Finance Stoch
DOI 10.1007/s00780-008-0071-y

Sensitivity estimates for portfolio credit derivatives using Monte Carlo

Zhiyong Chen · Paul Glasserman

Received: 24 November 2006 / Accepted: 6 May 2008

Abstract Portfolio credit derivatives are contracts that are tied to an underlying portfolio of defaultable reference assets and have payoffs that depend on the default times of these assets. The hedging of credit derivatives involves the calculation of the sensitivity of the contract value with respect to changes in the credit spreads of the underlying assets, or, more generally, with respect to parameters of the default-time distributions. We derive and analyze Monte Carlo estimators of these sensitivities. The payoff of a credit derivative is often discontinuous in the underlying default times, and this complicates the accurate estimation of sensitivities. Discontinuities introduced by changes in one default time can be smoothed by taking conditional expectations given all other default times. We use this to derive estimators and to give conditions under which they are unbiased. We also give conditions under which an alternative likelihood ratio method estimator is unbiased. We illustrate the application and verification of these conditions and estimators in the particular case of the multifactor Gaussian copula model, but the methods are more generally applicable.

Keywords Sensitivity calculation · Credit derivatives · Monte Carlo simulation · Efficiency · Pathwise method

Mathematics Subject Classification (2000) 91B28 · 65C05

JEL Classification G13 · G32 · C15

Z. Chen (✉)
Financial Analytics and Structured Transactions Group, Bear Stearns and Company Inc., New York, NY 10179, USA
e-mail: zhiyong.chen@gmail.com

P. Glasserman
Graduate School of Business, Columbia University, New York, NY 10027, USA
e-mail: pg20@columbia.edu

1 Introduction

Portfolio credit derivatives are contracts that are tied to an underlying portfolio of defaultable assets and have payoffs that depend on the default times of these assets. These types of derivatives facilitate the buying and selling of protection against the credit risk in a portfolio. Examples of portfolio (or multi-name) credit derivatives include basket default swaps and collateralized debt obligations (CDOs); see, e.g., Bruyère et al. [3] and Schönbucher [17] for general background on these types of contracts.

The payoff of a portfolio credit derivative may be viewed, abstractly, as a function of the default times of the underlying assets. Valuing the derivative entails calculating the expectation of the discounted payoff over the joint distribution of the default times. The marginal distribution of each default time (under a pricing measure) is typically inferred from the market prices of assets linked to an individual obligor, such as a bond or a credit default swap (see, e.g., Duffie and Singleton [7]); the joint distribution may then be specified through a copula function, as in Li [16]. Simple cases of this approach lead to pricing through transform inversion and numerical integration techniques, or other numerical methods, as in Andersen et al. [1], Hull and White [13], and Laurent and Gregory [15]. But more general cases require Monte Carlo simulation.

The hedging of portfolio credit derivatives requires the calculation of sensitivities of the contract value to parameters of the underlying default-time distributions. More specifically, default-time distributions are usually specified through their hazard rates, and hedging focuses on the effect of changes in these hazard rates—that is, on the *delta* with respect to hazard rates or parameters of hazard rates.

Estimating these deltas accurately and efficiently by Monte Carlo presents a challenge. Estimating price sensitivities by Monte Carlo is typically more difficult than estimating the prices themselves, and this is particularly true for payoffs with discontinuities. In the context of credit derivatives, discontinuities arise because small changes in default times can produce large changes in a contract's payoff. For example, in a first-to-default swap, a small change in a default time may introduce or eliminate a default within the life of the swap, or it may change the identity of the first name that defaults. Either of these changes can produce a jump in the swap's cashflows. In the case of a CDO, a discontinuity in cashflows arises when a small change in a default time causes the default to cross a coupon date.

Finite difference approximation is the most straightforward approach to estimating sensitivities by Monte Carlo. In the credit context, this means perturbing a hazard rate and resimulating to compute the change in price. But finite difference estimates are particularly poor in the presence of discontinuities; see, e.g., the discussion in Sect. 7.1.2 of Glasserman [10]. The main alternatives are the *pathwise method* and the *likelihood ratio method*, as discussed in Broadie and Glasserman [4] and Chap. 7 of Glasserman [10].

Joshi and Kainth [14] apply these techniques to the hedging of nth-to-default swaps in the Gaussian copula model. The likelihood ratio method is unaffected by payoff discontinuities and its application in this context is relatively straightforward. We provide rigorous support for its application by giving conditions under which it produces unbiased estimators.

In its simplest form, the pathwise method is inapplicable to discontinuous payoffs—the interchange of derivative and expectation required to make the method unbiased typically fails to hold in the presence of discontinuities. Methods for smoothing discontinuities that result from changes in the order of events have been studied in the literature on the simulation of queuing networks and other discrete-event systems; see, in particular, Suri and Zazanis [18], Gong and Ho [11], Chap. 7 of Glasserman [9], and Fu and Hu [8]. These methods use conditional expectations to smooth the effect of changes in the order of events (e.g., arrivals to or departures from a queue). Joshi and Kainth [14] arrive at similarly smoothed estimators independently and by a rather different route, using somewhat informal calculations with delta functions and stopping short of providing rigorous support for their methods. Here, we avoid the use of delta functions and instead derive smoothed pathwise estimators as limits of conditional expectations. We give conditions—for the most part, modest regularity conditions—under which these estimators are unbiased. As a byproduct of our analysis, we identify a term missing in Joshi and Kainth [14]: In the setting of nth-to-default swaps, they combine their estimators with a method that forces at least n defaults to occur on every path; we show that for the calculation of deltas, it is necessary to consider paths on which $n-1$ defaults occur as well.

The rest of this paper is organized as follows. Section 2 provides background on portfolio credit derivatives and the Gaussian copula model. Sections 3 and 4 derive the sensitivity estimators. For purposes of illustration, Section 5 presents the particular cases of basket default swaps and CDOs. In Section 6, we improve the performance of the sensitivity estimators using a variance reduction technique. Section 7 concludes the paper. Most proofs are contained in Appendix A.

2 Problem description and background

This section describes the class of credit derivatives we consider in this paper, and it reviews the popular Gaussian copula model which we use for illustration.

2.1 Portfolio credit derivatives

We consider credit derivatives tied to a basket (or portfolio) of N underlying names, such as bonds, loans, or credit default swaps. The number of underlying names typically ranges from 5 to 200. We use T to denote the life of the contract.

We denote by τ_i the default time of the ith asset, $i = 1, \ldots, N$, taking $\tau_i = \infty$ if the ith asset never defaults. The default times $\tau_1, \ldots, \tau_N$ are positive random variables with a joint density function $f(t_1, \ldots, t_N)$. We denote by $f_i(t_i)$ the marginal density of τ_i, and we denote by $f_i(t_i|\tau_1, \ldots, \tau_{i-1}, \tau_{i+1}, \ldots, \tau_N)$ the conditional density function of τ_i given the times of the other defaults. The marginal distribution of each τ_i is typically extracted from the market prices of credit default swaps or bonds; these market prices are used to construct a hazard rate function λ_i from which we get the distribution

$$F_i(t) = P(\tau_i \le t) = 1 - \exp\left(-\int_0^t \lambda_i(s)\,ds\right),$$

$$f_i(t) = \lambda_i(t)\exp\left(-\int_0^t \lambda_i(s)\,ds\right).$$

We consider credit derivatives with discounted payoff $V(\tau_1, \ldots, \tau_N)$, a function of the default times. We focus on sensitivities of prices with respect to the hazard rates of underlying names. The numbering of the underlying names is arbitrary, so there is no loss of generality in considering sensitivities associated with the first asset. Thus, suppose the hazard rate function λ_1 depends on a parameter h (as well as on time t), i.e., $\lambda_1 = \lambda_1(t, h)$. We consider sensitivities with respect to the parameter h as it varies over an open interval $\mathcal{H}$. In the simplest case, $\lambda_1(t, h) = \lambda_1(t) + h$ for all t, so that changes in h correspond to parallel shifts in the hazard rate function, but we consider more general parameterizations. The sensitivity (called *delta*) with respect to h is given by

$$\frac{\partial \mathbf{E}(V)}{\partial h} = \lim_{\epsilon \to 0} \frac{\mathbf{E}(V(\tau_1(h+\epsilon), \ldots, \tau_N)) - \mathbf{E}(V(\tau_1(h-\epsilon), \ldots, \tau_N))}{2\epsilon},$$

where we have attached $h \pm \epsilon$ to τ_1 to indicate the parameter values. We include h as an argument whenever we need to emphasize a change in the value of this parameter.

We use the conditions below in our analysis. After listing the conditions, we discuss their interpretation and scope. In conditions (A3), (A5) and elsewhere, when we refer to a property holding almost everywhere (a.e.), we mean with respect to Lebesgue measure on $\Re$ or $\Re^N$. Also, in stating the conditions, we use $\tau_1, \tau_2, \ldots$ to denote the random variables representing default times, and we use $t_1, t_2, \ldots$ as real variables representing possible outcomes of the default times. Thus, in (A4), $f(t_1, \ldots, t_N)$ is the joint density of the default times $\tau_1, \ldots, \tau_N$ evaluated at $t_1, \ldots, t_N$.

(A1) The discounted payoff $V = V(\tau_1, \ldots, \tau_N)$ is a bounded function of the default times.

(A2) Fix any positive $t_2, \ldots, t_N$ and let $t_{N+1} = T$. For any $i \ge 2$, we define $j = \operatorname{argmin}_{k \ge 2}\{t_k - t_i : t_k > t_i\}$. Then V is Lipschitz with respect to t_1 in the interval (t_i, t_j); i.e., for any t_1 and $t_1 + \Delta t$ in (t_i, t_j), there exists a $K_1(t_2, \ldots, t_N) < \infty$ such that

$$\left|V(t_1 + \Delta t, t_2, \ldots, t_N) - V(t_1, t_2, \ldots, t_N)\right| \le |\Delta t| K_1(t_2, \ldots, t_N).$$

(A3) The first derivative $\partial\lambda_1/\partial h$ exists and is nonnegative almost everywhere in $\mathcal{H}$.

(A4) The default times $\tau_1, \ldots, \tau_N$ admit a density $f(t_1, \ldots, t_N)$ and a conditional density $f(t_1|\tau_2, \ldots, \tau_N)$ given $\tau_2, \ldots, \tau_N$, a.s.

(A5) For every $h \in \mathcal{H}$ and almost every $t_1, \ldots, t_N \in (0, \infty)$, the partial derivative $\partial f(t_1, \ldots, t_N; h)/\partial h$ exists. Furthermore, there exists a nonnegative function g such that for any $h \in \mathcal{H}$ and all sufficiently small $|\epsilon|$,

$$\left|f(t_1, \ldots, t_N; h+\epsilon) - f(t_1, \ldots, t_N; h)\right| \le |\epsilon| g(t_1, \ldots, t_N), \quad \text{a.e.},$$

with

$$\int_0^\infty \cdots \int_0^\infty g(t_1, t_2, \ldots, t_N)\, dt_1 \cdots dt_N < \infty.$$

(A6) For every $h \in \mathcal{H}$ and almost every $t_1 \in (0, \infty)$, the partial derivative $\partial f_1(t_1; h|\tau_2, \ldots, \tau_N)/\partial h$ exists, a.s. Furthermore, there exists a nonnegative function g_1 such that for all sufficiently small $|\epsilon|$ and almost every t_1,

$$\left| f_1(t_1; h+\epsilon|\tau_2, \ldots, \tau_N) - f_1(t_1; h|\tau_2, \ldots, \tau_N)\right| \le |\epsilon| g_1(t_1|\tau_2, \ldots, \tau_N), \quad \text{a.s.,}$$

with

$$\int_0^\infty g_1(t_1|\tau_2, \ldots, \tau_N)\, dt_1 < \infty, \quad \text{a.s.}$$

(A7) A family of random variables $\{\tau_1(h),\ h \in \mathcal{H}\}$ having densities $\{f_1(t_1; h), h \in \mathcal{H}\}$ can be realized as an almost surely strictly decreasing differentiable function of h.

(A8) For all sufficiently small $|\epsilon|$, there exists a random variable K_τ such that $\mathbf{E}(K_\tau) < \infty$ and for any $h, h+\epsilon \in \mathcal{H}$, $|\tau_1(h+\epsilon) - \tau_1(h)| \le K_\tau |\epsilon|$, a.s.

The boundedness assumption in (A1) is widely applicable to portfolio credit derivatives, because the maximum gain or loss on each underlying asset in this context is usually bounded. Condition (A2) makes precise the idea that the payoff is continuous (in fact Lipschitz) so long as the changes in default times are sufficiently small to leave the order of the default times unchanged. Condition (A3) applies, for example, when the parameter h linearly shifts the entire hazard function λ_1, which would be the most typical sensitivity considered in practice. The requirement in (A4) that the default times admit a joint density holds in essentially any nondegenerate setting—for example, in the Gaussian copula model, provided no two obligors are perfectly correlated. Conditions (A5) and (A6) impose some modest regularity conditions on the joint density and conditional density of the default times; in both cases, the condition imposed is similar to requiring the existence of an integrable derivative. Condition (A7) can often be satisfied by letting U be uniformly distributed on the unit interval and setting $\tau_1(h) = F_1^{-1}(U; h)$, where $F_1(t_1; h)$ is the cumulative distribution function obtained from the density $f_1(t_1; h)$, and F_1^{-1} denotes the inverse with respect to the first argument. For example, in the case of a constant hazard rate $\lambda_1(t, h) \equiv \lambda_1(h)$, each $\tau_1(h)$ is exponentially distributed and $\tau_1(h) = -\log(1-U)/\lambda_1(h)$. In this case, we get

$$\tau_1'(h) = -\frac{\tau_1(h)}{\lambda_1(h)} \lambda_1'(h), \tag{2.1}$$

and K_τ in (A8) can be taken as the supremum of $|\tau_1'(h)|$ for $h \in \mathcal{H}$. The monotonicity and smoothness properties required of τ_1 in (A7) follow from corresponding conditions in $\lambda_1(h)$. The assumption of strict monotonicity in (A7) ensures that the mapping from h to $\tau_1(h)$ is invertible, a.s.

