RAPID AND ACCURATE DEVELOPMENT OF PRICES AND GREEKS FOR NTH TO DEFAULT CREDIT SWAPS IN THE LI MODEL.

MARK S. JOSHI AND DHERMINDER KAINTH

Abstract. New techniques are introduced for pricing nth to default credit swaps in the Li model. We demonstrate the use of importance sampling to greatly increase the rate of convergence of Monte Carlo simulations for pricing. This technique is combined with the likelihood ratio and pathwise methods for computing the sensitivities of these products to changes in the hazard rates of the underlying obligors. In particular the extension of the pathwise method has wider significance in that it is shown that the method can be used even when the pay-off is discontinuous.

1. Introduction

Credit derivatives based on a basket of obligors have recently become popular instruments. Instruments that have recently become popular are the nth to default swap and their closely related (but far more significant in terms of notional) cousins the tranched CDO. We will initially focus on the case of baskets, although we do discuss in the final section of the paper the (trivial) extensions of our formalism to deal with tranched CDOs. In the case of an nth default swap, one party — the so called buyer of protection — pays out a stream of payments until either n obligors from a larger basket of N obligors have defaulted or deal maturity is reached, whichever is earlier. Conversely the seller of protection pays out the loss rate on the nth defaulting asset at the time of default. One popular model for pricing such swaps is the Li model, [12]. In this paper, we show how to apply importance sampling to the pricing of such swaps within the Li model and obtain stable and sizeable speed ups. We also examine the problem of computing sensitivities to the default rates of assets within the model, and in particular show how to apply both the likelihood and pathwise methods of Broadie and Glasserman to this case, [4]. Our extension of the pathwise method is quite general in that we show that it can be applied even when the pay-off is discontinuous, which is a new and significant result and one which could be applied across all asset classes.

We begin by recalling some definitions and fix some notation. Suppose we have N obligors. The nth to default swap has two legs: the *premium* leg contains a stream of payments, sometimes called spread payments, are paid by the purchaser of protection until either the nth default or the maturity time, T, whichever is earlier. The seller pays nothing

Date: November 6, 2003.

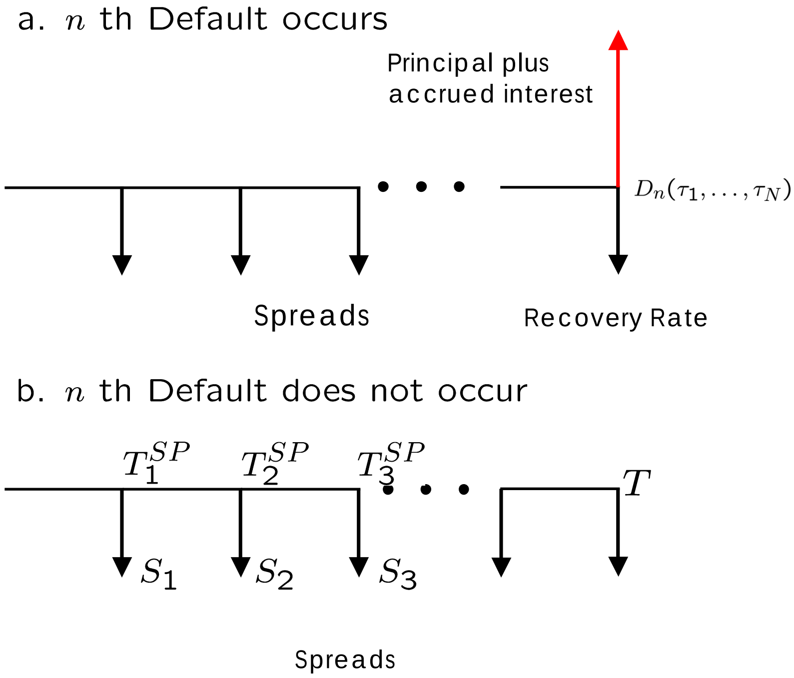

FIGURE 1. A diagrammatic sketch of the cash flows for an nth default swap. There are two possible scenarios: the nth default occurs before maturity in which case we have situation a, or it does not in which case we have the situation illustrated in fig. b.

unless n defaults occur before maturity. If n defaults do occur then at the nth default the purchaser pays the recovery rate on the nth default and any accrued spread payment (generally a linear accrual), and the seller pays the notional. The second leg is sometimes called the *value* leg. If there is no nth default there will, naturally, be no value leg.

Let τ_j and r_j denote the default times and recovery rate respectively of the jth obligor; $D_n(\tau_1, \ldots, \tau_N)$ denotes the time of the nth event, and let $r_n(\tau_1, \ldots, \tau_N)$ denote the recovery rate of the asset that causes the nth default. We will generally just write r_n in order to avoid overly cumbersome notation. (See fig. 1 for a diagrammatic representation of the cash-flows.) Furthermore, we denote the default-free discount rate out to time t by $P(t)$. The discounted pay-off for the value leg, V_{value} at time $D_n(\tau_1, \ldots, \tau_N)$ can then be written as:

$$V_{\text{value}} = (1 - r_n)H(T - D_n(\tau_1, \ldots, \tau_N))P(D_n(\tau_1, \ldots, \tau_n))$$

where H represents the Heaviside step function ($H(x) = 0$ for $x < 0$ and $H(x) = 1$ for $x \geq 0$) and T is the final maturity of the swap. Hence our pay-off for this leg has a discontinuity when the nth default time crosses time the maturity time horizon T. We will use V^u_{value} to denote the undiscounted value.

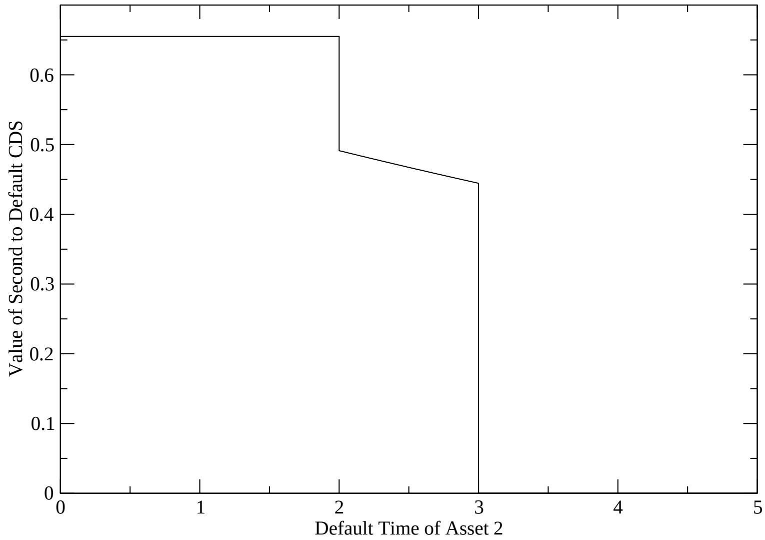

FIGURE 2. Illustration of the pay-off of a second to default swap on two assets as a function of the default time of the second asset, given that the first asset defaults at 2 years. The deal shown has a maturity of 3 years. We have two step like discontinuities: one at the maturity of the product and the second at 2 years i.e., when this asset switches from being the first to default to being the second. The step in the pay-off at the two year point arises because the recovery rates of the two assets are different.

This is illustrated in fig. 2.

