the start of the MPR may no longer be followed by collateral adjustment, they have the potential to either extinguish or exacerbate exposure. For this reason, the model assumptions with respect to the date when either party suspends trade flows are likely to have significant effect on the counterparty credit loss. In one common interpretation of the classical model, it is simply assumed that both B and C will continue to pay all trade flows during the entire MPR. As a consequence, the unpaid trade flow term $\mathrm{UTF}(t)$ in (3) will be 0, consistent with (4). For ease of reference, we will denote this version of the classical model "Classical+".

In another, less common, version of the classical model, the assumption is that both B and C will stop paying trade flows at the moment the MPR commences, i.e., at time $t-\delta$. In this case, we set the unpaid trade flows equal[6] to

$$\mathrm{UTF}(t) = \mathrm{TF}^{net}\left(t;(t-\delta,t]\right),$$

where $\mathrm{TF}^{net}\left(t;(t',t'']\right)$ is the time t value of all net trade flows scheduled to be paid on the interval $(t',t'']$.[7] We denote this version of the classical model "Classical−"; it is associated with an exposure definition of

$$E(t) = \left(V(t) + \mathrm{TF}^{net}\left(t;(t-\delta,t]\right) - c(t-\delta)\right)^{+}. \tag{6}$$

[5] The MPR floor must be increased in certain cases, e.g., for large netting sets, illiquid trades, illiquid collateral, and recent collateral disputes – however the increase is specified as a multiplier relative to the same default.

[6] We measure time in discrete units of business days, such that the notation $(u,s]$ is equivalent to $[u+1\text{ bday},s]$.

[7] If t is after the date of a margin flow, the trade flow value accrues forward from the payment date to t at a contractually specified rate.

In practice, neither the Classical+ (equation (4)) or Classical− (equation (6)) versions of the classical model are accurate representations of reality. Trade flows are likely paid at least by B in the beginning of the MPR, and are likely not paid by at least C at its end. For instance, due to the CSA protocol for collateral calculations (see Section 4), there is typically at least a 3 business day lag between the start of the MPR (the market observation date for the last full margin payment) and the date when B definitively observes that C has missed paying a margin flow; during this period B would always make all trade payments unless C commits any additional contract violations. Even after B has determined that C missed a margin payment, B's nominal legal right to suspend payments following the breach would, as mentioned earlier, not always be exercised aggressively. Legal reviews, operational delays, and grace periods can further delay the time when B would finally stop paying trade flows to C.

Another trade flow effect arises during the last 2-3 days of the MPR (just prior to termination), where C has already defaulted and neither party is likely making trade payments. Here, the the IMA stipulates that the missed trade flows in this period accrue forward at a contractually specified rate and become part of the bankruptcy claim. This gives rise to a termination period contribution to the $\mathrm{UTF}(t)$ term, in turn leading to an adjustment of the exposure.

4 Full Timeline of IMA/CSA Events

Loosely speaking, the IMA concerns itself with the events of default, termination, and close-out; and the CSA governs collateral exchanges, including the concrete rules for collateral amount calculations and posting frequencies. While we have touched on the workings of the IMA/CSA in previous sections, our model construction shall require more detailed knowledge of certain provisions regarding the normal exchange of collateral, the legal options available in case of a missed payment, and common bank policies with respect to availing itself of these options. A detailed exposition of the IMA and CSA legal complexities can be found in multiple sources, including at http://www.isda.org; here, we only provide a brief summary to the extent necessary to develop our model. Our focus is on the development of a plausible timeline of events taking place around a default and subsequent portfolio termination.

4.1 Events Prior to Default

Let us assume that bank B is the Calculation Agent[8] for computation of collateral amounts. As before, let A_B and A_C denote *prescribed* collateral amounts for B and C; as we discussed, these may differ from *actually* available collateral amounts M_B and M_C if one of the parties fails to make a margin flow or changes the prescribed amount.

The following list describes the complete sequence of events taking place at times $T_0, T_1, \ldots$. We will simplify and condense them into a tractable model in the next section.

1. Time T_0. Our timeline begins at T_0, the as-of date at which the value of the portfolio and its collateral is measured, for usage in the T_1 evaluation of the formulas for the Credit Support Amount (plainly, the amount of collateral). Typically T_0 is the close of business on the business day before T_1.

[8] To ease comparisons with actual contracts, this section uses capital letter for official legal terms.

2. Time T_1. For our purposes, we use T_1 to denote the last *undisputed and respected* Valuation Date prior to default. At time T_1, besides officially determining[9] $A_B(T_0)$ and $A_C(T_0)$, bank B calculates the incremental payment amounts to itself and to C as $m_B = A_B(T_0) - M_B(T_0)$ and $m_C = A_C(T_0) - M_C(T_0)$, respectively. Taking into account any Minimum Transfer Amounts, the transfer amounts m_B and m_C should normally be communicated by B to C prior to a Notification Time (e.g., 1PM local time).

3. Time T_2. After receiving notice of the calculated collateral amount, C must initiate transfers of sufficient amounts of eligible collateral on the Payment Date T_2. Assuming that B managed to get the collateral amount notification sent to C prior to the Notification Time, T_2 defaults to one business day after T_1. If B is late in its notification, T_2 would be two business days after T_1. We here assume that that the required amounts – which we recall were calculated at T_1 using market data from T_0 – are all settled without incident at T_2. However, T_2 will be the *last* time that margin flows settle normally before the default takes place.

4. Time T_3. We let T_3 denote the next scheduled Valuation Date after T_1. If α is the average scheduled time between collateral calculations, we have approximately (ignoring business calendar effects) $T_3 \approx T_1 + \alpha$. At T_3 (hopefully before the Notification Time), B will send a payment notice to C, but C has now gotten into financial stress and will not be able (or willing) to pay further margin flows. Should C simply fail to pay collateral outright at the next Payment Date, a *Credit Support Default* could be triggered shortly thereafter (non-payment of collateral is associated with a 2 business day grace period). To prevent this from happening, it is, as discussed earlier, likely that C would attempt to stall by *disputing* the result of the T_3 collateral calculation by B.[10]

5. Time τ. Exactly how long the margin dispute is allowed to proceed is largely a behavioral question that requires some knowledge of B's credit policies and its willingness to risk legal disputes with C. Additionally, one needs to consider to what extent C is able to conceal its position of financial stress by using dispute tactics or, say, blaming operational issues on its inability to pay collateral. Ultimately, however, B will either conclude that C is in default on its margin flows (a Credit Support Default), or C will commit a serious contract violation such as failing to make a trade-related payment. At that point B will conclude that a *Potential Event of Default* (PED) has occurred. We identify the time of this event as the true *default time*, τ.

6. Time T_4. Once the PED has taken place, B needs to formally communicate it to C, in writing. Taking into account mail/courier delays, legal reviews, and other operational lags, it is likely that the communication time, denoted T_4, takes place at slight delay to the PED.

7. Time T_5. After receipt of the PED notice, C will be granted a brief period of time to cure the PED. The length of this *cure period* is specified in the IMA and depends on both the type of PED and the specific IMA. For instance, if the PED in question is Failure to Pay,

[9] Notice that while calculations are formally made on T_1, we use T_0 as the time-argument on all margin amounts, to reflect the fact that the market data is observed at time T_0.

[10] One of the authors still has vivid memories of how traders at Long Term Capital Management suddenly started disputing even the most basic of swap pricing methodologies. Default followed shortly afterwards.

the default cure period[11] is 3 business days in the 1992 IMA and 1 business day in the 2002 IMA. At the end of the cure period – here denoted T_5 – an *Event of Default* (ED) formally crystallizes. We emphasize that we here do *not* associate T_5 (the "official" default time) with the true default time τ; instead we equate τ to the time of the actual event (the PED) that, after contractual formalities, lead to the default of C.

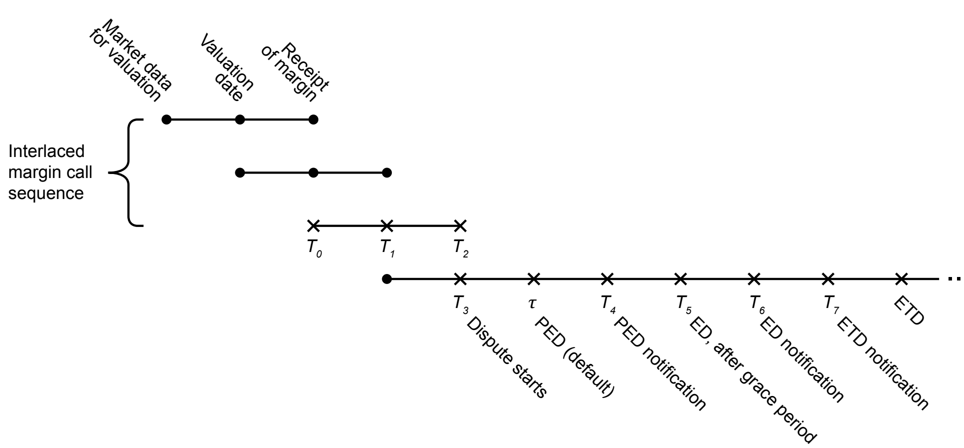
8. Time T_6 and T_7. After the ED has taken place, B will inform C of the ED at time $T_6 \geq T_5$ and may, at time $T_7 \geq T_6$, elect to designate an *Early Termination Date* (ETD).

9. Time T_8. The ETD is denoted T_8; per the IMA it is required that $T_8 \in [T_7, T_7{+}20 \text{ days}]$. The ETD constitutes the as-of date for the termination of C's portfolio and collateral position. Many banks will aim for a speedy resolution in order to minimize market risk, and will therefore aim to set the ETD as early as possible. There are, however, cases where this may not be optimal, as we discuss in Section 4.2 below.

10. *Subsequent events.* Once the portfolio claim has been established as of the ETD, the value of any collateral and unpaid trade flows held by C is added to the amount owed to B. Paragraph 8 of the CSA then allows B to liquidate any securities collateral in its possession and to apply the proceeds against the amount it is owed. Should the collateral be insufficient to cover what is owed to B, the residual amount will need to be submitted as a claim in C's insolvency. The claim is then usually challenged by the insolvency representative and, where parties cannot agree, may be referred to court. It can sometimes take a long time before the claim is resolved by bankruptcy courts and the realized recovery becomes known. The interest on the recovery amount for this time is added to the awarded amount. Note that in this paper we focus exclusively on modeling the magnitude of the exposure and bankruptcy claim, and do not challenge the established way of modelling the amount and timing of the eventual recovery using a loss-given-default (LGD) fraction.

The full timeline of IMA/CSA events is illustrated in Fig.1.