2.2 The Gaussian copula model

The joint distribution of the default times, $f(t_1, \ldots, t_N)$, has not yet been specified. The Gaussian copula (as in Li [16], Gupton et al. [12]) is a widely used mechanism for specifying a joint distribution for the default times consistent with given marginals, and will provide a useful and illustrative example. The dependence among $\tau_1, \ldots, \tau_N$ is determined by underlying jointly normal random variables $W_1, \ldots, W_N$. Each W_i has a standard normal distribution Φ, so $\Phi(W_i)$ is uniformly distributed on (0, 1) and $\tau_i = F_i^{-1}(\Phi(W_i))$ has distribution F_i. However, $W_1, \ldots, W_N$ are correlated, with covariance matrix $\mathbf{\Sigma}$, and this introduces (and, indeed, completely characterizes) dependence among the default times $\tau_1, \ldots, \tau_N$. We will make the simplifying assumption that $\mathbf{\Sigma}$ has full rank so that no asset has its default time completely determined by those of the other assets.

More generally, saying that $\tau_1, \ldots, \tau_N$ have copula function $\mathcal{C}$ means that

$$\mathbf{P}(\tau_1 \le t_1, \ldots, \tau_N \le t_N) = \mathcal{C}\big(F_1(t_1), \ldots, F_N(t_N)\big),$$

for any $t_1, \ldots, t_N$. The joint density function of $\tau_1, \ldots, \tau_N$ is then given by

$$\begin{aligned} f(t_1, \ldots, t_N) &= \frac{\partial \mathcal{C}(u_1, \ldots, u_N; \mathbf{\Sigma})}{\partial u_1 \ldots \partial u_N} \frac{\partial u_1}{\partial t_1} \cdots \frac{\partial u_N}{\partial t_N} \\ &= c(u_1, \ldots, u_N; \mathbf{\Sigma}) \prod_{i=1}^{N} f_i(t_i), \end{aligned} \tag{2.2}$$

where $u_i = 1 - \exp(-\int_0^{t_i} \lambda_i(s)\, ds)$.

For the Gaussian copula,

$$c(u_1, \ldots, u_N; \mathbf{\Sigma}) = \frac{1}{|\mathbf{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}\big(\Phi^{-1}(\mathbf{u})\big)^{\top}\big(\mathbf{\Sigma}^{-1} - \mathbf{I}\big)\Phi^{-1}(\mathbf{u})\right], \tag{2.3}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{u}$ is an $N \times 1$ vector with $\mathbf{u}_i = u_i$.

3 Estimating sensitivities: likelihood ratio method

There are three primary methods for estimating sensitivities by Monte Carlo: the finite difference method, the likelihood ratio method, and the pathwise method. The finite difference method is superficially easier to understand and implement; but it produces biased estimates and its efficient use requires a difficult balance between bias and variance in the selection of the perturbation ϵ. The method is particularly problematic with discontinuous payoffs. The likelihood ratio method avoids difficulties resulting from discontinuities by differentiating a probability density rather than the payoff. In contrast, the pathwise method differentiates each simulated outcome with respect to the parameter of interest and (in its simplest form) is limited to payoffs without discontinuities. The application of the likelihood ratio method and the pathwise method requires interchange in the order of differentiation and integration, an interchange that requires verification. For an introduction to these methods and issues, see Asmussen and Glynn [2], Broadie and Glasserman [4] or Chap. 7 of Glasserman [10].

In the setting of portfolio credit derivatives, the likelihood ratio method (which we address in this section) is relatively straightforward. The pathwise method (which we take up in the next section) requires a more extensive analysis.

The likelihood ratio method estimates sensitivities through the derivative of a probability density—in our setting, the density of the default times. It does not require any smoothness in the discounted payoff. It is therefore widely applied to different models and credit products. The next result confirms its applicability in our context.

Theorem 3.1 *Under conditions* (A1) *and* (A5), *the estimator of the delta given by the likelihood ratio method*,

$$V(\tau_1,\dots,\tau_N)\frac{\partial \ln f(\tau_1,\dots,\tau_N;h)}{\partial h}, \tag{3.1}$$

is unbiased, *i.e.*,

$$\frac{\partial \mathbf{E}(V)}{\partial h} = \mathbf{E}\left(V(\tau_1,\dots,\tau_N)\frac{\partial \ln f(\tau_1,\dots,\tau_N;h)}{\partial h}\right).$$

In the particular case of the Gaussian copula model, by (2.2) and (2.3), we have

$$\begin{aligned}\ln f(\tau_1,\dots,\tau_N) &= \ln c(u_1,\dots,u_N;\boldsymbol{\Sigma}) + \sum_{i=1}^N \ln f_i(\tau_i)\\ &= -\frac{1}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}\big(\Phi^{-1}(\mathbf{u})\big)^\top\big(\boldsymbol{\Sigma}^{-1}-\mathbf{I}\big)\Phi^{-1}(\mathbf{u})\\ &\quad + \sum_{i=1}^N\left(\ln\lambda_i - \int_0^{\tau_i}\lambda_i(s)\,ds\right),\end{aligned}$$

and

$$\frac{\partial \ln f(\tau_1,\dots,\tau_N)}{\partial h} = -\sum_{j=1}^N\big(\boldsymbol{\Sigma}^{-1}-\mathbf{I}\big)_{1j}\Phi^{-1}(u_j)\frac{\partial\Phi^{-1}(u_1)}{\partial u_1}\frac{\partial u_1}{\partial h} + \frac{1}{\lambda_1}\frac{\partial\lambda_1}{\partial h} - \tau_1\frac{\partial\lambda_1}{\partial h},$$

where

$$\lambda_1 = \lambda_1(\tau_1,h),\qquad \frac{\partial u_1}{\partial h} = \frac{\partial\lambda_1}{\partial h}\tau_1(1-u_1),\qquad \frac{\partial\Phi^{-1}(u_i)}{\partial u_i} = \sqrt{2\pi}e^{\Phi^{-1}(u_i)^2/2}.$$

All these expressions can be computed easily. For the case of constant hazard rates, this estimator appears in Joshi and Kainth [14].

4 The pathwise method

4.1 Discontinuities and jump terms

Before undertaking a detailed derivation of the pathwise estimator, we motivate our analysis with a discussion of the problem of discontinuities and how this problem can be circumvented.

In its usual form, the pathwise estimator of the sensitivity of the expected payoff $\mathbf{E}(V(\tau_1,\dots,\tau_N))$ to a parameter h of τ_1 would be

$$\frac{\partial V(\tau_1,\dots,\tau_N)}{\partial h}=\frac{\partial V(\tau_1,\dots,\tau_N)}{\partial \tau_1}\frac{d\tau_1(h)}{dh},$$

with the derivative of τ_1 as in (A8) and (2.1). This estimator captures the effect of local changes in the default time τ_1 resulting from small changes in the hazard rate λ_1. For example, it captures the effect of the timing of cashflows that might result from a small change in τ_1. However, it does not capture the effect of any discontinuity in V introduced by a change in the order of defaults.

To stress this point, consider a simple (if artificial) payoff of the form

$$V(\tau_1,\tau_2)=c_1\mathbf{1}(\tau_1<\tau_2)+c_2\mathbf{1}(\tau_2\le\tau_1),$$

for some constants c_1 and $c_2\neq c_1$. If $\tau_1\neq\tau_2$, then a sufficiently small change in h will not change the order of events, so the pathwise derivative is zero with probability 1. If τ_1 is decreasing in h and τ_1 is initially greater than τ_2, then an increase in h may eventually cause an interchange in the order of defaults, at which point V will jump by c_1-c_2. This jump is not reflected in the pathwise estimator.

To capture the effect of the discontinuity, we condition on all default times other than τ_1, which in this example means conditioning on τ_2:

$$\mathbf{E}(V|\tau_2)=c_1\mathbf{P}(\tau_1<\tau_2|\tau_2)+c_2\big(1-\mathbf{P}(\tau_1<\tau_2|\tau_2)\big).$$

Differentiating the conditional expectation yields

$$\frac{\partial\mathbf{E}(V|\tau_2)}{\partial h}=-f_1(\tau_2|\tau_2)\tau_1'\cdot(c_1-c_2),$$

where $f_1(\cdot|\tau_2)$ is the conditional density of τ_1 given τ_2, introduced in (A4), and τ_1' is evaluated at $\tau_1=\tau_2$. Anticipating the general form we derive in (4.2), we can write this estimator as

$$\frac{\partial V}{\partial h}-f_1(\tau_2|\tau_2)\tau_1'\cdot(c_1-c_2),$$

though the first (local) term is identically zero in this case.

The second term—the jump term—contains the essential features of the general cases we derive. First, we have the conditional density $f_1(\cdot|\tau_2)$ evaluated at τ_2 itself, giving the “probability” (in the sense of a density) that τ_1 occurs just at the same time as τ_2. This density is multiplied by $-\tau_1'$ (to be evaluated at $\tau_1=\tau_2$), and the product gives the rate at which τ_1 crosses τ_2 (from the right), given τ_2. Finally, this jump rate is multiplied by $c_1 - c_2$, the size of the jump in V when τ_1 crosses τ_2.

The rest of this section is devoted to generalizing and justifying this idea. The final estimator (4.2) includes a jump term for each possible discontinuity; each of these results from τ_1 crossing some τ_i, $i \neq 1$, or crossing the end of the contract $T \equiv \tau_{N+1}$. Each jump term is the product of a jump rate and a jump magnitude. To make these ideas precise, we first need to justify differentiating the conditional expectation, which we do in Proposition 4.1. We then need to show that the jump rate has the asserted form and that we can ignore terms resulting from τ_1 crossing more than one other default time; this we do in Lemma 4.2. We assemble the estimator in the steps leading to (4.2) and then verify unbiasedness in Theorem 4.3.

4.2 Derivation of the estimator

Our first step is to reduce the problem to one of estimating the sensitivity of the conditional expectation given $\tau_2, \ldots, \tau_N$. Recall (see Condition (A7)) that the parameter h affects only τ_1. The following result is proved in Appendix A.2.

Proposition 4.1 *Under conditions* (A1) *and* (A3)–(A6),

(a) $\partial \mathbf{E}(V)/\partial h$ *and* $\partial \mathbf{E}(V|\tau_2, \ldots, \tau_N)/\partial h$ *exist*
(b) $\partial \mathbf{E}(V)/\partial h = \mathbf{E}(\partial \mathbf{E}(V|\tau_2, \ldots, \tau_N)/\partial h)$.

Because the joint distribution of $\tau_1, \ldots, \tau_N$ admits a density, $\tau_i \neq \tau_j$ and $\tau_i \neq T$ with probability 1, for any $i \neq j$, $i, j = 2, \ldots, N$. Therefore,

$$\mathbf{E}\left(\frac{\partial \mathbf{E}(V|\tau_2, \ldots, \tau_N)}{\partial h}\right) = \mathbf{E}\left(\frac{\partial \mathbf{E}(V|\tau_2, \ldots, \tau_N, \tau_i \neq \tau_j, \tau_i \neq T, i \neq j, i, j = 2, \ldots, N)}{\partial h}\right).$$

In the following discussion, we consider only the case that $\tau_i \neq \tau_j$, $\tau_i \neq T$, $i \neq j$, $i, j = 2, \ldots, N$. So, our problem is reduced to finding an estimator of

$$\frac{\partial \mathbf{E}(V|\tau_2, \ldots, \tau_N, \tau_i \neq \tau_j, \tau_i \neq T, i \neq j, i, j = 2, \ldots, N)}{\partial h}. \tag{4.1}$$

We restrict attention to the almost-sure event

$$\{\tau_i \neq \tau_j, \tau_i \neq T, i \neq j, i, j = 2, \ldots, N\}$$

without making this restriction explicit in the following discussion.

The essential idea of the pathwise method is to interchange the order of differentiation and expectation. This requires that the expectation in (4.1) be taken in a measure that is independent of h. As in Condition (A7), we assume that the dependence on h enters through $\tau_1 = \tau_1(h)$, a random function of h. This makes V dependent on h through its dependence on τ_1.

Since $\partial \mathbf{E}(V|\tau_2, \ldots, \tau_N)/\partial h$ exists, we have

$$\frac{\partial \mathbf{E}(V|\tau_2, \ldots, \tau_N)}{\partial h} = \lim_{\epsilon \downarrow 0} \frac{\mathbf{E}(V^{(+\epsilon)}|\tau_2, \ldots, \tau_N) - \mathbf{E}(V^{(-\epsilon)}|\tau_2, \ldots, \tau_N)}{2\epsilon}$$
$$= \lim_{\epsilon \downarrow 0} \frac{\mathbf{E}(\Delta V^{(\epsilon)}|\tau_2, \ldots, \tau_N)}{2\epsilon},$$

where ϵ is positive and sufficiently small, we set $\tau_1^{(\pm\epsilon)} = \tau_1(h \pm \epsilon)$, $V^{(\pm\epsilon)} \equiv V(\tau_1^{(\pm\epsilon)}, \ldots, \tau_N)$, and $\Delta V^{(\epsilon)} \equiv V^{(+\epsilon)} - V^{(-\epsilon)}$. With the assumption of a.s. continuity of τ_1 as a function of h,

$$\lim_{\epsilon \downarrow 0} \tau_1^{(-\epsilon)} = \lim_{\epsilon \downarrow 0} \tau_1^{(+\epsilon)} = \tau_1.$$

Therefore, the left and right limits of V exist, a.s.; denote them by $V^{\pm} = \lim_{\epsilon \downarrow 0} V^{(\pm\epsilon)}$. Let ΔV denote $\lim_{\epsilon \downarrow 0} \Delta V^{(\epsilon)}$. On the event $\{\tau_1 = t_1\}$, for fixed t_1, we write ΔV_{t_1} for the value of ΔV to stress the dependence on τ_1.

