A Mixed PDE-Monte Carlo Approach for Pricing Credit Default Index Swaptions

Vlad Bally[*], Lucia Caramellino[†] and Antonino Zanette[‡]

Abstract

The problem of numerically pricing credit default index swaptions on a large number of names is considered. We place ourselves in a stochastic intensity framework, where Ornstein-Uhlenbeck type correlated processes are used to model both the firms distance to default and a macroeconomic state variable. The default of the firms follows here the reduced-form approach and the (random) intensity of the default depends on the behavior of the diffusion processes. We propose here a numerical method based on both a Monte Carlo and a deterministic approach for solving PDEs by finite difference. Numerical tests show the efficiency and the robustness of the proposed procedure.

2000 MSC: 91B28, 60H30, 65C05.

Keywords: credit default swaps and credit default swaptions; Feynman-Kac formula; Monte Carlo methods; numerical approximation of PDEs solutions.

Acknowledgement. The authors are grateful to Darrell Duffie for having suggested the problem and for his encouraging help. The authors also thank the editor Marco Li Calzi and two anonymous referees for their useful comments and suggestions.

Corresponding author: Lucia Caramellino.

Address: Dipartimento di Matematica,
Università di Roma *Tor Vergata*,
via della Ricerca Scientifica 1,
I-00133 Roma, Italy.

Email: `caramell@mat.uniroma2.it`
Phone number: + 39 06 72594709
Fax number: + 39 06 72594699

[*]Laboratoire d'Analyse et de Mathématiques Appliquées, Université de Marne-la-Vallèe, 77454 Champs-sur-Marne, France; e-mail: *bally@univ-mlv.fr*

[†]Dipartimento di Matematica, Università di Roma *Tor Vergata*, via della Ricerca Scientifica 1, I-00133 Roma, Italy; e-mail: *caramell@mat.uniroma2.it*

[‡]Dipartimento di Finanza dell'Impresa e dei Mercati Finanziari, Università di Udine, Via Tomadini 30/A, I-33100 Udine, Italy; e-mail *antonino.zanette@uniud.it*

1 Introduction

The credit risk market has been growing rapidly in the last years. In particular, credit default swaps (CDSs) are one of the most actively traded credit derivatives (see e.g. Duffie [2], Duffie, Saita and Wang [4], Schönbucher [9]).
A CDS provides protection in the event of default (called a credit event) of a specific company (called the reference entity). It is then an agreement between two counterparties, allowing the first one to be "long" a third-party credit risk, while the second counterparty to be "short" the credit risk. More precisely, one has first to introduce a reference asset, which is typically a credit risky bond issued by a third party corporation. Now, the two counterparties, say A and B, enter into an agreement: A pays to B a fixed periodic coupon for the specified life of the reference asset. Party B makes no payments unless the default event occurs for the reference asset. If such a credit event occurs, B makes a payment to party A, and the swap then terminates. The size of the payment is usually linked to the decline in the reference asset's market value following the credit event.
In this paper, we consider the problem of pricing options on a portfolio of a large number of names (typically 100), where the underlying instrument for each name is a CDS. Such kind of risky financial instruments are known in the literature as *index default swaptions* or *credit default index swaptions*, or also *CDS index swaptions*, the last one being the denomination we will use all over this paper (see e.g. Jackson [6]).
A CDS index swaption is an option to buy or sell the underlying CDSs at a specified date. A payer swaption gives the holder of the option the right to buy protection (pay premium) and a receiver swaption gives the holder of the option the right to sell protection (receive premium). This basic definition of a CDS index swaption is similar to that of a CDS option, an option on a single-entity CDS, but a CDS index swaption is significantly different from a CDS option. In the case of an option on a single-entity, if the reference entity defaults before the options expiry, the option will be knocked out and becomes worthless. For a CDS index swaption, when a reference entity defaults before the options expiry, the loss will be paid by the protection seller to the protection buyer when the option is exercised. Even if there is only one entity in the portfolio, a CDS index swaption is still different from a single-entity CDS option: if the entity defaults before the options expiry, the options seller will pay to the protection buyer the lost amount at expiry. Clearly, a CDS index swaption is always more valuable than a single-entity CDS option.
In this paper, we allow the underlying possible defaults to follow an intensity-based model, in which the default time is a stopping time with a given intensity process (as in Duffie and Singleton [5] or also Schönbuncher [9]; different approaches are available in literature, see e.g. Jackson [6]). In our application, the default intensities depend on the firms distance to default and on some observable variable which is linked with the likelihood of default, such as a macroeconomic variable related to the business cycle. In our framework, both the firms distance to default and the macroeconomic variable are modelled through correlated Ornstein Uhlenbeck processes, with suitable coefficients. The model, which will be mathematically described later (Section 2), wad introduced firstly in Duffie and Wang [3] and in a second time it was generalized by Duffie, Saita and Wang [4].
Our main contribution is the construction of an algorithm for the pricing of CDS index swap-

