# Credit Exposure in the Presence of Initial Margin*†

Leif Andersen
*Bank of America Merrill Lynch*

Michael Pykhtin
*Federal Reserve Board*

Alexander Sokol
*CompatibL*

First version: July 22, 2016. This version: May 7, 2017

**Abstract**

We leverage the new framework for collateralized exposure modelling in [1] to analyze credit risk on positions collateralized with both variation and initial margin. Special attention is paid to the dynamic BCBS-IOSCO uncleared margin rules soon to be mandated for bilateral inter-dealer trading in OTC derivatives markets. While these rules set initial margin at a 99th 2-week percentile level and aim to all but eliminate portfolio close-out risk, we demonstrate that trade flow effects can result in exposures being reduced significantly less than expected. We supplement our analysis with several practical schemes for estimating IM on a simulation path, and for improving the speed and stability of exposure simulation. We also briefly discuss potential ways to adjust the margin framework to more effectively deal with exposures arising from trade flow events.

## 1 Introduction

Collateralization has long been a standard technique of mitigating counterparty risk in OTC bilateral trading. The most common collateral mechanism is *variation margin* (VM), which aims to keep the exposure gap between portfolio value and posted collateral below certain, possibly stochastic, thresholds. Even when the thresholds for VM are set to zero, however, there remains residual exposure to the counterparty's default resulting from a sequence of contractual[1] and operational time lags, starting from the last snapshot of the market for which the counterparty would post in full the required VM, to the termination date after the counterparty's default. The aggregation of these lags results in a time period, known as the *margin period of risk* (MPoR),

*Source code for the model in [1] is 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 2016 Initial Margin, XVA & KVA; 2016 Quant Summit Europe; and 2016 Global Derivatives. All remaining errors are our own.

[1]The various collateral mechanisms, including the precise definition of variation margin thresholds, are typically captured in an ISDA Credit Support Annex (CSA), a portfolio-level legal agreement that supplements an ISDA Master Agreement.




Electronic copy available at: https://ssrn.com/abstract=2806156

during which the gap between the portfolio value and the collateral can widen. The length of the MPoR is a critical input to any model of collateralized exposure.

Posting of *initial margin* (IM) to supplement[2] VM provides banks with a mechanism to reduce the residual exposure resulting from market risk over the MPoR. Historically, IM in bilateral trading has mostly been reserved for bank counterparties deemed as high-risk (e.g., hedge funds) and typically done as a trade level calculation, established in term sheets at the transaction time of each trade. This type of IM posting is normally deterministic and either stays fixed over the life of the trade or amortizes down according to a pre-specified schedule.

In the inter-bank bilateral OTC world, changes to long-standing VM and IM collateral practices are now imminent. In 2013, the BCBS and IOSCO proposed (see [2])) and in 2015 finalized (see [3]) new uncleared margin rules (UMR) for bilateral trading. Under UMR, VM thresholds are forced to zero, and IM must be posted bilaterally (into segregated accounts) at the netting set level, either through an internal model or by look-up in a standardized schedule. If an internal model is used, the IM must be calculated as the netting set value-at-risk (VaR) for a 99% confidence level. The horizon used in this calculation equals $9 + a$ business days, where $a$ is the re-margining period (1 business day under US rules). In these calculations, diversification across distinct asset classes is not recognized, and calibration of an IM internal model for each asset class must include a period of stress for that asset class[3].

Under UMR, required levels of IM continuously change as trade cash flows are paid, new trades are booked, or markets move, and banks regularly need to either call for more IM or to return excess IM. This dynamic aspect of IM requirements makes the modeling of future exposure a challenge.

In this paper, we discuss modeling credit exposure in the presence of dynamic IM and question the conventional wisdom that IM essentially eliminates counterparty risk. Leaning on recent results in [1], we start by formulating a general model for exposure in the presence of VM and/or IM. The resulting framework is first applied to a simple case where no trade flows take place within the MPoR. For processes with Gaussian increments (e.g., an Ito process), we derive a limiting scaling factor that converts IM-free expected exposure (EE) to IM-protected EE, for sufficiently small MPoR. This universal value depends only on the IM confidence level and the ratio of the IM horizon to the MPoR[4]; it equals 0.85% at the BCBS-IOSCO confidence level of 99%, provided that the IM horizon equals the MPoR. While some deviations from this universal limit value due to a non-infinitesimal MPoR are to be expected, the reduction of the EE by about two orders of magnitude is, as we demonstrate, generally about right when no trade flows are present within the MPoR.

For those periods where trade flows do take place within the MPoR, however, any scheduled trade payment flowing away from the bank will result in a spike in the EE profile. Without IM, these spikes can make a fairly significant contribution to the Credit Valuation Adjustment (CVA) – say, 20% of an interest rate swap's total CVA may originate with spikes – but the CVA would still mostly be determined by the EE level between the spikes. We show that while IM is effective

[2]While it is often believed that IM is posted strictly *in addition* to VM, many CSAs intermingle the two types of collateral by letting IM affect the threshold computation of VM.

[3]To reduce the potential for margin disputes and to increase overall market transparency, ISDA has proposed a standardized sensitivity-based IM calculator known as SIMM (Standard Initial Margin Model); see [4]. As a practical matter, it is expected that virtually all banks will use SIMM for their day-to-day IM calculations.

[4]While conceptually the two are identical, a prudent MPoR for internal calculations may differ from the regulatory minimum IM horizon.




Electronic copy available at: https://ssrn.com/abstract=2806156

in suppressing EE *between* spikes, it will often fail to significantly suppress the spikes themselves. As a result, the relative contribution of the spikes to CVA is greatly increased in the presence of IM – e.g., for a single interest rate swap, the spikes' contribution to CVA can be well above 90% for a position with IM. Accounting for spikes, we find that IM reduces CVA by much less than the two orders of magnitude one might expect, with the reduction for interest rate swaps often being less than a factor of 10.

In the final part of our paper, we discuss practical approaches to calculating EE profile in the presence of IM. The first step in this calculation is the estimation of IM on the simulation paths, which can be done either by parametric regression or by kernel regression. When IM covers an insignificant number of trades in the netting set, IM calculated on a path can then be subtracted from the no-IM exposure realized on that path to generate EE profiles. However, when most trades of the netting set are covered by IM, this approach can be problematic because of excessive simulation noise and other errors. We propose an alternative approach that dampens noise and is generally more accurate.

We conclude the paper by summarizing our results and briefly discussing possible modifications to trade and collateral documentation that would make IM more effective in reducing residual counterparty credit risk.

## 2 Exposure in the Presence of IM and VM

Let us consider a bank $B$ that has a portfolio of OTC derivative contracts traded with a counterparty $C$. Suppose for simplicity that the entire portfolio is covered by a single netting agreement, which is supported by a margin agreement that includes VM and may include IM on a subset of the portfolio. Quite generally, exposure of $B$ to default of $C$ measured at time $t$ (assumed to be the early termination time after $C$'s default) is given by

$$E(t) = (V(t) - \mathrm{VM}(t) + U(t) - \mathrm{IM}(t))^+ , \tag{1}$$

where $V(t)$ is the time $t$ portfolio value from B's perspective; $\mathrm{VM}(t)$ is the VM available to $B$ at time $t$; $U(t)$ is the value of trade flows scheduled to be paid by both $B$ (negative) and $C$ (positive) up to time $t$, yet unpaid as of time $t$; and $\mathrm{IM}(t)$ is the value of IM available to $B$ at time $t$. Notice that VM can be positive ($C$ posts VM) or negative ($B$ posts VM) from $B$'s perspective. On the other hand, IM is always positive as IM for both sounterparties is kept in segregated accounts, whence IM posted by $B$ does not contribute to $B$'s exposure to the default of $C$.

Equation (1) specifies exposure of $B$ to $C$ in a generic, model-free way. To add modeling detail, we will assume that $B$ and $C$ post VM with zero thresholds and are required to post BCBS-IOSCO compliant IM to a segregated account. We then deal with the modeling of each of the terms VM, $U$, and IM in turn.