Figure 1: Full Timeline of Events During MPR (Daily Re-Margining).



[11] Which may very well be overridden in the actual documents.

4.2 Some Behavioral and Legal Aspects

While we have now established our event timeline, it remains for us to tie it to a proper model for exposure. In order to do so we shall, as already mentioned, need to combine the timeline with coherent assumptions about bank and counterparty behavior in each sub-period. These assumptions should be determined not only by the rights available under IMA/CSA, but also by the degree of operational efficiency in serving notices and getting legal opinions, as well as by the levels of prudence injected into the assumptions about the bank's ability and willingness to strictly uphold contractual terms within each client group as it pertains to margin flows and disputes.

From a legal rights perspective, the most important observation is that once notice of a PED has been served (here: time T_4), the so-called "suspension rights" of the IMA (Section 2(a)(iii)) and the CSA (Paragraph 4(a)) allow B to suspend all trade- and collateral-related payments to C until the PED has been cured. The extent to which suspension rights are actually exercised, however, is quite situational. A particular danger is that B exercises its suspension rights due to a Potential Event of Default (PED), but that subsequently the PED is ruled not to be valid. Should this happen, the bank can inadvertently commit a breach of contract which, especially in the presence of cross-default provisions, can have serious consequences for the bank.

Another, somewhat counter-intuitive, reason for B not to enforce its suspension rights is tied to the IMA's Section 2(a)(iii) which sometimes can make it favorable for B to *never* designate an ETD. Indeed, if B owes C money, it would seem a reasonable course of action for B to simply a) never designate an ETD; and b) suspend all payments on the portfolio until the default gets "cured" – which most likely will never happen. This tactic basically allows B to walk away from its obligations on the portfolio when C defaults, effectively making B a windfall gain.

The strategy of delaying the ETD in perpetuity has been tested by UK courts and found legal[12]. In the US, however, local "safe haven" laws have been ruled to prevent ETD's for more than about one year. Still, a one-year delay may prove tempting if B has a big negative exposure to C and is unwilling to immediately fund the large cash outflow needed to settle. As most large banks are presumably unlikely to play legal games with the ETD, we shall not consider the topic further here, but just note that there is potentially room to make more aggressive model assumptions around ETDs than what is done here.

5 Simplified Timeline of IMA/CSA Events

It should be evident from the preceding section that the full timeline of IMA/CSA events reviewed in Section 4 is in many ways different, and more complex, than what is assumed in both the Classical+ or Classical− versions of the classical model. However, it is equally evident that the full timeline is too complex to be modeled in every detail. In this section, we will offer a simplification of the timeline designed to extract the events most important for exposure modeling. The resulting model offers several important improvements over the classical model, while remaining practical and computationally feasible.

[12]Contract language has been proposed by ISDA to prevent the issue.

5.1 Identification of Key Time Periods

To recap, recall first that the classical model considers only two dates in the timeline of default: the start and end of the MPR. The start of the MPR, denoted by $t - \delta$, is defined as the last observation date for which margin was settled in full (a few days after the observation date). The end of the MPR, denoted by t, is the observation date on which B's claim is established. Note that t coincides with the IMA's *Early Termination Date* (ETD) discussed in Section 4 (time T_8). An alternative name for ETD frequently used in counterparty credit risk modelling is "closeout date".

In the classical model, there is no clear distinction between observation and payment dates, making it difficult to cleanly capture trade flow effects. For instance, in the Classical− version of the model, $t - \delta$ denotes both the last margin observation date as well as the date on which all trade flows cease. In reality, the last margin observation date is unlikely to be contentious and trigger stoppage of trade flows, as the margin payment to which the observation corresponds will only be missed by C several business days later. Specifically, if the market data is observed on day 0, the valuation is performed on day 1, then only on day 2 (or 3, if the notification was late) is the initiation of actual payment expected to take place. The length of this lag is of the same order of magnitude as typical assumptions for the length of the MPR, and can be a source of considerable model error if not handled properly.

In the simplified timeline we propose here, we take care to keep track of the distinction between observation and payment dates, and also consider the possibility that B may take the action of stopping a particular type of flow at a different time than C does. Accordingly, the model includes two (potentially different) observation dates for which B and C later settle their margin flows in full for the last time; and two (potentially different) dates when they pay their respective trade flows for the last time. The end of the MPR is defined the same way as the classical model, to coincide with ETD. Table 1 summarizes the notation for these five dates in our simplified timeline.

Event	Date type	Notation
Observation date for the last margin flow by C	Observation	$t_C = t - \delta_C$
Observation date for the last margin flow by B	Observation	$t_B = t - \delta_B$
Date of last trade flow payment by C	Settlement	$t'_C = t - \delta'_C$
Date of last trade flow payment by B	Settlement	$t'_B = t - \delta'_B$
ETD	Observation	t

Table 1: Notation for the dates in the simplified timeline.

The start of the MPR in our model is $t - \delta$, which in the notation of Table 1 may be defined symmetrically as $\delta = \max(\delta_B, \delta_C)$. We always expect the defaulting party C to stop posting margin no later than the non-defaulting party B, therefore one would very likely have $\delta_C \geq \delta_B$ and $\delta = \delta_C$.

The second column of Table 1 specifies which of the dates are observation dates, and which are settlement or payment dates. According to the notation established in Table 1, δ_B and δ_C are the lengths of time preceding the ETD during which changes in portfolio value no longer result in collateral payments by B and C, respectively. Similarly, δ'_B and δ'_C are the lengths of time preceding ETD during which the respective party does not pay trade flows. In, say, a "classical" 10-day MPR model, we have $\delta_B = \delta_C = 10bd$; with $\delta'_B = \delta'_C = 0$ for Classical+, and $\delta'_B = \delta'_C =$

$10bd$ for Classical−.

5.2 Establishing the Sequence of Events

A priori, the four events in Table 1 between the start and end of the MPR can occur in any order. However, we will now explain why the table very likely shows the proper sequence of the events.

As we discussed earlier, missing trade flows is often considered a more severe breach of contractual obligations than missing margin flows – especially as the latter may take the form of a margin valuation dispute. Therefore, it is reasonable to assume that neither party will stop paying trade flows before stopping the payment of margin flows. Accounting for the margin settlement lag between the observation date and the margin payment date, this yields:

$$\begin{aligned} \delta'_C &\le \delta_C - \text{margin settlement lag}, \\ \delta'_B &\le \delta_B - \text{margin settlement lag}. \end{aligned} \tag{7}$$

It is also reasonable to assume that either of the two types of flows is first missed by the defaulting party C, and only then by the non-defaulting party B. This leads to the following additional constraints on the sequence of events within the timeline:

$$\begin{aligned} \delta_C &\ge \delta_B, \\ \delta'_C &\ge \delta'_B. \end{aligned} \tag{8}$$

Except in rare and unique situations (such as outright operational failures), B would not continue to pay margin flows once C commits a more serious violation by missing a trade flow, resulting in

$$\delta'_C \le \delta_B - \text{margin settlement lag} \tag{9}$$

Combining these inequalities results in the chronological order of the events, as shown in Table 1.

5.3 Evaluation of Survival Probability

As was the case for the classical model, our setup anchors the exposure date t at the termination date (ETD), at the very end of the MPR. The ETD is the same for both parties, and constitutes a convenient reference point for aligning the actions of one party to those of the other. We emphasize that the ETD for which exposure is evaluated does not coincide with the date at which survival probability is evaluated, e.g., for the computation of CVA. In our simplified timeline, the counterparty survival probability should be evaluated for[13] $t - \delta'_C$, the last date when C stops paying trade flows. Hence, if $EE(t)$ is the expected exposure anchored at the ETD t, then the incremental contribution to (unilateral) CVA from time t is, under suitable assumptions, $EE(t) \cdot d\mathbb{P}(t - \delta'_C)$ where $\mathbb{P}$ is the survival probability under the model's measure; see Section 7.3.3 for concrete examples.

Evaluating the default probability at the anchor date t rather than $t - \delta'_C$ will introduce a slight error in computing the survival probability. While this error is relatively small and is often ignored by the practitioners, it takes virtually no effort, and has no impact on model efficiency, to evaluate the survival probability at the right date.

[13] Effectively we assume that default is due to failure-to-pay.

6 Timeline Calibration

As we mentioned earlier, the specific IMA/CSA terms for a given counterparty should ideally always be examined in detail, so that any non-standard provisions may be analyzed in terms of their impact on the timeline. For those cases when such bespoke timeline construction is not practical (typically for operational reasons), we will here propose two standard ("reference") parameterizations of our timeline. This will also allow us to demonstrate the thought processes behind timeline calibration, and will provide some useful base cases for our later numerical tests.

While factors such as portfolio size and dispute history with the counterparty should, of course, be considered in establishing the MPR, we believe that an equally important consideration in calibrating the model is the nature of the expected response by B to missed margin or trade flows by C. Even under plain-vanilla IMA/CSA terms, experience shows that reactions to contract breaches are subject to both human and institutional idiosyncrasies, rendering the MPR potentially quite variable. Recognizing that "one size does not fit all", we shall therefore consider two different calibrations: one "aggressive", which assumes best case scenario for rapidly recognizing the impending default and taking swift action; and one "conservative", which takes into account not only a likely delay in recognizing that counterparty default is imminent, but also the possibility that the bank may not aggressively enforce its legal rights afforded under IMA and CSA in order to avoid damaging its reputation. In both scenarios, we also assume daily remargining – if a CSA calls for less frequent margin calls than this, the MPR must be lengthened accordingly.

6.1 Aggressive Calibration

The Aggressive calibration applies to the trading relationship between two counterparties that both have strong operational competence, and where there is little reputational risk associated with swift and aggressive enforcement of the non-defaulting party's legal rights against the defaulting party. A good example would be trading between two large dealers, both willing to aggressively defend against a possible credit loss. We would here assume that credit officers are diligent in their monitoring of the counterparty, and generally able to see a default developing, rather than be caught by surprise.

Under Aggressive calibration, the event of C missing (or disputing) a margin call by any non-trivial amount will, given C's sophistication, immediately alert B that an impending default is likely. B will not be misled by claims of a valuation dispute or other excuses, and will send a notice of a Credit Support Default under the IMA/CSA the next business day after the breach of the margin agreement. At the same time, to further protect itself, B will stop both margin flows and trade flows. The counterparty C is assumed to simultaneously stop paying margin and trade flows as well, so that no further payments of any kind are exchanged by the parties.

The simultaneous action by both parties in our Aggressive scenario to stop paying trade flows at the earliest possible moment results in elimination of all *settlement risk* – the possibility that the bank may continue paying on its trade flow obligations while not receiving promised payments in return. In the context of cross-currency trades, this type of settlement risk is frequently referred to as "Herstatt" risk, after the bank which famously caused large counterparty losses in this manner[14]. Such risk shall be captured in our Conservative calibration case below, and shall be discussed in more detail in Section 7.2.