tions on a large number of names. It is based on a mixing between the classical numerical approximation by partial differential equations (PDEs) and the Monte Carlo approaches. For the model considered here, this algorithm offers a very efficient alternative to a pure vanilla Monte Carlo method, which is the classical tool to handle such kind of derivatives but, at the same time, becomes really unfeasible from a computational point of view when the number of the underlying firms is large. In fact, the comparisons with the results turning out by a pure vanilla Monte Carlo method are here presented mainly in order to benchmark the results of our algorithm. Let us add that it would be interesting to set up other pure and improved Monte Carlo techniques, such as control variates, which might possibly lead to better computational efficiency. But, to our knowledge, no results are available in this direction at this stage, although this could be an interesting research problem.
The advantage of our algorithm is that it circumvents the nested simulation problem of a standard Monte Carlo algorithm by approximating conditional expectations using a PDEs approach. In fact, the PDEs approach is particulary suited to problems which are controlled by 2-dimensional diffusions, and can be efficiently handled by standard numerical techniques, such as finite differences. For the sake of clearness, it has to be remarked that the model we are going to use, and in particular the modelled correlations among the underlying stochastic processes, actually allows one to reduce to a number of different 2-dimensional problems equal to the number of the considered firms.
Finally, let us remark that even if the main goal of our work is to introduce a new methodology (the mixed PDE-Monte Carlo algorithm) constructed *ad hoc* for the pricing of a complex credit product (CDS index swaption), nevertheless we hope that this contribution helps to paving the road for future work addressing more challenging credit risk problems in high dimension in a dynamic setting.
The paper is organized as follows: in Section 2 we introduce the model of the stochastic intensity process for the default, a description of CDSs and CDS index swaptions; Section 3 is devoted to the presentation of the mixed PDE-Monte Carlo approach for pricing CDS index swaptions; numerical results are presented in Section 4.

2 The model

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_t, \mathbb{P})$ a filtered probability space, denoting the "risk neutral world", where the following (diffusion) processes are defined:

- $(Y_t)_{t\in[0,T]}$, modelled as
$$dY_t = \kappa_Y(\theta_Y - Y_t)dt + \sigma_Y dW_t^Y; \tag{1}$$
- as $i = 1, \ldots, n$, $(D_{i,t})_{t\in[0,T]}$, modelled as
$$dD_{i,t} = \kappa_D(\theta_{D,i} - D_{i,t})dt + \sigma_D(\rho dW_t^c + \sqrt{1-\rho^2}dW_t^i). \tag{2}$$

In (1) and (2), the processes $W^Y, W^c, W^1, \ldots, W^n$ denote independent Brownian motions; ρ stands for a correlation coefficient, thus belonging to $(-1, 1)$; $\kappa_Y, \theta_Y, \sigma_Y$ and $\kappa_D, \theta_{D,i}, \sigma_D$ are all constant. Therefore, Y and $D_1, \ldots, D_n$ are Ornstein-Uhlenbeck type (correlated) diffusions.

Notice that Y is independent of $D_1, \dots, D_n$, while $D_1, \dots, D_n$ depend each other and such a dependence is given by the correlation ρ.
The referring filtration is then $\mathcal{F}_t = \sigma((W_s^Y, W_s^c, W_s^1, \dots, W_s^n)\,;\, s \le t)$.
The intuitive meaning of the process Y arises in the fact that it models a macroeconomic state variable, i.e. the U.S. personal income growth, while $D_1, \dots, D_n$ model the firms distance to default (for details, see Duffie and Wang [3] or Duffie, Saita and Wang [4]). Let us recall that the distances to default are correlated: in fact, the common Brownian motion W^c, which appears in all the SDEs giving $D_1, \dots, D_n$ (see (2)), captures the correlation amongst the individual distances to default.
We suppose that the time of default follows here the intensity-based approach: setting τ_i as the default instant for the ith firm, then its (random) failure intensity λ^i depends on both Y and D_i as follows

$$\lambda_t^i = \Lambda(Y_t, D_{i,t}), \quad \text{where } \Lambda(y,d) = \exp(\mu_0 + \mu_1 y + \mu_2 d) \tag{3}$$

Notice that the constant parameters μ_0, μ_1 and μ_2 are common to all firms.
Let us recall that the above definitions mean that, for $s < t$,

$$\mathbb{P}(\tau_i > t \,|\, \mathcal{F}_s) = \mathbb{E}(e^{-\int_s^t \Lambda(Y_u, D_{i,u})du} \,|\, \mathcal{F}_s), \quad \text{on the set } \{\tau_i > s\},$$

where, from now on, the symbol $\mathbb{E}$ denotes the expectation under $\mathbb{P}$.
For the construction of τ_i, we refer to Duffie and Singleton [5], Schönbucher [9] or also Lando [7]. We only recall here that this defines the default time as the first jump instant of a Cox process with intensity process λ^i, that is $\tau_i = \inf\{t \ge 0 : \int_0^t \lambda_s^i ds = Z\}$, being Z an exponential random variable, of parameter 1, independent of all the Brownian motions already defined.

Remark 2.1 *Let us make some remarks about our choice for the parameters modelling. Indeed, concerning the distance to default processes, we have assumed a common mean reversion κ_D and a common volatility σ_D (see (2)); moreover, the intensity failures are defined through a function with common parameters, i.e. μ_0, μ_1 and μ_2 (see (3)). Such kind of choice might appear very restrictive, but this parsimonious model tries to overcome the problem of an extremely high-dimensional state-vector, consisting in one macroeconomic covariate, personal income growth Y, and the distance to default D_i for each firm i among the n total ones. On the other hand, a relevant parameter concerning the distance to default, that is the long-run mean parameter $\theta_{D,i}$ (see (2)), varies firm by firm. Finally, the assumed homogeneity of correlation across different firms makes the model numerically tractable.*

Let us now mathematically describe a CDS index swaption. This is an option to buy protection on the CDS index at rate K, which pays at time t

$$\max\Big(\sum_{i=1}^n V_{i,t}, 0\Big)$$

where $V_{i,t}$ is the market value at time t of a default swap on name i at rate K (notice that here K does not depend on i). Let us recall that a CDS is a credit derivative that protects its owner from the default event of the issuer of the underlying CDS.