### 2.1 Modeling VM

The length of the MPoR, denoted by $\delta_C$, defines the last portfolio valuation date $t_C = t - \delta_C$ prior to the termination date $t$ (after $C$'s default) for which $C$ delivers VM to $B$. A common assumption, which we here denote the *Classical model*, assumes that $B$ stops posting VM to $C$ at the exact same time $C$ stops posting to $B$. That is, $\mathrm{VM}(t)$ in (1) is the VM prescribed by the margin agreement for portfolio valuation date $t_C = t - \delta_C$. Ignoring minimum transfer amount and rounding, the

prescribed VM in the Classical model is thus simply

$$\mathrm{VM}_{\mathrm{cl}}(t) = V(t_C) = V(t - \delta_C). \tag{2}$$

In the *Advanced model* of [1], operational aspects (and gamesmanship) of margin disputes are considered in more detail, leading to the more realistic assumption that $B$ may continue to post VM to $C$ for some period of time, even after $C$ stops posting. The model introduces another time parameter $\delta_B \leq \delta_C$ that specifies the last portfolio valuation date $t_B = t - \delta_B$ for which $B$ delivers VM to $C$. For portfolio valuation dates $T_i \in [t_C, t_B]$, $B$ would post VM to $C$ when the portfolio value decreases, but will receive no VM from $C$ when the portfolio value increases. This results in VM of

$$\mathrm{VM}_{\mathrm{adv}}(t) = \min_{T_i \in [t_C, t_B]} V(T_i). \tag{3}$$

Equation (3) of course reduces to (2) when one sets $\delta_B = \delta_C$.

### 2.2 Modeling $U$

In the most conventional version of the Classical model, here denoted "Classical+", it is assumed that all trade flows are paid without incident by both $B$ and $C$ for the entire MPoR up to and including the termination date (i.e. in the time interval[5] $(t_C, t]$). This assumption simply amounts to setting

$$U_{\mathrm{cl+}}(t) = 0. \tag{4}$$

One of the prominent features of the Classical+ model is that the time 0 expectation of $E(t)$, denoted $\mathrm{EE}(t)$, will contain upward spikes whenever there is a possibility of trade flows from $B$ to $C$ within the interval $(t_C, t]$. The spikes appear because, by the Classical model's assumption, $C$ makes no margin payments during the MPoR and consequently would fail to post an offsetting VM amount to $B$ after $B$ makes a trade payment to[6] $C$; see also Section 4.1. For banks having a sparse fixed exposure time grid, the alignment of grid nodes relative to trade flows will add numerical artifacts to genuine spikes, causing spikes in the EE profile appear and disappear as the calendar date moves. As a consequence, an undesirable instability in EE and CVA is introduced.

An easy way to eliminate exposure spikes is to assume that neither $B$ nor $C$ makes any trade payment within the MPoR. The resulting model, here denoted Classical−, consequently assumes that

$$U_{\mathrm{cl-}}(t) = \mathrm{TF}^{net}\left(t; (t_C, t]\right), \tag{5}$$

where $\mathrm{TF}^{net}\left(t; (s, u]\right)$ denotes the time $t$ value of all net trade flows payable on the interval $(s, u]$. The Classical− model avoids the EE and CVA instabilities in the Classical+ model, and is used by a significant number of banks.

It should be evident that neither Classical+ nor Classical− assumptions on trade flows are entirely realistic: in the beginning of the MPoR both $B$ and $C$ are likely to make trade payments, while at the end of the MPoR neither $B$ nor $C$ are likely to make trade payments. To capture this behavior, [1] introduce two more time parameters in the model, $\delta'_C$ and $\delta'_B \leq \delta'_C$, that specify the

[5] We measure time in discrete business day units, so $(u, s]$ is equivalent to $[u + 1\mathrm{bday}, s]$.

[6] In the Classical model, $B$ will also not post VM to $C$ in the event of a trade payment from $C$ to $B$, which results in a negative jump in exposure. However, these scenarios do not fully offset the scenarios where $B$ makes a trade payment because the zero floor in the exposure definition effectively limits the size of the downward exposure jump.

last dates, $t'_C = t - \delta'_C$ for $C$ and $t'_B = t - \delta'_B$ for $B$, at which trade payments are made prior to closeout at $t$. This results in an unpaid trade flows term of

$$U_{\text{adv}}(t) = \text{TF}^{C\to B}\left(t; (t'_C, t'_B]\right) + \text{TF}^{net}\left(t; (t'_B, t]\right), \tag{6}$$

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

The EE profiles obtained with the Advanced model contain spikes that typically are narrower and have a more complex structure than spikes under the Classical+ model; see [1] and Section 4.1 for more details. Rather than being unwelcome noise, it is argued in [1] that spikes in EE profiles are important features that represent actual risk. To eliminate any numerical instability associated with spikes, an approximation was proposed in [1] for calculation of EE on a daily time grid without daily re-valuation of the portfolio.

### 2.3 Modeling IM

Following the BCBS-IOSCO restrictions on diversification, we define IM for a netting set as the sum of IMs over $K$ asset classes[7]:

$$\text{IM}(t) = \sum_{k=1}^{K} \text{IM}_k(t). \tag{7}$$

Let $V_k(t)$ denote the value at time $t$ of all trades in the netting set that belong to asset class $k$ and are subject[8] to IM requirements. For an asset class $k$, we define IM as the quantile at confidence level $q$ of the "clean" portfolio value increment over the BCBS-IOSCO IM horizon $\delta_{\text{IM}}$ (which may or may not coincide with $\delta_C$), conditional on all the information available at time $t_C$. That is,

$$\text{IM}_k(t) = \text{Q}_q\left(V_k(t_C + \delta_{\text{IM}}) + \text{TF}_k^{net}\left(t_C + \delta_{\text{IM}}; (t_C, t_C + \delta_{\text{IM}}]\right) - V_k(t_C) | \mathcal{F}_{t_C}\right), \tag{8}$$

where $\text{Q}_q\left(\cdot|\mathcal{F}_s\right)$ denotes the quantile at confidence level $q$, conditional on information available at time $s$. Note that (8) assumes that $C$ stops posting IM at the same time it stops posting VM; hence IM is calculated as of $t_C = t - \delta_C$.

### 2.4 Summary and Calibration

To summarize, we have outlined three different models for the generic exposure calculation (1): Classical+, Classical−, and Advanced. Collecting results, we have

$$E_{\text{cl+}}(t) = \left(V(t) - V(t - \delta_C) - \text{IM}(t)\right)^+, \tag{9}$$

$$E_{\text{cl−}}(t) = \left(V(t) - V(t - \delta_C) + \text{TF}^{net}\left(t; (t_C, t]\right) - \text{IM}(t)\right)^+, \tag{10}$$

$$E_{\text{adv}}(t) = \left(V(t) - \min_{T_i \in [t_C, t_B]} V(T_i) + \text{TF}^{C\to B}\left(t; (t'_C, t'_B]\right) + \text{TF}^{net}\left(t; (t'_B, t]\right) - \text{IM}(t)\right)^+, \tag{11}$$

where for all three models $\text{IM}(t)$ is computed as in (7)-(8).

In practice, calibration of the time parameters of the Advanced model should be informed by both the bank's legal rights and its aggressiveness in pursuing them. For the former, we refer to

[7] Current rules have $K = 4$: Rates/FX, Credit, Equity, and Commodity.

[8] Note that only trades executed after the UMR go-live date will be covered by IM. The case where only a subset of a netting set is covered by IM will therefore be common in the near future.

[1] for a full discussion of ISDA Master Agreements, but note that once notice of a potential event of default (PED) has been served, the "suspension rights" of the ISDA Master Agreement (Section 2(a)(iii)) and its accompanying Credit Support Annex (Paragraph 4(a)) allow a bank to suspend all trade- and collateral-related payments to its counterparty until the PED has been cured. The extent to which suspension rights are actually exercised in practice, however, depends on the bank and its operating model, as well as the specifics of each PED. One particular danger in aggressive and immediate enforcement of suspension rights is that the original PED might be ruled unjustified. 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.