We set $\tau_{N+1} \equiv T$ and sort $\{\tau_2, \ldots, \tau_{N+1}\}$ as

$$\tau_{(2)} \leq \cdots \leq \tau_{(N+1)}.$$

Let I_i denote the open interval $(\tau_{(i)}, \tau_{(i+1)})$, for $i = 2, \ldots, N$, and set $I_1 = (0, \tau_{(2)})$ and $I_{N+1} = (\tau_{(N+1)}, \infty)$. For convenience, we use $\tau_{(1)}$ and $\tau_{(N+2)}$ to denote 0 and ∞, so that $I_i = (\tau_{(i)}, \tau_{(i+1)})$ for $i = 1, \ldots, N+1$. For any $\epsilon > 0$, the possible positions of $\tau_1^{(+\epsilon)}$ and $\tau_1^{(-\epsilon)}$ fall into four cases:

- They are in the same interval I_i.
- They are in two successive intervals, i.e., $\tau_1^{(+\epsilon)} \in I_i$ and $\tau_1^{(-\epsilon)} \in I_{i+1}$.
- They are in two intervals which are not successive, i.e., $\tau_1^{(+\epsilon)} \in I_i$ and $\tau_1^{(-\epsilon)} \in I_j$, where $j > i + 1$.
- At least one of them coincides with a point in $\tau_{(2)}, \ldots, \tau_{(N+1)}$.

The fourth case happens with probability 0. By omitting the fourth case, we have

$$\begin{aligned}
&\mathbf{E}\big(\Delta V^{(\epsilon)}\big|\tau_2, \ldots, \tau_N\big) \\
&\quad= \sum_{i=1}^{N+1} \mathbf{E}\big(\Delta V^{(\epsilon)}\big|\tau_1^{(+\epsilon)}, \tau_1^{(-\epsilon)} \in I_i, \tau_2, \ldots, \tau_N\big)\mathbf{P}\big(\tau_1^{(+\epsilon)}, \tau_1^{(-\epsilon)} \in I_i\big|\tau_2, \ldots, \tau_N\big) \\
&\qquad+ \sum_{i=1}^{N}\Big(\mathbf{E}\big(\Delta V^{(\epsilon)}\big|\tau_1^{(+\epsilon)} \in I_i, \tau_1^{(-\epsilon)} \in I_{i+1}, \tau_2, \ldots, \tau_N\big) \\
&\qquad\times \mathbf{P}\big(\tau_1^{(+\epsilon)} \in I_i, \tau_1^{(-\epsilon)} \in I_{i+1}\big|\tau_2, \ldots, \tau_N\big)\Big) \\
&\qquad+ \sum_{\substack{j>i+1 \\ i,j=1}}^{N+1} \Big(\mathbf{E}\big(\Delta V^{(\epsilon)}\big|\tau_1^{(+\epsilon)} \in I_i, \tau_1^{(-\epsilon)} \in I_j, \tau_2, \ldots, \tau_N\big) \\
&\qquad\times \mathbf{P}\big(\tau_1^{(+\epsilon)} \in I_i, \tau_1^{(-\epsilon)} \in I_j\big|\tau_2, \ldots, \tau_N\big)\Big).
\end{aligned}$$

Then,

$$\lim_{\epsilon\downarrow 0}\frac{\mathbf{E}(\Delta V^{(\epsilon)}|\tau_2,\dots,\tau_N)}{2\epsilon}$$

$$=\sum_{i=1}^{N+1}\lim_{\epsilon\downarrow 0}\frac{\mathbf{E}(\Delta V^{(\epsilon)}|\tau_1^{(+\epsilon)},\tau_1^{(-\epsilon)}\in I_i,\tau_2,\dots,\tau_N)\mathbf{P}(\tau_1^{(+\epsilon)},\tau_1^{(-\epsilon)}\in I_i|\tau_2,\dots,\tau_N)}{2\epsilon}$$

$$+\sum_{i=1}^{N}\lim_{\epsilon\downarrow 0}\frac{1}{2\epsilon}\big(\mathbf{E}\big(\Delta V^{(\epsilon)}\big|\tau_1^{(+\epsilon)}\in I_i,\tau_1^{(-\epsilon)}\in I_{i+1},\tau_2,\dots,\tau_N\big)$$

$$\times\mathbf{P}\big(\tau_1^{(+\epsilon)}\in I_i,\tau_1^{(-\epsilon)}\in I_{i+1}\big|\tau_2,\dots,\tau_N\big)\big)$$

$$+\sum_{\substack{j>i+1\\ i,j=1}}^{N+1}\lim_{\epsilon\downarrow 0}\frac{1}{2\epsilon}\big(\mathbf{E}\big(\Delta V^{(\epsilon)}\big|\tau_1^{(+\epsilon)}\in I_i,\tau_1^{(-\epsilon)}\in I_j,\tau_2,\dots,\tau_N\big)$$

$$\times\mathbf{P}\big(\tau_1^{(+\epsilon)}\in I_i,\tau_1^{(-\epsilon)}\in I_j\big|\tau_2,\dots,\tau_N\big)\big).$$

We verify these limits and analyze the three sums on the right in turn.

In the first sum, inside any I_i, conditions (A2), (A7) and (A8) together allow us to apply the dominated convergence theorem to interchange expectation and the limit as $\epsilon\downarrow 0$ to get

$$\lim_{\epsilon\downarrow 0}\frac{\mathbf{E}(\Delta V^{(\epsilon)}|\tau_1^{(+\epsilon)},\tau_1^{(-\epsilon)}\in I_i,\tau_2,\dots,\tau_N)\mathbf{P}(\tau_1^{(+\epsilon)},\tau_1^{(-\epsilon)}\in I_i|\tau_2,\dots,\tau_N)}{2\epsilon}$$

$$=\lim_{\epsilon\downarrow 0}\frac{\mathbf{E}(\Delta V^{(\epsilon)}\mathbf{1}(\tau_1^{(+\epsilon)},\tau_1^{(-\epsilon)}\in I_i)|\tau_2,\dots,\tau_N)}{2\epsilon}$$

$$=\mathbf{E}\left(\lim_{\epsilon\downarrow 0}\frac{\Delta V^{(\epsilon)}}{2\epsilon}\mathbf{1}\big(\tau_1^{(+\epsilon)},\tau_1^{(-\epsilon)}\in I_i\big)\middle|\tau_2,\dots,\tau_N\right)$$

$$=\mathbf{E}\left(\frac{\partial V}{\partial h}\mathbf{1}(\tau_1\in I_i)\middle|\tau_2,\dots,\tau_N\right)=\mathbf{E}\left(\frac{\partial V}{\partial \tau_1}\tau_1'(h)\mathbf{1}(\tau_1\in I_i)\middle|\tau_2,\dots,\tau_N\right),$$

where $\mathbf{1}$ is the indicator function. Because V is Lipschitz in I_i, it follows that, with probability 1, the derivative $\partial V/\partial\tau_1$ exists almost everywhere in I_i. For the second and third sums, we need the following lemma.

Lemma 4.2 *Under conditions* (A7) *and* (A8),

(a) $\mathbf{P}(\tau_1^{(+\epsilon)}\in I_i,\tau_1^{(-\epsilon)}\notin I_i|\tau_2,\dots,\tau_N)\to 0$, *a.s., when* $\epsilon\downarrow 0$
(b) $\mathbf{P}(\tau_1^{(+\epsilon)}\in I_i,\tau_1^{(-\epsilon)}\in I_j, j>i+1|\tau_2,\dots,\tau_N)/\epsilon\to 0$, *a.s., when* $\epsilon\downarrow 0$
(c) $\mathbf{P}(\tau_1^{(+\epsilon)}\in I_i,\tau_1^{(-\epsilon)}\in I_{i+1}|\tau_2,\dots,\tau_N)/(2\epsilon)\to -\tau_1'(h_{(i)})f_1(\tau_{(i)};h|\tau_2,\dots,\tau_N)$, *a.s., when* $\epsilon\downarrow 0$, *where* $h_{(i)}$ *satisfies* $\tau(h_{(i)})=\tau_{(i)}$.

This lemma is proved in Appendix A.3.

Note that $\Delta V^{(\epsilon)}$ is bounded since V is bounded, so the limit ΔV is bounded. By the lemma,

$$\begin{aligned}
&\lim_{\epsilon\downarrow 0}\frac{1}{2\epsilon}\big(\mathbf{E}\big(\Delta V^{(\epsilon)}\big|\tau_1^{(+\epsilon)}\in I_i,\tau_1^{(-\epsilon)}\in I_{i+1},\tau_2,\ldots,\tau_N\big)\\
&\quad\times\mathbf{P}\big(\tau_1^{(+\epsilon)}\in I_i,\tau_1^{(-\epsilon)}\in I_{i+1}\big|\tau_2,\ldots,\tau_N\big)\big)\\
&=-\Delta V_{\tau_{(i)}}\frac{\partial\tau_1}{\partial h}\bigg|_{h=h_{(i)}} f_1(\tau_{(i)};h|\tau_2,\ldots,\tau_N)\\
&=-\Delta V_{\tau_{(i)}}\tau_1'(h_{(i)})f_1(\tau_{(i)};h|\tau_2,\ldots,\tau_N),
\end{aligned}$$

and for $j>i+1$,

$$\begin{aligned}
&\lim_{\epsilon\downarrow 0}\frac{1}{2\epsilon}\big(\mathbf{E}\big(\Delta V^{(\epsilon)}\big|\tau_1^{(+\epsilon)}\in I_i,\tau_1^{(-\epsilon)}\in I_j,\tau_2,\ldots,\tau_N\big)\\
&\quad\times\mathbf{P}\big(\tau_1^{(+\epsilon)}\in I_i,\tau_1^{(-\epsilon)}\in I_j\big|\tau_2,\ldots,\tau_N\big)\big)=0.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{\partial\mathbf{E}(V|\tau_2,\ldots,\tau_N)}{\partial h}&=\lim_{\epsilon\downarrow 0}\frac{\mathbf{E}(\Delta V^{(\epsilon)}|\tau_2,\ldots,\tau_N)}{2\epsilon}\\
&=\sum_{i=1}^{N+1}\mathbf{E}\left(\frac{\partial V}{\partial\tau_1}\frac{\partial\tau_1}{\partial h}\mathbf{1}(\tau_1\in I_i)\middle|\tau_2,\ldots,\tau_N\right)\\
&\quad-\sum_{i=2}^{N+1}\Delta V_{\tau_{(i)}}\tau_1'(h_{(i)})f_1(\tau_{(i)};h|\tau_2,\ldots,\tau_N)\\
&=\mathbf{E}\left(\frac{\partial V}{\partial h}\middle|\tau_2,\ldots,\tau_N\right)-\sum_{i=2}^{N+1}\Delta V_{\tau_{(i)}}\tau_1'(h_{(i)})f_1(\tau_{(i)};h|\tau_2,\ldots,\tau_N).
\end{aligned}$$

Now we conclude that

$$\begin{aligned}
\frac{\partial\mathbf{E}(V)}{\partial h}&=\mathbf{E}\left(\frac{\partial\mathbf{E}(V|\tau_2,\ldots,\tau_N)}{\partial h}\right)\\
&=\mathbf{E}\left(\mathbf{E}\left(\frac{\partial V}{\partial h}\middle|\tau_2,\ldots,\tau_N\right)-\sum_{i=2}^{N+1}\Delta V_{\tau_{(i)}}\tau_1'(h_{(i)})f_1(\tau_{(i)};h|\tau_2,\ldots,\tau_N)\right)\\
&=\mathbf{E}\left(\frac{\partial V}{\partial h}-\sum_{i=2}^{N+1}\Delta V_{\tau_{(i)}}\tau_1'(h_{(i)})f_1(\tau_{(i)};h|\tau_2,\ldots,\tau_N)\right).
\end{aligned}$$

We therefore obtain from the (smoothed) pathwise method the unbiased estimator

$$\frac{\partial V}{\partial h}-\sum_{i=2}^{N+1}\Delta V_{\tau_{(i)}}\tau_1'(h_{(i)})f_1(\tau_{(i)};h|\tau_2,\ldots,\tau_N).\tag{4.2}$$

We have thus proved the following:

Theorem 4.3 *Under conditions* (A1)–(A8), *the estimator* (4.2) *of delta given by the* (*smoothed*) *pathwise method is unbiased.*

This estimator may be interpreted as follows. The first term is the (unsmoothed) pathwise derivative; it captures the effect of local changes in τ_1 (through changes in h) that do not introduce changes in the order of the defaults (or changes in which defaults occur before T). The sum captures the effect of possible order changes. Term i is the product of the rate at which τ_1 and $\tau_{(i)}$ change order, and $\Delta V_{\tau_{(i)}}$ is the size of the jump when the order change occurs. The rate of the order change is the product of the probability (density) that τ_1 falls exactly on $\tau_{(i)}$, multiplied by the derivative of τ_1 with respect to h when $\tau_1 = \tau_{(i)}$. The particular form of the jump ΔV depends on the payoff; in this sense, the pathwise method requires more information about a particular derivative contract than does the likelihood ratio method. We illustrate the estimators with an example in the next section.

5 Application to portfolio credit derivatives

This section discusses the sensitivities with respect to hazard rates in the cases of basket default swaps and collateralized debt obligations (CDOs). It illustrates the application of the general estimators and the verification of the conditions used in establishing unbiasedness.