[14] See, e.g., https://en.wikipedia.org/wiki/Settlement_risk .

Despite B's immediate (and rather aggressive) response, the MPR will still be fairly long due to the way the IMA/CSA operates in practice. In particular, notice that the first period in our simplified timeline is between the last observation date for which margin was fully settled, and the first date on which C misses a margin flow. As it takes (at least) 2 business days to settle a margin payment, plus 1 business day between the last margin that was settled successfully and the first margin payment that was not, a minimum of 3 business days will always accrue from the start of the MPR to a margin-related PED takes place. Further, once the margin flow is missed, C must send two notices and permit a grace period (usually 2 business days) to cure the violation, before an event of default (ED) has officially taken place and an ETD has been designated. Since an ETD cannot be designated prior to the event of default, it is unlikely that the MPR can ever be less than 7 business days. It is remarkable that even under the most aggressive set of assumptions, MPR is still only 3 business days shorter than the classical 2-week MPR.

The detailed taxonomy of our Aggressive timeline is listed in Table 2 in Section 6.3, and essentially splits the MPR into two sections: a margin delay period of 3 business days, and a default resolution period of 4 business days. During the latter period, B and C cease paying on the first day, leaving a period of 3 business days where neither party makes any payments. Notice that we assume that the ETD is declared to coincide with the ED, i.e., the bank will terminate as quickly as legally possible.

6.2 Conservative Calibration

The Conservative calibration is intended to cover the situation when the bank's enforcement of its rights under the IMA/CSA is deliberate and cautious, rather than swift. There may be several reasons for such a situation, sometimes acting in tandem.

First, a bank, if overly trigger-happy, can gain a market-wide reputation as being rigid and litigious, potentially causing clients to seek other trading partners. In fact, should aggressive legal maneuvers be applied to counterparties that may be considered "unsophisticated," there is even the potential for the bank to be perceived as predatory by the larger public. Second, there are situations where exercising legal rights would cause an unattractive leakage of information into the broader market. As mentioned, this may for instance happen if the formal collateral dispute methodology in Paragraph 5 of the ISDA CSA is activated: the market poll mechanism inherent to the methodology would inevitably reveal the positions held with the counterparty to competing banks. Third, sometimes an aggressive interpretation of legal rights can backfire in the form of lawsuits and counter-measures by the counterparty. For example, even when a bank may have the right to withhold payments (e.g., under Section 2(a)(iii)), it would often elect not to exercise this right immediately out of concern that a counter-ED would be raised against it or that withholding payments would exacerbate the liquidity situation of the counterparty, potentially exposing the bank to liability and lawsuits. As mentioned, a particular danger is that the bank exercises its suspension rights due to a Potential Event of Default (PED), but that subsequently the PED is ruled not to be valid. Should this happen, the bank can inadvertently commit a breach of contract.

Of course, even if a bank may potentially be willing to aggressively exercise its rights, it might not have the operational capacity to do so quickly. For example, the bank may not be able to perform the requisite legal review on a short notice, or may not have the efficiency needed to always get notices mailed out a the earliest possible date. On top of this, there is always the potential for human or technology-related errors and oversights.

While it is harder to get concrete data to estimate a reasonable the timeline for the Conservative case (this case being dependent not only on IMA/CSA details but also the specific banks' reputational considerations), under a perfectly reasonable set of assumptions the MPR ends up being more than twice as long as for our Aggressive case above. Under this calibration choice, the Conservative scenario assumes that the totality of margin dispute negotiations, operational delays, human errors, legal reviews, etc., adds to 8 business days, yielding a total MPR of 15 business days. One plausible scenario with daily re-margining could be:

$t-15$**:** B observes the portfolio value as needed for margin transfer amount #1 as of $t-15$.

$t-14$**:** B sends margin call #1 to C; B observes margin transfer amount #2.

$t-13$**:** B sends margin call #2; C honors margin call #1; B observes margin transfer amount #3.

$t-12$**:** C fails to honor margin call #2 and initiates dispute; B tries to resolve the dispute, while still paying and calculating margin.

$t-7$**:** C fails to make a trade payment.

$t-6$**:** B stops paying margin and sends the PED notice.

$t-5$**:** C receives PED; B keeps making trade payments.

$t-3$**:** The PED is not cured.

$t-2$**:** B stops trade payments and sends ED notice to C, designating t as the ETD.

t**:** ETD.

Notice that we have a number of different margin transfers active simultaneously (denoted #1, #2, and #3), reflecting the interlacing nature of daily margin calls. Notice also that unlike the earlier Aggressive calibration, the above scenario explicitly involves settlement risk, as a time period exists where only B pays trade flows (from $t-7$ to $t-3$, both dates inclusive).

To translate the scenario above into the notation of Section 5, first notice that $\delta_C = 15$ as the observation date of the last margin call (#1) honored by C is $t-15$. Second, as B makes its last possible margin payment at time $t-7$, based on an observation at time $t-9$, we have $\delta_B = 9$. Third, as C fails to make a trade payment at $t-7$, C's last payment date is $t-8$ and therefore $\delta'_C = 8$. And finally, since B stops its trade payments at $t-2$ we must have $\delta'_B = 3$.

6.3 Summary and Comparison of Timelines

In the notation of Section 5, our Aggressive and Conservative scenarios are presented in Table 2 below. For reference, the Classical+ and Classical− versions of the classical model (see Section 3) are presented in the table as well. Note that the 10bd assumption of the classical MPR model lies between the two calibration choices we propose, and is closer to the Aggressive scenario.

Our Aggressive and Conservative parameter choices represent two opposite types of bank-counterparty relationship, and may also be used as two limit scenarios for materiality and model risk analysis. Of course, the best approach would always be to set model parameters based on prudent analysis of the firm's historical default resolution timelines, to the extent that this is practically feasible. The model could also conceivably treat the various time lags as random variables to be simulated as part of exposure computations; yet it is debatable whether increasing the number of model parameters in this way is warranted in practice.

Parameter	Conservative	Aggressive	Classical+	Classical−
δ_C	15bd	7bd	10bd	10bd
δ_B	9bd	6bd	10bd	10bd
δ'_C	8bd	4bd	0bd	10bd
δ'_B	3bd	4bd	0bd	10bd

Table 2: MPR Periods for CSAs with Daily Re-margining

7 Advanced Model for Collateralized Exposure

To formulate our model in more precise mathematical terms, let us return to equation (3) and consider how to draw on our analysis in Sections 4, 5, and 6 to reasonably specify both the collateral amount $K(t)$ as well as as the value $\mathrm{UTF}(t)$ of net unpaid trade flows.

7.1 Unpaid Margin Flows and Margin Flow Gap

As in the classical model, we assume that the MPR starts at time $t_C = t - \delta_C$, the portfolio observation date associated with the last regular collateral posting by C. Recall that the classical model further assumes that B will stop posting collateral simultaneously with C, so that (compare to equation (4))

$$K(t) = c(t_C), \tag{10}$$

where $c(t_C)$ denotes the CSA-prescribed collateral support amount calculated from market data observed at time t_C.

In contrast to (10), our model assumes that B will continue posting and returning collateral to C for all contractual margin observation dates t_i, whenever required by the CSA, after $t_C = t - \delta_C$ and up to and including the observation date $t_B = t - \delta_B$. We will refer to the presence of an observation period of non-zero length for which B is posting and returning collateral while C does not as "margin flow gap".

Here, we always expect $t_B \geq t_C$, and therefore in effect assume the possibility of a time interval $(t_C, t_B]$ where only B honors its margin requirements. In this interval, B can match contractually stipulated amounts $c(t_i)$ only when they involve transfers from B to C. This asymmetry results in B holding at time t the *smallest* collateral computed on the observation interval $[t_C, t_B]$, i.e.,

$$K(t) = \min_{t_i \in [t_C, t_B]} c(t_i). \tag{11}$$

The "worst-case" form of (11) clearly provides a less optimistic view on available collateral than the classical model, resulting in larger exposure whenever there are multiple collateral observation dates in $[t_C, t_B]$.[15] All other things being equal, the difference in exposure relative to the classical model will increase with $\delta_C - \delta_B$. If $\delta_C - \delta_B$ is kept constant, the difference will increase with more frequent re-margining. Note that (11) matches (10) when $\delta_B = \delta_C$.

7.2 Unpaid Trade Flows and Trade Flow Gap

According to our assumptions in Section 5, the last date when C is still paying trade flows is $t'_C = t - \delta'_C$ and the last date when B is still paying trade flows is $t'_B = t - \delta'_B \geq t'_C$. We will refer

[15]It is assumed that one of the observation dates t_i always coincides with the start of the MPR, t_C.

to the period when B is still paying trade flows while C does not as the "trade flow gap".

The value of the net trade flows unpaid by the termination date t can be expressed using the notation of Section 3.2 as

$$\begin{aligned} \mathrm{UTF}(t) &= \mathrm{TF}^{C\to B}\left(t;(t'_C,t]\right) + \mathrm{TF}^{B\to C}\left(t;(t'_B,t]\right) \\ &= \mathrm{TF}^{C\to B}\left(t;(t'_C,t'_B]\right) + \mathrm{TF}^{net}\left(t;(t'_B,t]\right) \end{aligned} \tag{12}$$

where an arrow indicates the direction of the trade flows and $C \to B$ ($B \to C$) trade flows have positive (negative) sign.

In calculating (12), care needs to be taken in how trade flows are aggregated and accrued to the termination date t:

- Cash flows of opposite direction scheduled to be paid in the same currency on the same date[16] in period $(t'_B, t]$ are aggregated (netted) at the cash flow date, therefore only the aggregated amount (their difference) enters into equation (12). The aggregated amount of missed cash flows should be accrued to time t at the interest rate of the currency in question, and then converted to B's domestic currency.
- Cash flows in opposite directions scheduled to be paid in *different* currencies on the same date[17] in period $[t'_B, t]$ *are not* netted at the cash flow date. The missed cash flow amounts in each currency should be accrued to time t at the relevant interest rates, and then converted to B's domestic currency.
- The value of each asset flow[18] should be obtained through pricing at time t of the undelivered asset in B's domestic currency. Generally, asset flows are not aggregated.