With four time parameters, the Advanced model allows a bank great flexibility in modeling its risk tolerance and procedures for exercising its suspension rights. A bank may, in fact, calibrate these parameters differently for different counterparties, to reflect its risk management practices towards counterparties of a given type[9]. Additional discussion can be found in [1], which also provides some prototypical parameter settings: the Aggressive calibration ($B$ can always sniff out financial distress in its counterparties and is swift and aggressive in enforcing its legal rights) and the Conservative calibration ($B$ is deliberate and cautious in enforcing its rights, and acknowledges the potential for operational errors and for rapid, unpredictable deterioration in counterparty credit). For the numerical results in this paper, we mostly set the values of the time parameters to be in-between Aggressive and Conservative.

## 3 The impact of IM: No Trade Flows within the MPoR

In this section, we examine the impact of IM on EE when there are no trade flows within the MPoR. For simplicity, we work with the Classical model[10], and assume that the entire netting set is covered by IM and is comprised of trades all belonging to the same asset class.

In the absence of trade flows on $(t_C, t]$, (9)-(10) show that the Classical model computes expected exposure as

$$\mathrm{EE}(t) = \mathrm{E}\left[\left(V(t) - V(t_C) - \mathrm{Q}_q\left(V_k(t_C + \delta_{\mathrm{IM}})|\mathcal{F}_{t_C}\right)\right)^+\right], \quad t_C = t - \delta_C, \tag{12}$$

where $\mathrm{E}(\cdot)$ is the expectation operator. In the absence of of IM this expression would be $\mathrm{EE}_0(t) = \mathrm{E}([V(t) - V(t_C)]^+)$; we are interested in establishing meaningful estimates of the IM "efficiency ratio"

$$\lambda(t) \triangleq \frac{\mathrm{EE}(t)}{\mathrm{EE}_0(t)}. \tag{13}$$

### 3.1 Local Gaussian Approximation

Suppose that the portfolio value $V(t)$ follows an Itô process:

$$dV(t) = \mu(t)\,dt + s(t)^\top\,dW(t)$$

where $W(t)$ is a vector of independent Brownian motions, and $\mu(t)$ and $s(t)$ are well-behaved processes (with $s(t)$ being vector-valued) adapted to $W(t)$. Notice that both $\mu$ and $s$ may depend

[9] One may also make the time parameters stochastic, e.g., by making the lags a function of exposure magnitude. This way, one can, say, model the fact that a bank might tighten its operational controls when exposures are high.

[10] Results for the Advanced model are similar. See Section 4.3 for an example.

on the evolution of multiple risk factors prior to time $t$. For convenience, denote $\sigma(t) = |s(t)|$. Then, for some sufficiently small horizon $\delta$, the increment of portfolio value over $[t_C, t_C+\delta]$, conditional on $\mathcal{F}_{t_C}$, is well-approximated by a Gaussian distribution with mean $\mu(t_C)\delta$ and standard deviation $\sigma(t_C)\sqrt{\delta}$.

Assuming $\sigma(t_C) > 0$, we may ignore the drift-term for small $\delta$. Under the Gaussian approximation above, it is then straightforward to approximate the expectation in (12) in closed form:

$$\mathrm{EE}(t) \approx \mathrm{E}\left[\sigma(t_C)\right]\sqrt{\delta_C}\left[\phi\left(z(q)\right) - z(q)\Phi\left(-z(q)\right)\right], \quad z(q) \triangleq \sqrt{\frac{\delta_{\mathrm{IM}}}{\delta_C}}\Phi^{-1}(q),$$

where $\phi$ and $\Phi$ are the standard Gaussian PDF and CDF, respectively. Similarly,

$$\mathrm{EE}_0(t) \approx \mathrm{E}\left[\sigma(t_C)\right]\sqrt{\delta_C}\phi(0)$$

so that $\lambda$ in (13) is approximated by

$$\lambda(t) \approx \frac{\phi\left(z(q)\right) - z(q)\Phi\left(-z(q)\right)}{\phi(0)}. \tag{14}$$

Interestingly, the multiplier in (14) is independent of $t$ and depends only on two quantities: the confidence level $q$ used for specifying IM, and the ratio of the IM horizon to the MPoR. We emphasize that (14) was derived under weak assumptions, relying only on a local normality assumption for the portfolio value increments. Otherwise, the ratio is model-free: no further assumptions are made on the distribution of the portfolio value or on the dependence of the local volatility on risk factors.

A special case for which (14) becomes exact is when the portfolio process follows a Brownian motion, a case discussed in [6]. Our point is that (14) constitutes a small-$\delta$ limit, and useful approximation, for a much broader class of processes. In fact, we can extend it to jump-diffusion processes, provided that portfolio jumps are (approximately) Gaussian.

The ratio (14) is shown in Panel (a) of Figure 1 for the case $\delta_{\mathrm{IM}} = \delta_C$, graphed as a function of $q$. For the value of the confidence level $q = 99\%$ specified by the BCBS-IOSCO framework, (14) results in a value of $\lambda = 0.85\%$, i.e. IM is anticipated to reduce expected exposure by a factor of 117.

### 3.2 Numerical Tests

We recall that the approximation (14) hinges on the MPoR being small, but it is ex-ante unclear if, say, the commonly used value of 10 business days is small enough. To investigate the potential magnitude of errors introduced by non-infinitesimal MPoR, assume now (in a slight abuse of notation) that $V(t)$ follows a geometric Brownian motion with constant volatility $\sigma$, such that

$$dV(t)/V(t) = \mathcal{O}(dt) + \sigma\, dW(t),$$

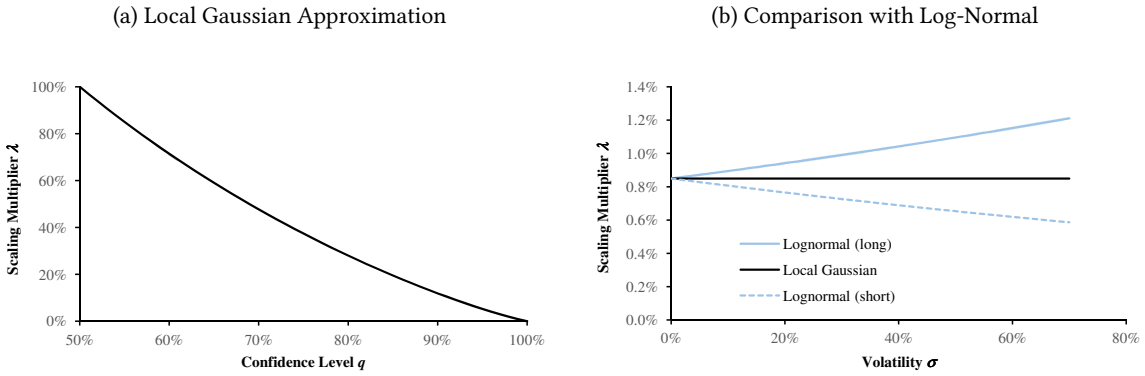
where $W(t)$ is one-dimensional Brownian motion. Compared to a Gaussian distribution, the distribution of $V(t)$ over finite time intervals is skewed left for $V(0) < 0$ and right for $V(0) > 0$, with $V$ never crossing the origin. While this specification may seem restrictive, its only purpose is to test the accuracy of the local Gaussian approximation.