In the following discussion, we denote by $D(t)$ the discount factor for the interval from 0 to t and take this to be deterministic; in the simplest case, $D(t) = \exp(-rt)$ for some fixed rate r. We assume $D(t)$ to be Lipschitz in t. We suppose the default times $\tau_1, \dots, \tau_N$ are correlated through a Gaussian copula function as described before. The hazard function of the first asset is $\lambda_1(t,h)$; we let h vary over an open set $\mathcal{H} = \{h : h > h^*\}$. It is natural to require that $\inf_{t,h} \lambda_1(t,h) > 0$; this ensures that $\mathbf{E}(\tau_1(h)) < \infty$, for all h. We also assume that $\lambda_1(t,h)$ is a strictly increasing differentiable function of h, and $\lambda_1(t,h)/\partial h$ is bounded for all h and t. Under these conditions, we verify in Appendix A.4 that the conditions listed in Sect. 2 are satisfied for basket default swaps and CDOs.

5.1 Basket default swaps

Basket default swaps are derivative securities tied to an underlying portfolio of corporate bonds or other assets subject to credit risk. A basket default swap provides protection against the nth default in the basket, with n smaller than N and typically much smaller. This type of nth-to-default swap is less expensive than insuring each asset separately and may provide adequate protection if multiple defaults are unlikely. Its cashflows are as follows. At dates $0 < T_1 < T_2 < \cdots < T_m \le T$, the protection *buyer* is scheduled to make fixed payments of $s_1, \dots, s_m$ to the protection *seller*. However, if the nth default occurs before T, these payments cease and the protection seller makes a payment to the protection buyer. This payment is determined by the identity of the nth asset to default, but is otherwise fixed. If the ith asset is the nth to default, the payment is $1 - r_i$, where r_i is the *recovery rate* and 1 is the (normalized) asset value. (Differences in asset values can be absorbed into differences in recovery rates.) We denote by R the recovery rate for the nth asset to default.

As in Chen and Glasserman [5], or in Joshi and Kainth [14] with signs reversed, we write the discounted payoff of the swap as the difference between the discounted payoffs of the payments made between the parties, called the protection leg and the value leg:

$$\widehat{V}(\tau_1, \dots, \tau_N) = V_{\texttt{value}}(\tau_1, \dots, \tau_N) - V_{\texttt{prot}}(\tau_1, \dots, \tau_N).$$

Let τ denote the time of the nth default. Then

$$V_{\texttt{value}}(\tau_1, \dots, \tau_N) = (1 - R)D(\tau)\mathbf{1}(\tau \le T),$$

with $\mathbf{1}(\tau \le T)$ the indicator of the event that the nth default occurs before T; and

$$V_{\texttt{prot}}(\tau_1, \dots, \tau_N) = \begin{cases} \sum_{i=1}^{j} s_i D(T_i) + s_{j+1} D(\tau) \frac{\tau - T_j}{T_{j+1} - T_j}, & \text{if } T_j \le \tau \le T_{j+1}, \\ \sum_{i=1}^{m} s_i D(T_i), & \text{if } \tau > T. \end{cases}$$

The first term indicates that upon the nth default, the protection buyer makes an accrued payment to the protection seller; these payments accrue linearly between the dates T_j.

We can subtract the deterministic component of the swap and define

$$V(\tau_1, \dots, \tau_N) = \widehat{V}(\tau_1, \dots, \tau_N) + \sum_{i=1}^{m} s_i D(T_i),$$

so it suffices to compute $\mathbf{E}(V(\tau_1, \dots, \tau_N))$ for pricing.

5.1.1 The likelihood ratio method estimator

The likelihood ratio estimate is straightforward and has the same form as in the generic case (3.1), where

$$\frac{\partial \ln f(\tau_1, \dots, \tau_N)}{\partial h} = -\sum_{j=1}^{N} \big(\boldsymbol{\Sigma}^{-1} - \mathbf{I}\big)_{1j} \Phi^{-1}(u_j) \sqrt{2\pi} e^{\Phi^{-1}(u_1)^2/2} \frac{\partial \lambda_1}{\partial h} \tau_1 (1 - u_1) + \frac{1}{\lambda_1} \frac{\partial \lambda_1}{\partial h} - \tau_1 \frac{\partial \lambda_1}{\partial h}.$$

Joshi and Kainth [14] derived the same formula in the case of constant hazard rates. In that case, the likelihood ratio estimator is

$$V(\tau_1, \dots, \tau_N) \left[\frac{1}{\lambda_1} - \tau_1 - \sqrt{2\pi} e^{\Phi^{-1}(u_1)^2/2} \tau_1 (1 - u_1) \sum_{j=1}^{N} \big(\boldsymbol{\Sigma}^{-1} - \mathbf{I}\big)_{1j} \Phi^{-1}(u_j) \right].$$

5.1.2 The pathwise estimator

Next, we consider the pathwise estimator for an nth-to-default swap. In this setting, if τ_1 is not the nth default time, a small change in τ_1 does not change V, so $\partial V/\partial\tau_1 = 0$. If τ_1 is the nth default time but $\tau_1 > T$, then $V = 0$ and again $\partial V/\partial\tau_1 = 0$. Thus,

$$\frac{\partial V}{\partial h} = \frac{\partial V}{\partial h}\mathbf{1}(\tau_1 = \tau \le T),$$

where, as before, τ is the time of the nth default.

To make the second part of the estimator more explicit, we write $h_T = \tau_1^{-1}(T)$ and $h_{(i)} = \tau_1^{-1}(\tau_{(i)})$. Then

$$\begin{aligned}
&\sum_{i=2}^{N+1} \Delta V_{\tau_{(i)}} \tau_1'(h_{(i)}) f_1(\tau_{(i)}; h|\tau_2, \ldots, \tau_{N+1}) \\
&\quad = \Delta V_{\tau_{(n-1)}} \mathbf{1}(\tau_{(n-1)} < T)\tau_1'(h_{(n-1)}) f_1(\tau_{(n-1)}; h|\tau_2, \ldots, \tau_N) \\
&\qquad + \Delta V_{\tau_{(n)}} \mathbf{1}(\tau_{(n)} < T)\tau_1'(h_{(n)}) f_1(\tau_{(n)}; h|\tau_2, \ldots, \tau_N) \\
&\qquad + \Delta V_T \mathbf{1}(\tau_{(n-1)} \le T \le \tau_{(n)})\tau_1'(h_T) f_1(T; h|\tau_2, \ldots, \tau_N).
\end{aligned}$$

The jump terms are as follows:

$$\begin{aligned}
\Delta V_{\tau_{(n-1)}} \mathbf{1}(\tau_{(n-1)} < T) &= D(\tau_{(n-1)})(r_1 - r_{(n-1)})\mathbf{1}(\tau_{(n-1)} < T), \\
\Delta V_{\tau_{(n)}} \mathbf{1}(\tau_{(n)} < T) &= D(\tau_{(n)})(r_{(n)} - r_1)\mathbf{1}(\tau_{(n)} < T), \\
\Delta V_T \mathbf{1}(\tau_{(n-1)} \le T \le \tau_{(n)}) &= D(T)(1 - r_1)\mathbf{1}(\tau_{(n-1)} \le T \le \tau_{(n)}).
\end{aligned}$$

In the first case, τ_1 and $\tau_{(n-1)}$ change order, changing the identity of the nth default; in the second case, τ_1 and $\tau_{(n)}$ change order, changing the identity of the nth default; and in the third case, τ_1 crosses T, causing the time of the nth default to cross the end of the life of the swap.

Combining these terms, we get the estimator

$$\begin{aligned}
&\frac{\partial V}{\partial h}\mathbf{1}(\tau_1 = \tau < T) \\
&\quad - D(\tau_{(n-1)})(r_1 - r_{(n-1)})\mathbf{1}(\tau_{(n-1)} < T)\tau_1'(h_{(n-1)}) f_1(\tau_{(n-1)}; h|\tau_2, \ldots, \tau_N) \\
&\quad - D(\tau_{(n)})(r_{(n)} - r_1)\mathbf{1}(\tau_{(n)} < T)\tau_1'(h_{(n)}) f_1(\tau_{(n)}; h|\tau_2, \ldots, \tau_N) \\
&\quad - D(T)(1 - r_1)\mathbf{1}(\tau_{(n-1)} \le T \le \tau_{(n)})\tau_1'(h_T) f_1(T; h|\tau_2, \ldots, \tau_N).
\end{aligned} \tag{5.1}$$

Joshi and Kainth [14] arrived at essentially the same estimator using delta functions. The sign of the first two terms in (8.2) of Joshi and Kainth [14] appears to be incorrect. In their jump terms (8.4) and (8.5), it appears that the swap lifetime T should be replaced by the default time τ_j.

5.1.3 Numerical examples

We illustrate the likelihood ratio method and the pathwise method discussed so far, as well as the finite difference method, with some numerical results. In these examples, we take an interest rate of $r = 5\%$. For simplicity, we assume just a single protection payment (i.e., $m = 1$) of $s = .10$, paid at maturity if fewer than n defaults have occurred. In the finite difference method, we use $\epsilon = 0.001$ as the increment. These parameters will be used in subsequent sections as well, and all numerical results are based on 2×10^5 replications.

Basket I–Swap A1

As a first illustration, we consider a basket of $N = 10$ independent assets, with constant hazard rates (0.1, 0.02, 0.015, 0.025, 0.1, 0.3, 0.01, 0.25, 0.15, 0.03). The recovery rates are (0.3, 0.1, 0.2, 0.1, 0.3, 0.1, 0.2, 0.2, 0.1, 0.3). Swap A1 is a fourth-to-default swap in Basket I.

Basket II–Swap A2

Basket II contains $N = 10$ assets with (the same) constant hazard rates $(0.1, 0.02, 0.015, 0.025, 0.1, 0.3, 0.01, 0.25, 0.15, 0.03)$. The recovery rates are also the same as above, namely (0.3, 0.1, 0.2, 0.1, 0.3, 0.1, 0.2, 0.2, 0.1, 0.3). They are correlated and the correlation matrix $\mathbf{\Sigma}$, has a three-factor structure. To generate such a correlation matrix, one can first randomly generate a 10×3 matrix $\mathbf{A}$, compute the supplementary vector $\mathbf{B}$ such that $B_i = \sqrt{1 - \sum_{j=1}^{3} a_{ij}^2}$, then let $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}' + \mathbf{B}\mathbf{B}'$. The matrix $\mathbf{A}$ we use is

$$
\mathbf{A} = \begin{pmatrix}
0.0815 & -0.4105 & -0.4589 \\
-0.5187 & -0.4044 & -0.1367 \\
0.0449 & -0.4735 & -0.6456 \\
0.5795 & 0.5493 & 0.3349 \\
-0.4976 & -0.3295 & -0.3440 \\
-0.4963 & 0.3291 & -0.1737 \\
0.2841 & 0.2423 & 0.3518 \\
0.5203 & -0.4565 & -0.3577 \\
0.5943 & -0.0715 & -0.5576 \\
-0.4050 & -0.0124 & 0.3809
\end{pmatrix}
$$

Swap A2 is a fourth-to-default swap in Basket II.

The results are displayed in Figs. 1 and 2, which plot the estimated delta and variance against maturity for swaps A1 and A2. In the right panel, the dashed line shows the variance of the finite difference estimator. The dotted line shows the variance of the likelihood ratio method estimator, and the solid line shows the variance of the pathwise estimator.

Figures 1 and 2 show that the finite difference method is very inefficient and unstable. The likelihood ratio method and the pathwise method perform much better than the finite difference method. Although the pathwise method takes slightly longer (about 1.1–1.2 times) than the likelihood ratio method, it has much smaller variance.

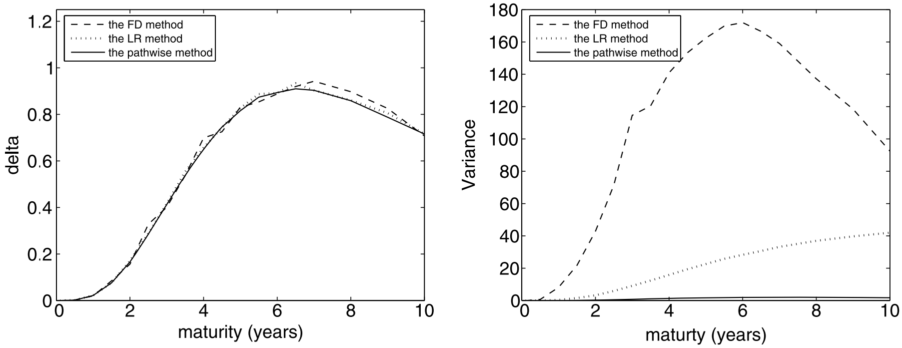

Fig. 1 The estimated delta (*left panel*) and its variance (*right panel*) of Swap A1

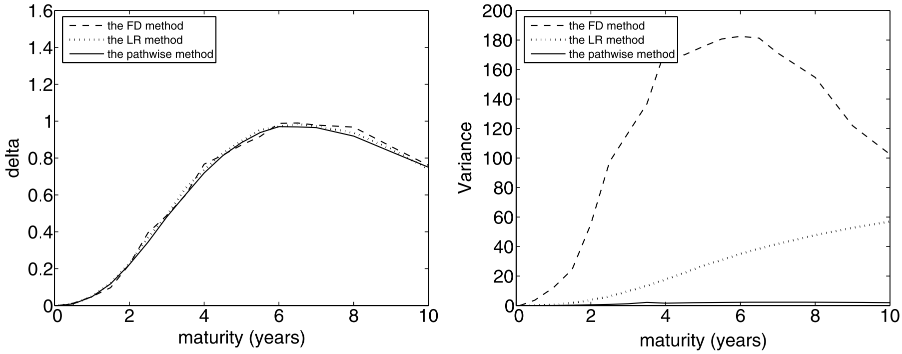

Fig. 2 The estimated delta (*left panel*) and its variance (*right panel*) of Swap A2

5.2 Collateralized debt obligations

Another popular type of security backed by a pool of defaultable assets is collateralized debt obligations (CDOs). In a CDO, credit loss on the pool is tranched and passed to different investors. Lower seniority tranches act as cushions against the loss in higher seniority tranches. When the ith asset defaults, it causes a normalized constant loss of $l_i = 1 - r_i$ in the portfolio, which is called the loss given default (LGD) of the ith asset. A tranche of a CDO absorbs losses from an attachment point S_ℓ to a detachment point S_u. The cashflows of this tranche are as follows. At dates $0 < T_1 < \cdots < T_m \le T$, the tranche holder receives payment proportional to the notional principal left in the tranche. If there are default losses in the portfolio, the tranche covers the cumulative portfolio loss in excess of S_ℓ and up to S_u. For simplicity, we assume that the net default payments occur only at the coupon dates $T_1, \ldots, T_m$.