To analyze the impact of the assumptions in (12), let us for simplicity consider a zero-threshold margin agreement with no MTA/rounding. Then, from (11), collateral available to B at the termination date can be written as

$$K(t) = V(t_{col}), \quad t_{col} \triangleq \arg \min_{t_i \in [t_C, t_B]} V(t_i). \tag{13}$$

Substituting (12) and (13) into (3) yields, for the simple CSA considered,

$$E(t) = \left[V(t) - V(t_{col}) + \mathrm{TF}^{C\to B}\left(t;(t'_C,t'_B]\right) + \mathrm{TF}^{net}\left(t;(t'_B,t]\right)\right]^+ . \tag{14}$$

Equation (14) implies that trading flows from B to C can occur within the MPR. These trade flows have the potential to generate large spikes in the exposure profile – especially in the presence of a trade flow gap, when only B pays trade flows. To see this, we drill further into the exposure components of (14) as follows. First, ignoring minor discounting effects inside the MPR, let us represent the portfolio value at time t_{col} as the sum of the portfolio's forward value V^F to time t and the value of all the trade flows taking place *after* t_{col} and *up to and including* t:

$$V(t_{col}) = V^F(t_{col}; t) + \mathrm{TF}^{net}(t_{col}; (t_{col}, t]) . \tag{15}$$

[16] For instance: the two legs of an ordinary single-currency interest rate swap.
[17] For instance: the two legs of a cross-currency interest rate swap.
[18] For instance: a swap that would result from exercising a physically settled swaption.

Here we may further write

$$
\begin{aligned}
\mathrm{TF}^{net}\left(t_{col};(t_{col},t]\right) = \mathrm{TF}^{net}&\left(t_{col};(t_{col},t'_C]\right) \\
&+ \mathrm{TF}^{C\to B}\left(t_{col};(t'_C,t'_B]\right) \\
&+ \mathrm{TF}^{B\to C}\left(t_{col};(t'_C,t'_B]\right) \\
&+ \mathrm{TF}^{net}\left(t_{col};(t'_B,t]\right),
\end{aligned}
$$

which, together with (15), allows us to restate (14) in the following illuminating form:

$$
\begin{aligned}
E(t) = \Big[&V(t) - V^F\left(t_{col};t\right) \\
&+ \mathrm{TF}^{C\to B}\left(t;(t'_C,t'_B]\right) - \mathrm{TF}^{C\to B}\left(t_{col};(t'_C,t'_B]\right) \\
&+ \mathrm{TF}^{net}\left(t;(t'_B,t]\right) - \mathrm{TF}^{net}\left(t_{col};(t'_B,t]\right) \\
&- \mathrm{TF}^{net}\left(t_{col};(t_{col},t'_C]\right) \\
&- \mathrm{TF}^{B\to C}\left(t_{col};(t'_C,t'_B]\right)\Big]^+ .
\end{aligned} \tag{16}
$$

Here, we have arranged the terms in (14) into five separate lines, corresponding to five contributions to exposure:

- *Line 1*: The change of the portfolio forward value to time t, driven by the change in market factors between t_{col} and t. This term is driven by the volatility of market factors between t_{col} and t; it produces no spikes in the expected exposure profile.
- *Line 2*: The change of the value of the trade flows scheduled to be paid (but actually unpaid) by C in the interval $(t'_C,t'_B]$, resulting from the change of market factors between t_{col} and t. This term is driven by the volatility of market factors between t_{col} and t; it produces no spikes in the expected exposure profile.
- *Line 3*: The change of the value of the net trade flows between C and B scheduled to be paid (but actually unpaid) in the interval $(t'_B,t]$, resulting from the change of market factors between t_{col} and t. This term likewise produces no spikes in the expected exposure profile.
- *Line 4*: The negative value of the net trade flows between C and B scheduled to be paid (and actually paid) in the interval $(t_{col},t'_C]$. Paths where B is the net payer (so that TF is negative) contribute to upward spikes in the EE profile.
- *Line 5*: The negative value of the trade flows scheduled to be paid (and actually paid) by B to C in the interval $(t'_C,t'_B]$. Whenever such trade flows are present, B is always the payer, leading to upward spikes in the EE profile. Furthermore, in some cases, the spikes arising from this term can be of extreme magnitude, e.g., a scheduled notional exchange in a cross-currency swap when B pays the full notional, but receives nothing.

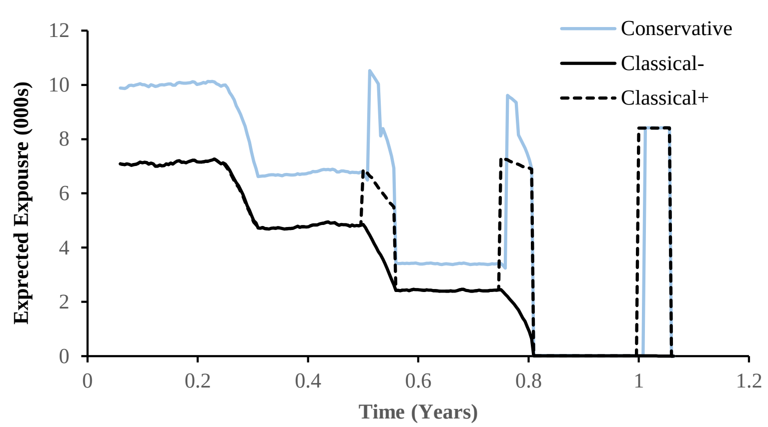
7.3 Numerical Examples

To gain intuition for our model, we now present exposure profiles and CVA metrics for several trade and portfolio examples, using both Aggressive and Conservative calibrations. We focus on ordinary and cross-currency swaps, as these instruments are the primary sources of exposure in most banks. For all numerical examples, we drive the stochastic yield curve scenarios with a one-factor Hull-White model; for our cross-currency swap examples, the foreign exchange rate is assumed to follow a Black-Scholes model.

7.3.1 *Individual Swap Results*

First, let us examine how our model differs from the classical exposure approach. In Figure 2 we have used Monte Carlo simulation on a USD10MM 1-year par-valued vanilla interest rate swap to compare exposures of our Conservative calibration with those computed from the Classical+ and Classical− models (see Table 2). To make comparisons meaningful, we have overridden the default setting of 10 business days for the Classical method, and instead set it equal to 15 business days, the length of the MPR for our Conservative calibration.

Figure 2: Expected Exposure for 1-Year Vanilla Swap



USD10MM 1-year payer swap, 2% coupon; fixed and floating coupons paid quarterly. Margin agreement: daily margin transfers, no thresholds, and no MTA/rounding. Expected exposures are computed by brute-force daily simulation (5,000 paths), using Classical+, Classical−, and Conservative models, where the MPR of the Classical methods is set to 15 bdays. Scenarios computed by a Hull-White model, with the initial yield curve flat at 2%, Gaussian (basis point) volatility of 1%, and mean reversion speed of 5%.

In Figure 2, note that The Classical− calibration is, of course, the least conservative setting, as it ignores both the effects of trade flows and margin asymmetry. The Classical+ calibration tracks the Classical− results at most times, but contains noticeable spikes around the last three[19] cash flow dates. The spikes originate from scenarios where B pays a net positive coupon on a cash flow date, yet, by assumption, receives no margin in return. The spikes commence whenever the exposure date (the ETD in our convention) hits a coupon payment date, and cease 15 business days later, the length of the MPR. Our Conservative calibration results also contain spikes around cash flow dates, although these differ from those of the Classical+ calibration in several ways. First, the Conservative calibration correctly recognizes that there will always be a part towards the end of the MPR (after time t'_B) where B and C will have both stopped paying margin and

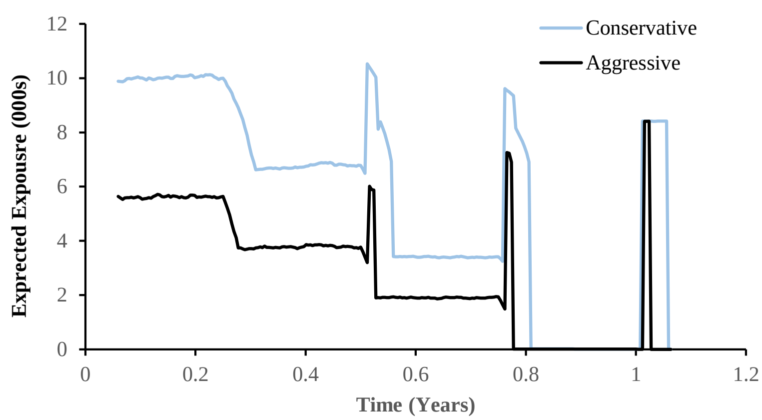
[19]No spike here occurs on the first quarterly cash flow date, as we assume that the floating rate is fixed at 2%, making the net cash flow zero in all scenarios.

coupons; as a result, the spikes of the Conservative calibration start later (here: 3 business days) than those of the Classical+ calibration. Second, the initial part of the spike (in the period from t'_C to t'_B) is substantially higher for the Conservative calibration, due to the assumption of only B paying cash flows in this sub-period. The remainder of the spike is comparable in height to the Classical+ spike.

Between spikes, we note that our Conservative calibration produces higher exposures than both the Classical+ and Classical− methods, by around 40%. This is, of course, a consequence of the "worst case" margin asymmetry mechanism in (11), the effect of which will grow with the diffusion volatility of the rate process. Notice that the last coupon period has no exposure between spikes since the volatility of the swap price vanishes after the last floating rate fixing at 0.75 years.

Comparison of the Aggressive calibration to the Classical+ and Classical− calibrations is qualitatively similar to the results in Figure 2, so we skip a detailed comparison and just note that the pick-up in exposure from margin asymmetry falls to about 15%, rather than the 40% we observed for the Conservative calibration – a result of the fact that the "worst case" margin result is established over much fewer days for the Aggressive calibration. A comparison of exposure profiles for the Aggressive and Conservative calibrations can be found in Figure 3; as expected, the Conservative calibration leads to both bigger and wider exposure spikes, as well as to higher exposure levels between spikes.

Figure 3: Expected Exposure for 1-Year Vanilla Swap

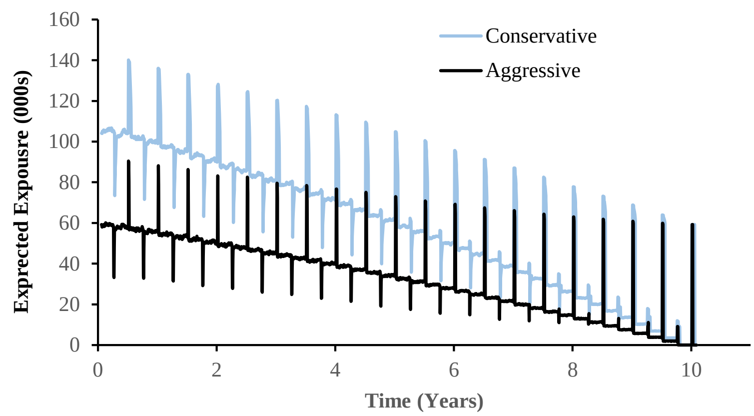


USD10MM 1-year payer swap, 2% coupon; fixed and floating coupons paid quarterly. Simulation and exposure setup as in Figure 2.