So, if r denotes the risk free interest rate, the market value at time 0 of such an option is given by

$$P_0 = e^{-rt}\mathbb{E}\Big(\max\Big(\sum_{i=1}^{n} V_{i,t}, 0\Big)\Big). \tag{4}$$

Let us now mathematically describe the quantities $V_{1,t}$,.., $V_{n,t}$. Their values depend on the fact that the default has or has not been observed at time t as follows:

- if $\tau_i \leq t$, that is name i has defaulted by time t, then $V_{i,t}$ is the loss L_i given default of name i:

$$V_{i,t} = L_i;$$

- if $\tau_i > t$, that is no default happened up to t, then

$$V_{i,t} = B_{i,t} - A_{i,t},$$

where $B_{i,t}$ is the value at time t of the future payments by the seller of protection at default and $A_{i,t}$ is the value at time t of the future payments of the buyer of protection.

To give an explicit expression for $B_{i,t}$ and $A_{i,t}$, one has to introduce further notations. Set $t = t_0 < t_1 < \cdots < t_N = T$ the premium payment dates, which are such that $t_{j+1} - t_j = \Delta t$ for any j (usually, Δt is equal to 3 months). Define

$$p_{i,j,t} = \mathbb{P}(\tau_i > t_j | \mathcal{F}_t) = \mathbb{E}\Big(\exp\Big(-\int_t^{t_j} \Lambda(Y_s, D_{i,s})ds\Big)\,\Big|\,\mathcal{F}_t\Big) \tag{5}$$

as the probability of survival from t to t_j given $\mathcal{F}_t$ on $\{\tau_i > t\}$. Finally, set

$$d_j = \exp(-r(t_j - t))$$

as the time discount factor from time t to time t_j. Then,

$$\begin{array}{l} B_{i,t} = \sum_{j=1}^{N} d_j (p_{i,j-1,t} - p_{i,j,t})\mathbb{E}(L_i) \\ A_{i,t} = \sum_{j=1}^{N} d_j p_{i,j,t} K. \end{array} \tag{6}$$

To resume, for any i, the market value $V_{i,t}$ at time t of a credit default swap on name i at rate K is given by

$$V_{i,t} = L_i \, \mathbf{1}_{\{\tau_i \leq t\}} + (B_{i,t} - A_{i,t}) \, \mathbf{1}_{\{\tau_i > t\}},$$

where $B_{i,t}$ and $A_{i,t}$ are defined in (6).

Remark 2.2 In principle, the loss L_i can be modelled in many different ways. Here, L_i is a random variable distributed between zero and one according to a specified law, which is the uniform one. It has to be said that more elaborate laws could be used in practice, but we should benchmark a simple case. It has been discussed allowing L_i to be correlated with other variables but, for the sake of simplicity, we assume here that L_i is independent of all else.

3 The mixed PDE-Monte Carlo numerical approach

Consider a derivative whose price P_0 is given by (4). To numerically compute P_0, consider a Monte Carlo approach, that is

$$P_0 = e^{-rt}\mathbb{E}\Big(\max\Big(\sum_{i=1}^{n} V_{i,t}, 0\Big)\Big) \simeq \frac{e^{-rt}}{M}\sum_{m=1}^{M}\max\Big(\sum_{i=1}^{n} V_{i,t}^{(m)}, 0\Big)$$

where the index m stands for one of the total M trials and $V_{i,t}^{(m)}$ denotes the mth simulation for $V_{i,t}$. Then, the problem is how to simulate

$$V_{i,t} = L_i\,\mathbf{1}_{\{\tau_i \le t\}} + (B_{i,t} - A_{i,t})\,\mathbf{1}_{\{\tau_i > t\}}. \tag{7}$$

Obviously, the r.v. $V_{i,t}$ cannot be exactly replicated, so we have to discuss firstly an approximation $\bar{V}_{i,t}$ of $V_{i,t}$.

3.1 Checking if default does or does not occur

First of all, we have to check if the default of the ith firm does or does not occur up to time t. In order to do this, we have to use the default probabilities, which of course depend on the paths followed by the process Y and D_i on $[0,t]$. Therefore, one has first to discretize the time interval $[0,t]$ in order to set up an Euler scheme for the two processes and consequently to compute the default probabilities. To be more precise, set $0 = s_0 < s_1 < \cdots < s_\ell = t$ a discretization of $[0,t]$ such that $s_{k+1} - s_k = \Delta s$ is "small enough". Let us define $\bar{Y}$ and $\bar{D}_{i,t}$ the (discrete) Euler schemes for Y and D_i respectively:

$$\begin{aligned}
&\bar{Y}_0 = Y_0,\ \bar{D}_{i,0} = D_{i,0} \ \text{ and for } k = 1,\dots,\ell:\\
&\bar{Y}_k = \bar{Y}_{k-1} + \kappa_Y(\theta_Y - \bar{Y}_{k-1})\Delta s + \sigma_Y \Delta W_k^Y\\
&\bar{D}_{i,k} = \bar{D}_{i,k-1} + \kappa_D(\theta_{D,i} - D_{i,k-1})\Delta s + \sigma_D(\rho\Delta W_k^c + \sqrt{1-\rho^2}\Delta W_k^i).
\end{aligned}$$

Here, $\Delta W_k^Y, \Delta W_k^c$ and ΔW_k^i denote the variation between s_{k-1} and s_k of the Brownian motions W^Y, W^c and W^i respectively. Let us stress that, in our notations, one has $\bar{Y}_k \simeq Y_{s_k}$ and $\bar{D}_{i,k} \simeq D_{i,s_k}$. Once the above schemes are set, we can approximate the event "the default has occurred up to time t" as follows. First, let us observe that the probability of no default of the firm i up to s_k given that the firm is still alive at time s_{k-1} can be approximated by using the approximation $(\bar{Y}, \bar{D}_i)$ as follows: on the set $\{\tau_i > s_{k-1}\}$ one has

$$\mathbb{P}(\tau_i > s_k \,|\, \mathcal{F}_{s_{k-1}}) = \mathbb{E}\Big(e^{-\int_{s_{k-1}}^{s_k}\Lambda(Y_s, D_{i,s})ds}\,\Big|\,\mathcal{F}_{s_{k-1}}\Big) \simeq e^{-\Lambda(\bar{Y}_{k-1}, \bar{D}_{i,k-1})\Delta s}.$$