The spreads $S_1, S_2, \ldots, S_P$ are paid at discrete intervals, $T_1^{Sp}, T_2^{Sp}, \ldots, T_p^{Sp}$. If the nth default occurs between two spread payment times, the linear accrual means that the value of protection leg, V_{Prot}, can be written:

(1.1)

$$V_{\text{Prot}}(D_n(\tau_1, \ldots, \tau_N)) = \begin{cases} \sum_{i=1}^{m} S_i P(T_i^{Sp}) + S_{m+1} \dfrac{D_n - T_m^{Sp}}{T_{m+1}^{Sp} - T_m} P(t) & \text{if } T_m^{Sp} < D_n < T_{m+1}^{Sp} \\ \sum_{i=1}^{p} S_i P(T_i^{Sp}) & \text{if } D_n > T \end{cases}$$

Note that if default occurs before the first time, the first sum is empty and T_m^{Sp} is zero.

If we have implied a joint density, ψ, for the default times from some model then the value of the product is

$$\mathbb{E}[V_{\text{Prot}} - V_{\text{Value}}] = \mathbb{E}[V_{\text{Prot}}(D_n(\tau_1, \ldots, \tau_N)) - P(D_n(\tau_1, \ldots, \tau_N))[(1-r_n)H(T - D_n(\tau_1, \ldots, \tau_N)]],$$

which can be written in terms of the default times density $\psi(\tau_1, \ldots \tau_n)$ as:

(1.2)

$$\int \{V_{\text{Prot}}(D_n(\tau_1, \ldots, \tau_N)) + P(D_n(\tau_1, \ldots, \tau_N)[(1-r_n)H(T - D_n(\tau_1, \ldots, \tau_N))]\} \psi(\tau_1, \ldots, \tau_N) d\tau_1 \ldots d\tau_N.$$

In the Li model, defaults are supposed to occur according to a Poisson process for each obligor. We suppose that these Poisson processes have deterministic time-dependent intensities, $h_j(t)$, known as hazard rates. We then have that the τ_j have a cumulative exponential distribution function

$$\mathbb{P}(\tau_j < T) = 1 - \exp\left(-\int_0^T h_j(s)ds\right). \tag{1.3}$$

The basis of the Li model is that these one-dimensional random variables are connected to each other by a multivariate normal copula. The correlation matrix, ρ, for this copula is then a model input.

Before we discuss our procedures for accelerating the computation of prices and sensitivities to hazard rates, we briefly reiterate details of the pricing algorithm in Li's model.

Let A be a pseudo-square root of the correlation matrix. Let $E(\tau, h)$ denote the cumulative exponential distribution function in τ for a fixed intensity h. Let $E^{-1}(u, h)$ denote the inverse function in the first variable holding the second variable fixed. Let $N(x)$ denote the cumulative normal function and $N^{-1}(x)$ its inverse; we can go from normals to uniforms by applying N and from uniforms to normals by applying N^{-1}. For each Monte Carlo path we do the following

(1) Draw n uniforms from a random number generator.
(2) Transform the uniforms into a vector of normals, Z.
(3) Set $W = AZ$.
(4) Set $u_i = N(W_i)$ for each i.
(5) Set $\tau_i = E^{-1}(u_i, h_i)$
(6) Compute the cash-flows implied by this set of default times and discount according to the discount curve.

Hence, we have at the final step
(1.4)
$$F(\tau_1, \dots, \tau_N) = V_{\text{Prot}}(D_n(\tau_1, \dots, \tau_N)) - P(D_n(\tau_1, \dots, \tau_N)[(1 - r_n)H(T - D_n(\tau_1, \dots, \tau_N))]$$

We assume that the recovery rate is constant (over time) for each obligor; however, we do not require that different obligors have the same recovery rate i.e., the baskets we analyse are *not* homogeneous. Note that the discounted pay-off, F, has a jump discontinuity when D_n crosses the product's final horizon time T. The average over many Monte Carlo paths is then an approximation to (1.2).

Before proceeding to our improved methods, we examine why Monte Carlo simulations in the Li model can be slow to converge. If no defaults occur before the maturity, then the default part of the product pays zero, and the only payments are the spread payments if any. Such paths therefore result in a fixed value.

If we consider a deal with maturity T, with n uncorrelated obligors each with default intensity h then the probability of all n defaulting is roughly

$$(hT)^n.$$

So if h is around two percent and T around 1, then even for small n, only a very small fraction of paths will result in a default pay-off. For a first to default swap, the situation is seemingly not so bad but even then only about hTn paths will result in a pay-off and the numbers again work against us for T small. An example of this failure to converge is illustrated in fig. 10. We therefore want to apply importance sampling to ensure that the region where the pay-off is zero is not sampled.

We discuss the details of our importance sampling algorithm for the Gaussian copula model in Sections 2, 3 and 4. Our arguments depend mainly on the fact for a multi-variate normal the joint distribution of any k projections conditioned on the other $N-k$ projections is still a multi-variate normal with easily computable covariances. Our computations are facilitated by using a Cholesky decomposition. Numerical results are demonstrated in Section 9.

The application of copula techniques to finance has been an active area of research over the past five years, one that has been given substantial impetus by Li's work on the use of the Gaussian copula for pricing nth default baskets. A number of good reviews have appeared recently — in particular the reader is directed to the excellent text by Schonbucher [17]. The problem considered in this paper — that of computing the Greeks of such products has, it seems, remained largely untouched in the literature. Textbook examples of the application of importance sampling to single name default swaps can, for example, be found in the text by Tavella [16]; however, the application to multiname products is not to our knowledge found in the literature. This is not to say that other solutions to this problem do not exist: in particular by using restricted, so called 'factor' forms for the correlation matrix, one can by following the formalism of Laurent and Gregory [18] compute the prices and hazard rate sensitivities of nth default swaps and tranched CDOs. These approaches are, however, restricted in the forms of the correlation and recovery rate dependencies that they can accomodate; there are no such restrictions on the methods described in the paper below — of considerable importance when considering real portfolios.

One approach to hedging such instruments relies on holding/selling delta amounts of the underlying vanilla default swaps — where the delta signifies the sensitivity of the price of the nth to default swap to changes in the underlying hazard rate of a particular obligor. When computing sensitivities to hazard rates there are additional difficulties, compounding the problems encountered in computing the price. The pay-off is discontinuous as a function of the default times: it jumps when the nth default time passes from being before the expiry of the product to being after it. When computing sensitivities by differencing using Monte Carlo, this means that only a tiny fraction of paths, for which the default time changes from being before expiry to after expiry when the hazard rate changes by a small amount, are the main contributor to the computation. This results in huge variance and renders the computation of the Greeks of these products via the naive Li algorithm almost practically impossible.

There are by now well-known methods for accelerating the covergence when computing Greeks by Monte Carlo. One is the likelihood ratio method of Broadie and Glasserman, [4], which involves multiplying the pay-off on each path by a weighting term. Another due to the same authors is the pathwise method which involves differentiating the pay-off. We show that both these methods can be used for computing sensitivities in the Li models, and that they can be combined with importance sampling to enable very rapid computation of Greeks. We develop expressions for the density which are necessary for both methods in Section 6. We study the likelihood ratio method in Section 7, and the pathwise method in Section 8.