Assuming for simplicity that $\delta_{\text{IM}} = \delta_C$, applying the calculations in Section 3.1 to the lognormal set-up, and neglecting terms of order $\delta$ or higher, results in

$$\lambda(t) = \psi \frac{1 - \Phi\left(\Phi^{-1}(q) - \psi\sigma\sqrt{\delta_C}\right) - (1-q)\exp\left(\psi\sigma\sqrt{\delta_C}\Phi^{-1}(q)\right)}{2\Phi\left(\frac{\sigma\sqrt{\delta_C}}{2}\right) - 1}, \quad \psi \triangleq \text{sign}\left(V(0)\right). \tag{15}$$

While the multiplier in (14) depends only on the confidence level $q$ when $\delta_{\text{IM}} = \delta_C$, the multiplier in (15) additionally depends on the product $\sigma\sqrt{\delta_C}$ and on the sign of the portfolio exposure; the limit for $\delta_C$ (or $\sigma$) approaching zero can be verified to equal (14). Panel (b) in Figure 1 compares (14) to (15) for various levels of $\sigma$.

Figure 1: Exposure Scale Factor $\lambda$




*Panel (a)*: $\lambda$ in local Gaussian approximation (14) when $\delta_{\text{IM}} = \delta_C$, as a function of the IM confidence level $q$. *Panel (b)*: $\lambda$ for geometric Brownian motion (15) as a function of the volatility level $\sigma$, when MPoR is $\delta_{\text{IM}} = \delta_C = 10$ business days and IM confidence level $q = 99\%$. In panel (b), both cases $\psi = 1$ ("long") and $\psi = -1$ ("short") are shown, with "Local Gaussian" denoting the limit result of (14).

As expected, we observe from Figure 1 that the deviation of the local lognormal multiplier from the local Gaussian one increases with the volatility $\sigma$ (actually, with the product $\sigma\sqrt{\delta_C}$). Furthermore, the multiplier for the case $\psi = 1$ ($\psi = -1$) is always greater (smaller) than the limit result of 0.85% – an easily seen consequence of the fact that the relevant distribution tail for the cases $\psi = 1$ ($\psi = -1$) is thicker (thinner) than the tail of the Gaussian distribution, as discussed earlier. Yet, overall, the results illustrate that, for the case of 10-day horizon, the local Gaussian approximation limit produces reasonable values of the scaling multiplier at most levels of relative volatility. Certainly, the results support the idea that introducing IM at the level of $q = 99\%$ should result in about two orders of magnitude reduction of EE, when no trade flows occur within the MPoR.

## 4 The impact of IM: Trade Flows within the MPoR

In this section, we consider the efficacy of IM in the more complicated case where trade flows are scheduled to take place within the MPoR. In practice, many portfolios produce trade flows nearly every business day, so this case is of considerable relevance. The introduction of cash flows will

make computation of EE both more complex and more model-dependent, so additional care is here needed to distinguish between the Classical± and Advanced model results.

### 4.1 Expected Exposure: Basics

Unless one operates under the Classical− model assumptions of no trade flows paid by either $B$ or $C$ within the MPoR, any possibility of $B$ making a trade payment to $C$ in the future results in a spike in the EE profile. These spikes appear because the portfolio value will jump following $B$'s payment, but $C$ would fail to post or return VM associated with this jump. By way of example, suppose that there is a trade payment due at future time $u$ where $B$ is the net payer. A spike in the EE profile originating with this cash flow will appear for a range of termination times $t$ such that i) $u$ lies within the MPoR (mathematically: $t \in [u, u + \delta_C)$); and ii) $B$ would actually make the payment. For the models we consider, we have:

- *Classical+ model*: the assumption is that $B$ (as well as $C$) would make all contractual trade payments for the entire MPoR, so a spike of width $\delta_C$ will appear in $\mathrm{EE}(t)$ for the range $t \in [u, u + \delta_C)$;
- *Advanced model*: it is assumed that $B$ would make contractual trade payments from the beginning $t - \delta_C$ of the MPoR to time $t - \delta'_B$, so a spike of width $\delta_C - \delta'_B$ will appear in $\mathrm{EE}(t)$ for the range $t \in [u + \delta'_B, u + \delta_C)$.

While spikes produced by the Advanced model would always be narrower than those of the Classical+ model, the former is often taller. In particular, the Advanced model, unlike the Classical+ model, contains a range of trade payment times $(t - \delta'_C, t - \delta'_B]$ where $B$ would make a contractual trade payment, but $C$ would not. This creates a narrow peak for the range $t \in [u+\delta'_B, u+\delta'_C)$ at the left edge of the spike. This peak can be very high when $B$'s and $C$'s payments do not net (e.g., because of payment in different currencies), and makes an extra contribution to EE that is not present in the Classical+ model.

We now introduce IM to the exposure computation. Specifically, we suppose that the netting set consists of trades belonging to a single asset class and, in addition to VM, is fully covered by dynamic IM as defined by UMR in (8). We already know, from Section 3, that in the areas *between* EE spikes, IM will reduce the EE by a factor approximately given by (14). The remaining question is how exposure spikes will be affected by IM. The answer to this question is: "it depends". As the simulation time passes, the portfolio will get closer to maturity and fewer and fewer trade payments will remain. Because of this "amortization" effect, the width of the distribution of the portfolio value increment on each path becomes smaller, and, being the VaR of increments, the IM is reduced. On the other hand, trade payments do not generally become smaller as the trade approaches its maturity, and in fact can often become larger due to risk factor diffusion. The effectiveness of IM in reducing EE spikes depends on the size of trade payments relative to the width of the portfolio value increment distribution and can therefore in many cases decline towards the maturity of the portfolio.

### 4.2 Expected Exposure: Numerical Example 1

As an example, consider a 2-year interest rate swap under a two-way CSA with zero-threshold daily VM, but without IM. Assume that $B$ pays a 2% fixed rate semi-annually, and receives a

floating rate quarterly. Setting the initial interest rate curve flat at 2% (quarterly compounding) with a 50% lognormal volatility, Panel (a) in Figure 2 shows exposure profiles for the Classical± and Advanced models.

As expected, both the Classical+ and the Advanced models produce spikes around the quarterly dates on which payments take place. The nature of the spikes depends on whether a fixed payment is made or not. Specifically:

- *Fixed payment time points (twice per year)*: as $C$ pays quarterly, on its semi-annual fixed payment days $B$ will on average pay twice as much as $C$, resulting in high upward spikes in the exposure profile. As discussed above, the width of the Classical+ model spikes is $\delta_C = 10$ business days, while the width of the Advanced model spikes is $\delta_C - \delta'_B = 6$ business days.

- *Floating-only payment points (twice per year)*: on quarterly payment dates where no fixed payment is due by $B$, $C$ still pays a floating rate for the quarter. This results in downward spikes in the EE profile: $C$ makes a trade payment and then defaults before $B$ must return VM. The Classical+ model assumes that trade payments are made by $C$ (and $B$) and margin payments are not made by $B$ (and $C$) for the entire MPoR. Under these (unrealistic) assumptions, the width of the downward spikes is $\delta_C = 10$ business days. The Advanced model, on the other hand, assumes that $C$ would make trade payments over the time interval $\delta_C - \delta'_C$, from which we should subtract the interval $\delta_C - \delta_B$ over which $B$ would pay VM to $C$. In the aggregate, the width of the downward spikes in the Advanced model is therefore $\delta_B - \delta'_C = 2$ business days.

Between the spikes, we observe from Figure 2 that the EE profile produced by the Advanced model is about 22% higher than the one produced by the Classical model. This difference originates with the VM specifications: the Classical models assume that both $B$ and $C$ stop paying VM at the beginning of the MPoR (see (2)), while the Advanced model assumes that $B$ would post VM for some time after $C$ stopped doing so (see (3)).