Then, if we set

$$\bar{q}_{i,k} = e^{-\Lambda(\bar{Y}_{k-1}, \bar{D}_{i,k-1})\Delta s},$$

- one first computes $\bar{q}_{i,0}$ and: with probability $\bar{q}_{i,0}$ one says that no default occurred up to s_1; with probability $1 - \bar{q}_{i,0}$ one says that default did occur up to s_1;

- if the default did not occurred up to s_{k-1}, then one computes $\bar{q}_{i,k}$ and: with probability $\bar{q}_{i,k}$ one says that no default occurred up to s_k; with probability $1-\bar{q}_{i,k}$ one says that default did occur up to s_k.

Let us recall that in practice this means to generate r.v.'s $Z_{i,k}$, uniformly distributed on $(0,1)$, independent each other and of any other random variable written above, and to check if $Z_{i,k}$ is or is not less than $\bar{q}_{i,k}$: if $Z_{i,k} < \bar{q}_{i,k}$ then no default is assumed to occur; if $Z_{i,k} > \bar{q}_{i,k}$ then the default is assumed to occur.
This procedure brings to an approximation $\bar{\mathbf{1}}_i$ of the random variable $\mathbf{1}_{\{\tau_i > t\}}$: we set $\bar{\mathbf{1}}_i = 1$ if no default has been observed up to $s_\ell = t$, otherwise $\bar{\mathbf{1}}_i = 0$. Once $\bar{\mathbf{1}}_i$ is obtained, we set:

- $V_{i,t} \simeq \bar{V}_{i,t} = L_i$ if $\bar{\mathbf{1}}_i = 0$;
- $V_{i,t} \simeq \bar{V}_{i,t} = \overline{(B-A)}_{i,t}$ if $\bar{\mathbf{1}}_i = 1$,

where $\overline{(B-A)}_{i,t}$ is a suitable approximation for $B_{i,t} - A_{i,t}$, to be discussed in Section 3.2 and 3.3.

3.2 Standard Monte Carlo approach if default does not occur

In this Section, we suppose to have observed no default up to t, that is to be on the set $\{\tau_i > t\}$. Then, we have to approximate $B_{i,t} - A_{i,t}$. By formula (6), we can rewrite

$$\begin{array}{rcl} B_{i,t} - A_{i,t} = & \sum_{j=1}^N d_j (p_{i,j-1,t} - p_{i,j,t}) \mathbb{E}(L_i) - \sum_{j=1}^N d_j p_{i,j,t} K \\ = & \sum_{j=1}^N c_{i,j} p_{i,j,t} + c_{i,0} \end{array} \tag{8}$$

where

$$c_{i,j} = \begin{cases} \mathbb{E}(L_i) d_1 & \text{if } j = 0 \\ -(K + \mathbb{E}(L_i)) d_j + \mathbb{E}(L_i) d_{j+1} & \text{if } 1 \le j \le N-1 \\ -(K + \mathbb{E}(L_i)) d_N & \text{if } j = N \end{cases} \tag{9}$$

(we have used above the notation $p_{i,0,t} = 1$).
This means that $B_{i,t} - A_{i,t}$ is a linear combination of the functions

$$p_{i,j,t} = \mathbb{E}\Big(\exp\Big(-\int_t^{t_j} \Lambda(Y_s, D_{i,s}) ds \Big) \,\Big|\, \mathcal{F}_t \Big).$$

Now, since the pair $(X_{i,s})_s = (Y_s, D_{i,s})_s$ is a diffusion, then the Markov property holds, so that

$$p_{i,j,t} = P_{i,j}(t, X_{i,t}), \quad \text{where } P_{i,j}(t,x) = \mathbb{E}_{t,x}\Big(\exp\Big(-\int_t^{t_j} \Lambda(X_{i,s}) ds \Big)\Big), \tag{10}$$

where as usual $\mathbb{E}_{t,x}$ denotes expectation among the paths starting at x at time t. Therefore, $p_{i,j,t}$ has a representation in terms of an expectation, so that one can think to compute it by the standard Monte Carlo method: setting $\bar{X}_{i,t} = (\bar{Y}_t, \bar{D}_{i,t})$ the approximation already observed for $X_{i,t}$, then

$$p_{i,j,t} \simeq P_{i,j}(t, \bar{X}_{i,t}) \simeq \frac{1}{M'} \sum_{m'=1}^{M'} \exp\Big(-\int_t^{t_j} \Lambda(\bar{X}_{i,s}^{(m')}) ds \Big) \tag{11}$$

where $(\bar{X}_{i,s}^{(m')})_{s\geq t}$ denotes a simulation of an approximation $(\bar{X}_{i,s})_{s\geq t}$ (e.g. by Euler scheme) for $(X_{i,s})_{s\geq t}$ on the time interval $[t,T]$ starting from $\bar{X}_{i,t}$ at time t (and obviously, the integral is approximated by Riemann sums).
Therefore, summing up, we have for any firm i:

- $m = 1,\dots,M$ simulations of the path on $[0,t]$, telling us if default does or does not occur, and if not, giving us a position, say $\bar{X}_{i,t}^{(m)}$, of the pair (Y, D_i) at time t;
- for any $m = 1,\dots,M$ giving no default, we have to start with M' more simulations of the pair $X_i = (Y, D_i)$ on $[t,T]$ starting at $\bar{X}_{i,t}^{(m)}$, in order to approximate the probabilities as in (11).