Let $L(t)$ be the cumulative loss on the collateral portfolio at time $t \le T$, i.e., $L(t) = \sum_{i=1}^N l_i \mathbf{1}(\tau_i \le t)$. The cumulative loss on the tranche at time t is

$$M(t) = \big(L(t) - S_\ell\big)^+ - \big(L(t) - S_u\big)^+.$$

As shown in Cherubini et al. [6], the discounted payoff of a CDO tranche could be written as the difference between the default payment leg and the premium payment leg,

$$V(\tau_1,\dots,\tau_N) = V_{\mathrm{def}}(\tau_1,\dots,\tau_N) - V_{\mathrm{pre}}(\tau_1,\dots,\tau_N).$$

Using the cumulative loss on the tranche, the payoff of the default payment leg can be expressed as

$$V_{\mathrm{def}}(\tau_1,\dots,\tau_N) = \sum_{i=1}^{m} D(T_i)\big(M(T_i) - M(T_{i-1})\big),$$

where $T_0 = 0$. If we ignore the accrual factor for payment days, the discounted payoff of the premium payment leg is

$$V_{\mathrm{pre}}(\tau_1,\dots,\tau_N) = c\sum_{i=1}^{m} D(T_i)\big(S_u - S_\ell - M(T_i)\big),$$

where the constant c is the *spread* or the coupon rate of this tranche. Thus, the payoff of the CDO tranche can be written as

$$V(\tau_1,\dots,\tau_N) = (1+c)\sum_{i=1}^{m} D(T_i)M(T_i) - \sum_{i=1}^{m-1} D(T_{i+1})M(T_i) \\ - c(S_u - S_\ell)\sum_{i=1}^{m} D(T_i).$$

We observe that $V(\tau_1,\dots,\tau_N)$ is a linear combination of the cumulative loss on the tranche $M(t)$ for $t = T_1,\dots,T_m$. So for the sensitivity

$$\frac{\partial \mathbf{E}(V)}{\partial h} = (1+c)\sum_{i=1}^{m} D(T_i)\frac{\partial \mathbf{E}(M(T_i))}{\partial h} - \sum_{i=1}^{m-1} D(T_{i+1})\frac{\partial \mathbf{E}(M(T_i))}{\partial h},$$

it is sufficient to compute

$$\frac{\partial \mathbf{E}(M(t))}{\partial h}, \quad \text{for } t = T_1,\dots,T_m.$$

5.2.1 *Sensitivity estimators*

The likelihood ratio estimator is straightforward and has the same form as in the generic case (3.1) and in the basket default swap case, but with a different expression for $V(\tau_1,\dots,\tau_N)$.

Next, we consider the pathwise method for a CDO tranche with attachment point S_ℓ and detachment point S_u. As pointed out in the previous section, the sensitivity of the tranche payoff with respect to h is a linear combination of the sensitivities of $M(t)$ for $t = T_1, \ldots, T_m$, with

$$M(t) = \left(\sum_{i=1}^{N} l_i \mathbf{1}(\tau_i \le t) - S_\ell\right)^+ - \left(\sum_{i=1}^{N} l_i \mathbf{1}(\tau_i \le t) - S_u\right)^+.$$

A change in τ_1 that is sufficiently small to leave the number of defaults in $[0, t]$ unchanged does not change $M(t)$. Therefore, the continuous part of the sensitivity estimate is

$$\frac{\partial M(t)}{\partial h} = \frac{\partial M(t)}{\partial \tau_1} \frac{\partial \tau_1}{\partial h} = 0.$$

To affect $M(t)$, a change in τ_1 must cause τ_1 to cross t. Moreover, if we have $\sum_{i=2}^{m} l_i \mathbf{1}(\tau_i \le t) \le S_\ell - l_1$ or $\sum_{i=2}^{m} l_i \mathbf{1}(\tau_i \le t) \ge S_u$, then $M(t) = 0$ or $M(t) = S_u$, accordingly, regardless of whether τ_1 is before or after t. Let $h_t = \tau_1^{-1}(t)$; then we can write the jump terms of the pathwise estimator as

$$\sum_{i=2}^{N+1} \Delta M_{\tau_{(i)}} \tau_1'(h_{(i)}) f_1(\tau_{(i)}; h|\tau_2, \ldots, \tau_{N+1}) = \Delta M_t \tau_1'(h_t) f_1(t; h|\tau_2, \ldots, \tau_{N+1}),$$

where

$$\begin{aligned}\Delta M_t = {} & \left[l_1 + L_2^m(t) - S_\ell - \left(l_1 + L_2^m(t) - S_u\right)^+ - \left(L_2^m(t) - S_\ell\right)^+\right] \\ & \times \mathbf{1}\left(S_\ell - l_1 < L_2^m(t) < S_u\right),\end{aligned}$$

and

$$L_2^m(t) = \sum_{i=2}^{m} \left(l_i \mathbf{1}(\tau_i \le t)\right).$$

This term describes the case that the cumulative portfolio loss jumps into or out of the tranche when τ_1 crosses t. Combining both parts, we obtain the pathwise estimator for the sensitivity of $\mathbf{E}(M(t))$ as

$$-\Delta M_t \tau_1'(h_t) f_1(t; h|\tau_2, \ldots, \tau_{N+1}).$$

The pathwise estimator for the tranche value $\mathbf{E}(V(\tau_1, \ldots, \tau_N))$ is

$$\begin{aligned}& -(1+c) \sum_{i=1}^{m} D(T_i) \Delta M_{T_i} \tau_1'(h_{T_i}) f_1(T_i; h|\tau_2, \ldots, \tau_{N+1}) \\ & \quad + \sum_{i=1}^{m-1} D(T_{i+1}) \Delta M_{T_i} \tau_1'(h_{T_i}) f_1(T_i; h|\tau_2, \ldots, \tau_{N+1}).\end{aligned}$$

5.2.2 Numerical examples

We consider two CDOs, C1 and C2, with tranches 0%–3%, 3%–7%, 7%–10%, 10%–15%, 15%–30%, and 30%–100%. Each CDO has a pool of $N = 200$ underlying assets and a maturity of 5 years. The coupon rate of each tranche is 3% and is paid quarterly. We use a risk-free rate of $r = 5\%$. All Monte Carlo simulation results are based on 10^6 replications.

CDO C1

CDO C1 is on a portfolio of 200 independent assets, which are divided into four groups. The assets in group 1, 2, 3, and 4 have constant hazard rates of 0.5, 0.1, 0.12, and 0.2 respectively, and have LGDs of 0.9, 0.6, 0.5 and 0.1.

CDO C2

The underlying portfolio of CDO C2 also contains 200 assets. The assets are divided into four groups with constant hazard rate of 0.5, 0.1, 0.12, 0.2, and LGDs of 0.9, 0.6, 0.5, 0.1, respectively. They are correlated and the correlation matrix $\mathbf{\Sigma}$ has a three-factor structure. The factor loading matrix for the four groups is given by

$$\begin{pmatrix} 0.0815 & -0.4105 & -0.4589 \\ -0.5187 & -0.4044 & -0.1367 \\ 0.0449 & -0.4735 & -0.6456 \\ 0.5795 & 0.5493 & 0.3349 \end{pmatrix}$$

We first estimate sensitivities of $\mathbf{E}(M(t))$ for super-senior tranches (30%–100%) at $t = T_1, \ldots, T_m$. The results are displayed in Figs. 3 and 4, which plot the estimated delta and variance against coupon dates for CDOs C1 and C2. We compare finite differences, pathwise estimates and likelihood ratio method estimates. In the right panel, the dashed line shows the variance of the finite difference estimator. The dotted line shows the variance of the likelihood ratio method estimator, and the solid line shows the variance of the pathwise estimator.

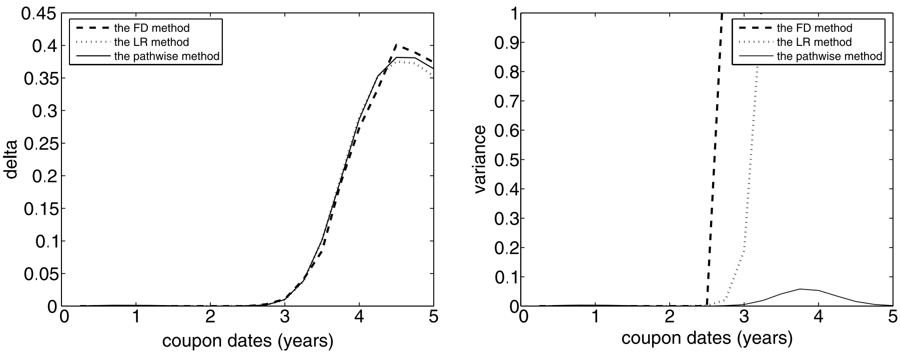

Fig. 3 The estimated delta of $M(t)$ (*left panel*) and its variance (*right panel*) of CDO C1

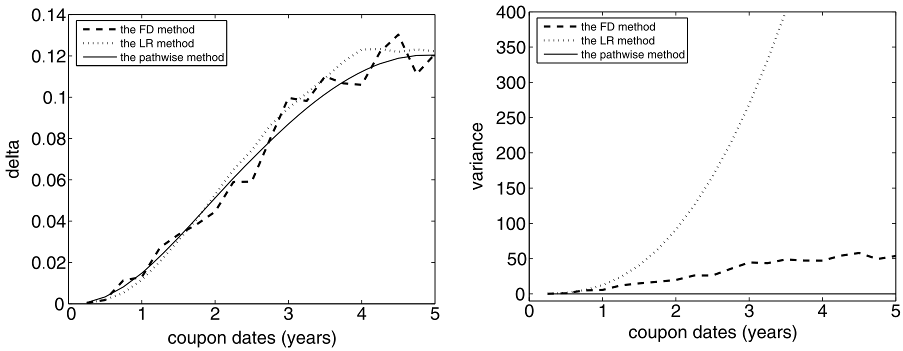

Fig. 4 The estimated delta of $M(t)$ (*left panel*) and its variance (*right panel*) of CDO C2

Table 1 Numerical results of CDO C1

	The FD method	The LR method	The pathwise method
0%–3%	0.0025 (0.0002)	−0.0212 (0.0182)	0.0024 (0.0361e−4)
3%–7%	0.0147 (0.0005)	0.0060 (0.0236)	0.0151 (0.0774e−4)
7%–10%	0.0189 (0.0006)	0.0265 (0.0170)	0.0188 (0.0987e−4)
10%–15%	0.0446 (0.0009)	0.0377 (0.0273)	0.0462 (0.1601e−4)
15%–30%	0.2390 (0.0022)	0.2365 (0.0688)	0.2385 (0.4787e−4)
30%–100%	0.3475 (0.0103)	0.3260 (0.0138)	0.3520 (0.5052e−4)

Table 2 Numerical results of CDO C2

	The FD method	The LR method	The pathwise method
0%–3%	0.0069 (0.0003)	−0.0046 (0.0221)	0.0072 (0.0081e−4)
3%–7%	0.0267 (0.0007)	−0.0078 (0.0279)	0.0273 (0.0276e−4)
7%–10%	0.0338 (0.00011)	0.0264 (0.0200)	0.0317 (0.0436e−4)
10%–15%	0.0763 (0.0022)	0.0536 (0.0314)	0.0759 (0.1132e−4)
15%–30%	0.4044 (0.0087)	0.3309 (0.0735)	0.3893 (0.3343e−4)
30%–100%	0.1439 (0.0060)	0.0972 (0.0331)	0.1415 (0.1437e−4)

Tables 1 and 2 show the performance of the three methods in estimating the tranche sensitivities. We give the estimated value and standard error for each delta estimate.

Figures 3 and 4 and Tables 1 and 2 show that the pathwise method produces much more precise estimates than the finite difference method and the likelihood ratio method. In fact, for the case of dependent assets, the likelihood ratio method is less precise than the finite difference method. The pathwise method takes about 2/3 of the computing time required for the finite difference method in the independent assets case, and about 2.5 times as much as in the dependent assets case. The likelihood ratio method takes roughly 3/5 of the computing time of the finite difference method in both cases. Considering both variance reduction and computing time, the pathwise method is much more effective than the other two methods. We should stress that the pathwise method would not be applicable to this problem without the inclusion of the jump terms.

6 Applying importance sampling methods

In calculating price sensitivities for credit derivatives, we encounter a difficulty that is also present in the estimation of the prices themselves: in baskets of high-quality credits, defaults are rare, making ordinary Monte Carlo inefficient.

Joshi and Kainth [14] and Chen and Glasserman [5] developed importance sampling techniques for basket default swaps that force at least n defaults to occur on every path. They accomplish this by sequentially increasing the default probabilities for the names in the basket. The methods differ in how they accomplish this.

Forcing n defaults means that the original probability measure is not absolutely continuous with respect to the new (importance sampling) probability. While this could lead to biased estimates in general, it introduces no bias in the pricing of nth-to-default swaps because $V = 0$ in the event that fewer than n defaults occur before time T.