While instructive, we note that our 1-year vanilla swap example is quite benign exposure-wise: not only is the instrument very short-dated, it also allows for netting of coupons on trade flow dates, thereby reducing the effects of trade flow spikes. We can relax both effects by increasing the maturity of the swap, and by making the fixed and floating legs pay on different schedules. Exposure results for such a case is shown in Figure 4.

In Figure 4 upward exposure spikes occur twice per year, whenever the bank must make a (semi-annual) fixed payment. On dates where the counterparty makes a (quarterly) floating rate

Figure 4: Expected Exposure for 10-Year Vanilla Swap



USD10MM 10-year payer swap, 2% coupon; fixed coupons paid semi-annually, floating coupons paid quarterly. Model and exposure setup otherwise as in Figure 2.

payment that is not accompanied by a fixed payment by the bank, also twice per year, a narrow *downward* spike now emerges, due to the schedule delay in transferring back the coupon to the counterparty through the margin mechanism. Notice that the exposure between spikes is much larger than in Figure 3, a consequence of the higher volatility of a 10-year swap compared to a 1-year swap. Of course, as the swap nears maturity, its duration and volatility die out, so the non-spike exposure profile predictably gets pulled to zero at the 10-year date. Also as predicted, the Aggressive calibration produces much lower exposures than the Conservative calibration, by nearly a factor of 2.

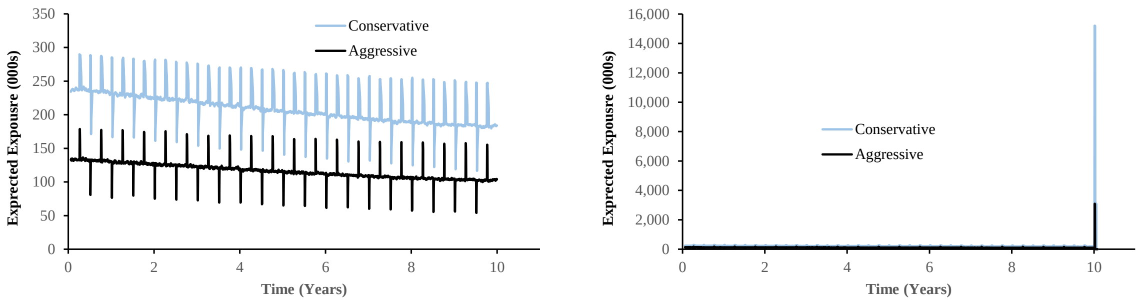
A more extreme form of trade flow spikes will occur for cross-currency swaps, where neither coupons nor the final notional payment can be netted. The notional payment, in particular, can induce a very significant exposure spike (Herstatt risk), whenever the exposure model allows for a trade flow gap. As we recall, our Conservative calibration has a trade flow gap, but our Aggressive calibration does not. As Panel (b) in Figure 5 confirms, the exposure for the Conservative calibration has a very large terminal spike that is not present in the Aggressive calibration. Like the Conservative calibration, the Aggressive calibration will, of course, still produce exposure spikes at cash flow dates, due to margin effects; see Panel (a) in Figure 5. For variation and for realism, note that we in Figure 5 have elected to use a seasoned deal traded at a time when the rate levels and the EUR/USD FX rate were all higher than at present. As a consequence, the principal exchange is likely far away from break-even, result in a large exposure spike at maturity. Although smaller than for the Conservative calibration, notice that a spike at maturity is also present for the Aggressive calibration: while both B and C pay the principal exchange, C does not make the margin transfer for the balance of the principal payments.

7.3.2 Portfolio Results

For individual trades, the presence of localized spikes in the exposure profiles may ultimately have a relatively modest impact on credit risk metrics, such as the CVA – after all, the likelihood

Figure 5: Expected Exposure for 10-Year Cross-Currency Swap

(a) Exclude exposure at maturity (b) Include exposure at maturity



USD10MM 10-year USD-EUR cross-currency payer swap, fixed 3% EUR semi-annual coupon against a quarterly USD floating rate with 15mm USD notional and 10mm EUR notional. The margin agreement uses daily margin transfers, no thresholds, and no MTA/rounding. Expected exposures are computed by brute-force daily simulation (5,000 paths), with default exposures captured in Aggressive and Conservative calibrations, as specified on the graphs. Scenarios computed by uncorrelated Hull-White models for the interest rates, with the initial USD and EUR yield curves flat at 2% and 1%, respectively, Gaussian (basis point) volatilities of 1%, and mean reversion speeds of 5%. The exchange rate is log-normal, with USD/EUR spot at 1.2 and a constant FX diffusion volatility[20] of 10%. For clarity, Panel (a) excludes the final spike, to see more detail in the pre-maturity profile; the spike is shown in Panel (b).

of a counterparty default in a narrow time interval around a quarterly or semi-annual cash flow event is typically low. For a *portfolio* of swaps, however, the spikes will add up and can affect the net exposure profile nearly everywhere. To show this, we picked 50 interest rate swaps with quarterly floating rate payments and semi-annual fixed rate payments of 2%. The terms of the swaps were randomized as follows:

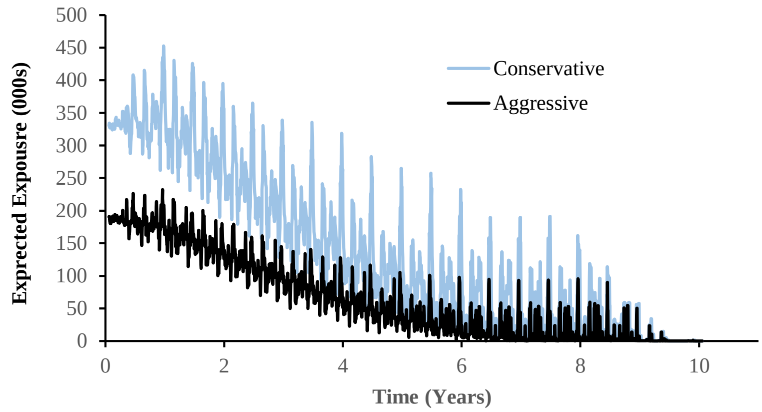
- Notionals are sampled uniformly on the interval from 0 to USD10MM.
- Direction of fixed leg payments (payer or receiver) is random.
- Start date of each swaps is subject to a random offset to avoid complete MPR overlaps.
- Swap maturities are sampled uniformly on the interval from 1 to 10 years.

The resulting expected exposure profiles are shown in Figure 6. Notice how both the Conservative profile and, to a lesser extent, the Aggressive profile include frequent spikes around trade flow times above the "baseline" exposure level. As we will see in Section 7.3.3, these spikes make a significant contribution to CVA metrics. As before, the exposure under Conservative calibration is about twice as large than under the Aggressive calibration.

To repeat the portfolio results with cross currency swaps, we constructed a 50-deal portfolio by randomization, using the following rules:

- EUR notionals are sampled uniformly on the interval from 0 to USD10MM.
- USD notionals are 1.5 times EUR notionals.

Figure 6: Expected Exposure for Vanilla Swap Portfolio



Portfolio of 50 interest rate swaps, as defined in text; fixed coupons (2%) paid semi-annually, floating coupons paid quarterly. Simulation and exposure setup otherwise as in Figure 2, except 1,000 paths was used for all portfolio examples.

- EUR leg has fixed semiannual coupon of 3%, USD leg floating quarterly coupon.
- Direction of fixed leg payments (payer or receiver) is random.
- Start date of each swaps is subject to a random offset to avoid complete MPR overlaps.
- Swap maturities are sampled uniformly on the interval from 1 to 10 years.

As for Figure 5, we generated the swaps within the portfolio such that the principal exchange and fixed coupon are not at-the-money, to mimic the typical situation for a portfolio of seasoned trades. As shown in Figure 7, the exposure for the Conservative calibration is, as expected, here dominated by a series of Herstatt risk spikes, one per swap in the portfolio. As before, we generated the swaps within the portfolio such that the principal exchange and fixed coupon are not at the money, to mimic the typical situation for a portfolio of seasoned trades done at the time of different market rates.

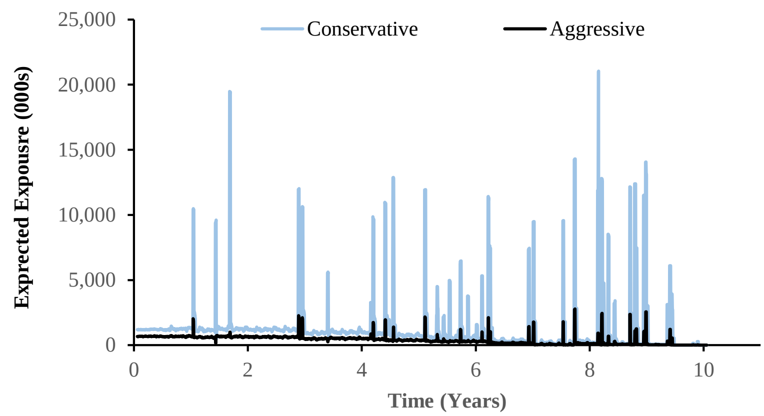
7.3.3 CVA Results

As mentioned earlier, a common use of expected exposure results is the computation of CVA. Under suitable assumptions, B's unilateral CVA may be computed from the expected exposure (EE) profile as

$$\mathrm{CVA} = (1 - R) \int_0^\infty P(u + \delta'_C)\mathrm{EE}(u + \delta'_C)dX(u) \tag{17}$$

where R is the recovery rate, $P(t)$ is the time 0 discount factor to time t, and $X(t)$ is the time 0 survival probability of C to time t. Notice that we, as discussed in Section 5.3, offset the exposure profile by δ'_C to properly align exposures with default events.

Figure 7: Expected Exposure for Cross-Currency Swap Portfolio



Portfolio of 50 cross-currency rate swaps, as defined in text; fixed EUR coupons (3%) paid semi-annually, floating USD coupons paid quarterly. Simulation and exposure setup otherwise as in Figure 5, except 1,000 paths was used for all portfolio examples.

The CVA metric serves as a convenient condensation of the exposure profiles in Sections 7.3.1 and 7.3.2 into single numbers, and Table 3 lists CVA results for the instruments/portfolios in Figures 4-7. We discretized the integral in (17) on a daily grid, assumed $R = 40\%$, and let the forward default intensity be constant at 2.5% such that $X(t) = \exp(-0.025t)$. Note that the table for reference also includes results for the Classical method, with MPR lengths equal to both that of the Aggressive calibration (7 bdays) and of the Conservative calibration (15 bdays).