This can give rise to $M \times M'$ simulations, for any $i = 1,\dots,n$, so the total number can arrive to $M \times M' \times n$. In other words, an algorithm of this kind is too expensive from a computational point of view. One could think to use some variance reduction techniques, for example a control variate. But this approach would not overcome the problem of the further M' simulations, leading us to look for an alternative method.

3.3 PDE approach if default does not occur

The idea starts from the fact that deterministic methods to approximate expectations by numerical solutions of PDEs, work quite well if the dimension is small.
Let us come back to $B_{i,t} - A_{i,t}$ and let us rewrite it as

$$B_{i,t} - A_{i,t} = \varphi_i(t, X_{i,t}) \tag{12}$$

where (see (8), (9) and (10))

$$\varphi_i(t,x) = \sum_{j=1}^{N} c_{i,j} P_{i,j}(t,x) + c_{i,0}. \tag{13}$$

Now, we propose here to numerically compute $\varphi_i(t, X_{i,t})$ as the solution of a parabolic PDE on $\mathbb{R}^2$, which in turn is evaluated by finite difference methods. To this purpose, using the Feynman-Kac formula one can first compute $P_{i,j}(t, X_{i,t})$ as follows:

Proposition 3.1 *One has $P_{i,j}(t, X_{i,t}) = v_{i,j}(t, X'_{i,t})$, where $X'_{i,t} = (Y'_t, D'_{i,t})$ is the 2-dimensional diffusion solving the SDE*

$$\begin{aligned} dY'_t &= -\kappa_Y Y'_t dt + \sigma_Y dW_t^Y, \quad Y'_0 = Y_0 \\ dD'_{i,t} &= -\kappa_D D'_{i,t} dt + \sigma_D(\rho dW_t^c + \sqrt{1-\rho^2} dW_t^i), \quad D'_{i,0} = D_{i,0}. \end{aligned} \tag{14}$$

and $v_{i,j}$ solves

$$\begin{aligned} &\partial_t v_{i,j}(s,x) + \mathcal{L} v_{i,j}(s,x) - g_i(s,x) v_{i,j}(s,x) = 0 \quad (s,x) \in (0,t_j) \times \mathbb{R}^2 \\ &v_{i,j}(t_j, x) = 1 \end{aligned} \tag{15}$$

being (recall that $x = (y, d)$)

$$\begin{aligned}
&\mathcal{L} = \tfrac{1}{2}\sigma_Y^2\partial_{yy}^2 + \tfrac{1}{2}\sigma_D^2\partial_{dd}^2 - \kappa_Y y\partial_y - \kappa_D d\partial_d,\\
&g_i(s,x) = q_i(s)\exp(\mu_1 y + \mu_2 d), \quad \textit{with}\\
&q_i(s) = \exp\Big(\mu_0 + \mu_1\theta_Y(1-e^{-s\kappa_Y}) + \mu_2\theta_{D,i}(1-e^{-s\kappa_D})\Big)
\end{aligned}$$

Remark 3.2 There exists a very simple relation connecting the processes $X_i = (Y, D_i)$ and $X_i' = (Y', D_i')$: for any s,

$$\begin{aligned}
&Y_s' = Y_s - \theta_Y(1-e^{-\kappa_Y s}) \quad \textit{and}\\
&D_{i,s}' = D_{i,s} - \theta_{D,i}(1-e^{-\kappa_D s}).
\end{aligned}$$

The proof is straightforward (but see also the next proof).

Proof of Proposition 3.1. It is well known that the pair $(X_{i,s})_s = ((Y_s, D_{i,s}))_s$ has an explicit expression given by

$$\begin{aligned}
Y_s = \;& Y_0\, e^{-\kappa_Y s} + \theta_Y(1-e^{-\kappa_Y s}) + \sigma_Y\textstyle\int_0^s e^{-\kappa_Y(u-s)}\, dW_u^Y\\
D_{i,s} = \;& D_{i,0}\, e^{-\kappa_D s} + \theta_{D,i}(1-e^{-\kappa_D s}) +\\
&+\sigma_D\textstyle\int_0^s e^{-\kappa_D(u-s)}\, d(\rho W_u^c + \sqrt{1-\rho^2}\, dW_u^i)
\end{aligned}$$

Then, setting $Y_s' = Y_s - \theta_Y(1-e^{-\kappa_Y s})$ and $D_{i,s}' = D_{i,s} - \theta_{D,i}(1-e^{-\kappa_D s})$, it is straightforward to see that the pair $(X_{i,s}')_s = ((Y_s', D_{i,s}'))_s$ solves (14). Now, for any s,

$$\begin{aligned}
&\Lambda(Y_s, D_{i,s}) = \exp\Big(\mu_0 + \mu_1 Y_s + \mu_2 D_{i,s}\Big) = \exp\Big(\mu_1 Y_s' + \mu_2 D_{i,s}'\Big)\cdot q_i(s)\\
&\text{where} \quad q_i(s) = \exp\Big(\mu_0 + \mu_1\theta_Y(1-e^{-s\kappa_Y}) + \mu_2\theta_{D,i}(1-e^{-s\kappa_D})\Big)
\end{aligned}$$

Then, for $x = (y, d)$, if $g_i(s,x) = q_i(s)\exp(\mu_1 y + \mu_2 d)$, one has $\Lambda(X_{i,s}) = g_i(s, X_{i,s}')$ and therefore,

$$\mathbb{E}_{t,X_{i,t}}\Big(\exp\Big(-\int_t^{t_j}\Lambda(X_{i,s})ds\Big)\Big) = \mathbb{E}_{t,X_{i,t}'}\Big(\exp\Big(-\int_t^{t_j} g_i(s, X_{i,s}')ds\Big)\Big).$$

In other words, we have $P_{i,j}(t, X_{i,t}) = v_{i,j}(t, X_{i,t}')$, where we have put

$$v_{i,j}(t,x) = \mathbb{E}_{t,x}\Big(\exp\Big(-\int_t^{t_j} g_i(s, X_{i,s}')ds\Big).$$

Now, by the Feynman-Kac formula, it immediately follows that $v_{i,j}$ solves the PDE problem (15). □

Using Proposition 3.1, one can represent the function φ_i in (12) and (13) as the solution of a PDE problem, as follows.