Combining this approach to importance sampling with pathwise sensitivity estimation, however, requires care; indeed, a straightforward combination produces incorrect results, a phenomenon apparently overlooked in Joshi and Kainth [14]. The issue may be understood as follows. One of the jump terms in the pathwise method arises from the possibility that a small change in h may cause the number of defaults to increase from $n-1$ to n. An importance sampling technique that forces n defaults to occur on every path fails to capture this term. We correct this by modifying the importance sampling scheme to force at least $n-1$ defaults, rather than just n or more defaults.

We refer to the importance sampling method of Joshi and Kainth [14] as the JK method. Chen and Glasserman [5] call their version the Conditional Probability (CP) method. The JK method uses somewhat arbitrary default probabilities for importance sampling; in the CP method, the default probability for each name is set equal to its conditional probability of default, given that at least n names default. Chen and Glasserman [5] show that these importance sampling probabilities are, in a sense, optimal and guarantee variance reduction.

6.1 Importance sampling in the likelihood ratio method

Recall the likelihood ratio method estimator of delta in (3.1). Also, $V(\tau_1, \dots, \tau_N) = 0$ when $\tau > T$, and $\partial \ln f(\tau_1, \dots, \tau_N)/\partial h$ exists almost everywhere in $\mathcal{H}$, almost surely. Thus

$$V(\tau_1, \dots, \tau_N)\frac{\partial \ln f(\tau_1, \dots, \tau_N)}{\partial h} = 0 \quad \text{when } \tau > T.$$

Therefore, the CP method (and the JK method) can be applied with the likelihood ratio method estimator by taking $V(\tau_1, \dots, \tau_N)\partial \ln f(\tau_1, \dots, \tau_N)/\partial h$ as the new "payoff."

6.2 Importance sampling in the pathwise method

The pathwise estimator in (5.1) is zero whenever fewer than $n-1$ of $\tau_2, \ldots, \tau_N$ are less than T. However, we cannot ignore the event that exactly $n-1$ of these defaults occur before T, as we could in pricing the swap; doing so would fail to capture the second term and the last term in (5.1). Thus, we apply the CP method forcing $n-1$ defaults on every path, rather than n defaults on every path.

6.3 Numerical examples

We compare the likelihood ratio method and the pathwise method with and without importance sampling numerically using the test cases Swaps A1 and A2, using 2×10^5 replications.

Figure 5 illustrates the error that results from a straightforward combination of the JK method with the pathwise estimator. In that figure, the solid line shows the correct value (produced by the pathwise estimator), the dashed line with circles is produced by the pathwise method combined with a direct application of the JK method, and the solid line with circles shows the sum of the left out value and the (biased) estimates produced by a direct application of the JK method with the pathwise estimator.

Figures 6 and 7 show substantial variance reduction achieved by applying the importance sampling method proposed by Chen and Glasserman [5]. Both the likelihood ratio method estimator and the pathwise estimator benefit substantially from the use of importance sampling; the pathwise exhibits lower variance than the likelihood ratio method estimator, both with and without importance sampling.

7 Concluding remarks

In this article, we have derived and analyzed estimators of delta—the sensitivity of the price with respect to the hazard rate of an underlying asset—for a general class of portfolio credit derivatives. These estimators build on work of Joshi and Kainth [14].

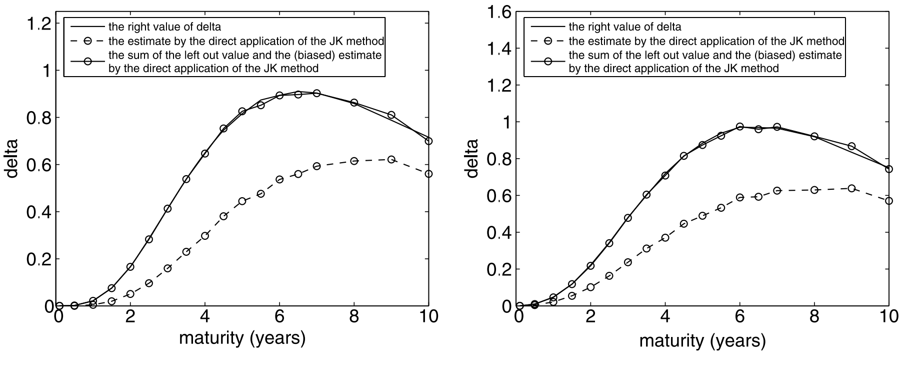

Fig. 5 The direct application of the JK method in the pathwise method provides incorrect estimates (*left panel*: Swap A1, *right panel*: Swap A2)

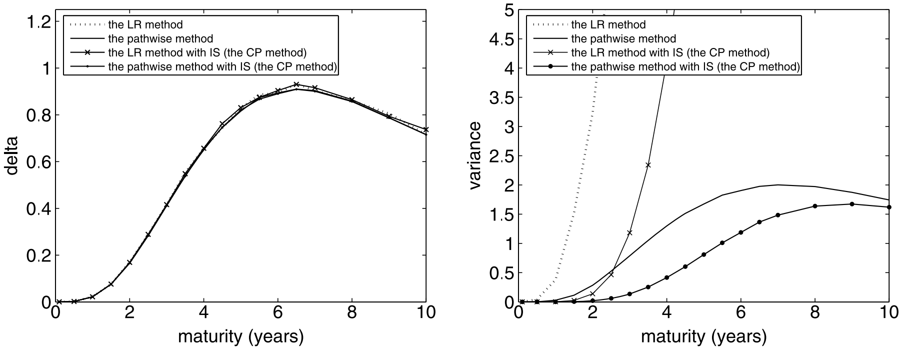

Fig. 6 The estimated delta (*left panel*) and its variance (*right panel*) of Swap A1

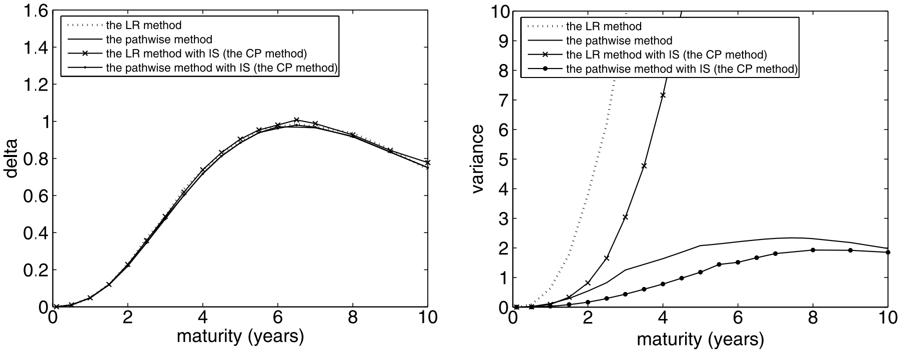

Fig. 7 The estimated delta (*left panel*) and its variance (*right panel*) of Swap A2

We have provided broadly applicable conditions for the application of likelihood ratio method and pathwise estimators. An important feature of the pathwise estimators derived and analyzed here is that they use conditional expectations to smooth the effect of changes in the order of defaults.

Combining the estimators with importance sampling produces substantial variance reduction. However, the combination of the pathwise estimator with importance sampling requires care: We show that in the case of an nth-to-default swap forcing n defaults on every path—which is effective in pricing—introduces an error in delta estimates. This error is eliminated by forcing just $n-1$ defaults instead. This adjustment is closely connected to the smoothing of jumps.

In addition to basket default swaps and CDOs, the methods can be applied to more complicated products, including virtually any portfolio credit derivative for which the payoff V is a function of the underlying default times. In our examples, we have focused on sensitivities of individual contracts; however, the methods can be applied easily to sensitivities calculated at the book level, provided the various deals in the book are valued and hedged in a consistent model of the joint distribution of default times. In deriving and implementing sensitivity estimators, the only distinction is that

V should now be interpreted as the payoff of a portfolio of contracts, rather than the payoff of a single contract.

Acknowledgements The authors thank Mark Joshi and Dherminder Kainth for helpful discussions about their work, and we thank the referees for their comments and suggestions. This work is supported by a grant from Moody's Corporation and by NSF grants DMI0300044 and DMS0410234.

Appendix A: Proofs

A.1 Proof of Theorem 3.1

A fairly generic result on the unbiasedness of the likelihood ratio method is given in Asmussen and Glynn [2], Proposition 7.3.5. However, in specific cases, one still needs to verify conditions for interchanging derivative and expectation.

The delta, if it exists, is given by

$$\begin{aligned}
\frac{\partial \mathbf{E}(V)}{\partial h} &= \lim_{\epsilon\to 0} \frac{\mathbf{E}(V(\tau_1(h+\epsilon),\ldots,\tau_N)) - \mathbf{E}(V(\tau_1(h-\epsilon),\ldots,\tau_N))}{2\epsilon} \\
&= \lim_{\epsilon\to 0} \frac{1}{2\epsilon} \int_0^\infty \cdots \int_0^\infty \big[V(t_1,\ldots,t_N) \\
&\quad \times \big(f(t_1,\ldots,t_N;h+\epsilon) - f(t_1,\ldots,t_N;h-\epsilon)\big)\big]\, dt_1\cdots dt_N.
\end{aligned}$$

The last equality holds because the payoff function V does not depend on h explicitly. Because V is bounded and

$$\big|\big(f(\tau_1,\ldots,\tau_N;h+\epsilon) - f(\tau_1,\ldots,\tau_N;h-\epsilon)\big)\big/(2\epsilon)\big|$$

is bounded by an integrable function, we can apply the dominated convergence theorem to interchange the integral and the limit as $\epsilon \to 0$ and conclude that $\partial\mathbf{E}(V)/\partial h$ indeed exists and is given by

$$\begin{aligned}
\frac{\partial \mathbf{E}(V)}{\partial h} &= \lim_{\epsilon\to 0} \frac{\mathbf{E}(V(\tau_1(h+\epsilon),\ldots,\tau_N)) - \mathbf{E}(V(\tau_1(h-\epsilon),\ldots,\tau_N))}{2\epsilon} \\
&= \int_0^\infty \cdots \int_0^\infty \lim_{\epsilon\to 0} \bigg[\frac{(f(t_1,\ldots,t_N;h+\epsilon) - f(t_1,\ldots,t_N;h-\epsilon))}{2\epsilon} \\
&\quad \times V(t_1,\ldots,t_N)\bigg]\, dt_1\cdots dt_N \\
&= \int_0^\infty \cdots \int_0^\infty V(t_1,\ldots,t_N) \frac{\partial f(t_1,\ldots,t_N)}{\partial h}\, dt_1\cdots dt_N.
\end{aligned}$$

We can obtain an unbiased estimator in the likelihood ratio method by observing that

$$\begin{aligned}\frac{\partial \mathbf{E}(V)}{\partial h} &= \int_0^\infty \cdots \int_0^\infty V(t_1,\ldots,t_N)\frac{\partial f(t_1,\ldots,t_N)}{\partial h}\,dt_1\cdots dt_N \\ &= \int_0^\infty \cdots \int_0^\infty V(t_1,\ldots,t_N)\frac{\partial \ln f(t_1,\ldots,t_N)}{\partial h} f(t_1,\ldots,t_N)\,dt_1\cdots dt_N \\ &= \mathbf{E}\biggl(V(\tau_1,\ldots,\tau_N)\frac{\partial \ln f(\tau_1,\ldots,\tau_N)}{\partial h}\biggr).\end{aligned}$$

A.2 Proof of Proposition 4.1

(a) We have shown in Appendix A.1 that $\partial \mathbf{E}(V)/\partial h$ exists. Similarly, by conditions (A1) and (A5), we can apply the dominated convergence theorem to interchange expectation and the limit as $\epsilon \to 0$ to conclude that

$$\frac{\partial \mathbf{E}(V|\tau_2,\ldots,\tau_N)}{\partial h} = \int_0^\infty V(t_1,\tau_2,\ldots,\tau_N)\frac{\partial f_1(t_1|\tau_2,\ldots,\tau_N)}{\partial h}\,dt_1.$$

Therefore, $\partial \mathbf{E}(V)/\partial h$ and $\partial \mathbf{E}(V|\tau_2,\ldots,\tau_N)/\partial h$ exist.

(b) We have

$$\begin{aligned}\frac{\partial \mathbf{E}(V)}{\partial h} &= \int_0^\infty \cdots \int_0^\infty V(t_1,\ldots,t_N)\frac{\partial f(t_1,\ldots,t_N)}{\partial h}\,dt_1\cdots dt_N \\ &= \int_0^\infty \cdots \int_0^\infty V(t_1,\ldots,t_N)\frac{\partial (f_1(t_1|t_2,\ldots,t_N) f(t_2,\ldots,t_N))}{\partial h}\,dt_1\cdots dt_N \\ &= \int_0^\infty \cdots \int_0^\infty V(t_1,\ldots,t_N)\frac{\partial f_1(t_1|t_2,\ldots,t_N)}{\partial h} f(t_2,\ldots,t_N)\,dt_1\cdots dt_N \\ &= \int_0^\infty \cdots \int_0^\infty \frac{\partial \mathbf{E}(V|t_2,\ldots,t_N)}{\partial h} f(t_2,\ldots,t_N)\,dt_2\cdots dt_N \\ &= \mathbf{E}\biggl(\frac{\partial \mathbf{E}(V|\tau_2,\ldots,\tau_N)}{\partial h}\biggr),\end{aligned}$$

where the third equality holds because $f(t_2,\ldots,t_N)$ is independent of h.

A.3 Proof of Lemma 4.2

(a) This part is a direct application of parts (b) and (c).