Securities	Conservative	Aggressive	Classical+(15)	Classical−(15)	Classical+(7)	Classical−(7)
10Y Swap	7,785	4,187	5,552	5,242	3,739	3,561
Swap Portfolio	29,399	14,459	19,615	17,961	13,069	12,246
10Y XCCY Swap	18,131	8,524	12,219	10,122	8,131	6,889
XCCY Swap Portfolio	167,066	49,825	78,083	56,807	50,257	38,785

Table 3: CVA Results for Conservative, Aggressive, and Classical Calibrations

Results in Table 3 confirm what we have seen earlier. For instance, the CVA for the Aggressive calibration is 50% to 70% smaller than CVA for the Conservative calibration. In addition the CVA of the Conservative calibration is between 50% to 100% larger than that of the Classical+ calibration (at similar MPR), which in turn is larger than the CVA of the Classical− calibration by around 5% to 25%. Not surprisingly, the CVA results for the XCCY portfolio are particularly high in the Conservative calibration, due to Herstatt risk.

8 Improvement of Computation Times

In exposure calculations for realistic portfolios, horizons can be very long, often exceeding 30 years. For such lengthy horizons, brute-force Monte Carlo simulation of exposures on a daily, or even weekly, time grid will often be prohibitively slow. It is therefore common to use daily simulation steps only for the earliest parts of the exposure profile (e.g. the first month), and then gradually increase the time step length over time to monthly or quarterly, in order to keep the total number of simulation dates manageable. Unfortunately, such a coarsening of the time grid will inevitably fail to capture both the “worst case” margin effect and the trade spikes that are key to our exposure model.

In this section, we discuss ways to capture exposure without having to resolve to brute-force daily simulation. We first review a common speed-up technique for the Classical model (the Coarse Grid Lookback method), highlighting its shortcomings and pitfalls. We then propose an improved practical technique based on a Brownian bridge.

8.1 The Coarse Grid Lookback Method and Its Shortcomings

Assume that portfolio simulation is not done daily, but instead only on a coarse grid $\{s_j\}$ where j runs from 1 to J. We use s rather than t to distinguish the model grid from the daily margin calculation grid.

In the classical model, the collateral depends only on the portfolio value at the start and at the end of the MPR, i.e., $s_j - \delta$ and s_j (see eqs. (4) and (5)), where the MPR is usually around $\delta = 10$ business days for CSAs with daily re-margining. To achieve acceptable computational performance, the time step of the coarse model grid, $s_j - s_{j-1}$, must be significantly greater than the length of the MPR. This, however, would preclude one from establishing the portfolio value at $s_j - \delta$. The Coarse Grid Lookback method deals with this issue by simply adding a second ”lookback” time point $s_j - \delta$ to all “primary” measurement times s_j, in effect replacing each node of the coarse model grid by a pair of closely spaced nodes. For each simulated portfolio path, the portfolio value at a lookback point is then used to determine the collateral available at the corresponding primary time point.

The Coarse Grid Lookback scheme causes, at worst, a factor of $\times 2$ slowdown relative to valuing the portfolio once per node of the coarse model grid. If even a $\times 2$ performance loss is not acceptable, a Brownian bridge constructed between the primary coarse grid nodes can be used to interpolate the value of the portfolio at each lookback point, see, for example, [8]. Notice that the use of Brownian bridge for this purpose should not be confused with its use in Section 8.2.

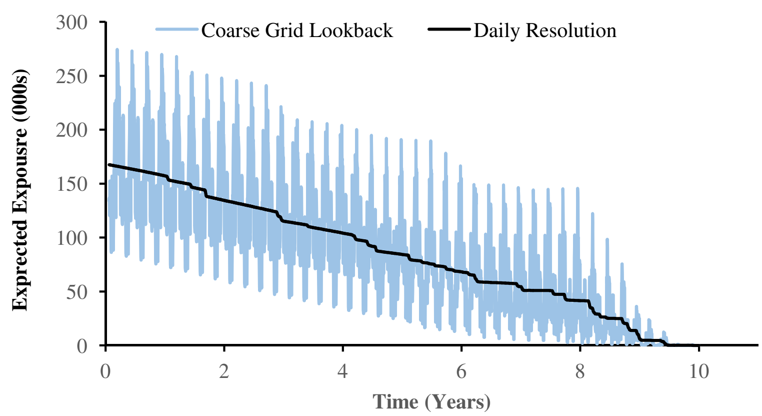
The Coarse Grid Lookback method is a common way of addressing the mismatch between the long time step of the coarse model grid and the much shorter MPR. Similarly to commonly used models of uncollateralized exposure, the method produces accurate (with respect to the underlying assumptions of the Classical model) exposure numbers at the coarse time grid points, but provides no information on exposure between the grid points. For uncollateralized positions, the exposure profiles are reasonably smooth, so one can safely interpolate between the grid points when calculating integral quantities, such as CVA. In the collateralized case, however, one cannot rely on such an interpolation because the true exposure profile, as we have seen, is likely to have spikes and jumps between the grid points. The Coarse Grid Lookback method has no means of determining the position or magnitude of these irregularities between the grid points and, thus, is not suitable for CVA or capital calculations.

To briefly expand on this, consider the Classical+ version of the classical model. Here it is assumed that all trade flows are paid within the MPR where, as we have shown in Section 7, trade flows often result in exposure spikes. Exposure profile computed with daily time steps would consequently show spikes from all trade flows until the maturity of the portfolio. In contrast, in a typical implementation with sparsely spaced MPRs, only trade flows that happen to be within one of the actually simulated MPRs may result in spikes; the exposure profile would then miss the spikes related to all other trade flows. Furthermore, as the location of simulation points likely will change with the advancement of calendar time, trade flows would move in and out of the simulated MPRs, and the exposure profile spikes that one will report on any given day may very well differ significantly from those that were reported the day before. This, in turn, causes CVA or risk capital to exhibit significant, and entirely spurious, oscillations.

While the Classical− exposure model does not exhibit outright spikes, its exposure profiles still exhibit jumps around significant trade flows; see Figure 2. The classical coarse-grained implementation would not be able to resolve the position of these jumps, instead only showing the cumulative jump between two adjacent exposure measurement points often separated by many months. This creates another source of instability, present in both Classical− and Classical+ versions of the classical model.

To illustrate the effects described above, let us define the concept of time t "forward" CVA, denoted CVA^t, obtained by i) changing the lower integration limit in (17) from 0 to t; and ii) dividing the result by $P(t)X(t)$. Using the same portfolio of 50 EUR-USD cross currency swaps as in Figure 7, Figure 8 below shows the t-dependence of CVA^t, on a daily grid to portfolio maturity.

Figure 8: Forward CVA for a Portfolio of 50 Cross-Currency Swaps



CVA^t for Classical+ method, as a function of time (t). Coarse Grid Lookback method (quarterly sampling) against brute force (daily sampling), for a portfolio of 50 cross currency swaps. Portfolio parameters and model setup are as in Figure 7.

As CVA is an integral of exposure, spikes in exposure profile should result in jumps rather than oscillations in CVA^t. However, when one of the Coarse Grid Lookback method's sparsely located "MPR windows" moves over a large trade flow, the contribution to CVA temporarily increases, only to drop back when the window moves past the large trade flow. This results in large oscillations of CVA^t shown in Figure 8. Such oscillations are spurious and their presence is highly unattractive when CVA is computed and reported as part of daily P&L.

8.2 Brownian Bridge Method

Overcoming the deficiencies outlined in Section 8.1 through brute-force daily simulation is, unfortunately, prohibitively expensive for large portfolios, mostly due to the expense of re-pricing the entire portfolio at each simulation path and each observation date. On the other hand, merely simulating risk factors at a daily resolution is generally feasible, as the number of the simulated risk factors is typically relatively small (e.g., several hundred) and the equations driving risk factor dynamics are usually simple. Furthermore, having produced risk factor paths on a daily grid, one can normally also produce all realized trade flows along each path because trade flows, unlike trade prices, are usually simple functions of the realized risk factors.

Based on these observations, we propose the following algorithm for generating paths of portfolio values and trade flows on a daily time grid:

1. Simulate paths of market risk factors with daily resolution.

2. For each path m, use the simulated risk factors to calculate trade flows on the path with a daily resolution.

3. For each path m and each coarse portfolio valuation time point s_j ($j = 1, \ldots, J$), use the simulated risk factors to calculate portfolio value on the path $V_m(s_j)$.

4. For each path m and each time point s_j, use the trade flows realized on the path between times s_{j-1} and s_j to calculate the "forward" to s_j portfolio values $V'_m(s_{j-1}; s_j)$:

$$V'_m(s_{j-1}; s_j) = V_m(s_{j-1}) - \mathrm{TF}_m^{net}\left(s_j; (s_{j-1}, s_j]\right) \tag{18}$$

   Note that $V'_m(s_{j-1}; s_j)$ is not a true forward value because we are subtracting the realized trade flows rather than forward trade flow values calculated at s_{j-1}.

5. For each path m and each exposure measurement time point s_j, compute the local variance $\sigma_m^2(t_{j-1})$ for the portfolio value "diffusion" $V_m(s_j) - V'_m(s_{j-1}; s_j)$ via a kernel regression estimator [21] conditional on the realized value of $V'_m(s_{j-1}; s_j)$. Here we use the term "diffusion" to indicate that the portfolio value change has been defined to avoid any discontinuities resulting from trade flows.

6. For each path m and each exposure measurement time point s_j, simulate an independent, daily sampled, Brownian bridge process[22] that starts from value $V'_m(s_{j-1}; s_j)$ at time s_{j-1} and ends at value $V_m(s_j)$ at time s_j. The volatility of the underlying Brownian motion should be set equal to $\sigma_m(s_{j-1})$.

7. For each path m and each exposure measurement time point s_j, the portfolio values for each time u of the daily grid in the interval (s_{j-1}, s_j) are approximated from the simulated Brownian bridge $\mathrm{BB}_m(u)$ by adding the trade flows realized along the path m between times u and s_j:

$$V_m^{approx}(u) = \mathrm{BB}_m(u) + \mathrm{TF}_m^{net}\left(s_j; (u, s_j]\right). \tag{19}$$

[21] E.g., the Nadaraya-Watson Gaussian kernel estimator, see [7] or [12]. The selection of bandwidth for the kernels is covered in, e.g., [6]. In our numerical results, we use *Silverman's Rule of Thumb*, see [11].

[22] See, for instance, Chapter 3 of [5].