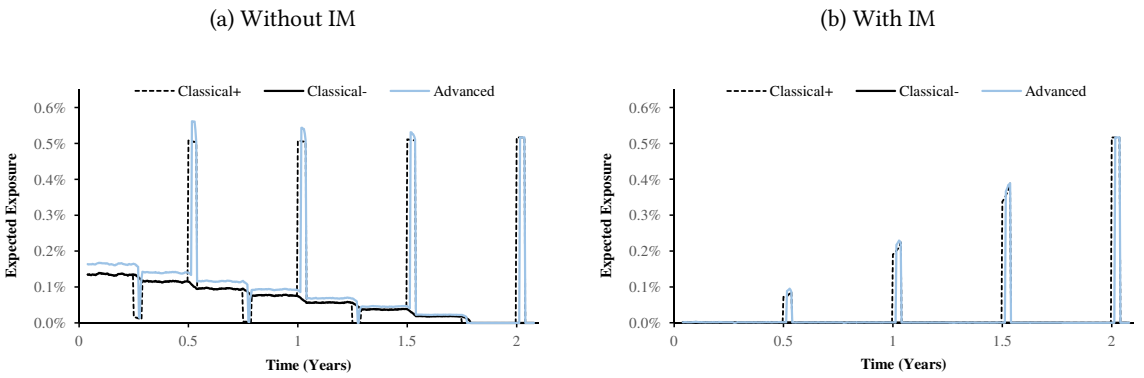
Let us consider the impact of IM on the EE profile for this interest rate swap. We retain all assumption about the MPoR and CSA, and set the IM horizon $\delta_{\text{IM}} = \delta_C = 10$ business days. Since we model the dynamics of the swap value with a single factor model, we can conveniently estimate IM on each path exactly by stressing the risk factor on each path to the 99% level over the IM horizon. Panel (b) in Figure 2 shows the resulting EE profile on the same scale as the no-IM result in Panel (a). Notice that the only part of the EE profile visible in Panel (b) on this scale are the upward spikes from $B$'s semi-annual payments of the fixed rate.

To analyze the IM results in Figure 2, we break the exposure profile into two categories:

- *Between spikes*: Between the spikes, the EE produced by the Classical model with IM is 1.06% of the EE produced by the Classical model without IM. Since the interest rate is modeled with a lognormal process, this number is closer to the value of 1.10% predicted by (15) (the lognormal case) than the value of 0.85% given by (14) (the local Gaussian approximation). Under the Advanced model, the IM scales down EE to 1.00% of its value under VM alone, which is very similar to the Classical model case.

- *Spikes*: As discussed, the degree of suppression for the spikes decreases with the simulation time, as the IM on all simulation paths shrinks over time. Because payments at the end of

Figure 2: Expected Exposure Profiles for 2-Year Swap




EE profiles of 2-year interest rate swap, in percent of swap notional. $B$ pays 2% fixed (semi-annually), receives floating (quarterly). Interest rate curve flat at 2% quarterly compounded, with 50% lognormal volatility. CSA: zero-threshold VM, no MTA, no rounding, daily posting. MPoR assumptions: $\delta_C = 10$ bd, $\delta_B = 8$ bd, $\delta'_C = 6$ bd, $\delta'_B = 4$ bd. Exposure computed from 5,000 Monte Carlo simulations on a daily simulation grid. *Panel (a)*: Without IM. Note that the Classical− EE profile has no spikes; away from spikes, the Classical+ and Classical− EE profiles perfectly coincide. *Panel (b)*: With IM, at the confidence level $q = 99\%$ for $\delta_C = \delta_{\text{IM}} = 10$ bdays.

a period are determined at the beginning of the period, the final payment is known exactly at time 1.75 years, so there is, in fact, no IM requirement in the period from 1.75 to 2 years. As the result, the final spike is not reduced at all.

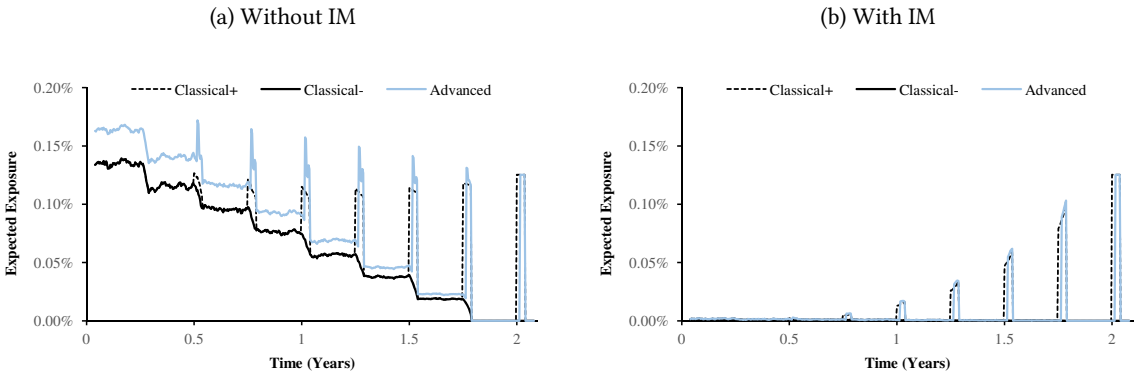
### 4.3 Expected Exposure: Numerical Example 2

To some extent, the spikes in Figure 2 in the presence of IM are particularly pronounced because of the unequal payment frequency[11] on the fixed and floating legs on the swap. Indeed, on the semi-annual payment dates the fixed leg pays, on average, twice as much as the floating leg, so $B$ is the net payer on the vast majority of scenarios, which in turn results in sizable semi-annual upward spikes even when no IM is present.

To examine the extent to which the unequal payment frequency is responsible for spike dominance in the presence of IM, let us move on to a 2-year interest rate swap with quarterly payment frequencies on both legs, with all other model assumptions being the same as in the previous example. For this swap, $B$ is the net payer on the payment dates only on approximately half of the scenarios, so the spike height will necessarily be reduced. This is confirmed in Figure 3, which shows the EE profiles for the quarterly-quarterly swap under the Classical± and Advanced models. It is evident from Panel (a) that, while upward EE spikes are present around the payment dates in the absence of IM, the height of the spikes relative to the level of the EE between the spikes is, as expected, significantly lower than in the example shown in Figure 2. Nevertheless, Panel (b) of Figure 3 shows that IM remains incapable of completely suppressing the spikes.

[11] Semi-annual fixed against quarterly floating is the prevailing market standard in the US.

Figure 3: Expected Exposure Profiles for 2-Year Swap




EE profiles of 2-year interest rate swap, in percent of swap notional. *B* pays 2% fixed (quarterly), receives floating (quarterly). Everything otherwise as in Figure 2. *Panel a)*: Without IM. *Panel b)*: With IM.

### 4.4 The Impact of IM on CVA

CVA constitutes a convenient condensation of EE profiles into a single number, of considerable practical relevance. To demonstrate the impact of IM on CVA for the swap examples above, Table 1 shows CVA calculated[12] from the EE profiles in Panels (a) and (b) of Figure 2 and 3. Since the Classical− model does not have spikes, CVA in this model is reduced by two orders of magnitude by IM, as one would expect from (14). The presence of spikes, however, reduces the effectiveness of IM significantly: for the case of semi-annual fixed payments, the CVA with IM is about 24% of the CVA without IM for the Classical+ model and 15% for the Advanced model. For the case of quarterly fixed payments, the reduced height of the spikes renders IM noticeably more effective in reducing CVA: the CVA with IM is here about 9% of the CVA with IM for the Classical (+) model and about 5% for the Advanced model. Nonetheless, when IM is present, EE spikes still dominate CVA, as can be verified from the ratio of the CVA for the Classical+ or Advanced model to the CVA for the Classical− model (in the second row of Table 1).

Table 1: CVA of 2-Year Swap

| | Fixed S/A | | | Fixed Q | | |
|---|---|---|---|---|---|---|
| Quantity | Classical− | Classical+ | Advanced | Classical− | Classical+ | Advanced |
| CVA (no IM) | 0.267 | 0.391 | 0.408 | 0.209 | 0.237 | 0.274 |
| CVA (with IM) | 0.003 | 0.094 | 0.061 | 0.002 | 0.022 | 0.015 |
| Ratio (IM to no IM) | 1.05% | 24.0% | 14.8% | 1.05% | 9.2% | 5.5% |

The impact of IM on CVA under different models of collateralized exposure. All CVA values are given relative to the CVA value for the case of no VM and no IM. The third row lists the ratio of CVA with IM to CVA without IM. For CVA calculations, the recovery rate was 50% and the credit intensity was 150 basis points. Exposure profiles for the CVA calculations were imported from Figures 2 and 3. "Fixed S/A" and "Fixed Q" denote the swap with semi-annual and quarterly payments, respectively.