Proposition 3.3 *Consider the following backward set of functions $\psi_{i,N}(s,x)$, ..., $\psi_{i,1}(s,x)$ (recall that $\mathcal{L}$ and g_i are defined in Proposition 3.1):*

- $\psi_{i,N}(s,x)$ *solves:*

$$\partial_t \psi_{i,N}(s,x) + \mathcal{L}\psi_{i,N}(s,x) - g_i(s,x)\psi_{i,N}(s,x) = 0 \quad (s,x) \in (0,t_N) \times \mathbb{R}^2$$
$$\psi_{i,N}(t_N,x) = c_{i,N};$$

- *as* $j = N-1,\dots,1$, $\psi_{i,j}(s,x)$ *solves:*

$$\partial_t \psi_{i,j}(s,x) + \mathcal{L}\psi_{i,j}(s,x) - g_i(s,x)\psi_{i,j}(s,x) = 0 \quad (s,x) \in (0,t_j) \times \mathbb{R}^2$$
$$\psi_{i,j}(t_j,x) = c_{i,j} + \psi_{i,j+1}(t_j,x)$$

Then, $\varphi_i(t,X_{i,t}) = \psi_{i,1}(t,X'_{i,t}) + \mathbb{E}(L_i)d_1$, *where* X' *is defined through (14).*

Proof. Consider first $\psi_{i,N}$: it solves the same PDE as in (15) with $j = N$, with constant (final) Cauchy condition equal to $c_{i,N}$. Then, it holds

$$\psi_{i,N}(s,x) = c_{i,N} v_{i,N}(s,x), \quad s \le t_N.$$

Now, for $j = N-1,\dots,1$, again $\psi_{i,j}$ solves the same PDE as in (15), with (final) Cauchy condition equal to $c_{i,j} + \psi_{i,j+1}(t_j,x)$. It then follows that

$$\psi_{i,j}(s,x) = c_{i,j} v_{i,j}(s,x) + \psi_{i,j+1}(s,x), \quad s \le t_j.$$

Therefore, by iteration,

$$\psi_{i,1}(s,x) = \sum_{j=1}^{N} c_{i,j} v_{i,j}(s,x), \quad s \le t_1,$$

so that as $s = t (< t_1)$ one has

$$\psi_{i,1}(t,X'_{i,t}) = \sum_{j=1}^{N} c_{i,j} v_{i,j}(t,X'_{i,t}) = \sum_{j=1}^{N} c_{i,j} P_{i,j}(t,X_{i,t}),$$

the last equality being proved in Proposition 3.1. Therefore, by (13), one has

$$\varphi_i(t,X_t) = \psi_{i,1}(t,X'_{i,t}) + c_{i,0} = \psi_{i,1}(t,X'_{i,t}) + \mathbb{E}(L_i)d_1,$$

and the statement holds. □

Remark 3.4 Let us observe that one could also avoid considering the new process X'_i and use directly the original process X_i. In such a case, one would obtain quite the same results, but of course the operator $\partial_t + \mathcal{L} - g_i$ has to be replaced by $\partial_t + \mathcal{L}_i - \Lambda$. It has to be stressed that the advantage of our representation is that the differential operator $\mathcal{L}$ *is the same for any* i and only g_i changes as i varies. Now, this is important from a computational point of view since this allows to reduce the number of numerical operations (e.g. in an explicit finite differences scheme).

Now, the result in Proposition 3.3 is used in practice in the following way. We first compute an approximation $\bar{\varphi}_i$ of φ_i by using deterministic numerical methods for solutions of PDEs (here, for a fixed (firm) i, the associated PDE is on $\mathbb{R}^2$, so deterministic methods efficiently work). Therefore, we can finally set the approximation

$$B_{i,t} - A_{i,t} \simeq \overline{(B-A)}_{i,t} = \bar{\varphi}_i(t, \bar{X}_{i,t}).$$

The algorithm can be resumed as:

a) fix the firm i and the mth simulation ($i = 1, \ldots, n$ and $m = 1, \ldots, M$);

b) following Section 3.1, check if default does or does not occur on $[0, t]$, and set the mth simulation $\bar{V}_{i,t}^{(m)}$ of $V_{i,t}$ as:

   b1) if default is observed, then $\bar{V}_{i,t}^{(m)} = L_i$;

   b2) if no default is observed, then $\bar{V}_{i,t}^{(m)} = \bar{\varphi}_i(t, \bar{X}_{i,t}^{(m)})$

c) compute the price by averaging:

$$P_0 \simeq \bar{P}_0 = \frac{e^{-rt}}{M} \sum_{m=1}^{M} \max\Big(\sum_{i=1}^{n} \bar{V}_{i,t}^{(m)}, 0 \Big)$$

It is worth to stress some remarks concerning item b2) above.
We numerically solve the PDEs in Proposition 3.3 for each name i with a finite difference method. The numerical procedure consists in the following five steps.