In a nutshell, the algorithm above uses a Brownian bridge process to interpolate portfolio values from a coarse grid in a manner that ensures that intermediate trade flow events are handled accurately. The algorithm produces paths of portfolio values and trade flows on a daily time grid, wherefore exposure can be calculated as described in Section 7.3 with daily resolution and overlapping MPRs. Furthermore, daily sampling allows for further refinements of the proposed model by consistently incorporating thresholds, minimum transfer amount, and rounding.

A key assumption made by the Brownian bridge algorithm is that the portfolio value process within the interpolation interval is a combination of an approximately normal "diffusion" overlaid by the realized trade flows. For Wiener process models without risk factor jumps, this assumption is accurate in the limit of infinitesimal interpolation interval, and is often a satisfactory approximation for monthly or even quarterly interpolation steps. Nevertheless, the presence of trade flows that depend on values of risk factors between the end points introduces two types of error:

1. Suppose that there is a trade flow between the end points that depend on a risk factor value at the date when it is paid. The independence of the Brownian bridge process from the risk factor processes that drive that trade flow would result in an error in the expected exposure profile around the trade flow date. This error is the largest for trade flows in the middle of the interpolation interval and disappears for trade flows near the ends of the interval.

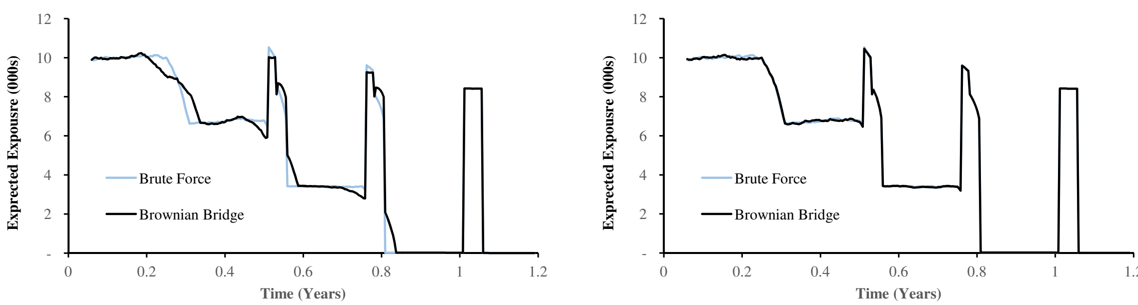
2. Suppose that there is there is a trade flow that occurs after the end point of the interpolation interval, but whose value entirely depends on the value of a risk factor within the interpolation interval. A typical example would be a vanilla interest rate swap when the floating leg payment being paid after the end of the interpolation interval depends on the interest rate on a date within the interval. Even in the absence of a trade flow within the interpolation interval, the volatility of swap value drops at the floating rate fixing date as some of the uncertainty is resolved. Thus, the "true" swap value process has two volatility values: a higher value before the rate fixing date and a lower value after the rate fixing date. In contrast, the approximation algorithm assumes a single value of volatility obtained via kernel regression between the end points. Similarly to the decorrelation error discussed above, the error resulting from this volatility mismatch is the largest for fixing dates in the middle of the interpolation interval and disappears for fixing dates near the end points.

To illustrate the two errors above, Panel (a) in Figure 9 shows the expected exposure profile for a one-year interest rate swap when a monthly grid for full valuation is situated so that the payment/fixing dates sit roughly in the middle of the interpolation interval, thus maximizing the error of the Brownian bridge algorithm. While there, as expected, are some errors around the trade flow dates, they are acceptable in magnitude and overall unbiased, in the sense that over-estimation of exposure is about as frequent as under-estimation of exposure. For, say, CVA purposes, the Brownian bridge results would therefore be quite accurate.

Panel (b) in Figure 9 shows expected exposure profiles when the monthly valuation points are aligned with rate fixing/payment dates. In this case, the Brownian bridge approximation is nearly exact. Of course, in practice such alignment is only possible for a single trade or a small netting set, and not for a large portfolio where trade flows will occur daily. Yet, even for large netting sets the calculation accuracy will improve if interpolation pillars are aligned with the largest trade flows (e.g. principal exchange date for the largest notional amounts). In practice,

Figure 9: Expected Exposure for 1-Year Interest Rate Swap

(a) Grid not aligned with trade flows (b) Grid aligned with trade flows



USD10MM 1-year payer swap, 2% coupon; fixed and floating coupons paid quarterly. Simulation and exposure setup as in Figure 2, with Conservative calibration used everywhere. Panel (a) places the full valuation grid approximately half-way between trade flow dates; Panel (b) aligns the full valuation grid on trade flow dates. The Brownian bridge method uses monthly valuation points; kernel regression for volatility estimation in the bridge is based on a Gaussian kernel with bandwidth determined by Silverman's rule of thumb (see [11]).

errors can typically be expected to be somewhere between the two extremes of panels (a) and (b) in Figure 9.

While the exact speedup provided by the Brownian bridge method depends on the implementation, for most portfolios the overhead of building the Brownian bridge at daily resolution is negligible compared to computing the exposure on the model's coarse grid. In this case the computational effort of the daily Brownian bridge method is about half of the computational effort of the Coarse Grid Lookback method, as the former method does not require adding a "lookback" point to each of the primary coarse grid points. Thus the Brownian bridge method is both faster and significantly more accurate than the standard Coarse Grid Lookback method.

9 Initial Margin

The posting of initial margin (IM), in addition to regular variation margin (VM) collateral, provides banks with a mechanism to gain additional default protection. The practice of posting IM has been around for many years, typically with IM being computed at trade inception on a travel-level basis. This type of IM is entirely deterministic and normally either stays fixed over the life of the trade or amortizes down according to a pre-specified schedule. As a consequence, modeling the impact on exposure is trivial: for the exposure points of interest, all trade-level IM amounts are summed across the netting set and the total (which is the same for all paths) is subtracted from the portfolio value on each path.

A more interesting type of IM is dynamically refreshed to cover portfolio-level close-out risk at some high percentile, often 99%. This type of margin is routinely applied by Clearinghouses (CCPs) and by margin lenders, and will also soon be required by regulators for inter-dealer OTC transactions. In particular, in 2013 the BCBS and IOSCO issued (see [1]) a final framework on margin requirements, under which any two covered entities[23] that are counterparties in non-

[23]Covered entities include all financial firms and systemically important non- financial firms. Central banks and

centrally cleared derivatives are required to i) exchange VM under a zero threshold margin agreement; and ii) to post IM to each other without netting the amounts. IM must be held in a default-remote way, e.g., by a custodian, so that IM posted by a counterparty should be immediately available to it should the other counterparty default.

Under the BCBS and IOSCO rules, regulatory IM can be calculated either by an internal model or by look-up in a standardized schedule. If an internal model is used, the calculation must be made at the netting set level as a the value-at-risk for a 99% confidence level. The horizon used in this calculation equals 10 business days for daily exchange of VM or 10 business days plus the re-margining period for less frequent exchange of VM. Diversification across distinct asset classes is not recognized, and the IM internal model must be calibrated to a period of stress for each of the asset class. The required levels of the IM will change as cash flows are paid, new trades are booked or markets move. To accommodate this, banks would either call for more IM or return the excess IM.

For trades done with CCPs or under BCBS and IOSCO rules, one must find a way to estimate the future IM requirements for each simulated path. No matter how simple the IM VaR model is, it will likely be difficult to perform such calculation in practice if one wants to incorporate all the restrictions and "twists" of the IM rules: stress calibration, limited diversification allowance, and, for CCPs, add-ons for credit downgrades and concentration risk. However, it is possible to utilize our model to calculate counterparty exposure if one ignores these complications. Note that ignoring such complications is conservative, as it would virtually always lead to a *lower* level of IM and, therefore, to a *higher* level of exposure.

To calculate exposure at time t, we have assumed that the last observation date for which C would deliver VM to B is $t_C = t - \delta_C$. It is reasonable to assume that this date will also be the last observation date for which C would deliver IM to a custodian. To simplify modeling, we assume that the custodian would not return any amount of IM to C for observation dates after $t - \delta_C$. Thus, to calculate exposure at time t, IM on a path has to be estimated from the dynamics of the exposure model as of time $t - \delta_C$.

Assuming, as is common in practice, that portfolio values are locally Gaussian, it suffices to know the local volatility of the portfolio value for the period $[t - \delta_C, t]$ estimated at $t - \delta_C$. Denoting the IM horizon by δ_{IM} and the local volatility of portfolio value at time u on path m via $\sigma_m(u)$, the IM available to B at the ETD date t on path m is given by

$$\text{IM}_m(t - \delta) = \sigma_m(t - \delta)\sqrt{\delta_{\text{IM}}}\Phi^{-1}(q) \tag{20}$$

where q is a confidence level (often 99%) and $\Phi^{-1}(\cdot)$ is the inverse of the standard normal cumulative distribution function.

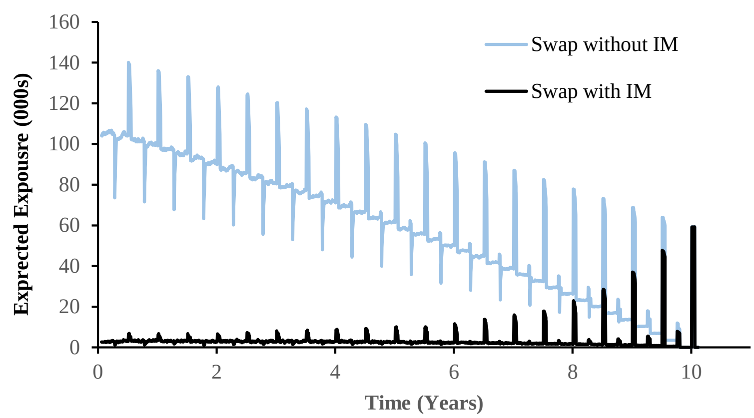
Estimation of the local volatility $\sigma_m(t-\delta)$ can be done via kernel regression, as in Section 8.2. If the portfolio value is simulated at both $t - \delta_C$ and t, the kernel regression for $\sigma_m(t - \delta)$ could be run on the "P&L" $V(t) - V(t - \delta_C) + \text{TF}^{net}(t; (t - \delta_C, t])$ conditional on realization of portfolio value on path m at the beginning of the MPR, $V_m(t - \delta_C)$. If one does not calculate portfolio value at the beginning of the MPR, but uses the fast approximation of Section 8.2 instead, $\sigma_m(t - \delta)$ can be set equal to the local volatility estimated for the time interval that encloses the given MPR $[t - \delta_C, t]$. Thus, our Brownian bridge framework can now produce not only the collateralized exposure under VM alone, but also a reasonable estimate of the collateralized exposure under a combination of VM and IM.

sovereigns are not covered entities.