[12]To show the impact of VM on uncollateralized CVA, all CVA numbers are shown relatively to the CVA of an otherwise identical uncollateralized swap.

Overall, our swap examples demonstrate that when trade flows within the MPoR are properly modeled, IM at the 99% VaR level may not be sufficient to achieve even a one order of magnitude reduction of CVA. Also, since spikes dominate CVA in the presence of IM, accurately modeling the trade flows paid within the MPoR is especially important when IM is present. Here, the Advanced model is clearly preferable, as neither the Classical+ nor the Classical− model produce reasonable CVA numbers for portfolios with IM.

Finally, we should note that even if one uses only the Advanced model, the impact of IM on CVA may vary significantly, depending on the trade/portfolio details and the model calibration. In particular, we can make the following general observations:

- *Payment frequency*: Higher frequency of trade payments results in more EE spikes, thus reducing the effectiveness of IM.
- *Payment size*: Higher payment size relative to trade/portfolio value volatility results in higher EE spikes, thus reducing the effectiveness of IM.
- *Payment asymmetry*: EE spikes are especially high for payment dates where only $B$ pays or where $B$ is almost always the net payer. The presence of such payment dates reduces the effectiveness of IM.
- *Model calibration*: In the Advanced model, the width of the spikes is determined by the time interval within the MPoR where $B$ makes trade payments, i.e. $\delta_C - \delta'_B$. Thus, larger $\delta_C$ (i.e., the MPoR) and/or the smaller $\delta'_B$ (i.e., the time interval within the MPoR when $B$ does not make trade payments) would result in wider spikes and, thus, reduced effectiveness of IM.

We will briefly illustrate the last point. In the swap examples above, we used the "Baseline" timeline calibration of the Advanced model ($\delta_C = 10$ bd, $\delta_B = 8$ bd, $\delta'_C = 6$ bd, $\delta'_B = 4$ bd) with the MPoR being equal to the BCBS-IOSCO-mandated IM horizon $\delta_{\text{IM}} = 10$ bd. Table 2 shows CVA numbers and their ratios for two alternative calibrations of the Advanced model mentioned earlier.

## 5 Numerical Techniques

### 5.1 Daily Time Grid

Our discussion so far has made obvious the importance of accurately capturing exposure spikes from trade flows. To achieve this, one would generally need to calculate the EE with a daily resolution, something that, if done by brute force methods, likely will not be feasible for large portfolios. In [1], the authors discuss a fast approximation that produces a reasonably accurate EE profile without a significant increase of computation time relative to standard (coarse grid) calculations. The method requires simulation of risk factors and trade flows with a daily resolution, but the portfolio valuations (which are normally the slowest part of the simulation process) can be computed on a much coarser time grid. Portfolio values on the daily grid are then obtained by Brownian bridge interpolation between the values at the coarse grid points, with careful accounting for trade flows.

Table 2: CVA of 2-Year Swap under Alternative Time Line Calibrations

| | Fixed S/A | | | Fixed Q | | |
|---|---|---|---|---|---|---|
| Quantity | Classical− | Classical+ | Advanced | Classical− | Classical+ | Advanced |
| | Aggressive Calibration (MPoR=7bd) | | | | | |
| CVA (no IM) | 0.224 | 0.314 | 0.297 | 0.175 | 0.196 | 0.210 |
| CVA (with IM) | 0.001 | 0.065 | 0.029 | 0.000 | 0.014 | 0.007 |
| Ratio (IM to no IM) | 0.27% | 20.6% | 9.8% | 0.27% | 7.3% | 3.2% |
| | Conservative Calibration (MPoR=15bd) | | | | | |
| CVA (no IM) | 0.325 | 0.505 | 0.622 | 0.254 | 0.294 | 0.393 |
| CVA (with IM) | 0.011 | 0.147 | 0.130 | 0.008 | 0.037 | 0.035 |
| Ratio (IM to no IM) | 3.23% | 29.2% | 21.0% | 3.23% | 12.6% | 9.0% |

The impact of IM on CVA under different models and different calibrations. All CVA values are given relative to the CVA value for the case of no VM and no IM. For CVA calculations, the recovery rate was 50% and the credit intensity was 150 basis points. "Aggressive" calibration: $\delta_C = 7$ bd, $\delta_B = 6$ bd, $\delta'_C = 4$ bd, $\delta'_B = 4$ bd. "Conservative" calibration: $\delta_C = 15$ bd, $\delta_B = 9$ bd, $\delta'_C = 8$ bd, $\delta'_B = 3$ bd. "Fixed S/A" and "Fixed Q" denote the swap with semi-annual and quarterly payments, respectively.

### 5.2 Calculation of Pathwise IM

Per (7)-(8), calculation of IM requires dynamic knowledge of the distribution of portfolio value increments ("P&L"), across $K$ distinct asset classes. Since the conditional distributions of P&L are generally not known, one must rely on numerical methods to calculate IM. Regression approaches are useful for this, although the selection of regression variables can be difficult for large, diverse portfolios where IM would depend on an often impractically large number of risk factors.

To simplify the regression approach, we here choose portfolio value $V_k(t_C)$ as the single regression variable for asset class $k$. Mathematically, we are replacing conditioning on $\mathcal{F}_{t_C}$ in (8) with conditioning on $V_k(t_C)$ and hope that this projection would not have a material impact on the result. For Monte Carlo purposes, these assumptions approximate (8) with

$$\mathrm{IM}_k^{(m)}(t_C) \approx \mathrm{Q}_q\left(V_k(t_C+\delta_{\mathrm{IM}}) + \mathrm{TF}_k\left(t_C+\delta_{\mathrm{IM}};(t_C,t_C+\delta_{\mathrm{IM}}]\right) - V_k(t_C) | V_k(t_C) = V_k^{(m)}(t_C)\right) \quad (16)$$

where superscript $m$ designates the $m$th simulation path and $\mathrm{TF}_k^{(}\cdot)$ is the time $t_C+\delta_{\mathrm{IM}}$ value of all net trade flows scheduled to be paid on the interval $(t_C, t_C+\delta_{\mathrm{IM}}]$, realized along simulation path $m$.

Following [1], we ignore discounting effects and assume that the P&L under the quantile in the right-hand side of (16), conditional on $V_k(t_C) = V_k^{(m)}(t_C)$, is Gaussian[13] with zero drift, such that simply

$$\mathrm{IM}_k^{(m)}(t_C) \approx \sigma_k^{(m)}(t_C)\sqrt{\delta_{\mathrm{IM}}}\Phi^{-1}(q), \quad (17)$$

[13] [5] explored non-Gaussian assumptions, using kernel regression to estimate the first four conditional moments of the P&L. As kernel regression for the third and fourth moments is prone to instability, the approach was deemed insufficiently robust.

where

$$\sigma_k^{(m)}(t_C)^2 \triangleq \delta_{\text{IM}}^{-1}\text{E}\left(\left(V_k(t_C+\delta_{\text{IM}}) + \text{TF}_k^{(m)}\left(t_C+\delta_{\text{IM}};(t_C,t_C+\delta_{\text{IM}}]\right) - V_k(t_C)\right)^2 | V_k(t_C) = V_k^{(m)}(t_C)\right). \tag{18}$$

Estimation of the conditional expectation in (18) may be computed by, say, parametric regression or kernel regression. In a parametric regression, the conditional variance is approximated as a parametric function (e.g., a polynomial) of $V_k(t_C)$, with function parameters estimated by least squares regression; see, for instance, [7]. In a kernel regression, the conditional expectation is computed non-parametrically, through regression on kernel weights. This is the approach taken in [1] and in this paper. Specifically, we use the Nadaraya-Watson Gaussian kernel estimator, with bandwidth determined by applying Silverman's Rule of Thumb to the local (i.e., path-specific) standard deviation of $V_k(t_C)$. In defining the local standard deviation on path $m$, we follow [8] and scale the unconditional standard deviation of $V_k(t_C)$ by the ratio of two probability densities: in the denominator is the probability density of $V_k(t_C)$ on path $m$; in the numerator is probability density that a normally distributed random variable with the same mean and standard deviation as $V_k(t_C)$ would have on path $m$. The actual calculation of the local standard deviation is done via Equation (27) of [8].