1. The set of linear parabolic problems in Proposition 3.3 have to be localized to a bounded domain in space, that is $\mathbb{R}^2$ is in practice replaced by $\Omega_l = (y'_{\min}, y'_{\max}) \times (d'_{\min}, d'_{\max})$. The choice of Ω_l must answer two main purposes. First, Ω_l must be large enough to ensure the convergence of the approximating function to the true one. Second, the values of $(\bar{Y}_t'^{(m)}, \bar{D}_{i,t}'^{(m)})_m$ (obtained from $(\bar{Y}_t^{(m)}, \bar{D}_{i,t}^{(m)})_m$ by using the transformation as in Remark 3.2) have to be contained in Ω_l. Then, for each name i, we first simulate all the paths and secondly construct the grid. Afterwards, we solve the PDEs with a finite difference method.

2. The approximating solutions $\bar{\psi}_{i,N}, \ldots, \bar{\psi}_{i,1}$ are computed by means of finite difference methods involving discrete functions. For our numerical studies, we use an explicit finite difference method (see Wilmott, Dewynne and Howison [10], Villeneuve and Zanette [11]) and we suitably add, as usually done, Neumann homogenous artificial boundary conditions.

3. We construct for this purpose a time-space grid. Each time interval $(t_{j-1}, t_j]$ is split in subintervals and we consider a mesh of the space domain $[y'_{\min}, y'_{\max}] \times [d'_{\min}, d'_{\max}]$ consisting of subintervals (y'_k, y'_{k+1}) and (d'_k, d'_{k+1}), where $y'_{\min} = y'_0 < \ldots < y'_{M_Y} = y'_{\max}$ and $d'_{\min} = d'_0 < \ldots < d'_{M_D} = d'_{\max}$.

4. Since we need $\bar{\varphi}_i$ at time t, we solve backwardly the PDEs in Proposition 3.3 up to t, so we will have (see Remark 3.2)

$$\bar{\varphi}_i(t, y_{j_1}, d_{j_2}) = \bar{\psi}_{i,1}(t, y'_{j_1}, d'_{j_2}) + \mathbb{E}(L_i)d_1, \quad \text{with}$$
$$y'_{j_1} = y_{j_1} - \theta_Y(1 - e^{-\kappa_Y t}) \quad \text{and} \quad d'_{j_2} = d_{j_2} - \theta_{D,i}(1 - e^{-\kappa_D t})$$

5. Now, in order to estimate $\bar{\varphi}_i(t, \bar{Y}_t^{(m)}, \bar{D}_{i,t}^{(m)})$, we look for the four points of the space grid which are neighborhood points for $(\bar{Y}_t'^{(m)}, \bar{D}_{i,t}'^{(m)})$ and then we approximate $\bar{\varphi}_i(t, \bar{Y}_t^{(m)}, \bar{D}_{i,t}^{(m)})$ using bilinear interpolation of $\bar{\psi}_{i,1}$ in these points.

Remark 3.5 *Let us recall that the use of PDEs techniques have been done in order to overcome the problem of the further nested M' simulations for each firm i. This problem could be handled also by using regression in place of PDEs, as it has been done for numerically pricing American options (see Carriere [1] and Longstaff and Schwartz [8]), a topic which is worthy to be investigated. However, the use of regression for approximating conditional expectations is a very good idea for American option pricing purposes, where the time interval is split into small time subintervals. In fact, this ensures that the conditional expectations is done following the law of a process in a small time interval, and this makes the regression approximation work very well. Our context is actually different (see e.g. (10)), so that a method involving regression is not trivial and should be analyzed in detail.*

4 Numerical results

In this Section we numerically illustrate the efficiency of the PDE-Monte Carlo approach introduced in Section 3.3 and we compare, for benchmark purposes, this approach with the vanilla Monte Carlo approach of Section 3.2. All the computations have been performed in double precision on a PC Pentium 1.73 GHz.
We treat the problem of pricing CDS index swaptions in the model introduced in Section 2 and we consider the following values of the parameters :

- varying number of firms: $n = 50, 100, 125$;
- option parameters: times $t = 1$, $T = 6$ then $t_j = t + 0.25\,j$, $j = 1, \ldots, 20$; L_i uniformly distributed on $(0, 1)$; interest rate $r = 0.05$; rate $K = 0.001$;
- default probability parameters: $\mu_0 = -4.2017$, $\mu_1 = -0.4597$, $\mu_2 = -0.4411$

Concerning the diffusion parameters, we consider two test cases, named A and B :

- $\boxed{\text{case A:}}$ $\kappa_Y = 0.6524$, $\theta_Y = 1.8901$, $\sigma_Y = 0.8888$, $Y_0 = 3$; $\kappa_D = 0.1185$, $\theta_{D,i} = 4.32 + i/n$, $\sigma_D = 0.9657$, $D_{i,0} = 3$; correlation coefficient $\rho = 0.2684$;
- $\boxed{\text{case B:}}$ $\kappa_Y = 0.1137$, $\theta_Y = 0.1076$, $\sigma_Y = 1$, $Y_0 = 3$; $\kappa_D = 0.0355$, $\theta_{D,i} = 4.32 + i/n$, $\sigma_D = 0.346$, $D_{i,0} = 3$; correlation coefficient $\rho = 0.2684$;

The choice of these parameters was inspired by a similar combination in Duffie and Wang [3]. We are going to compare the results obtained by the mixed PDE-Monte Carlo approach and the (standard) Monte Carlo algorithm, in which the following settings are taken into account.

(i) As for the mixed PDE-Monte Carlo approach, we take the number of simulation $M = 10.000$, the number of discretization step $\ell = 32$ in which the time interval $[0, t]$ is split. The PDEs are numerically solved through an explicit finite difference method, with time discretization step set equal to 40, while the spatial discretization step is obtained via the stability condition. We add Neumann homogenous artificial boundary conditions.