Our results also extend the work of Broadie and Glasserman: our application of the pathwise method is more general than the cases that he discusses in that we show that it can still be used even when the pay-off has a jump discontinuity. It has commonly been argued previously that the pathwise method is not applicable in this case, see for example [2] p35. Our arguments depend on ideas from distribution theory. In particular, the differentiation in that case results in delta distributions; whilst these are hard to sample by Monte Carlo, they are trivial to evaluate analytically, and we show how the difficulties can be overcome. This technique will have widespread applications to other models for derivatives pricing. For a rigorous introduction to distribution theory see [9].

Although this paper is principally about the Li model, there is considerable evidence as to the inadequacy of the normal copula for the modelling of asset returns. Whether or not the distribution of default times conforms to the same type of correlation as that of the assets is a moot point; certainly a number of authors have discussed (see, for example, the works by Breymann *et al.* [3] and Mashal and Zeevi [14]) the use of other copulas, in particular the student T, for the modelling of asset correlation and the consequent effects on pricing of basket credit derivatives. It can be shown that our results below can be extended in a straightforward manner to all elliptical copulas: we will show explicitly the extensions for our importance sampling algorithm for elliptic copulas in Section 5. We do not address here the extension of the likelihood ratio and pathwise methods to the more general elliptical case, because the results are dependent on the particular form of the density function. However, there is a clear recipe to follow: compute the density function and then differentiate appropriately, as in the Gaussian copula case — all of which is described in detail below. Expressions for other copula density functions are readily derived and for the Student T case, for example, are readily available in [1].

In conclusion, we have shown that a judicious combination of importance sampling, standard techniques for computing Monte Carlo Greeks and distribution theory, allows rapid and accurate computation of prices and Greeks using the Li model for nth to default swaps.

2. The important region

In order to ease the discussion we assume for the moment that we have a product that results in zero value unless k defaults occur before time T. Figure 3 illustrates a calculation of default times generated by the copula model for a basket comprising two assets. Assuming that the length of the deal is 5 years it is clear that for the majority of paths generated by our simulation we do not have a default in the relevant time; consequently we receive a fixed sum — the total value of all the spread payments or in this case 0. It is clear that we wish to sample more thoroughly in the regions where defaults occur. Our objective then is to sample the set of interesting points alone.

Going via the cumulative exponential function and inverse cumulative normal function, we can translate the condition $\tau_i < T$ into a condition on the correlated variate W_i. We define x_i to be the number such that $\tau_i < T$ if and only if

$$W_i < x_i.$$

For the importance sampling, we therefore work purely with the normal variates.

We now assume that the pseudo-square root is lower triangular, with positive diagonal entries. Such a decomposition always exists, see for example [15], and is known as the Cholesky decomposition. This will allow us to successively rescale draws. Writing this in a more concrete fashion:

$$\rho = AA^T \qquad A = [a_{ij}] \tag{2.1}$$

For simplicity, we temporarily restrict to the case where $k = 1$. We proceed by making the ith asset default before time T with probability $\frac{1}{n+1-i}$, provided the 1 through $i-1$ assets have not defaulted. This ensures that we will always have at least one asset default — thereby ensuring that every path is important. Since we have altered the probabilities we will require an importance adjustment to reflect this.

For our first asset, we have $W_1 < x_1$ is equivalent to $Z_1 < x_1/a_{11}$. Let $p_1 = N(x_1/a_{11})$. [1]

If $u_1 < \frac{1}{n}$ we set

$$v_1 = np_1u_1,$$

and let $Z_1 = N^{-1}(v_1)$ thus making the first asset default. We multiply the pay-off's value for the path by np_1 to reflect the extra sampling. Let u_i be as in the pricing algorithm specified in the introduction. If $u_1 > \frac{1}{n}$, we set

$$v_1 = p_1 + \frac{1-p_1}{1-\frac{1}{n}}\left(u_1 - \frac{1}{n}\right),$$

[1] In fact, $a_{11} = 1$.

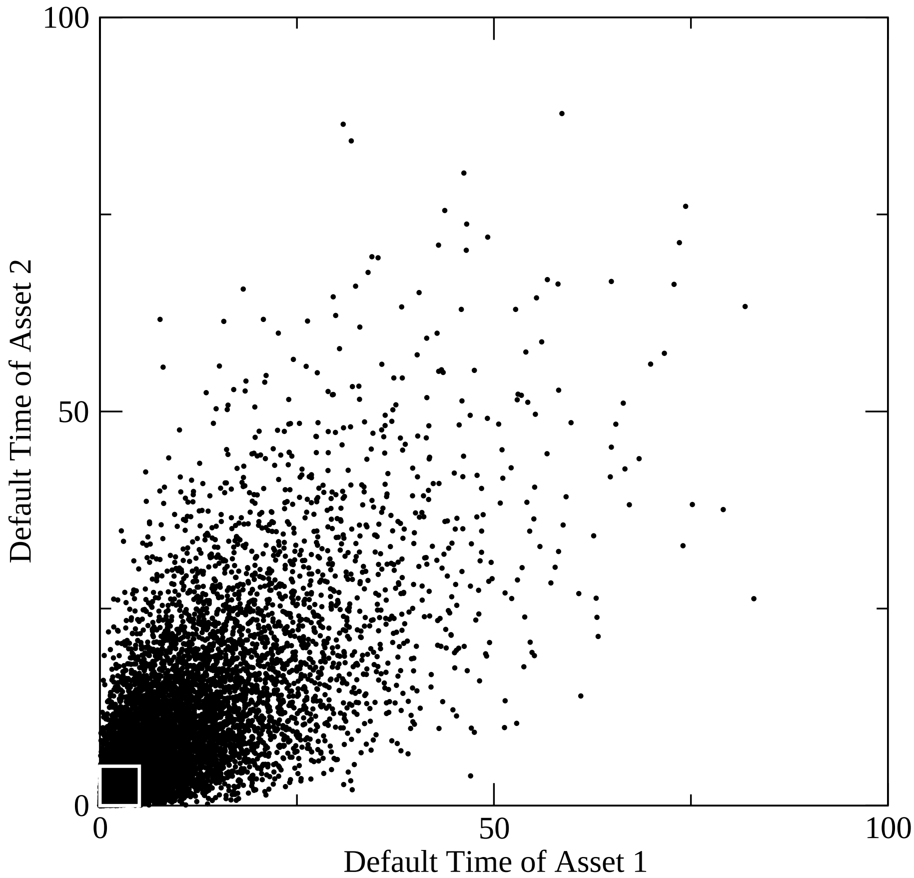

FIGURE 3. Default times generated using a Gaussian copula for two assets. Assuming that the deal has a length of 5 years then only those points which fall in the small white square in the bottom left corner are "important" to the Monte Carlo. This particular set of default times were generated assuming a flat hazard rate for both assets of 0.1; the correlation between the two assets was assumed to be 0.5.