(b) With the assumption of the almost sure differentiability of τ_1 and the mean-value theorem, there exist $0 \le \epsilon_1, \epsilon_2 \le \epsilon$ such that

$$\tau_1^{(+\epsilon)} = \tau_1(h+\epsilon) = \tau_1(h) + \epsilon\tau_1'(h+\epsilon_1), \qquad \tau_1^{(-\epsilon)} = \tau_1(h-\epsilon) = \tau_1 - \epsilon\tau_1'(h-\epsilon_2).$$

Thus,

$$\bigl|\tau_1^{(+\epsilon)} - \tau_1^{(-\epsilon)}\bigr| = \epsilon\bigl|\tau_1'(h+\epsilon_1) + \tau_1'(h-\epsilon_2)\bigr|.$$

By condition (A8), there is a K_τ such that $|\tau_1'(h)| \le K_\tau$, and then

$$\bigl|\tau_1^{(+\epsilon)} - \tau_1^{(-\epsilon)}\bigr| = \epsilon\bigl|\tau_1'(h+\epsilon_1) + \tau_1'(h-\epsilon_2)\bigr| \le 2\epsilon K_\tau.$$

Suppose that to improve look ℓ is the length of the smallest of the intervals I_k, i.e., $\ell = \min_{k=2,\ldots,N+1}(\tau_{(k)} - \tau_{(k-1)})$. Then,

$$\begin{aligned}
&\mathbf{P}\big(\tau_1^{(+\epsilon)} \in I_i, \tau_1^{(-\epsilon)} \in I_j, j > i+1 \big| \tau_2, \ldots, \tau_N\big) \\
&\quad = \mathbf{P}\big(\tau_1^{(+\epsilon)} \in I_i, \tau_1^{(-\epsilon)} \in I_j, j > i+1, \big|\tau_1^{(+\epsilon)} - \tau_1^{(-\epsilon)}\big| \geq \ell \big| \tau_2, \ldots, \tau_N\big) \\
&\quad \leq \mathbf{P}\big(\big|\tau_1^{(+\epsilon)} - \tau_1^{(-\epsilon)}\big| \geq \ell \big| \tau_2, \ldots, \tau_N\big) \\
&\quad \leq \mathbf{P}(2\epsilon K_\tau \geq \ell | \tau_2, \ldots, \tau_N) = \mathbf{P}\left(K_\tau \geq \frac{\ell}{2\epsilon}\middle| \tau_2, \ldots, \tau_N\right).
\end{aligned}$$

Because $\mathbf{E}(K_\tau) < \infty$, then as $\epsilon \downarrow 0$, we have, almost surely,

$$\frac{\ell}{2\epsilon}\mathbf{P}\left(K_\tau \geq \frac{\ell}{2\epsilon}\middle| \tau_2, \ldots, \tau_N\right) \to 0,$$

or

$$\frac{1}{\epsilon}\mathbf{P}\left(K_\tau \geq \frac{\ell}{2\epsilon}\middle| \tau_2, \ldots, \tau_N\right) \to 0.$$

Along with $\mathbf{P}(\tau_1^{(+\epsilon)} \in I_i, \tau_1^{(-\epsilon)} \in I_j, j > i+1 | \tau_2, \ldots, \tau_N) \geq 0$, we conclude $\mathbf{P}(\tau_1^{(+\epsilon)} \in I_i, \tau_1^{(-\epsilon)} \in I_j, j > i+1 | \tau_2, \ldots, \tau_N)/\epsilon \to 0$, when $\epsilon \downarrow 0$.

(c) With the assumption of the continuity of τ_1, $\tau_1^{(\pm\epsilon)} \to \tau_1$. Since $\tau_1^{(+\epsilon)} \in I_i$ and $\tau_1^{(-\epsilon)} \in I_{i+1}$ for all $\epsilon > 0$, then $\tau_1 = \tau_{(i)}$ with probability 1. Suppose $h_{(i)}$ is the value that gives $\tau(h_{(i)}) = \tau_{(i)}$, then

$$\begin{aligned}
&\lim_{\epsilon\downarrow 0} \frac{1}{2\epsilon}\mathbf{P}\big(\tau_1^{(+\epsilon)} \in I_i, \tau_1^{(-\epsilon)} \in I_{i+1} \big| \tau_2, \ldots, \tau_N\big) \\
&\quad = \lim_{\epsilon\downarrow 0} \frac{1}{2\epsilon}\mathbf{P}\big(\tau_1^{(+\epsilon)} \in I_i, \tau_1^{(-\epsilon)} \in I_{i+1}, \tau_1^{(+\epsilon)} \leq \tau_{(i)} \leq \tau_1^{(-\epsilon)} \big| \tau_2, \ldots, \tau_N\big) \\
&\quad = -\tau_1'(h_{(i)}) f_1(\tau_{(i)} | \tau_2, \ldots, \tau_N).
\end{aligned}$$

A.4 Verifying the conditions

In this section, we verify that the assumptions made for our main results hold for basket default swaps and CDOs in the Gaussian copula model.

(A1): For any given $t_1, \ldots, t_N$ and $t_{N+1} = T$ let $t_{(1)} \leq \cdots \leq t_{(N+1)}$ denote their sorted values. Then in the case of basket default swaps, the discounted payoff function V can be expressed as a function of $t_{(n)}$:

$$\begin{aligned}
&V(t_1, \ldots, t_N) \\
&\quad = \begin{cases} \sum_{i=j+1}^m s_i D(T_i) + (1-R)D(t_{(n)}) - s_{j+1}D(t_{(n)})\frac{t_{(n)}-T_j}{T_{j+1}-T_j}, & T_j < t_{(n)} \leq T_{j+1}; \\ 0, & t_{(n)} > T. \end{cases}
\end{aligned}$$

This V has no explicit dependence on h.

If $t_{(n)} \geq T$, $V = 0$; otherwise, V is continuous with respect to $t_{(n)} \in (0, T)$. Also note that

$$\lim_{t_{(n)} \to 0} |V| = (1 - R) + \sum_{i=1}^{m} s_i D(T_i) < \infty,$$
$$\lim_{t_{(n)} \to T} |V| = (1 - R) D(T) < \infty.$$

Therefore, $V(t_1, \ldots, t_N)$ is bounded.

In the case of CDOs, the discounted payoff function V can be expressed as a linear combination of $M(t; t_1, \ldots, t_N)$, where

$$M(t; t_1, \ldots, t_N) = \left(\sum_{i=1}^{N} l_i \mathbf{1}(t_i \leq t) - S_\ell \right)^+ - \left(\sum_{i=1}^{N} l_i \mathbf{1}(t_i \leq t) - S_u \right)^+,$$

for $t = T_1, \ldots, T_m$. Since $0 \leq M(t; t_1, \ldots, t_N) \leq S_u - S_\ell$, V is then bounded.

(A2): We first consider the case of basket default swaps. Since $t_i < t_1$ and $t_1 + \Delta t < t_j$, $t_1 + \Delta t$ is in the same position as t_1 when these times are sorted in the ascending order $t_{(1)} \leq \cdots \leq t_{(N+1)}$. If $t_1 \neq t_{(n)}$, the change in t_1 does not affect the value of V, i.e.,

$$V(t_1, t_2, \ldots, t_N) = V(t_1 + \Delta t, t_2, \ldots, t_N),$$

and V is obviously Lipschitz with respect to t_1. If $t_1 = t_{(n)}$, suppose $T_p \leq t_1 \leq T_{p+1}$, $T_k \leq t_1 + \Delta t \leq T_{k+1}$; then,

$$V(t_1, t_2, \ldots, t_N) = \sum_{i=p+1}^{m} s_i D(T_i) + (1 - R) D(t_1) - s_{p+1} D(t_1) \frac{t_1 - T_p}{T_{p+1} - T_p},$$

$$\begin{aligned} &V(t_1 + \Delta t, t_2, \ldots, t_N) \\ &\quad = \sum_{i=k+1}^{m} s_i D(T_i) + (1 - R) D(t_1 + \Delta t) - s_{k+1} D(t_1 + \Delta t_1) \frac{t_1 + \Delta t - T_k}{T_{k+1} - T_k}. \end{aligned}$$

Because $D(t)$ is assumed Lipschitz, there exists a constant C_1 such that $|D(t + \Delta t) - D(t)| \leq C_1 |\Delta t|$. If $p = k$,

$$\begin{aligned} &\big| V(t_1 + \Delta t, t_2, \ldots, t_N) - V(t_1, t_2, \ldots, t_N) \big| \\ &\quad = \bigg| -\frac{s_{p+1}}{T_{p+1} - T_p} \big((t_1 + \Delta t - T_p) D(t_1 + \Delta t) - (t_1 - T_p) D(t_1) \big) \\ &\qquad + (1 - R) \big(D(t_1 + \Delta t) - D(t_1) \big) \bigg| \\ &\quad \leq \frac{s_{p+1}}{T_{p+1} - T_p} \big| (t_1 + \Delta t - T_p) D(t_1 + \Delta t) - (t_1 - T_p) D(t_1) \big| \\ &\qquad + (1 - R) \big| D(t_1 + \Delta t) - D(t_1) \big| \end{aligned}$$

$$\le \left(\frac{s_{p+1}}{T_{p+1}-T_p}(t_1-T_p)+(1-R)\right)\left|D(t_1+\Delta t)-D(t_1)\right|$$
$$+\frac{s_{p+1}}{T_{p+1}-T_p}|\Delta t|D(t_1+\Delta t)$$
$$\le \left((s_{p+1}+1-R)C_1+\frac{s_{p+1}D(T_p)}{T_{p+1}-T_p}\right)|\Delta t|.$$

If $|k-p|=1$, without loss of generality, suppose $k=p+1$, i.e., $\Delta t>0$ and $T_p\le t_1\le T_{p+1}(=T_k)\le t_1+\Delta t\le T_{k+1}$. Then

$$\left|V(t_1+\Delta t,t_2,\dots,t_N)-V(t_1,t_2,\dots,t_N)\right|$$
$$=\left|-\frac{t_1+\Delta t-T_k}{T_{k+1}-T_k}s_{k+1}D(t_1+\Delta t)+\frac{t_1-T_p}{T_k-T_p}s_kD(t_1)-s_kD(T_k)\right.$$
$$\left.+(1-R)\big(D(t_1+\Delta t)-D(t_1)\big)\right|$$
$$\le \frac{s_{k+1}}{T_{k+1}-T_k}|t_1+\Delta t-T_k|D(t_1+\Delta t)+(1-R)\left|D(t_1+\Delta t)-D(t_1)\right|$$
$$+\frac{s_k}{T_k-T_p}\left|\big(D(t_1)-D(T_k)\big)(t_1-T_p)-D(T_k)(T_k-t_1)\right|$$
$$\le \frac{s_{k+1}D(T_k)}{T_{k+1}-T_k}|\Delta t|+s_k\left|D(t_1)-D(T_k)\right|+D(T_k)|\Delta t|$$
$$+(1-R)\left|D(t_1+\Delta t)-D(t_1)\right|$$
$$\le \left(\left(\frac{s_{k+1}}{T_{k+1}-T_k}+1\right)D(T_k)+(s_k+1-R)C_1\right)|\Delta t|.$$

If $|k-p|>1$, $|\Delta t|\ge|T_k-T_p|$. Then, recalling that V is bounded,

$$\left|V(t_1+\Delta t,t_2,\dots,t_N)-V(t_1,t_2,\dots,t_N)\right|\le|\max V-\min V|$$
$$\le\frac{|\max V-\min V|}{|T_k-T_p|}|\Delta t|.$$

Thus, V is Lipschitz in the interval (t_i,t_j).

For CDOs, it is sufficient to prove that $M(t;t_1,\dots,t_N)$ is Lipschitz with respect to t_1. Since a small change Δt in t_1 does not change the order of $t_1,\dots,t_{N+1}$, it does not affect the value of $M(t;t_1,\dots,t_N)$, i.e.,

$$M(t;t_1+\Delta t,\dots,t_N)-M(t;t_1,\dots,t_N)=0.$$

The Lipschitz continuity follows directly.

(A3): As in the cases of basket default swaps and CDOs, this is a condition on how the intensity λ is parameterized by h.

(A4): In the Gaussian copula model, $\tau_1,\dots,\tau_N$ have joint density

$$f(t_1,\dots,t_N)=\frac{\partial\mathcal{C}(u_1,\dots,u_N;\mathbf{\Sigma})}{\partial u_1\dots\partial u_N}\frac{\partial u_1}{\partial t_1}\cdots\frac{\partial u_N}{\partial t_N}=c(u_1,\dots,u_N;\mathbf{\Sigma})\prod_{i=1}^N f_i(t_i).$$

Suppose $\bar{\mathbf{\Sigma}}$ is the correlation matrix of $\tau_2,\dots,\tau_N$ in the Gaussian copula model; the joint density function of $\tau_2,\dots,\tau_N$ is then

$$f(t_2,\dots,t_N)=\frac{\partial\mathcal{C}(u_2,\dots,u_N;\bar{\mathbf{\Sigma}})}{\partial u_2\dots\partial u_N}\frac{\partial u_2}{\partial t_2}\cdots\frac{\partial u_N}{\partial t_N}=c(u_2,\dots,u_N;\bar{\mathbf{\Sigma}})\prod_{i=2}^N f_i(t_i),$$

with

$$c(u_2,\dots,u_N;\bar{\mathbf{\Sigma}})=\frac{1}{|\bar{\mathbf{\Sigma}}|^{1/2}}\exp\left[-\frac{1}{2}\big(\Phi^{-1}(\bar{\mathbf{u}})\big)^\top\big(\bar{\mathbf{\Sigma}}^{-1}-\mathbf{I}\big)\Phi^{-1}(\bar{\mathbf{u}})\right],$$

where $\bar{\mathbf{u}}$ is an $(N-1)\times 1$ vector with $\bar{\mathbf{u}}_{i-1}=u_i=1-e^{-\int_0^{t_i}\lambda_i(s)\,ds}$, for $i=2,\dots,N$. Then we have

$$\begin{aligned}f_1(t_1|\tau_2,\dots,\tau_N)&=\frac{f(t_1,\tau_2,\dots,\tau_N)}{f(\tau_2,\dots,\tau_N)}\\&=\frac{|\bar{\mathbf{\Sigma}}|^{1/2}}{|\mathbf{\Sigma}|^{1/2}}\exp\Bigg\{-\frac{(\mathbf{\Sigma}^{-1})_{11}-1}{2}\big(\Phi^{-1}(u_1)\big)^2\\&\quad-\sum_{i=2}^N\big(\mathbf{\Sigma}^{-1}\big)_{1j}\Phi^{-1}(u_j)\Phi^{-1}(u_1)\Bigg\}\lambda_1 e^{-\int_0^{t_1}\lambda_1(s,h)\,ds},\end{aligned}$$

where $u_1=1-e^{-\int_0^{t_1}\lambda_1(s,h)\,ds}$ and $u_i=1-e^{-\int_0^{\tau_i}\lambda_i(s)\,ds}$, for $i>1$.