We note in passing that if a bank uses an out-of-model margin calculator (e.g., the SIMM method in [4]), an adjustment will be needed to capture the difference between the in-model IM computed (as above) by regression methods and the out-of-model margin calculator actually used. Many possible methods could be contemplated here, e.g., a multiplicative adjustment factor that aligns the two margin calculations at time 0.

### 5.3 Calculation of Pathwise Exposure

Once IM on a path is calculated[14], exposure on that path can, in principle, then be obtained directly by subtracting the calculated IM value from the no-IM exposure (see Section 2.4). This approach works well when the trades covered by IM represent a reasonably small fraction of the netting set. However, when the netting set is dominated by IM-covered trades, this approach suffers from two issues that have an impact on the accuracy of EE (and, therefore, CVA) calculations:

- *Simulation noise*: Suppose that all trades of the netting set are covered by IM and that $\delta_{\text{IM}} = \delta_C$. If all trades belong to the same asset class, then non-zero exposure will be realized on average only 1% of the time between exposure spikes, as IM by design covers 99% of the P&L increase. If multiple asset classes are present in the netting set, the percentage of non-zero realizations will be even less because of the disallowed diversification across asset classes. Thus, EE calculated by direct exposure simulation will be extremely noisy between spikes.

- *Non-normality*: IM is calculated under the assumption that, conditionally on a path, P&L over the MPoR is Gaussian. When all (or most of) the netting set is covered by IM, the EE between the spikes calculated by a direct simulation of exposure is very sensitive to deviations from local normality. If the conditional P&L distribution has heavier (lighter) upper

[14] While not of central importance to this paper, we note that pathwise IM results allow for straightforward computation of Margin Valuation Adjustments (MVA); see, e.g., [9].

tail than the Gaussian distribution, IM between the spikes will be understated (overstated), and EE will therefore be overstated (understated).

Both of these issues can be remedied by calculating the time $t_C$ pathwise *expected* exposure for a time $t$ default, rather than the exposure itself. If our target exposure measure is the unconditional (time 0) EE, this substitution is valid, of course, by the Law of Iterated Expectations. To proceed with this idea, let us first simplify the Advanced model slightly and assume (as the Classical model does) that $B$ and $C$ stop posting margin simultaneously (i.e., that $\delta_B = \delta_C$). Then,

$$E^{(m)}_{\text{adv}}(t) = \left( V^{(m)}(t) - V^{(m)}(t_C) + \text{U}^{(m)}_{\text{adv}}(t) - \sum_{k=1}^{K} \text{IM}^{(m)}_k(t) \right)^+ \tag{19}$$

where $\text{U}^{(m)}_{\text{adv}}(t)$ is the realization of the right-hand side of (6) on path $m$. The expectation of this exposure conditional on $V(t_C) = V^{(m)}(t_C)$ is, in the Advanced model,

$$\text{EE}^{(m)}_f(t; t_C) \triangleq \text{E}\left( E_{\text{adv}}(t) | V(t_C) = V^{(m)}(t_C) \right). \tag{20}$$

Averaging equations (19) and (20) over all Monte Carlo paths will lead to the same result for $\text{EE}(t)$, up to Monte Carlo sample noise.

To calculate the right-hand side of Equation (20) analytically, we assume that the increment of the portfolio value over the MPoR is Gaussian. Specifically, given $V(t_C) = V^{(m)}(t_C)$, we assume that $\text{TF}^{(m)}(t; (t_C, t])$ is known at time $t_C$ and equal to its realization on the path, and that

$$V(t) - V^{(m)}(t_C) \sim \mathcal{N}\left( \text{TF}^{(m)}(t, (t_C, t]), \sigma^{(m)}(t_C)\sqrt{\delta_C} \right), \tag{21}$$

where $\text{TF}^{(m)}(t, (t_C, t])$ represents the sum of all scheduled portfolio payments on $(t_C, t]$ on path $m$, and $\sigma^{(m)}(t_C)$ is defined as

$$\sigma^{(m)}(t_C)^2 = \delta_C^{-1} \text{E}\left( \left( V(t) + \text{TF}^{(m)}(t; (t_C, t]) - V^{(m)}(t_C) \right)^2 | V(t_C) = V^{(m)}(t_C) \right). \tag{22}$$

We emphasize that (22) differs from (18) in several important ways. First, the portfolio used in (22) spans the entire netting set, rather than just the sub-portfolio covered by IM and associated to asset class $k$. Second, the length of the time horizon is $\delta_C$, which may be different from $\delta_{\text{IM}}$. As before, $\sigma^{(m)}(t_C)$ can be calculated by parametric or kernel regression.

Using this assumption, one can calculate (20) to obtain, in the Advanced model,

$$\begin{aligned} \text{EE}^{(m)}(t) &= \sigma^{(m)}(t_C)\sqrt{\delta_C} \left( d^{(m)}(t) \Phi\left( d^{(m)}(t) \right) + \phi\left( d^{(m)}(t) \right) \right), \\ d^{(m)}(t) &\triangleq \frac{-\text{PTF}^{(m)}_{\text{adv}}(t, (t_C, t]) - \sum \text{IM}^{(m)}_k(t)}{\sigma^{(m)}(t_C)\sqrt{\delta_C}}, \end{aligned} \tag{23}$$

where (omitting arguments) $\text{PTF}^{(m)}_{\text{adv}} = \text{TF}^{(m)} - \text{U}^{(m)}_{\text{adv}}$ are the net trade flows on the MPoR *actually* paid, according to the Advanced model.

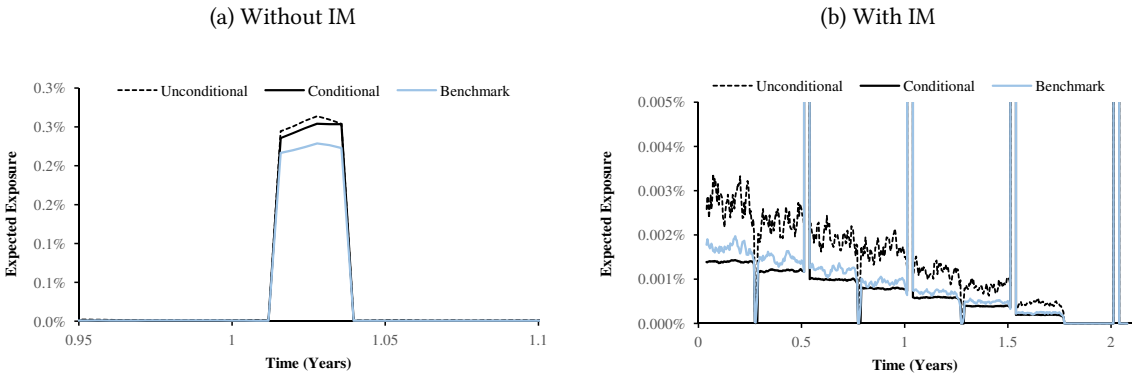
We recall that (23) was derived in the Advanced model, for the simplified case where $\delta_B = \delta_C$. Applying it directly to the more realistic case of $\delta_B < \delta_C$ would, however, result in a significant understatement of EE between the spikes (e.g., for our earlier swap example, the understatement

would be by around 22%). Directly extending (23) to cover $\delta_B < \delta_C$ is, however, not straightforward, so we instead propose a simple scaling algorithm. Here, paths of exposure *without* IM are first simulated for the Advanced model (i.e., (11) with $\mathrm{IM}(t) = 0$). Then, for each exposure measurement time $t$ and each simulation path $m$, the exposure value is multiplied by the ratio $\mathrm{EE}^m(t)/\mathrm{EE}_0^m(t)$, where $\mathrm{EE}^m(t)$ is computed from (23), and $\mathrm{EE}_0^m(t)$ is the special case of (23) where $\mathrm{IM}(t) = 0$.