to obtain the full range of possible non-default times. This is illustrated in fig. 4. Again, we have to scale the product's value for the path appropriately; in this case we multiply by

$$\frac{1-p_1}{1-\frac{1}{n}}.$$

Now suppose we have done the first $j-1$ assets. If an asset has defaulted in the requisite time-frame, we allow the jth asset to behave as in the original algorithm that is we set Z_j to the inverse cumlative normal of u_j. Otherwise, we make the jth asset default with probability

$$q_j = \frac{1}{n+1-j}.$$

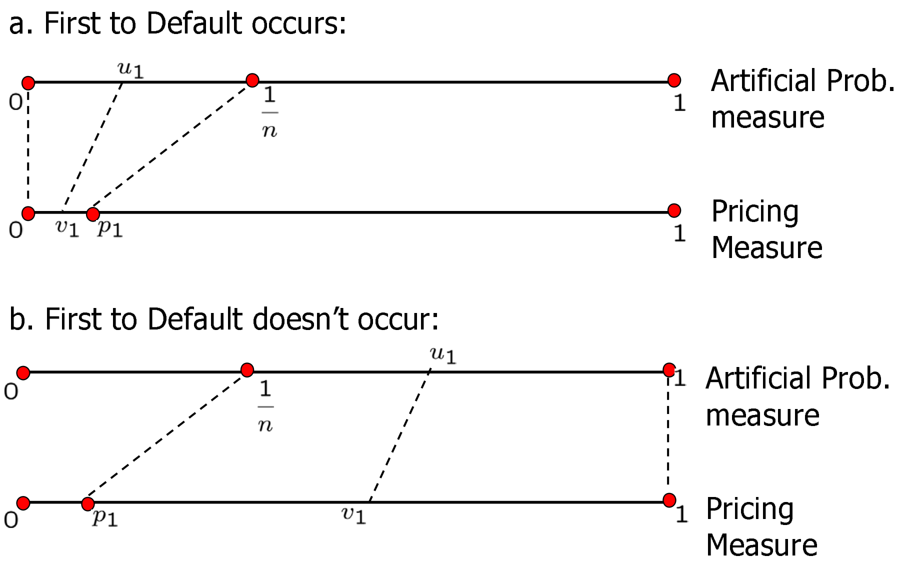

FIGURE 4. A diagrammatic representation of the mappings used to derive the importance sampling reweighted probabilities for the case of a first to default.

The difference now is that the unmassaged default probability will depend on Z_i, for $i < j$. In fact, we have that

$$W_j < x_j \text{ if and only if } \sum_{i<j} a_{ij} Z_i + a_{jj} Z_j < x_j.$$

This is equivalent to

$$Z_j < \frac{x_j - \sum_{i<j} a_{ij} Z_i}{a_{jj}}.$$

We therefore define for $j > 1$

$$p_j = \frac{x_j - \sum_{i<j} a_{ij} Z_i}{a_{jj}},$$

and we have

$$W_j < x_j \text{ if and only if } Z_j < p_j.$$

We now just rescale u_j to get v_j in the same way, we got v_1 from u_1.

So if $u_j < q_j$, we put

$$v_j = \frac{p_j u_j}{q_j},$$

and scale the pay-off by a further p_j/q_j. Otherwise, we set

$$v_j = p_j + \frac{1 - p_j}{1 - q_j}(u_j - q_j),$$

whilst scaling the pay-off by $\frac{1-p_j}{1-q_j}$.

To get Z_j, we take the inverse cumulative normal of v_j. Having obtained the vector of (Z_j) we then proceed as in the original case, multiplying the final pay-off according to the accumulated weights.

Note that if no defaults have occurred for $j < n$ then we have that $q_n = 1$, thereby guaranteeing that at least one default occurs. Our choice of an ascending q_j has guaranteed that at least one default occurs without favouring any particular asset. In particular, when hazard rates are small, the chance of each asset defaulting for a given path is approximately $1/n$. Note that we have allowed the remaining assets to default with their natural probabilities when the requisite number of defaults has already been obtained. Alternate strategies would be to make them not default or require them to default after the 1st default.

3. Multiple defaults

We now discuss how to carry importance sampling when the discounted pay-off is zero unless $k > 1$ defaults occur. Our algorithm is similar. We simply have to change the probabilities so that the extra defaults are guaranteed. Our principal change is that if i defaults occur in the first $j-1$ assets then we set

$$q_j = \frac{k-i}{n-(j+1)}.$$

If q_j is less than or equal to zero then we set $v_j = u_j$, otherwise we scale as before.

Note that if $i < k$ defaults have occured in the first $n-(k-i)$ assets, we have that q_j is equal to 1 for $j > n-(k-i)$. Thus we are guaranteed that at least k defaults occurs and every path makes a non-trivial contribution.

4. Fixed past some point

So far we have assumed that our product is of zero value unless n defaults occur before some fixed time. We now show that the techniques apply equally well to products that have fixed discounted pay-off if insufficient defaults occur before some time. Note that here as elsewhere in the paper, we have assumed deteminstic interest rates.

For example, an nth to default credit default swap will always pay the same, if the nth default is after the maturity of the deal. This value will be the discounted value of spread payments. In this case, importance sampling is more tricky in that it initially appears that we still need some sampling of the extra set. We can, however, reduce to the previous case.

Let V be the discounted value of the payments which occur if less than n defaults occur before maturity. We therefore divide the product into two pieces. We write

$$\text{Discounted-Payoff} = (\text{Discounted-Payoff} - V) + V$$

The discounted value of the second term is always V so its discounted expectation is also V. For the first term, we have reduced to the previous case. Thus our importance sampling method works equally well in this case.

5. Elliptic copulas

The importance sampling techniques that we have described above are not limited only to the Gaussian copula case but can be extended to general elliptic copulas in a straightforward manner. The key observation is that after drawing a random variable which specifies the variance, we are back in the Gaussian case.

Recall that the non-importance sampling algorithm to construct uniform variates from such an elliptic copula is

(1) Draw a random variable, V, from some distribution for the variance, (e.g. chi-squared with ν degrees of freedom (χ^2_ν) for the student-T distribution.)
(2) Draw a vector of independent normal variates $Z = (Z_j)$.
(3) Set $W = \frac{\sqrt{k}}{\sqrt{V}} AZ'$ where A is a pseudo-square root of the correlation matrix and k is some constant (For example with the t distribution we set $W = \sqrt{\frac{\nu}{s}} AZ'$, where s is our chi-squared variate with ν degrees of freedom).
(4) Let $U_j = \mathrm{Cum}(W_j)$, (where Cum is the one-dimensional cumulative distribution.)

At stage 3 we are precisely in the same situation in the Gaussian case so we can repeat the same arguments described above but with A replaced by $\frac{A}{\sqrt{V/k}}$.

6. The default density

We turn now to investigate the computation of the sensitivities of nth to default swaps to changes in the hazard rates of individual obligors. As discussed in the introduction, the computation of sensitivities is more challenging than the computation of the price because the discontinuities in the pay-off lead to large variances. These problems are severe enough to render the computation of sensitivities of e.g., fourth to default baskets with a short maturity almost impractical using naive Monte Carlo.

As discussed above, we will adapt the likelihood ratio and pathwise methods of Broadie and Glasserman to this problem.

For both the likelihood ratio method and the pathwise method, we shall need the joint density of the default times explicitly. In this section, we develop a formula for it under the Li model. We also compute its logarithm.

Recall, [6], that if random variables, x_i, are joined via a copula $\mathbf{C}$, then their joint density function, f is given by:

$$\mathbf{f}(x_1, x_2, \dots x_n) = \frac{\partial \mathbf{C}(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n} \frac{\partial u_1}{\partial x_1} \dots \frac{\partial u_N}{\partial x_N}$$

$$= \mathbf{c}(u_1, \dots, u_N) \prod_{i=1}^{N} f_i(x_i) \tag{6.1}$$

where f_i is the probability density of x_i. In what follows the x_i will be the default times, τ_i, while the f_i will be the marginal densities of the exponential distributions.