In calculating the impact of IM, an important consideration is the timing and mechanics of adjustment to the IM when C misses a margin flow or a trade flow. For instance, when a large trade reaches maturity, the portfolio VaR may be reduced, in which case some of the IM posted by C must be refunded. The issue of whether this refund can be delayed due to an ongoing margin dispute is not yet fully resolved. To simplify the calculations, we have assumed that no part of IM is returned to C during the MPR.

To show some numerical results, we can consider the individual trades and portfolios in Section 7.3. Our first example uses the 10y vanilla swap from Figure 4, for which the impact of IM on exposure is shown in Figure 10.

Figure 10: Expected Exposure for 10-Year Vanilla Swap



USD10MM 10-year payer swap, 2% coupon; fixed coupons paid semi-annually, floating coupons paid quarterly. Model and exposure setup as in Figure 4, except that the graph "Swap with IM" uses risk-based IM to protect the swap, per Equation (20) with $q = 99\%$. The Conservative exposure model calibration is used everywhere.

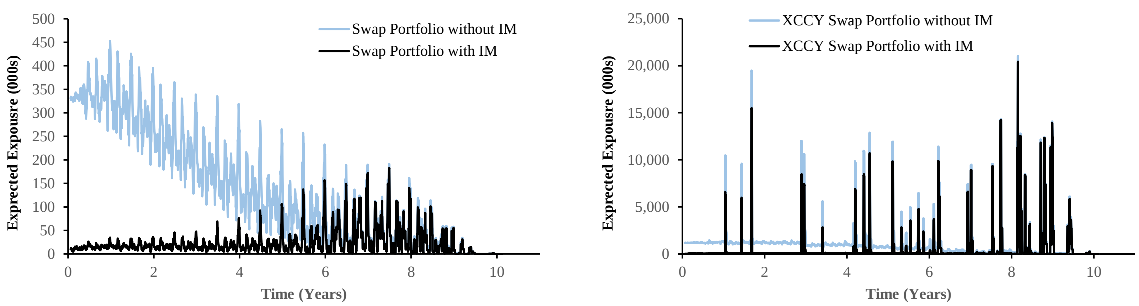
As evident from the figure, the IM mechanism strongly reduces exposure away from trade flows, but near trade flow dates the protection gets progressively weaker toward trade maturity, and disappears almost completely for the last couple of trade flows. The reason for this uneven benefit of IM on this trade is that the 10-day VaR of the trade bears no direct relationship to the size of trade flows that determine exposure spikes in our model. The variance of the "P&L" is reduced as the swap approaches maturity, so the amount of IM on a given path is also reduced. However, the size of the trade flows is not reduced, but can actually grow with the simulation time (as larger and larger realizations of the floating rate are possible). Thus, when the swap approaches maturity, the amount of IM is greatly reduced relative to the trade flows, so exposure spikes grow larger, while the "diffusion" component of exposure becomes smaller.

The impact of IM on the vanilla swap and cross currency swap portfolios described in Section 7.3 are shown in Figure 11. As in Figure 10, the IM strongly suppresses the "diffusion" component of portfolio value changes, but proves inadequate in reducing the spikes of exposure for both the single currency and, especially, the cross currency portfolio.

Figure 11: Expected Exposure for Swap and XCCY Swap Portfolios

(a) Regular Interest Rate Swaps

(b) Cross Currency Swaps



Exposure Profiles for the Portfolios in Figures 6 and 7, respectively. Each panel contains exposures both with and without IM protection.

10 Conclusion

Industry standard models for collateralized credit risk are well-known to produce non-negligible counterparty credit exposure, even under full variation margin collateralization. This exposure essentially arises due to inevitable operational and legal delays (margin period of risk, or MPR) that are "baked into" the workings of the ISDA contracts that govern OTC trading.

In the most common industry implementation, the length of the MPR, and what precisely transpires inside it, is, however, often treated in highly stylized fashion. Often the MPR is set equal to 10 business days for little reason other than tradition, and often counterparties are assumed to have oddly synchronized behavior inside the MPR. For instance, one common approach (denoted Classical−) assumes that margin and trade flows by both counterparties terminate simultaneously, at the beginning of MPR. Another approach (Classical+) assumes that margin flows terminate at the beginning of the MPR, but trade flows terminate (simultaneously) at the end of the MPR. Surprisingly, the Classical+ and Classical− approaches continue to co-exist in the market, and neither have become the sole market practice. We speculate that one reason for this state of affairs is that the two models correspond to different choices for a trade-off between implementation complexity and model stability: Classical+ is easier to implement but it is prone to spurious spikes in daily CVA P&L (as demonstrated in Section 8.1), whereas the Classical− model is more difficult to implement but is free from such spikes.

Ultimately, a model should, of course, be selected not on the basis of implementation ease or the properties of a specific numerical technique, but on the basis of how well[24] the model captures the legal and behavioral aspects of events around a counterparty default. To this end, even a cursory analysis suggests that the perfect synchronicity of the Classical± methods cannot be supported in reality. For instance, due to the ways a CSA works in practice, a non-defaulting party will need at least 3 business days after a portfolio valuation date to determine for sure that the corresponding margin payment by its counterparty will not be honored.

[24]The term "well" can mean different things in different applications of the exposure model. For regulatory capital purposes, prudence and conservatism may, for instance, be as important as outright precision.

In this paper, we carefully dissected the MPR into a full timeline around the the default event, starting with a missed margin call and culminating on the post-default valuation date on which the termination value of the portfolio is established. For modeling purposes, we have condensed this timeline into four[25] model parameters, each specified as the number of business days prior to termination for the following events:

1. The last market data measurement for which the margin flow is received (δ_C) and paid (δ_B) as prescribed.
2. The last date when the defaulting party (δ'_C) and the bank (δ'_B) make trade flow payments as prescribed.

As we show, each of these four parameters has a precise legal and/or operational interpretation, enabling calibration from the terms of the CSA and from the operational setup of the bank. Note that the proposed model parameterization includes Classical+ and Classical− models as limit cases.

For indicative purposes, we describe two particular calibration choices for our model, denoted Aggressive and Conservative. The former assumes that the non-defaulting bank always operates at an optimal operational level and will will enforce the legal provisions of the ISDA legal contracts as strictly as possible. The latter allows for some slack in the operations of the bank, to allow for manual checks of calculations, legal reviews, "gaming" behavior of the counterparty, and so forth. The Conservative model setting obviously produces higher exposure than the Aggressive setting, for the following reasons:

1. The Conservative setting has a longer overall length of MPR.
2. The Conservative setting has a longer margin flow gap period where the bank pays, but does not receive, margin flows.
3. The Conservative setting, unlike the Aggressive setting, contains a trade flow gap period where the bank pays but does not receive trade flows.

In our numerical tests, the first two factors cause the Conservative setting to have approximately twice the exposure of both the the Aggressive setting and the Classical± settings away from the dates of large trade flows. The last factor (i.e. the presence of a trade flow gap) may cause exposure spikes of extremely large magnitude under the Conservative calibration. Despite the fairly short duration of these spikes, they may easily add up to very significant CVA contributions, especially for cross currency trades with principal exchange ("Herstatt risk"). Credit losses due to such trade flow gaps materialized in practice during the financial crisis (especially due to the Lehman Brothers default), so their incorporation into our exposure model is both prudent and realistic.

Detailed tracking of margin and trade flow payments requires stochastic modeling of the trade portfolio on a daily grid. As brute-force simulations at such a resolution is often impractically slow, it is important that numerical techniques be devised to speed up calculations. While the focus of our paper was mainly on establishing fundamental principles for margin exposure, we also proposed an acceleration method based on kernel regression and a Brownian bridge applied to portfolio values "stripped" of cash flows. For ordinary and cross-currency swaps, we demonstrate that this method is both accurate and much faster than either brute-force simulation

[25] In contrast, the classical model has only one parameter, the full length of the MPR.

or standard acceleration techniques for the classical model. Further improvements of acceleration techniques, and expansion of applicability into more exotic products, is an area of future research.

Under suitable assumptions, kernel regression may also be used to embed risk-based initial margin into the exposure simulation. As we demonstrated in the final section of the paper, initial margin at a 99% level succeeds in greatly reducing the bilateral exposure for the Classical− model calibration. For all other calibration choices, and especially for the Conservative setting, the reduction in counterparty exposure afforded by the initial margin fails around the time of large trade flows, when the sudden change of exposure following a trade flow exceeds the initial margin level. Note that the (already inadequate) level of IM protection around trade flows deteriorates further towards the maturity of portfolio, where the local volatility of portfolio value decreases, but trade flows do not. Overall, accurate modeling of events within the MPR becomes critically important for portfolios covered by dynamic IM.

References

[1] Basel Committee on Banking Supervision and International Organization of Securities Commissions (2013), "Margin Requirements for Non-Centrally Cleared Derivatives,"September.

[2] Böcker, K. and B. Schröder (2011), "Issues Involved in Modelling Collateralized Exposure," Presentation at Risk Minds conference, Geneva, December.

[3] Brigo, D., A. Capponi, A. Pallavicini and V. Papatheodorou (2011), "Collateral Margining in Arbitrage-Free Counterparty Valuation Adjustment including Re-Hypotecation and Netting," Working Paper (http://arxiv.org/pdf/1101.3926v1.pdf).

[4] Gibson, M. (2005), "Measuring Counterparty Credit Exposure to a Margined Counterparty," in *Counterparty Credit Risk Modelling* (M. Pykhtin, ed.), Risk Books

[5] Glasserman, P. (2004), *Monte Carlo Methods in Financial Engineering*, Springer Verlag.

[6] Jones, M., J. Marron and S. Sheather (1996), "A Brief Survey of Bandwidth Selection for Density Estimation," *Journal of the American Statistical Association*, 91 (433), pages 401-407.

[7] Nadaraya, E. A. (1964), "On Estimating Regression,"*Theory of Probability and its Applications*, 9(1), pages 141-142.

[8] Pykhtin, M. (2009), "Modeling Credit Exposure for Collateralized Counterparties,"*Journal of Credit Risk*, 5 (4) (Winter), pages 3-27.

[9] Pykhtin, M. (2010), "Collateralized Credit Exposure," in *Counterparty Credit Risk*, (E. Canabarro, ed.), Risk Books.

[10] Pykhtin, M. and A. Sokol (2013), "Exposure under Systemic Impact," *Risk Magazine*, September, pages 88-93.

[11] Silverman, B. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London.

[12] Watson, G. S. (1964), "Smooth Regression Analysis", *Sankhya: The Indian Journal of Statistics*, Series A 26(4), pages 359-372.