### 5.4 Numerical Example

To illustrate the benefits of the conditional EE simulation method described above, we turn to the 2-year interest rate swap (with unequal payment frequencies) considered earlier. EE profiles from several computational approaches are shown in Figure 4.

Figure 4: Expected Exposure Profiles for 2-Year Swap




EE profiles of 2-year interest rate swap in the Advanced model, in percent of swap notional. Swap terms, CSA, and model setting as in Figure 2. Graphs contain three different exposure computations: "Unconditional": (19) with 5,000 daily simulation paths, and the IM on each path computed from (17) by kernel regression; "Conditional": (20) with 5,000 daily simulation paths, and the IM on each path computed from (17) by kernel regression; and "Benchmark": (19) with 50,000 simulations and IM computed exactly, using the same technique as in Figure 2. *Panel (a)*: Focus on the 1-year spike. *Panel (b)*: Focus on exposure between spikes.

In Panel (a) of Figure 4, attention is focused on the upward EE spike produced by trade payments at the 1-year point. Both unconditional and conditional EE estimators here produce an almost identical spike, slightly exceeding the benchmark spike height – a consequence of using kernel regression and a Gaussian distribution to estimate IM. Panel (b) shows the EE profiles at a fine exposure scale, allowing for clear observation of EE between the spikes. The advantages of the conditional EE approach can then be clearly seen:

- *Reduction in simulation noise*: While both methods used Monte Carlo simulation with 5,000 paths, the simulation noise in the conditional EE approach is substantially less than that in the unconditional EE approach. In fact, the conditional EE noise is even less than in the benchmark EE results that were calculated with 50,000 paths.

- *Reduced error from non-Gaussian dynamics*: We used a high-volatility lognormal interest rate model in our example, to produce significant deviations from local normality over the

10-day horizon. This non-normality is the main reason for the deviations of the conditional and unconditional EE curves from the benchmark. In estimating (12), the conditional estimator uses a Gaussian distribution to approximate *both* IM and the portfolio increment, resulting in partial error cancellation. Such error cancellation does not take place for the unconditional estimator which uses the empirical (here: lognormal) distribution for the portfolio increment, yet estimates IM from a Gaussian distribution. Between spikes, the EE errors for the unconditional estimator are therefore here significantly larger than for the conditional estimator.

## 6 Conclusion

There is universal agreement that the new BCBS-IOSCO IM rules will lead to very substantial amount of margin postings into segregated accounts, accompanied by inevitable increases in the funding costs (MVA) banks will face when raising funds for IM. According to conventional wisdom, these postings, while expensive, should effectively eliminate counterparty credit risk.

In this article, we have examined the degree to which bilateral IM required by the BCBS-IOSCO margin rules suppresses counterparty exposure. As we have shown in [1], any trade flow to the defaulting party for which it does not return margin during MPoR causes a spike in exposure profile. These spikes are often ignored by banks as "spurious" or as being part of "settlement risk". In reality, these spikes are integral part of the exposure profile and represent real risk that has previously materialized in many well-documented incidents, notably the non-payment by Lehman of reciprocal margin to trade payments that arrived around the time of the bankruptcy filing.

We have shown that, under very general assumptions, the BCBS-IOSCO IM specified as the 99% 10-day VaR reduces exposure between the spikes by a factor of over 100, but fails to suppress the spikes to a comparable degree. This happens because IM is calculated without reference to trade payments, and is based only on changes of the portfolio value resulting from risk factor variability. As an example, we showed that IM reduces the CVA of a 2-year interest rate swap with VM by only a factor of around 7. While VaR based IM fails to fully suppress the contribution of exposure spikes to CVA and EAD, increasing IM to always exceed peak exposure would be impractical, and would require moving large amounts of collateral back and forth in a matter of days.

Another important property of CVA under full IM coverage is that it is dominated by exposure spikes: in our 2-year swap example, spikes' contribution to CVA is about 95% in the presence of IM (compared to about 20% without IM). Thus, in the presence of IM, the focus of exposure modeling should be on capturing the impact of trade payments, which involves making realistic assumptions on what payments the bank and the counterparty are expected to make contingent on the counterparty's default. Furthermore, to accurately calculate CVA mostly produced by narrow exposure spikes, one needs to produce exposure on a daily time grid. A method for producing daily exposure without daily portfolio revaluations was discussed above, along with other useful numerical techniques.

A natural question to ask is why similar payment effects have not been recognized in trading through central counterparties (CCPs), which also require IM posting that is typically based on 99% VaR over the MPoR. As it turns out, CCPs already use a mechanism that amounts to netting of trade and margin payments. Unfortunately, the same approach cannot be adopted in bilateral trading as it would require changing all of the existing trade documentation, which is a practical

impossibility.

While a trade payment and a reciprocal lagged margin payment cannot be netted in bilateral trading, the lag can be eliminated and two payments made to fall on the same day by making a simple change in the CSA. Specifically, if CSA is amended to state that known trade payments due to arrive prior to the scheduled margin payment date must be subtracted from portfolio valuation for the purposes of margin (technically, this amendment effectively sets VM based on a 2-day portfolio forward value), then the call for reciprocal margin will happen ahead of time, and it will arrive on the same day as the trade payment (a "no lag margin settlement").

From an IT and back office perspective, this change in the CSA is relatively easy to align with existing mark-to-market and cash flow processes, and is beneficial in several ways. First, it shortens the duration of exposure spikes and MPoR overall, reducing counterparty risk. Second, it makes margin follow MtM without a two-day lag, thereby eliminating the need to use outside funds to fund hedging during this two-day period. Finally, with reciprocal trade and margin payments falling on the same day, payment-versus-payment (PvP) services such as CLS Bank (see [10], [11] and [12]) may be able to settle trade and margin payments together, reducing residual counterparty risk even further.

## References

[1] Andersen, L., M. Pykhtin, and A. Sokol (2017), "Rethinking the Margin Period of Risk," *Journal of Credit Risk*, 13(1), 1-45.

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

[3] Basel Committee on Banking Supervision and International Organization of Securities Commissions (2015), "Margin Requirements for Non-Centrally Cleared Derivatives," March.

[4] International Swap and Derivatives Association (2016), "ISDA SIMM Methodology," Version 3.15, April.

[5] Andersen, L. and M. Pykhtin (2015), "Accounting for Dynamic Initial Margin in Credit Exposure Models," Presentation at RiskMinds International 2015.

[6] Gregory, J. (2015), *The XVA Challenge: Counterparty Credit Risk, Funding, Collateral, and Capital*, 3rd Edition, Wiley.

[7] Anfuso, F., D. Aziz, P. Giltinan and K. Loukopoulos (2016), "A Sound Modelling and Backtesting Framework for Forecasting Initial Margin Requirements," Working Paper, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2716279.

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

[9] Andersen, L., D. Duffie, and Y. Song (2016), "Funding Value Adjustment," Working Paper, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2746010.

[10] Galati, G. (2002), "Settlement Risk in Foreign Exchange Markets and CLS Bank," *BIS Quarterly Review*, December, 55-66. Available at http://www.bis.org/publ/qtrpdf/r_qt0212f.pdf.

[11] Lindley, R. (2008), "Reducing Foreign Exchange Settlement Risk," *BIS Quarterly Review*, September, 53-65. Available at http://www.bis.org/publ/qtrpdf/r_qt0809g.pdf.

[12] Brazier, J. (2015), "CLS, Markit Launch Cross Currency FX Settlement Service," *Waters Technology*, available at http://www.waterstechnology.com/sell-side-technology/news/2436534/cls-markit-launch-cross-currency-fx-settlement-service .