For the Gaussian copula we have that

$$\mathbf{c}(u_1, \dots, u_N; \rho) = \frac{1}{|\rho|^{1/2}} \exp\left[-\frac{1}{2}\eta^T(\rho^{-1} - \mathbf{1})\eta\right] \tag{6.2}$$

with $\eta_n = \phi^{-1}(u_n)$. Here ϕ denotes the cumulative normal function.

In what follows, we will take the hazard rates to be constant, for simplicity. The extension to be piece-wise constant hazard rates is straight-forward but fiddly.

We thus have that the joint density, ψ, of the default times is given by

$$\psi(\tau_1, \dots \tau_n) = \frac{1}{|\rho|^{1/2}} \exp\left[-\frac{1}{2}\eta^T(\rho^{-1} - \mathbf{1})\eta\right] \prod_{i=1}^{N} h_i e^{-h_i \tau_i}. \tag{6.3}$$

Taking logs, we obtain

$$\log \psi(\tau_1, \dots \tau_n) = -\frac{1}{2} \log|\rho| - \frac{1}{2}\left[\eta^T(\rho^{-1} - \mathbf{1})\eta\right] + \sum_{i=1}^{N} (\log h_i - h_i \tau_i) \tag{6.4}$$

7. The likelihood ratio method

We begin by considering the likelihood ratio method. The value of a general path-dependent derivative with no early exercise decisions can be written in the form

$$V = \int B(\tau_1, \dots, \tau_N) \psi(\tau_1, \dots, \tau_N) d\tau_1 \dots d\tau_N. \tag{7.1}$$

where B is the discounted pay-off and ψ is the joint density of the default times. Here ψ has an implicit dependence on the default intensities whilst B does not.

In computing the sensitivity of the value of the credit default swap to the hazard rate of the i th asset we differentiate under the integral sign with respect to the hazard rate of the i th asset, h_i. This gives:

$$\frac{\partial V}{\partial h_i} = \int B(\tau_1, \dots, \tau_N) \frac{\partial \psi(\tau_1, \dots, \tau_N)}{\partial h_i} d\tau_1 \dots d\tau_N. \tag{7.2}$$

The problem in doing this is that we have altered the form of the Monte Carlo — we are no longer integrating against a density function. The key observation, [4], is that we can reintroduce the density by writing:

$$\frac{\partial V}{\partial h_i} = \int B(\tau_1, \ldots, \tau_N) \frac{\partial \psi(\tau_1, \ldots, \tau_N)}{\partial h_i} \frac{1}{\psi(\tau_1, \ldots, \tau_N)} \psi(\tau_1, \ldots, \tau_N) d\tau_1 \ldots d\tau_N. \tag{7.3}$$

Simplifying,

$$\frac{\partial V}{\partial h_i} = \int B(\tau_1, \ldots, \tau_N) \frac{\partial \log \psi(\tau_1, \ldots, \tau_N)}{\partial h_i} \psi(\tau_1, \ldots, \tau_N) d\tau_1 \ldots d\tau_N. \tag{7.4}$$

Thus in order to develop an expression for the hazard rate sensitivity we multiply the integrand used in computing the price by the derivative of the log of the density function with respect to the hazard rate and then carry out the Monte Carlo as before.

We have developed the log of the joint density of the default times, (6.4). Differentiating with respect to the ith hazard rate gives:

$$\frac{\partial \log f}{\partial h_i} = -\sum_j (\rho^{-1} - \mathbf{1})_{ij} \eta_j \frac{\partial \eta_i}{\partial u_i} \frac{\partial u_i}{\partial h_i} + \frac{1}{h_i} - \tau_i \tag{7.5}$$

We can compute $\frac{\partial \eta_i}{\partial u_i}$ explicitly by noting that

$$\phi^{-1}(u) = \sqrt{2} \text{ erf }^{-1}(2x - 1)$$

Carrying out the differentiation, we see that:

$$\frac{\partial \eta_i}{\partial u_i} = \sqrt{2\pi} e^{\frac{1}{2}\phi^{-1}(u)^2}$$

These expressions are easily computed and thus computing derivatives with the likelihood ratio is easy to implement. The full power of the method, however, requires combination with importance sampling. We have the additional difficulty that the weighted pay-off will not be constant in the non-default domain even if the pay-off is.

However, as in section 4, if we subtract a constant, V, so that $B - V$ is zero if an insufficient number of defaults occur, then we will have also that $(B - V)\frac{\partial \psi}{\partial h_i}$ is zero is unless the requisite number of defaults occurs. We can therefore apply importance sampling as before. Note that as V is constant across all paths, its subtraction from the pay-off will not affect the value of the Greek.

8. The pathwise method

We refer the reader to the original work by Broadie and Glasserman [4], and the discussion by Jaeckel [11]. In what follows let each hazard rate be constant. Let $F(\tau_1, \ldots, \tau_n)$ denote the discounted pay-off of the product, and let $E(\tau, h)$ denote the cumulative exponential distribution with parameter h. Suppose we try the naive approach to compute a delta by bumping the relevant parameter, h_j, but using the same random number path.

The pricing algorithm, from the introduction, only changes at the stage where the default times are computed. If ψ denotes the density of the uniforms implied by the copula, we are effectively evaluating

$$\int G(u, h, \epsilon)\psi(u)du$$

where

$$G(u, h, \epsilon) = \frac{1}{\epsilon}(F(E^{-1}(u_1, h_1), \dots, E^{-1}(u_j, h_j + \epsilon), \dots, E^{-1}(u_N, h_N)) \\ - F(E^{-1}(u_1, h_1), \dots, E^{-1}(u_j, h_j), \dots, E^{-1}(u_N, h_N))). \tag{8.1}$$

As we are interested in small ϵ, we consider the limit as ϵ goes to zero, and obtain

$$\frac{\partial F}{\partial \tau_j}\frac{\partial E^{-1}}{\partial h_j}(u_j, h_j).$$

What is this first term? If we are in a domain where

$$\tau_j \neq D_n(\tau_1, \dots, \tau_N)$$

we obtain zero. Otherwise, we obtain (in a distributional sense)

$$\frac{\partial F}{\partial \tau_j} = \frac{\partial P}{\partial t}(\tau_j)[H(T - \tau_j)(1 - r_n)] \\ -P(\tau_j)[\delta(\tau_j - T)(1 - r_n) + H(\tau_j - T) \quad \frac{\partial}{\partial t}(1 - r_n)|_{t=\tau_j}] \\ +\frac{\partial}{\partial t}(V_{\text{Prot}}(t))|_{t=\tau_j} \tag{8.2}$$

where δ denotes the Dirac delta distribution. Here P is as in the introduction.

The first three terms in the expression above arise from the differentiation of the value leg with respect to the time of the jth default; the last term comes from the differentiation of the protection leg. We are using the fact that the distributional derivative of a Heaviside function is a delta distribution, see [9], and the Leibniz product rule for differentiation. We will discuss how to handle the integration of each of these terms against the multivariate density of the default times in greater detail below; however, for the moment we highlight some of the features of the terms above.

The delta distribution (the second term on the right-hand side of eq. 8.2) means that the Monte Carlo will be slow to converge for small ϵ. It reflects the fact in a zero interest rate environment that the only paths which will pick up a change in price after the ϵ change in hazard rate will be those for which the time D_n moves from being after time T to before time T, and these paths will be of magnitude $1/\epsilon$. This results in high variance. In the case of non-zero interest rates, these paths will have a similar effect. We can see this dependence in our graphs of the pay-off function (fig. 2).