(A5): With the assumption of the existence of $\partial\lambda_1/\partial h$, in the Gaussian copula model,

$$\begin{aligned}&\frac{\partial f(t_1,t_2,\dots,t_N)}{\partial h}\\&=\frac{1}{|\mathbf{\Sigma}|^{1/2}}\exp\left[-\frac{1}{2}\big(\Phi^{-1}(\mathbf{u})\big)^\top\big(\mathbf{\Sigma}^{-1}-\mathbf{I}\big)\Phi^{-1}(\mathbf{u})\right]\\&\quad\times\left(-\sqrt{2\pi}\sum_{j=1}^N\big(\mathbf{\Sigma}^{-1}-\mathbf{I}\big)_{1j}\Phi^{-1}(u_j)e^{\Phi^{-1}(u_1)^2/2}(1-u_1)t_1+\frac{1}{\lambda_1}-t_1\right)\\&\quad\times\frac{\partial\lambda_1}{\partial h}\prod_{i=1}^N f_i(t_i).\end{aligned}$$

In Lemma A.1 in Appendix A.5, we show that for $t_i>0$, $i=1,\dots,N$, and $h\in\mathcal{H}$,

- $\exp[-\frac{1}{2}(\Phi^{-1}(\mathbf{u}))^\top(\mathbf{\Sigma}^{-1}-\mathbf{I})\Phi^{-1}(\mathbf{u})]$ is bounded.
- $\Phi^{-1}(u_1)e^{\Phi^{-1}(u_1)^2/2}(1-u_1)$ and $e^{\Phi^{-1}(u_1)^2/2}(1-u_1)$ are bounded.

Along with the assumption that $\partial\lambda_1/\partial h$ is bounded, there exist M_1 and M_2 such that almost everywhere in $\mathcal{H}$,

$$\left|\frac{\partial f(t_1, t_2, \ldots, t_N)}{\partial h}\right| \le \left(M_1 t_1 + \frac{M_2}{\lambda_1}\right)\prod_{i=1}^{N} f_i(t_i).$$

Define the positive function

$$g(t_1, \ldots, t_N) = \left(M_1 t_1 + \frac{M_2}{\lambda_1}\right)\prod_{i=1}^{N} f_i(t_i).$$

Then

$$\begin{aligned}
&\int_0^\infty \cdots \int_0^\infty g(t_1, t_2, \ldots, t_N)\, dt_1 \cdots dt_N \\
&\quad = M_1 \int_0^\infty \cdots \int_0^\infty t_1 \prod_{i=1}^{N} f_i(t_i)\, dt_1 \cdots dt_N + \int_0^\infty \cdots \int_0^\infty \frac{M_2}{\lambda_1}\prod_{i=1}^{N} f_i(t_i)\, dt_1 \cdots dt_N \\
&\quad \le M_1 \mathbf{E}(\tau_1) + \frac{M_2}{\inf_{t,h}\lambda_1(t,h)} < \infty.
\end{aligned}$$

(A6): Because

$$f_1(t_1|\tau_2, \ldots, \tau_N) = \frac{f(t_1, \tau_2, \ldots, \tau_N)}{f(\tau_2, \ldots, \tau_N)},$$

$\partial f(t_1, \tau_2, \ldots, \tau_N)/\partial h$ exists, and $f(\tau_2, \ldots, \tau_N)$ does not depend on h,

$$\frac{\partial f_1(t_1|\tau_2, \ldots, \tau_N)}{\partial h} = \frac{\partial f(t_1, \tau_2, \ldots, \tau_N)}{\partial h}\frac{1}{f(\tau_2, \ldots, \tau_N)}.$$

Define

$$g_1(t_1|\tau_1, \ldots, \tau_N) = \frac{g(t_1, \tau_2, \ldots, \tau_N)}{f(\tau_2, \ldots, \tau_N)}.$$

Then

$$\left|\frac{\partial f_1(t_1|\tau_2, \ldots, \tau_N)}{\partial h}\right| \le g_1(t_1|\tau_1, \ldots, \tau_N),$$

and

$$\begin{aligned}
&\int_0^\infty g_1(t_1|\tau_1, \ldots, \tau_N)\, dt_1 \\
&\quad = \frac{\prod_{i=2}^{N} f_i(\tau_i)}{f(\tau_2, \ldots, \tau_N)} \int_0^\infty \left(M_1 t_1 + \frac{M_2}{\lambda_1}\right) f_1(t_1)\, dt_1 \\
&\quad \le \frac{\prod_{i=2}^{N} f_i(\tau_i)}{f(\tau_2, \ldots, \tau_N)}\left(M_1 \mathbf{E}(\tau_1) + \frac{M_2}{\inf_{t,h}\lambda_1(t,h)}\right) < \infty, \quad \text{a.s.}
\end{aligned}$$

(A7): To realize the required construction let U be uniformly distributed in $(0, 1)$. Take the inverse of $F_1(t)$ and set $\tau_1(h) = F_1^{-1}(U; h)$; then $\tau_1(h)$ has density $f_1(\tau_1; h)$. Since λ_1 is a strictly increasing differentiable function of h, $\lambda_1(t, h) > 0$, and $F_1(\tau_1; h) = 1 - e^{-\int_0^{\tau_1} \lambda_1(s,h)\,ds}$, it is easy to see that $\tau_1(h) = F^{-1}(U; h)$ is a.s. a strictly decreasing differentiable function of h.

(A8): We have shown that $\tau_1(h)$ is a strictly decreasing function. To verify the Lipschitz property of $\tau_1(h)$, it is sufficient to show for a sufficiently small $\epsilon > 0$ that there exists K_τ such that $\mathbf{E}(K_\tau) < \infty$ and, a.s.,

$$\tau_1(U; h) - \tau_1(U; h + \epsilon) \le \epsilon K_\tau.$$

For any given U, let $\tau_1^{(0)} = \tau_1(U; h)$ and $\tau_1^{(+)} = \tau_1(U; h + \epsilon)$. Since $U = F_1(\tau_1^{(0)}; h) = F_1(\tau_1^{(+)}; h + \epsilon)$,

$$\int_0^{\tau_1^{(0)}} \lambda_1(s, h)\,ds = \int_0^{\tau_1^{(+)}} \lambda_1(s, h + \epsilon)\,ds,$$

so

$$\int_{\tau_1^{(+)}}^{\tau_1^{(0)}} \lambda_1(s, h)\,ds = \int_0^{\tau_1^{(+)}} \big(\lambda_1(s, h + \epsilon) - \lambda_1(s, h)\big)\,ds.$$

Let $\nu = \inf_{t,h} \lambda_1(t, h)$, and $\tau_1^* = \tau_1(U; h^*)$, where h^* is the lower bound of $\mathcal{H}$ as defined before. Then

$$\begin{aligned}
\tau_1^{(0)} - \tau_1^{(+)} &\le \frac{1}{\nu} \int_0^{\tau_1^{(+)}} \big(\lambda_1(s, h + \epsilon) - \lambda_1(s, h)\big)\,ds \\
&\le \frac{1}{\nu} \int_0^{\tau_1^{*}} \big(\lambda_1(s, h + \epsilon) - \lambda_1(s, h)\big)\,ds.
\end{aligned}$$

Because $\partial\lambda_1(t, h)/\partial h$ exists and is bounded for all h and t, there exists a constant M such that

$$\lambda_1(t, h + \epsilon) - \lambda_1(t, h) \le M\epsilon.$$

Therefore,

$$\tau_1(U; h) - \tau_1(U; h + \epsilon) = \tau_1^{(0)} - \tau_1^{(+)} \le M\epsilon\tau_1^*/\nu,$$

and

$$\mathbf{E}(M\epsilon\tau_1^*/\nu) < \infty.$$

A.5 Lemma A.1 and its proof

Lemma A.1 *Suppose* $\mathbf{\Sigma}$ *is an* $N \times N$ *matrix and* $\mathbf{u}$ *is an* $N \times 1$ *vector. If* $\mathbf{\Sigma}$ *is positive semidefinite with diagonal elements equal to* 1 *and* $0 < u_1 < 1$,

1. $\exp[-\frac{1}{2}(\Phi^{-1}(\mathbf{u}))^\top(\mathbf{\Sigma}^{-1}-\mathbf{I})\Phi^{-1}(\mathbf{u})]$ *is bounded.*
2. $\Phi^{-1}(u_1)e^{\Phi^{-1}(u_1)^2/2}(1-u_1)$ *and* $e^{\Phi^{-1}(u_1)^2/2}(1-u_1)$ *are bounded.*

Proof 1. It is enough to show that $(\mathbf{\Sigma}^{-1}-\mathbf{I})$ is positive semidefinite. Since $\mathbf{\Sigma}$ is positive semidefinite with diagonal elements equal to 1, its eigenvalues are all nonnegative and less than or equal to 1. This means the eigenvalues of $(\mathbf{\Sigma}^{-1}-\mathbf{I})$ are then nonnegative. Therefore $(\mathbf{\Sigma}^{-1}-\mathbf{I})$ is positive semidefinite, and

$$\exp\left[-\frac{1}{2}\big(\Phi^{-1}(\mathbf{u})\big)^\top\big(\mathbf{\Sigma}^{-1}-\mathbf{I}\big)\Phi^{-1}(\mathbf{u})\right]\le 1.$$

2. Let $y=\Phi^{-1}(u_1)$. We have

$$\begin{aligned}\Phi^{-1}(u_1)(1-u_1)e^{\Phi^{-1}(u_1)^2} &= \big(1-\Phi(y)\big)ye^{\frac{y^2}{2}}(1-u_1)e^{\Phi^{-1}(u_1)^2}\\ &= \big(1-\Phi(y)\big)e^{\frac{y^2}{2}}.\end{aligned}$$

When $y\to\pm\infty$, i.e., $u\to 0$ or 1,

$$\big(1-\Phi(y)\big)ye^{\frac{y^2}{2}}\sim\frac{\phi(y)}{y}ye^{y^2/2}=\frac{1}{\sqrt{2\pi}},\quad \text{where } \phi(y)=\frac{1}{\sqrt{2\pi}}e^{-y^2/2},$$

hence

$$\lim_{y\to\pm\infty}\big(1-\Phi(y)\big)e^{\frac{y^2}{2}}=0. \qquad \square$$

References

1. Andersen, L., Sidenius, J., Basu, S.: All your hedges in one basket. Risk **16**, 67–72 (2003)
2. Asmussen, S., Glynn, P.W.: Stochastic Simulation. Springer, New York (2007)
3. Bruyère, R., Cont, R., Copinot, R., Fery, L., Jaeck, J., Spitz, T., Smart, G.: Credit Derivatives and Structured Credit: A Guide for Investors. Wiley, Chichester (2006)
4. Broadie, M., Glasserman, P.: Estimating security price derivatives using simulation. Manag. Sci. **42**, 269–285 (1996)
5. Chen, Z., Glasserman, P.: Fast pricing of basket default swaps. Oper. Res. **56**, 286–303 (2008)
6. Cherubini, U., Luciano, E., Vecchiato, W.: Copula Methods in Finance. Wiley, New York (2004)
7. Duffie, D., Singleton, K.: Credit Risk: Pricing, Measurement, and Management. Princeton Univ. Press, Princeton (2003)
8. Fu, M.C., Hu, J.Q.: Conditional Monte Carlo: Gradient Estimation and Optimization Applications. Kluwer Academic, Boston (1997)
9. Glasserman, P.: Gradient Estimation via Perturbation Analysis. Kluwer Academic, Norwell (1991)
10. Glasserman, P.: Monte Carlo Methods in Financial Engineering. Springer, New York (2004)
11. Gong, W.B., Ho, Y.C.: Smoothed (conditional) perturbation analysis of discrete event dynamical systems. IEEE Trans. Automat. Contr. AC-32 **10**, 858–866 (1987)
12. Gupton, G., Finger, C., Bhatia, M.: CreditMetrics Technical Document. Morgan & Co., New York (1997)
13. Hull, J., White, A.: Valuation of a CDO and an *n*th-to-default CDS without Monte Carlo simulation. J. Deriv. **12**(2), 8–23 (2004)
14. Joshi, M., Kainth, D.: Rapid and accurate development of prices and Greeks for *n*th-to-default credit swaps in the Li model. Quant. Finance **4**, 266–275 (2004)

15. Laurent, J.-P., Gregory, J.: Basket default swaps, CDOs, and factor copulas. J. Risk **7**(4), 103–122 (2005)
16. Li, D.: On default correlation: A copula function approach. J. Fixed Income **9**, 43–54 (2000)
17. Schönbucher, P.: Credit Derivatives Pricing Models: Models, Pricing and Implementation. Wiley, New York (2003)
18. Suri, R., Zazanis, M.: Perturbation analysis gives strongly consistent sensitivity estimates for the M/G/1 queue. Manag. Sci. **34**, 39–64 (1988)