(ii) As for the Monte Carlo algorithm, we take a number of simulations M of the first part equal to the number of simulations M' of the second part, both equal to 1000 or 10000; moreover, we split the time interval $[0, t]$ in $\ell = 32$ or 64 subintervals.

The results for the price and the estimated standard deviation, divided by the square root of the number of simulations, for case A and B are presented in next tables, in *Price* and *St.Dev.*/$\sqrt{M}$ column respectively, according to the mixed PDE-Monte Carlo method summarized in (i) (in tables, PDE-MC) or the vanilla Monte Carlo method as in (ii) (in table, MC 1000×32 and MC 10000×64). Finally, reports are given in Table 1, 2 and 3, for a number n of firms equal to 50, 100 and 125 respectively.

CASE		A		B
	Price	*St.Dev.*/$\sqrt{M}$	*Price*	*St.Dev.*/$\sqrt{M}$
PDE-MC	0.038	0.002	0.094	0.003
MC 1000×32	0.036	0.005	0.082	0.010
MC 10000×64	0.036	0.001	0.095	0.003

Table 1: Number of firms $n = 50$: price and standard deviation for the mixed PDE-Monte Carlo method (PDE-MC) and the full Monte Carlo one (MC).

CASE		A		B
	Price	*St.Dev.*/$\sqrt{M}$	*Price*	*St.Dev.*/$\sqrt{M}$
PDE-MC	0.029	0.001	0.145	0.004
MC 1000×32	0.032	0.006	0.144	0.010
MC 10000×64	0.024	0.001	0.141	0.005

Table 2: Number of firms $n = 100$: price and standard deviation for the mixed PDE-Monte Carlo method (PDE-MC) and the full Monte Carlo one (MC).

CASE		A		B
	Price	*St.Dev.*/$\sqrt{M}$	*Price*	*St.Dev.*/$\sqrt{M}$
PDE-MC	0.024	0.001	0.166	0.006
MC 1000×32	0.021	0.006	0.162	0.020
MC 10000×64	0.021	0.001	0.170	0.007

Table 3: Number of firms $n = 125$: price and standard deviation for the mixed PDE-Monte Carlo method (PDE-MC) and the full Monte Carlo one (MC).

Let us take the price turning out by the MC 10000×64 as a benchmark value. In such a case, we observe two different performances for our PDE-MC method. In fact, case B works very well, because the relative error is between 1% and 2.8%. Despite this, case A works poorer, giving an error higher than the one arising from case B. Nevertheless, it has to be remarked that in case A prices are much smaller, and it is well known that low values might be responsible of high relative errors.
Let us spend some final words about the computational cost in time of the PDE-MC algorithm. The time computation results for $n = 50, 100, 125$ are presented in next Table 4.

number of firms	$n = 50$	$n = 100$	$n = 125$
PDE-MC CPU time	10 sec.	19 sec.	23 sec.

Table 4: Time computation for the mixed PDE-Monte Carlo method (PDE-MC).

Obviously, comparisons with the time needed by the full Monte Carlo algorithms are not significant (just to give an idea, the range is in the between of some hours and some days, and in fact the use of the vanilla Monte Carlo approach has been done in order to benchmark our results). However, Table 4 shows very low CPU times spent by our algorithm. We can then conclude that the proposed mixed PDE-Monte Carlo method can be used to provide a fast and reasonably robust pricing of CDS index swaptions.

5 Conclusions

In this paper, we have investigated a new numerical method, a mixed PDE-Monte Carlo algorithm, to approximate the price of credit default index swaptions on a large number of names in presence of stochastic intensity.
The straightforward application of a pure Monte Carlo algorithm is too expensive from a computational of view. On the contrary, the main appeal of the proposed methodology is that it allows to obtain price results for this credit derivatives product with a cheap computational time cost. The numerical results confirm also the reliability of the method.
Based on these results, the mixed PDE-Monte Carlo algorithm presented in this work may be considered to tackle our credit derivative pricing problems in a dynamic setting.

References

[1] J.Carriere: Valuing of the early-exercise price for options using simulations and non parametric regression. *Insurance:Mathematcs and Economics*, **14**, 19-30, 2001.

[2] D. Duffie: Credit Swap Valuation. *Financial Analusts Journal*, January-February ,73-87, 1999

[3] D. Duffie and K. Wang: Multi-Period Corporate Failure Prediction with Stochastic Covariates. Preprint, 2004.

[4] D. Duffie, L. Saita and K. Wang: Multi-Period Corporate Failure Prediction with Stochastic Covariates. Preprint available at `www.standford.edu/∼duffie/ddkw.pdf` (last check: May 3, 2006).

[5] D. Duffie and K. Singleton: Modelling Term Structures of Defaultable Bonds. *Review of Financial Studies*, **12**, 687-720, 1999.

[6] A. Jackson: A New Method for Pricing Index Default Swaptions. Working paper, Citigroup, 2005.

[7] D. Lando: On Cox Proceses and Credit-Risky Securities. *Review of Derivatives Research*, **2**, 99-120, 1999.

[8] F. Longstaff and E. Schartz: Valuing American Options by Simulations: a Simple Least-Squares Approach. *Reviews of Financial Studies*, **14**, 19-30, 2001.

[9] P.J. Schönbucher: *Credit Derivatives Pricing Models: Models, Pricing, Implementation.* Wiley Finance, 2003.

[10] P. Wilmott, J. Dewynne and S. Howison: *Option Pricing: Mathematical Models and Computation* Oxford Financial Press, 1993.

[11] S. Villeneuve and A. Zanette: Comparison of Finite Difference Methods for Pricing American Options on Two Stocks. Working paper of the Dipartimento di Finanza dell'Impresa e dei Mercati Finanziari, Università di Udine, 1-2001