At first sight it may seem odd that we need to differentiate the recovery rates with respect to time. However, if the different bonds underlying the product have different recovery rates, then as bumping the hazard rate of asset j will cause it to default earlier, it may change from being the $(n+1)$th default to the nth default, which means that the product's pay-off will change. Note that this causes the hazard rate sensitivity to be recovery rate dependent.

The first term is easy to evaluate by Monte Carlo, it is essentially the same algorithm as before. The second is not so obvious by Monte Carlo. However, it is easy to analytically integrate against a delta distribution! It is just the function evaluated at $\tau_j = T$. The second term in our integral therefore becomes

$$-P(T)\frac{\partial E^{-1}}{\partial h_j}(E(T,h_j),h_j)\int \psi(\tau_1,\ldots,\tau_{j-1},T,\tau_{j+1},\ldots,\tau_N)d\tau_1\ldots d\tau_{j-1}d\tau_{j+1}\ldots d\tau_N.$$

We still need to evaluate the remaining dimensions. Let ψ_{n-1} denote the joint density function of τ_i for these dimensions. This integral is problematic because our density has changed; however, we can reintroduce the original density by rewriting the integral as

$$\int \frac{(I\psi)(\tau_1,\ldots,\tau_{j-1},T,\tau_{j+1},\ldots,\tau_N)}{\psi_{n-1}(\tau_1,\ldots,\tau_{j-1},\tau_{j+1},\ldots,\tau_N)} \psi_{n-1}(\tau_1,\ldots,\tau_{j-1},\tau_{j+1},\ldots,\tau_N)d\tau_1\ldots d\tau_{j-1}d\tau_{j+1}\ldots d\tau_N, \tag{8.3}$$

where the function I is one if $D_n(\tau_1,\ldots,\tau_{j-1},T,\tau_{j+1},\ldots,\tau_N)$ equals T and zero otherwize. These densities are straightforward to compute and we perform a Monte Carlo in all variables except τ_j to carry out the evaluation.

The third term gives rise to the recovery rate dependence of the hazard rate sensitivity. In order to calculate the magnitude of this contribution, assume that as in the Li model, given the hazard rates, we have drawn a set of default times. We then order the default times so as to find the nth to default asset; we also simultaneously order the recovery rates according to the default times. We can see from fig. 8 on bumping the jth hazard rate, that after sorting we will affect the value of the hazard rate sensitivity whenever the jth bond becomes the $(n-1)$ th or the nth to default. Suppose that bumping the jth hazard rate alters the $(n-1)$th bond after sorting (Contribution 1 in fig. 8.) Then it can be seen that the value of the product alters by:

$$\delta(\tau_{n-1}-T)[((1-r_n)-(1-r_{n-1}))P(T)] \tag{8.4}$$

We can also see from fig. 8 that there is a second contribution arising when altering the jth hazard rate alters the nth bond. In this case the value of the product alters by

$$\delta(\tau_n-T)[((1-r_{n+1})-(1-r_n))P(T)] \tag{8.5}$$

Summing these two parts together gives us the overall hazard rate sensitivity due to the recovery rates; the Monte Carlo integration to value this is straightforward.

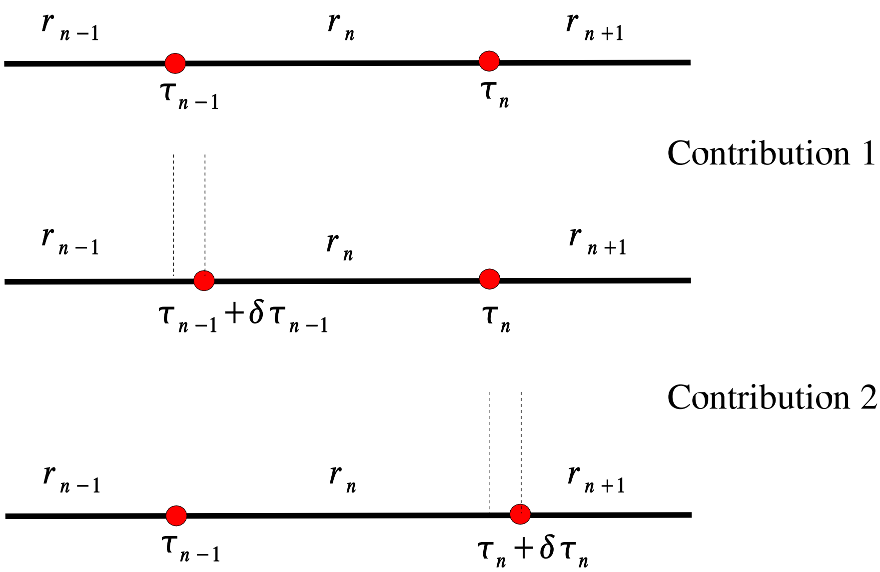

FIGURE 5. If different assets underlying an nth to default have different recovery rates then the hazard rate sensitivity will be recovery rate dependent. In this figure we illustrate the two contributions which arise when we bump the hazard rate of a single asset.

The integration of the protection leg is straightforward once we recall equation (1.1) which we restate here for the reader's convenience

(8.6)
$$V_{\text{Prot}}(t) = \begin{cases} \sum_{i=1}^{m} S_i P(T_i^{Sp}) + S_{m+1}\dfrac{t - T_m^{Sp}}{T_{m+1}^{Sp} - T_m^{Sp}} P(t) & \text{provided that } T_m^{Sp} < t < T_{m+1}^{Sp} \\ \sum_{i=1}^{p} S_i P(T_i^{Sp}) \quad \text{if no } n\text{th default occurs} & \end{cases}$$

Differentiating with respect to t and evaluating at τ_j leaves us only with a contribution due to the second term in the defaulting case:

$$\frac{\partial}{\partial t} V_{\text{Prot}}(t)P(t) = S_{m+1}\frac{1}{T_{m+1}^{Sp} - T_m} P(\tau_j) + S_{m+1}\frac{t - T_m^{Sp}}{T_{m+1}^{Sp} - T_m^{Sp}} \frac{\partial}{\partial t} P(t)|_{t=\tau_j} \tag{8.7}$$

We remark that this method is not dependent on the normal copula. We simply need to know the relevant density functions for whichever copula we are working with. As with the likelihood ratio method, the full benefits of this method are realized when it is combined with importance sampling.

9. NUMERICAL RESULTS

We present some numerical results. Our measure of goodness is the standard deviation of the simulation as a fraction of the limit. Note also that we have plotted the standard

Maturity	Without importance	With Importance
0.02	23.1	1.06
0.04	16.5	1.02
0.06	13.4	1.01
0.08	11.6	0.996
0.1	10.4	0.988
0.2	7.32	0.967
0.4	5.12	0.953
0.6	4.24	0.950
0.8	3.65	0.951
1	3.27	0.953
2	2.29	0.977
3	1.85	1.01
4	1.58	1.04
5	1.40	1.06
6	1.26	1.09
7	1.15	1.12
8	1.06	1.15
9	0.99	1.18
10	0.93	1.21

FIGURE 6. Normalized standard deviation of a (quasi) Monte Carlo simulation used to compute the price of a first to default swap on a basket of four names. Notice that the normalized standard deviation for the importance sampled case is small and constant regardless of maturity.

deviation not the standard error. The standard error can be obtained by dividing the the square root of the number of paths. We present results purely for the value protection leg to avoid cancellation effects.

We take continuously compounding rate $r = 5\%$, four credits with constant hazard rates, $0.05, 0.01, 0.02$ and 0.02. The recovery rates are $0.2, 0.7, 0.5$ and 0.3. We took a constant correlation of 0.2. A Monte Carlo simulation was run to estimate the mean and variance of the price of the protection leg. We used 2^{19} paths and Sobol numbers. We give the standard deviation of the simulation as a fraction of price with and without importance sampling. We first give the first to default case in figure 6. We also give the fourth to default case in figure 7. There are no entries for less than 0.5 years without importance sampling because every path gave zero. Note that in both cases, the importance sampling error is stably around 1. This means that we can get expect a price within one percent of the true value with ten thousand paths in a straight Monte Carlo simulation, and an even more accurate answer using Sobol numbers.

Maturity	Without importance	With Importance
0.02		0.718
0.04		0.709
0.06		0.703
0.08		0.699
0.1		0.696
0.2		0.685
0.4		0.674
0.6	458	0.666
0.8	363	0.661
1.0	258	0.658
2.0	96.3	0.646
4.0	41.6	0.639
6.0	24.3	0.639
8.0	17.1	0.643
10.0	13.1	0.650

Figure 7. Normalized standard deviation of the price of (quasi) Monte Carlo simulation used to compute the price of a fourth to default swap on a basket of four names. Notice that the normalized standard deviation for the non importance sampled case blows up for short maturities.

The pathwise method performs significantly better than finite differencing as a method for computing hazard rate sensitivities; indeed we demonstrate that it even out performs the likelihood ratio method by quite a margin for computing sensitivities. The reason for this marked decrease in the variance is the fact that we have analytically integrated out the delta distribution arising in the derivative of these products. We present graphs of the normalized standard deviation across deal maturities in fig. 8 and fig. 9. The first of these is without importance sampling, whereas the second uses it. Note the vast difference in scales between these graphs. For these simulations, we have taken zero recovery rates, a constant correlation of 0.2 and constant hazard rates of 2%. Interest rates are a constant 5% continuously compounding rate.

Again the pathwise method does suffer something of a depreciation in performance if we analyse deals on the nth asset with short maturity; however, empirically we find that this depreciation in performance is nowhere near as marked as for finite differencing or for the likelihood ratio method. Fig. 10 illustrates a typical comparison between the pathwise method and finite differencing in computing hazard rate sensitivities.

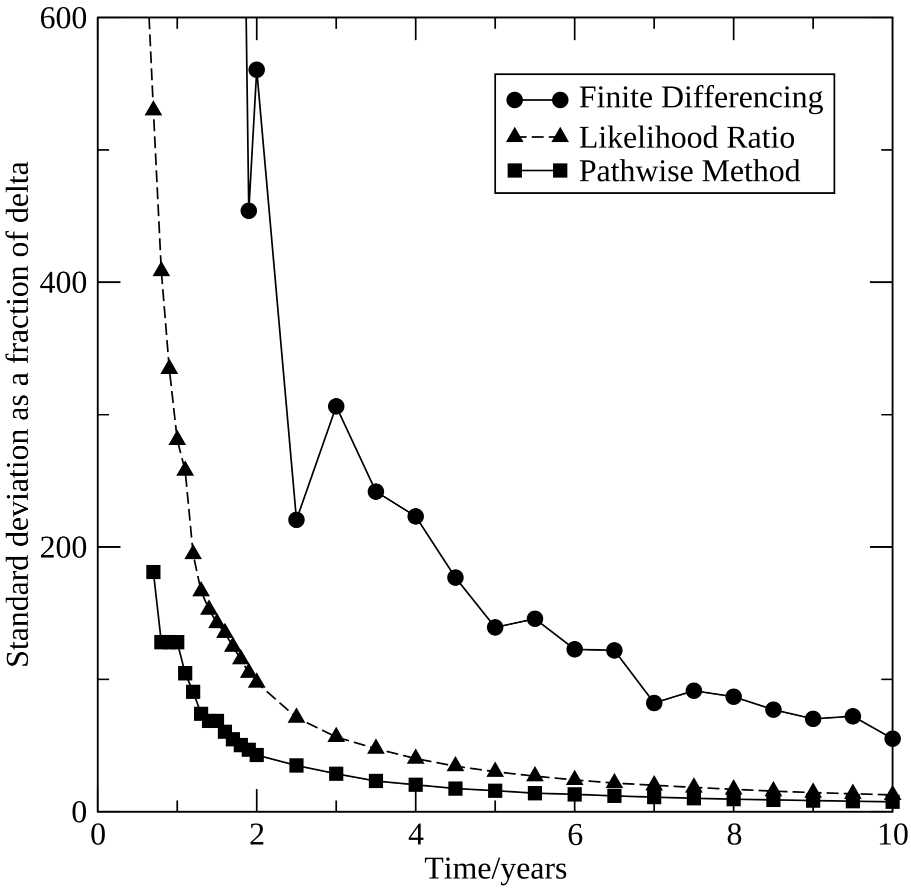

FIGURE 8. Normalized standard deviation across a range of deal maturities of the hazard rate sensitivity of a non-importance sampled (quasi) Monte Carlo simulation for a fourth to default swap on a basket of four names.

10. Other products

Whilst this paper is about nth to default basket credit derivatives, we give indications in this final section of how the techniques can be applied to other credit derivative products. One example of such a product is a tranched credit derivative where the holder receives a spread payment in return for paying all losses from a basket of names between two fixed levels.

Our techniques for importance sampling can be applied identically to any product which has fixed discounted pay-off unless a minimum of defaults occur. As recovery rates are deterministic in the Li model, this includes tranched credit derivatives as the minimum number of defaults is then the number required to breach the lower loss level.

The likelihood ratio method depends purely on the density and not on the pay-off. We can therefore use it identically to assess sensitivities for any product, including a tranched credit derivative, where the cash-flows generated are purely a function of the default times

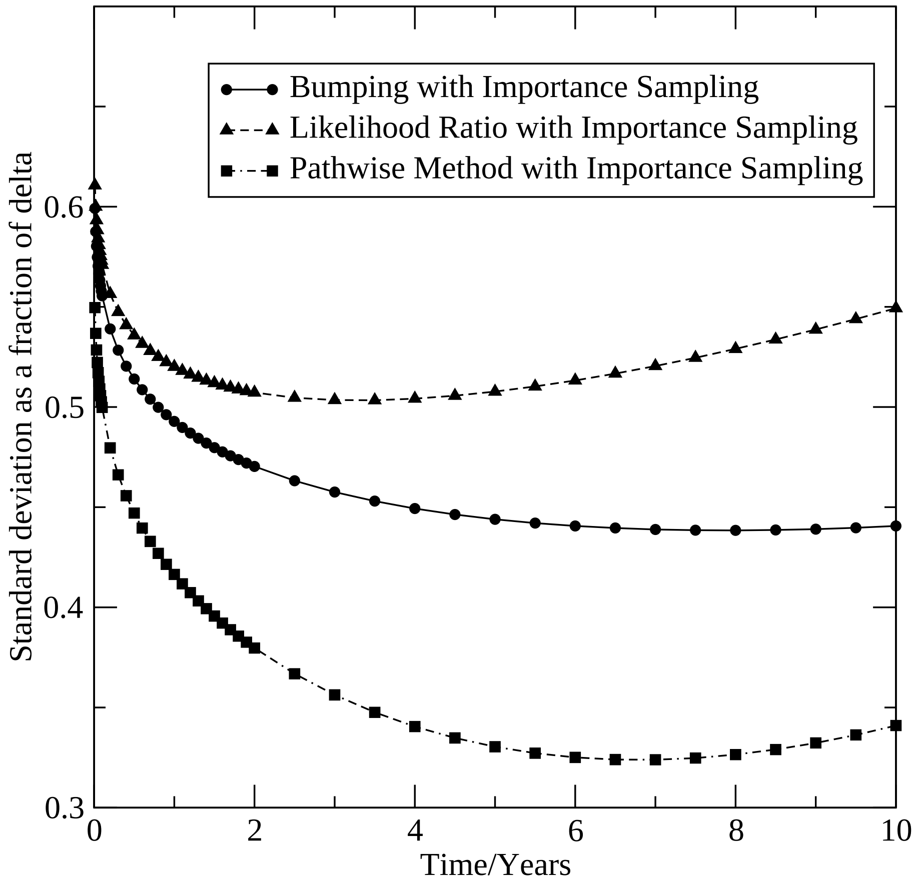

FIGURE 9. Normalized standard deviation across a range of deal maturities of the hazard rate sensitivity of an importance sampled (quasi) Monte Carlo simulation for a fourth to default swap on a basket of four names.

of the reference credits credits. The great advantage of the method is, indeed, that once it has been implemented for one product it automatically works for all products for which the pricer has been implemented.

The pathwise method can be applied to any product; however, the pay-off must be differentiated analytically for each product individually. In addition, the delta functions must be recognized and the integrals coded individually. So although it is often the most effective method, it is also often the most time-consuming to implement.

REFERENCES

[1] E. Bouye, V. Durrleman, A. Nikeghbai & T. Roncalli, Copulas for Finance — a reading guide and some applications Paper presented in the 17th International Conference in Finance organized by the French Finance Association, Paris (June 2000) Available from http://gro.creditlyonnais.fr/

[2] P. Boyle, M. Broadie, P. Glasserman, Monte Carlo methods for security pricing, in *Monte Carlo: Methods and Applications for Pricing and Risk Management*, edited by B. Dupire, Risk Books 1998, p15-44

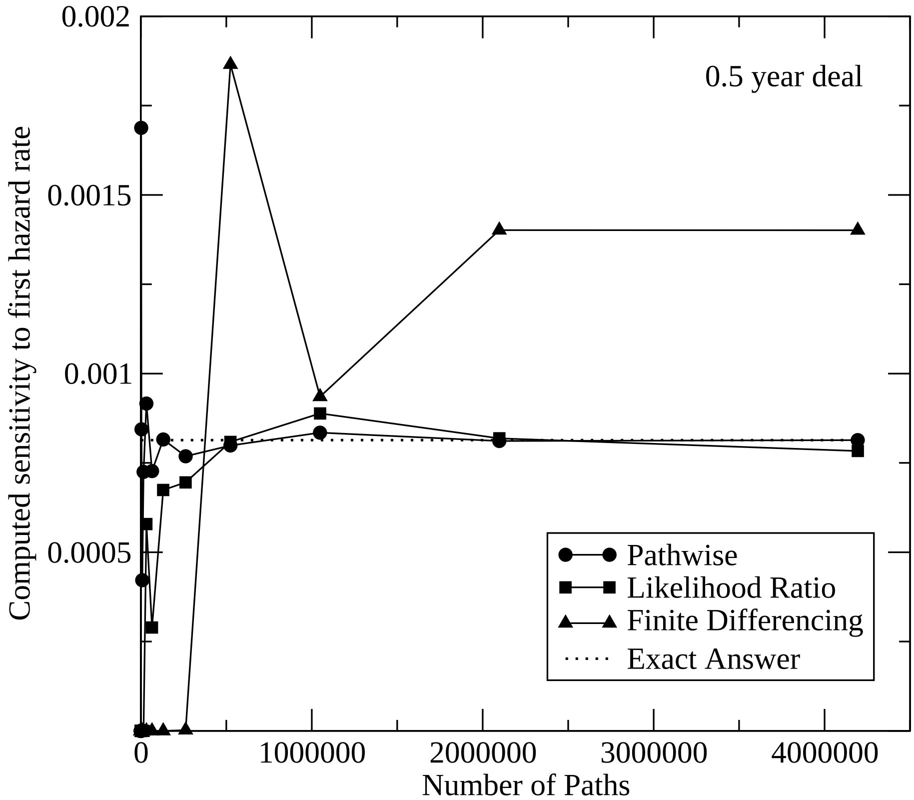

FIGURE 10. Convergence diagram for the computation of the sensitivity of a first to default credit default swap with 4 underlying credits to changes in the first (flat) hazard rate. The deal maturities shown is 0.5 years while the hazard rates for each of the credits is at 2%. We show comparisons between forward differencing by 1% of the hazard rate and, the likelihood ratio method, the pathwise method and the exact answer.

[3] W. Breymann, A. Dias, P. Embrechts, Dependence structures for multivariate high-frequency data in finance, Working paper, RiskLab, ETH Zurich

[4] M. Broadie, P. Glasserman, Estimating security derivative prices by simulation, *Management science*, 42, 269-285, (1996).

[5] P. Embrechts, F. Lindskog and A. McNeil, Modelling Dependence with Copulas and Applications to Risk Management www.math.ethz.ch/ baltes/ftp/papers.html

[6] P. Embrechts, F. Lindskog and A. McNeil, Correlation:Pitfalls and Alternatives Risk Magazine May 1999, p69-71 Preprint, http://www.math.ethz.ch/ baltes/ftp/papers.html

[7] P. Georges, A. G. Lamy, E. Nicholas, G. Quibel and T. Roncalli Multivariate survival modelling: a unified approach with copulas Preprint, Credit Lyonnais http://gro.creditlyonnais.fr/ 2001

[8] E. Fournie, J.M. Lasry, J. Lebuchoux and N. Touzi, Applications of Malliavin Calculus to Monte Carlo Methods in Finance Finance & Stochastics 1998, vol 3, issue 4, p391-412

[9] F.G. Friedlander, M. Joshi, Introduction to the theory of distributions, 2nd Edition, Cambridge University Press 1999

[10] Peter Friz, Introduction to Malliavin Calculus Available at www.math.nyu.edu/phdstudents/frizpete(2002)

[11] P. Jaeckel, Monte Carlo Methods in Finance, Wiley Interscience 2002

[12] D. X. Li, On Default Correlation: A Copula Function Approach, RiskMetrics Working Paper
[13] R. Mashal and M. Naldi, Pricing Multiname Credit Derivatives: Heavy Tailed Hybrid Approach, Preprint, http://www.columbia.edu/ rm586, 2002
[14] R. Mashal, A. Zeevi, Beyond correlation: extreme co-movements between financial assets, Working paper, Columbia Business School
[15] W.H. Press, S.A. Teutolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in C*, 2nd edition, Cambridge University Press, 1992
[16] D. Tavella *Quantitative Methods in Derivatives Pricing : An Introduction to Computational Finance*, Wiley Finance, 2001
[17] P. J. Schonbucher *Credit Derivatives Pricing Models: Models, Pricing and Implementation*, Wiley Finance, 2003
[18] J. P. Laurent & J. Gregory, Basket Default Swaps, CDO's and Factor Copulas, Working Paper, www.maths.univ-evry.fr/mathfi/JPLaurent.pdf